



UNIVERSITÉ DE NANTES
ÉCOLE DOCTORALE SPI
SCIENCES POUR L'INGÉNIEUR

HABILITATION À DIRIGER DES RECHERCHES

Présentée et soutenue par
Florent AUTRUSSEAU

Perception des images numériques, de la sécurité des contenus à l'analyse des images médicales

Jury :

Rapporteurs :

Christophe CHARRIER - Maître de Conférences (HDR), GREYC, Caen

Olivier DÉFORGES - Professeur des Universités, IETR, Rennes

William PUECH - Professeur des Universités, LIRMM, Montpellier

Cathy CASTELAIN - Directrice de Recherche CNRS, LTeN, Nantes

Yves AMOURIQ - Professeur des Universités (PU/PH), RMeS, Nantes

JeanPierre GUÉDON - Professeur des Universités, LS2N, Nantes

Remerciements

Je tiens tout d'abord à remercier chaleureusement tous ceux qui m'ont soutenu au cours de ces nombreuses années dans le monde de la recherche, tous ceux qui m'ont aidé à fournir les travaux présentés dans ce mémoire.

Merci donc à tous mes co-auteurs, aux doctorants, post-docs, ingénieurs, stagiaires, mais aussi à tous les collègues, Français ou étrangers avec qui j'ai pu travailler sur divers projets de recherche toujours captivants.

Un grand merci aussi, bien sur, à tous les membres des différentes équipes que j'ai eu la chance de pouvoir intégrer, les membres d'IVC, du RMeS et aussi du LTeN.

Bien évidemment, je remercie aussi grandement tous les membres de mon jury, qui m'ont fait l'honneur d'accepter de m'accompagner pour cet ultime diplôme.

Finalement, je tiens à remercier du fond du coeur toute ma famille et en particulier Emmanuelle, Alexis et Coralie. Merci d'être là, tout simplement !

Table des matières

Liste d'abréviations	vii
1 Curriculum Vitae	1
1.1 Expérience Professionnelle	2
1.2 Formation	2
1.3 Projets de Recherche	3
1.4 Collaborations Académiques et Industrielles	3
1.5 Activités scientifiques diverses	4
1.6 Diffusion et dissémination	5
1.7 Activités d'Encadrement:	6
1.8 Liste des publications	10
2 Introduction	21
2.1 Positionnement de mon travail	21
2.1.1 Les thématiques de recherche	21
2.1.2 Résumé succinct des contributions	24
2.2 Contributions personnelles	25
2.2.1 La perception visuelle appliquée à l'image	25
2.2.2 La perception des couleurs	27
2.2.3 La transformation Mojette	28
2.2.4 L'imagerie médicale et biologique	29
3 La perception visuelle	33
3.1 La dissimulation de l'information	33
3.1.1 Application à la protection des droits d'auteurs des images numériques	33
3.1.2 Protection des vidéos numériques	41
3.2 L'évaluation de la qualité	44
3.2.1 Métriques de qualité pour le tatouage	45
3.2.2 La reconnaissance des images	51
4 La perception des couleurs	59
4.1 Adaptation temporelle des champs récepteurs +s/-s	59
4.2 Modélisation optique de l'oeil	62
4.3 La variabilité des observateurs	66
5 La transformation Mojette	71
5.1 Transformation Mojette	71
5.1.1 Transformation Directe	71
5.1.2 Inversion de la Transformation Mojette	72

5.2	Compression Mojette	76
5.3	Sécurité basée Mojette	83
5.3.1	Applications au tatouage	84
5.3.2	Applications au chiffrement	86
6	Imagerie Médicale et Biologique	91
6.1	Détection des dents	91
6.1.1	Approche “Computer Vision”	92
6.1.2	Approche “Deep Learning”	93
6.2	Détection de l’ostéoporose	95
6.2.1	Caractérisation osseuse	96
6.2.2	Reconnaissance du foramen mentonnier	97
6.3	Caractérisation de l’arbre vasculaire cérébral	98
6.3.1	Caractérisation des bifurcations	99
6.3.2	Estimation du risque de formation des anévrismes	102
7	Collaborations et projets collaboratifs	105
7.1	Implication dans des Projets de recherche	105
7.2	Collaborations académiques et industrielles	111
7.2.1	Collaborations nationales	111
7.2.2	Collaborations internationales	112
7.3	Transfert de Technologie	113
8	Conclusions et perspectives de travaux	115
8.1	La Perception visuelle	115
8.1.1	Principales contributions	116
8.1.2	Perspectives	117
8.2	La Mojette	117
8.2.1	Principales contributions	118
8.2.2	Perspectives	118
8.3	L’imagerie médicale et biologique	118
8.3.1	Principales contributions	119
8.3.2	Perspectives	119
8.4	Perspectives générales	119
9	Sélection d’articles	121
9.1	Chromatic and wavefront aberrations: L-, M- and S-cone stimulation with typical and extreme retinal image quality	122
9.2	Lossless image compression via predictive coding of discrete Radon projections	136
9.3	Perceptual DFT watermarking with improved detection and robust- ness to geometrical distortions	149

9.4 Characterization of 3D Bifurcations in Micro-scan and MRA-TOF Images of Cerebral Vasculature for Prediction of Intra-Cranial Aneu- rysms	162
Bibliographie	173

Liste d'abréviations

2AFC Two Alternative Forced Choice	52
ACP Analyse en Composante Principale	102
ACR Absolute Category Rating	52
AIC Anévrismes Intra Crânien	30
APP Agence de Protection des Programmes	113
CBCT “Cone Beam Computed Tomography”	24
CIE Commission Internationale de l’Éclairage	28
CMF Color Matching Functions	66
CPA Carosi, Pankajakshan & Atrusseau	44
CSF Contrast Sensitivity Function	34
DCP Décomposition en Canaux Perceptuels	35
DICOM “Digital imaging and communications in medicine”	23
DSCQS Double Stimulus Continuous Quality Scale	52
DSIS Double Stimulus Impairment Scale	48
DWT Discrete Wavelet Transform	26
DXA “Dual-energy X-ray Absorptiometry”	30
FBP “Filtered Back-Projection”	75
GMM Gaussian Mixture Model	93
HDR High Dynamic Range	26
IRCCyN Institut de Recherche en Communication et Cybernétique de Nantes	22
IRM Imagerie par Résonance Magnétique	24
ITU International Telecommunication Union	51
IVC Image et Vidéo-Communication	22
JND “Just Noticeable Difference”	34
LCA Longitudinal Chromatic Aberrations	28
LIOAD Laboratoire d’Ingénierie Ostéo-Articulaire et Dentaire	23
LS2N Laboratoire des Sciences du Numérique de Nantes	22
LTeN Laboratoire de Thermique et Énergie de Nantes	120
MOS Mean Opinion Score	47
MOSp Mean Opinion Score (predicted)	47
MTF Modulation Transfer Function	63

MVD Motion Vector Difference	43
NCC Normalized Cross-Correlation	55
PHRC Projet Hospitalier de Recherche Clinique	105
PSF Point Spread Function	61
PSNR Peak Signal to Noise Ratio	26
REGOS Regenerative Medicine of Bone Tissues	91
RMeS Regenerative Medicine and Skeleton	23
RMSE Root Mean Square Error	47
SC3M Service Commun de Microscopie électronique, Microcaractérisation et Morpho-histologie-imagerie fonctionnelle	23
SIBM Signaux et Images en Biologie et Médecine	23
SIFT Scale-Invariant Feature Transform	55
SSIM Structural Similarity Index	26
STEP Skeletal Physiopathology and Joint Regenerative Medicine	91
SURF Speeded-Up Robust Features	55
SVH Système Visuel Humain	25
TOF “Time Of Flight” : Angiographie par temps de vol.	31
VDP Visible Difference Predictor	26
VIF Visual Information Fidelity	44
VQEG Video Quality Experts Group	51
VSL Visual Science Laboratories	28

Table des figures

3.1	Décomposition du spectre de Fourier en canaux perceptuels.	36
3.2	Robustesse de la méthode dans les 3 bandes de fréquences face aux 89 attaques de Stirmark.	36
3.3	Matrices de corrélation pour deux tailles distinctes de patches insérés.	38
3.4	Robustesse face à Stirmark au seuil de visibilité.	38
3.5	Distributions $\mathcal{H}_0$ et $\mathcal{H}_1$ pour les trois méthodes testées.	39
3.6	Illustration du test de Grubbs, lors duquel une valeur de l'échantillon est vue comme un outlier dans une distribution Gaussienne. Nous présentons des exemples de Corrélations pour diverses méthodes de la littérature.	40
3.7	Les neuf modes de prédiction <i>intra</i> pour le décodage de la vidéo. . .	41
3.8	Analyse des hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$ pour deux facteurs de quantifications.	42
3.9	Comparaison des PSNR sur la méthode de [Noorkami 2007] et la méthode proposée dans [Chen 2014].	43
3.10	Qualité Subjective des séquences tatouées pour différents pas de quan- tifications.	43
3.11	PSNR et VIF sur les différentes séquences de test tatouées avec notre méthode.	44
3.12	Variation du contraste selon l'axe vertical, et variation de la fréquence spatiale sur l'axe horizontal.	46
3.13	Fonction de Sensibilité au Contraste en version 1D (et 2D dans l'encart).	47
3.14	Performances de la métrique CPA sur 4 bases de données subjectives.	48
3.15	Tracés des courbes MOS en fonction de MOSp pour six métriques sue la base [Atrusseau 2010b].	49
3.16	Comparaison de la fidélité de deux métriques (SSIM & CPA) pour estimer la qualité du chiffrement sélectif.	50
3.17	Pouvoir de discrimination des différentes métriques dans les 4 clusters (FP,FN,VP,VN).	50
3.18	Configuration du protocole de test à choix forcé, l'observateur a pour objectif de trouver la paire d'images.	53
3.19	Clustering par dendrogrammes sur les observateurs des 2 tests conduits à Nantes.	53
3.20	Corrélations (Pearson & Spearman)	55
3.21	Corrélation de Pearson et RMSE sur l'intégralité de la base de don- nées subjectives lorsqu'un nombre variable d'observateurs est considéré.	56
3.22	Corrélations de Pearson pour différentes tailles de bases de données lorsque 15 observateurs ont été retenus.	56
4.1	Perception altérée d'un anneau orange / rose intercalé entre des an- neaux vert-amande / violet.	60

4.2	Aperçu de l'inversion de phase que l'on peut rencontrer (phénomène d'assimilation) (a) et à l'inverse l'induction chromatique est induite en phase (contraste) en (b).	61
4.3	Comparaison entre l'aberration chromatique seule (a) et la prise en compte des aberrations de front d'ondes (b).	62
4.4	Génération de deux stimuli de diverses fréquences spatiales pour calculer la <i>MTF</i>	64
4.5	Absorption des trois types de cônes (L, M & S) pour 100 modèles optiques d'observateurs.	64
4.6	Différence de perception sur la lettre "E", lorsqu'uniquement la LCA est considérée, et quand les aberrations de front d'onde s'y ajoutent.	65
4.7	Comparaison du profil 1D sur une image stimulus de bandes verticales noires et blanches. On distingue les profils avec et sans aberrations de front d'onde et les travaux de [Marimont 1994].	65
4.8	Densité spectrale de puissance des deux moniteurs utilisés dans notre analyse.	68
4.9	Simulation des shifts chromatiques de sept stimuli par simple modification des courbes fondamentales d'absorption des cônes dans un diagramme chromatique (u', v') (voir texte pour les détails).	69
4.10	Box plots représentant la variabilité inter-observateurs pour l'expérience de "color matching".	69
5.1	(a) Angle θ_i formé par (p_i, q_i) , projection des pixels sur le bin b de la proj $proj_{p_i, q_i}(b)$, (b) Exemple de transformation Mojette sur une grille 3×3 suivant les directions $\{(1, 0), (1, 1), (-1, 1)\}$	72
5.2	Projection d'un pixel via une Transformation Mojette <i>Spline 0</i> selon les angles $(1, 0)$, $(2, 1)$ et $(1, 1)$	73
5.3	Quelques exemples de projections Mojette "Haar" sur une image de texture.	73
5.4	Trois premières étapes d'une reconstruction Mojette. 5.1.	74
5.5	Projections Mojette sur l'image "watch" selon des orientations $(1, 128)$, $(1, 129)$ et $(1, 130)$	77
5.6	Zoom sur les projections $(1, 128)$, $(1, 129)$, $(1, 130)$, différence intra-projection sur $(1, 130)$. Range des bins : $[11000, 11400]$	78
5.7	(a) Entropie pour la compression Intra- et Inter-projections comparée avec JPEG2000 pour les 11 images de test montrées en (b).	78
5.8	(a) Projections 1D de l'image "Lena" et (b) ces mêmes projections affichées sous forme 2D.	79
5.9	Résultats en termes d'entropie pour la compression intra-projections entre deux projections. Nous utilisons les méthodes de compression : DPCM-1, DPCM-3, LOCO, CALIC, et Glicbawls. Nous comparons ces résultats à un codage JPEG2000 et CALIC appliqués directement sur les images	80

5.10	Entropie tracée en fonction du nombre de projections, N , nécessaires pour la reconstruction. Les trois courbes représentent trois scénarios différents: 1) “sans projection redondante”, 2) “une projection supplémentaire” et finalement 3) “deux projections supplémentaires”.	80
5.11	Taux de compression moyen (sur 11 images de test) pour le scenario de codage intra-projections (a) et inter-projections (b).	81
5.12	Les lignes de directions des vecteurs (15, 2) (points) et (21, 2) (tirets) peuvent subir un sous-échantillonnage (3, 2).	82
5.13	Représentation de $\mathcal{M}(\mathcal{W})$, on observe la scalabilité des projections Mojette.	82
5.14	Propriété de redimensionnement sur une image au format 48×48 pixels, $\mathcal{W}$ représente l’opérateur de scaling.	82
5.15	Deux possibilités différentes de reconstruire une image à partir de ses projections scalables.	83
5.16	Convolution de 3 éléments structurants à 2 pixels ((1,0), (1,1) et (-1,1)) pour composer un fantôme et projections Mojette du fantôme le long de ces orientations.	84
5.17	Image originale et sa version tatouée, ainsi que le watermark dans sa version spatiale (images du haut) et dans le domaine de Fourier (images du bas).	85
5.18	Détection du fantôme sur les projections Mojette.	85
5.19	Rétro-projection en suivant un chemin spécifique de l’ensemble des bins à correspondance unitaire.	86
5.20	Codage Mojette intra-projection et inter-projections conduisant à la génération d’une projection de base et une projection de différences.	88
5.21	Cascade de projections Mojette avec codage intra- et inter-projections. L’ultime projection de base peut être chiffrée, rendant la reconstruction impossible.	88
5.22	Résultats fournis par la méthode de chiffrement par cascade de projection. Lorsque le nombre de boucles de la cascade augmente, il est nécessaire de chiffrer moins de données, mais au prix d’un taux de compression accru.	89
6.1	Un exemple de radiographie panoramique et de schéma dentaire.	92
6.2	Une radiographie panoramique avec une représentation d’un profil 1D de l’image (en bleu).	92
6.3	Schéma bloc représentant des différentes étapes du processus de détection des dents.	93
6.4	La plupart des dents sont détectées par l’un des 2 GMM	94
6.5	Radiographie panoramique avec le détournement manuel des dents.	94
6.6	Résultats du réseau de neurones U-Net sur la segmentation de la denture	95
6.7	Dimension fractale dans l’os cortical (à gauche) et dans l’os trabéculaire (à droite).	96

6.8	Le Foramen Mentonnier est marqué par le cercle rouge dans la coupe frontale (image en bas à droite).	97
6.9	Différents Foramens Mentonniers sur des acquisitions au CBCT.	98
6.10	Représentation graphique des mesures prises sur l'arbre vasculaire.	99
6.11	(a) Arbre vasculaire de souris après acquisition au micro-scanner, (b) squelette de la vascularisation obtenue par morphologie mathématique, et (c) graphe de connexion entre les branches du squelette.	100
6.12	(a) Bifurcation d'une TOF d'IRM sur patient humain, (b) squelette 3D et (c) graphe 3D.	100
6.13	Polygone de Willis sur un micro-scan de souris.	101
6.14	Projections Mojette selon différents angles d'une cross-section 2D, obtention des diamètres min et max (resp. 9 et 21 bins (aux angles $\theta = 153^\circ$ et $\theta = 44^\circ$).	101
6.15	Distributions des diamètres min et max d'après le modèle (en noir) et mesures manuelles par des experts (en gris).	102
6.16	(a) Bifurcation 3D, (b) Angles 3D formés par la bifurcation.	102
6.17	Calcul de la tortuosité sur les artères.	103
6.18	(a) BoxPlot des Angles, (b) Somme des angles, (c) section des artères, (d) Différences de diamètres (mère/fille), (e) distances géodésiques, (f) tortuosité.	104

Curriculum Vitae

Contents

1.1	Expérience Professionnelle	2
1.2	Formation	2
1.3	Projets de Recherche	3
1.4	Collaborations Académiques et Industrielles	3
1.5	Activités scientifiques diverses	4
1.6	Diffusion et dissémination	5
1.7	Activités d'Encadrement:	6
1.8	Liste des publications	10

Florent Autrusseau
Né le 06 Novembre 1975 à Clermont-Ferrand
Marié, 2 enfants
Ingénieur de Recherche première classe

Adresse Professionnelle :

RMeS-INSERM UMR 1229, UFR d'Odontologie
1, Place Alexis Ricordeau
44042, Nantes cedex 1, FRANCE
Tel. : 02.40.41.29.99

LTen UMR 6607, Polytech'Nantes
La Chantrerie, Rue Christian Pauc
BP50609, 44306, Nantes, Cedex 3, FRANCE
Tel. : 02.40.68.31.37

Florent.Autrusseau@univ-nantes.fr
<http://www.univ-nantes.fr/autrusseau-f>
<http://www.polytech.univ-nantes.fr/autrusseau-f/>

1.1 Expérience Professionnelle

- 2020- Ingénieur de Recherche à l'Ecole Polytechnique de l'Université de Nantes,
Laboratoire de Thermique et Energie de Nantes (LTeN) - UMR 6607
- 2016-2020 Ingénieur de Recherche à l'Ecole Polytechnique de l'Université de Nantes,
Laboratoire RMES, U1229. Imagerie médicale & biologique
(MEB/MET, microscopie confocale, microscanner, IRM, TOF, ...),
analyses statistiques de données biologiques...
- 2004-2016 Ingénieur de Recherche à l'Ecole Polytechnique de l'Université de Nantes,
Laboratoire IRCCYN. Traitement du signal et des images, biométrie,
transmission sécurisée, compression, évaluation de la qualité
des images, imagerie médicale...
- 2003-2004 Post-Doctorat en Sciences Visuelles et Cognitives au
Laboratoire de Sciences Visuelles de l'Université de Chicago, USA.
Etude de la perception de la couleur, mise en œuvre
d'un modèle optique de l'œil.
- 2003 Collaboration à un projet du Réseau MEGALIS.
Proposition d'une technique de chiffrement d'images
médicales pour une transmission sécurisée.

1.2 Formation

- 1999-2002 Thèse de Doctorat en Traitement d'Images, Ecole Polytechnique
de l'Université de Nantes, Laboratoire IRCCYN
Sujet : « Modélisation psychovisuelle pour le tatouage des images ».
Enseignements dispensés : C, Matlab, Java, Unix, Théorie
de l'Information, Motif, Traitement du Signal et des images, IHM
- 1998-1999 D.E.A. d'Electronique, option Image,
Laboratoire IRCCYN, IRESTE, NANTES.
Stage portant sur le tatouage numérique
(projet AQUAMARS).
- 1997-1998 Maîtrise E.E.A. (Electronique Electrotechnique Automatique)
option micro- électronique. (Université de Clermont-Ferrand).

Informations complémentaires

Informatique Programmation : C/C++, Python, Java, Tcl/Tk, Matlab/Octave.
 Logiciels : XCode, Dev-cpp, CodeBlocks, ImageJ, R, Gimp, Slicer...
 Systèmes : Mac OS, Linux, Unix, Windows.
 Administration du parc machines de l'équipe IRCCyN-IVC 2004-2016

Langues Français, Anglais

1.3 Projets de Recherche

Au fil des ans, j'ai participé à de nombreux projets de recherche. Pour la plupart de ces projets, j'étais soit le porteur du projet, soit le responsable scientifique pour mon laboratoire de rattachement. Au total, j'ai pu collecter plus de 800 K€. Je présente ici la liste de ces différents projets.

- I-Site NExT, Cluster "INNOVEZ", Projet "Odon" 2020 (55 K€).
- PHRC (Projet Hospitalier de Recherche Clinique), CHU Nantes, 2020-2021 (60 K€) dont encadrement d'un post-doc (41K€).
- I-Site NExT, Cluster "FAME", 2020 (3,6 K€).
- Financement d'un post-doc via crédits INSERM (Richard Redon) sur projet ANR-ICAN, 2018-2019 (40 K€).
- AAP Interdisciplinaire Univ-Nantes, 2017-2018 (14 K€).
- FWF (Autriche) Visual security, 2015-2018 (45 K€).
- Accompagnement de thèse CIFRE – Keosys, 2016 (15 K€).
- Contrat de maturation avec la SATT "Ouest-Valorisation", 2014 (59 K€).
- Contrat de faisabilité pour la Société DataSysCom, 2014 (3,6 K€).
- OSEO/Région "StreamMaster", 2012-2014 (68 K€).
- OSEO/Région "e-Map", 2010-2012 (80 K€).
- Thèse CIFRE - Technicolor, 2008-2011 (37 K€).
- Région Pays de la Loire - bourse post-doctorale, 2007-2008 (24 K€).
- ANR - CSOSG "SAFIMAGE", 2007-2010 (178 K€).
- Région Pays de la Loire "MILES", 2006-2009 (36 K€).
- ANR - SSIA "TSAR", 2005-2008 (101 K€).
- Rapport d'expertise Jessica-Ouest, biométrie, 2005 (3 K€ pour Polytech).
- ANR - RNRT "AQUAMARS" 1999-2002.

1.4 Collaborations Académiques et Industrielles

Ces différents projets, certains recrutements et encadrements, ainsi que des rencontres dans les colloques auxquels j'ai pu participer m'ont permis de m'impliquer dans plusieurs collaborations scientifiques tant au niveau national qu'international. Une majorité des collaborateurs avec lesquels j'ai pu m'associer est issue du monde

académique, mais j'ai aussi pu tisser quelques liens avec certains industriels. Je dresse ici une liste non exhaustive de ces collaborations.

- Patrizio Campisi (Uni Roma Tre, Italie),
- Andrew Kingston (ANU Canberra, Australie),
- Thomas Stütz (FHS, Salzburg, Autriche),
- Andreas Uhl (Univ. of Salzburg, Autriche),
- Heinz Hofbauer (Univ. of Salzburg, Autriche),
- Junyu Dong (Ocean Univ. of China, Qingdao, China),
- Steve Shevell (VSL, Univ. of Chicago, USA),
- Larry Thibos (School of optometry, Indiana, USA),
- Vinod Pankajakshan (IIT Roorkee, Inde),
- Zafar Shahid (IIT Abbottabad, Pakistan),
- Fan Zhang (Lenovo Research, Hong Kong),
- Olivier Déforges (IETR Rennes, France),
- François Cayre (GIPSA-lab Grenoble, France),
- William Puech (LIRMM, Montpellier, France),
- Teddy Furon (INRIA, Rennes, France),
- Patrick Bas (Ecole Centrale de Lille, France),
- Hubert Desal (Institut du Thorax, Nantes),
- Romain Bourcier (Institut du Thorax, Nantes),
- Gwenaél Doërr (Technicolor & Content Armor),
- Laurent Blondé (Technicolor, France),
- Sylvie Davila (Keosys, France),
- Anass Nouri (Université de Kénitra).

Je détaillerai plus avant mes diverses contributions dans des projets de recherche, ainsi que les collaborations qui y ont été associées dans le Chapitre 7.

1.5 Activités scientifiques diverses

Relecture d'articles dans des revues internationales :

- IEEE Trans on Image Processing (2008, 2010, 2012, 2014),
- IEEE Trans on Information Forensics and Security (2013),
- IEEE Trans on Multimedia (2013),
- IEEE Trans on Circuits and Systems for Video Technology (2009, 2010, 2013, 2014),
- Elsevier Information science (2007),
- Elsevier Signal Processing (2005, 2007, 2008, 2010),
- Elsevier Signal Processing : Image Communication (2008, 2009, 2010, 2011, 2012, 2013),
- Elsevier Digital Signal Processing (2014, 2020),
- Elsevier Pattern Recognition (2009, 2010, 2011, 2015, 2016, 2017),
- Elsevier Journal of Systems and Software (2008),

- SPIE Optical Engineering (2005, 2006, 2008, 2009),
- SPIE Journal of Electronic Imaging (2005, 2006, 2007, 2008, 2010, 2011),
- IET Image Processing (2008),
- Journal of Pattern Recognition Research (2011),
- Springer Signal, Image and Video Processing (2012).

Relecture d'articles dans des conférences internationales :

- ICIP (2014, 2015, 2017, 2018, 2019, 2020),
- ICASSP (2014, 2015, 2016),
- IH &MMSec (2013, 2014),
- CISP (2010, 2012, 2013),
- Eusipco (2007, 2011),
- QoEMCS (2012),
- VPQM (2010),
- IPTA (2010).

Membre de comités de programmes :

- ACM IH & MMSEC 2014,
- ACM IH & MMSEC 2013,
- IEEE CISP/BMEI 2013,
- QoEMCS 2012,
- IEEE CISP 2012,
- IEEE IPTA 2010,
- IEEE CISP 2010.

Reviewer pour projets de recherche :

- Programme CONTINT, ANR 2011,
- Projet de la *Czech Science Foundation*, 2005.

Special Session Chair pour la conférence :

- APSIPA ASC 2015.

Co-Editeur & Comité d'organisation :

- Springer DGCI 2016.

1.6 Diffusion et dissémination

- Dépôt d'un logiciel à l'Agence pour la Protection des Programmes "TIRPS : Tatouage d'Images Robuste au Print & Scan" (et dépôt de brevet associé), 2013.
- Dépôt d'un logiciel à l'Agence pour la Protection des Programmes "Détection et caractérisation automatiques d'anévrismes intra-crâniens dans des vascu-

- latures cérébrales 3D” (et dépôt de brevet associé), 2018.
- Diffusion de divers codes sources (en C/C++, Python, Matlab, Mathematica,...) : <http://www.polytech.univ-nantes.fr/autrusseau-f/>
- Diffusion d’un player de vidéos YUV : (3342 téléchargements en Septembre 2020) <https://sourceforge.net/projects/sdl-yuvplayer/>
- Diffusion de divers protocoles de tests subjectifs : <http://websubjexp.polytech.univ-nantes.fr/>
- Diffusion de données collectées lors de tests subjectifs : <http://www.polytech.univ-nantes.fr/autrusseau-f/Publications.html#DB>
- Diffusion d’une web-app à diverses entreprises & laboratoires académiques : <http://ivc.univ-nantes.fr/processing/WmkTIFS2014/> et diffusion des bibliothèques correspondantes (dlls).
- Mise à disposition de divers outils développés pour les membres du laboratoire RMeS : <http://www.polytech.univ-nantes.fr/autrusseau-f/RMeS/>

1.7 Activités d’Encadrement :

Etudiants en Thèse de doctorat :

- **Sara Chater**, “Estimation du risque de survenue des anévrismes intra-crâniens basée sur l’analyse des images d’IRM”, début de la thèse en Janvier 2020, co-encadrement avec Mr. Youssef el Merabet (Directeur de thèse), Mr. Anass Nouri Université Ibn Tofail, Kénitra, Maroc (YeM : 33%, AN : 33%, FA : 33%). Thèse en cours.
- **Ibtissam Essadik**, “Génération d’un atlas anatomique de la vasculature cérébrale pour la détection des anévrismes intra-crâniens”, début de la thèse en Janvier 2020, co-encadrement avec Mme. Raja Touahni (Directrice de thèse), Mr. Anass Nouri Université Ibn Tofail, Kénitra, Maroc (RT : 33%, AN : 33%, FA : 33%). Thèse en cours.
- **Laure Mahé**, Ph. D. “Développement d’outils de quantification innovants en études cliniques neuro-vasculaires” 2015. Co-encadrement avec Jeanpierre Guédon (Directeur de Thèse) (JPG : 50%, FA : 50%). Abandon de la thèse CIFRE après une année, suite à une restructuration de la recherche au sein de l’entreprise. Rachat de l’entreprise Keosys, l’encadrant chez Keosys a démissionné. Publications communes : [Mahé 2016]. Laure est à présent Chef de projet au GCS e-Santé Bretagne.
- **Abhijit Sarkar**, Ph.D. “From Color Science To Media And Entertainment Industry Applications, By Means Of Display, Vision And Quality Models”, 2008-2011. Co-encadrement avec Patrick Le Callet (FA : 50%, PLC : 50%). Publications communes : [Sarkar 2009b, Sarkar 2009a, Sarkar 2010a, Sarkar 2010b, Sarkar 2010c, Sarkar 2010d, Sarkar 2011a]. Abhijit Sarkar est à présent "Senior Display Engineer" chez Microsoft (USA).

Post-doctorants :

- **Nathan Lauzeral** “Cerebral vascular tree segmentation using neural networks trained on mice vasculature”, (Financement via PHRC INSERM) début du post-doc en Sept. 2020.
- **Anass Nouri** “Full Automatic 3D Bifurcations Detection on Cerebral Vasculatures” 2017-2018 (Financement INSERM), [Nouri 2019, Nouri 2018, Autrusseau 2020, Nouri 2020], Anass enseigne à présent à l’Univ. de Kénitra (Maroc).
- **Matthieu Urvoy** “Document/image watermarking, robustness against Print & Scan” - 2012-2014 (financement sur projet StreamMaster), [Autrusseau 2013d, Urvoy 2014a, Urvoy 2013, Urvoy 2014b], Matthieu est maintenant Ingénieur de Recherche chez Bluecime.
- **Dalila Goudia** “Wavelet domain watermarking, embed watermarks within low activity areas” - ATER : 2011-2012, post-doc/research engineer : 2013-2014. (Financement sur projet SATT), [Autrusseau 2013b, Autrusseau 2013c, Urvoy 2014b], Dalila est maintenant Innovation reliefs and incentives Supervisor chez KPMG.
- **Thomas Stuetz** “H.264 video watermarking” - 2011 (financement sur projet e-Map), [Stütz 2014, Stütz 2011, Chen 2014], Thomas est à présent enseignant chercheur à l’Université de Salzbourg.
- **Zafar Shahid** “Perceptual watermarking of H.264 videos” - 2011 (financement sur projet e-Map), [Chen 2014], Zafar est maintenant Media Architect chez Nvidia (USA).
- **Vinod Pankajakshan** “Watermarking based on HVS models” - 2009-2011 (financement sur projets Safim@ge & e-Map), [Stütz 2010, Autrusseau 2010b, Carosi 2010, Autrusseau 2011a, Pankajakshan 2010], Vinod enseigne à l’IIT Roorkee (Inde).
- **Andrew Kingston** “Lossless compression, Selective Encryption, Multi-resolution Mojette transform” - 2006-2007 (financement sur projet Région), [Kingston 2007, Kingston 2008b, Freuchet 2015, Kingston 2009a], Andrew est postdoc fellow au sein de l’ANU en Australie.

Stages de recherche de Mastère :

- **Charlotte Fosse**, “Recalage 3D d’IRM-TOF pour la constitution d’un atlas anatomique de la vasculature cérébrale” - stage M1 Ecole Centrale Nantes 2020,
- **Clément Morin**, “Caractérisation osseuse d’images de CBCT pour détecter le risque ostéoporotique” - stage M2 SIBM 2020,
- **Clément Morin**, “Détection de la denture et des soins dentaires sur des radiographies panoramiques” - Stage M2 Ingénieur HEI 2020,
- **Julia Guillou**, “Segmentation de l’arbre vasculaire cérébral via des critères de perception humaine” - M2 SIBM, 2019, [Guillou 2020],

- **Natalia De La Fourniere**, “Automatic Generation of the Dental Scheme based on 2D Radiographs” - M2 SIBM, 2018,
- **Shenbao Peng**, “RST invariant watermarking using chaotic maps” - M2 MDM 2015,
- **Ye Rui**, “Watermarking using chaotic maps” - M2 MDM 2013,
- **Wei Chen** “Video Watermarking” - M2 SEGE 2007, [Chen 2008, Chen 2014],
- **Alexandre Ninassi**, “Watermarking based Image quality assessment” - M2 ALD 2005, [Ninassi 2006, Autrusseau 2007b],
- **Pascal Lesage**, “Biometrics, signature recognition” - M2 ALD 2005.

Erasmus : (Niveau Mastère 2, collaboration avec Patrizio Campisi, Univ de Rome)

- **Enrico Marini** “Quality assessment of watermarked images”- 2006, [Marini 2007a, Marini 2007b],
- **Simone Colosimo** “Selective encryption based on the Mojette transform” - 2006, [Kingston 2007] .

Stage d’ingénieurs formation CNAM :

- **Mohamed Bouchakour** “DWT watermarking” - 2007, [Bouchakour 2008],
- **Thierry Hamon** “Security, distributed storage, PKI, Mojette encryption” - 2006, [Kingston 2009b, Evenou 2006].

Ingénieurs de Recherche :

- **Sylvain David** “Wavelet based Just Noticeable Difference Masks for digital watermarking” - 2009-2010 (financement sur projet Safim@ge), [Autrusseau 2011a, Autrusseau 2010a],
- **Maurizio Carosi** “Quality metrics for digital watermarking” - 2009 (financement Safim@ge), [Carosi 2010],
- **Wei Chen** “H.264 watermarking” - 2007-2008 (financement Safim@ge), [Chen 2008, Chen 2014],
- **Pierre Evenou** “Objective Quality metrics” - 2007 (financement TSAR), [Evenou 2006].

Stages d’Ecole d’Ingénieurs (Polytech’ Nantes) :

- Automatic generation of dental scheme based on radiographs - 2016,
- Depth estimation based on 2D video acquisition - 2015,
- Vascular lesion detection on IRM images - 2015,
- Fourier Watermarking for geometrical distortions compensation - 2010,
- Mojette based security (2 projets) - 2009 (PRD-Info5),
- Wavelet domain watermarking - 2009 (PRD-Info5),
- Mojette based geometrical distortions compensation - 2009,

- Mojette based watermarking & encryption - 2009,
- Lossless compression based on the Mojette transform (2 projets) - 2006,
- Grid computing (2 projets) - 2006,
- Hand detection and tracking (2 projets) - 2005,
- Video watermarking (H.264) - 2005,
- Quality assessment - 2005.

Participation au jury de Thèse de Jorge Roberto Manjarrez Sanchez le 26 Octobre 2009 sur le sujet "Recherche Parallèle par Contenu dans les Bases de Données Images", ainsi qu'au jury de thèse de Abhijit Sarkar le 26 Octobre 2011 "From Color Science To Media And Entertainment Industry Applications, By Means Of Display, Vision And Quality Models".

Résumé des publications

Voici un bref résumé de mes publications, les index (H-index et i10-index) ont été obtenus depuis la base Google Scholar. La liste complète de mes publications est donnée dans la section suivante.

H-index	17
i10-index	24
Publications	100
Revue internationale	15
Revue nationale	2
Chapitres d'ouvrages	4
Brevets	3
Proceeding	1
Publications en conférences internationales	48
Publications en conférences nationales	6
Abstracts	5
Rapports Techniques	2
Mémoires	2
Bases de données subjectives	12

Pour certaines de ces publications, ma contribution rentrait dans le cadre d'un accompagnement de la recherche, il s'agit des publications [9,11,74,75,76,91,92]. En effet, ces travaux ont été publiés ou présentés dans des articles ou conférences spécialisés dans le domaine médical ou biologique. Ma contribution s'est limitée au développement d'outils logiciels répondant à des besoins d'imagerie. Pour toutes les autres publications de cette liste, j'ai conduit des recherches portant sur des thématiques de traitement et analyse du signal et des images, ou en perception visuelle.

1.8 Liste des publications

Brevets

- [1] F. Autrusseau, A. Nouri, R. Bourcier : “Method for Segmenting Digital Images, Devices and Systems for the Same. ”, patent application Eur. 20305143.8, submitted on Feb. 14 2020.
- [2] A. Nouri, F. Autrusseau and R. Bourcier, “Method for Locating and Characterizing Bifurcations of a Cerebral Vascular Tree, Associated Methods and Devices”, Patent Nb EP18306612.5, sub. on Dec. 4th, 2018.
- [3] F. Autrusseau, PCT : “Method for inserting in the Fourier domain, a digital watermark in an image and corresponding method for detecting a digital watermark in an image to be analyzed”, patent application 1251325, FR 2986890, U.S. Pat No. 9401002 published on Aug. 08 2013.

Proceeding

- [4] N. Normand, JP. Guédon, F. Autrusseau, “ Discrete Geometry for Computer Imagery”, Apr 2016, Nantes, France. 9647, Springer International Publishing, pp.453, 2016, LNCS, ISBN : 978-3-319-32359-6.

Chapitres d’ouvrages

- [5] P. Le Callet, F. Autrusseau, P. Campisi, “Visibility control and Quality assessment of watermarking and data hiding algorithms”, Chap IX in Multimedia Forensics and Security, Idea Group Publishing, Editor : Chang-Tsun Li, pp. 163- 192, ISBN : 978-1-59904-869-7, April 2008.
- [6] A. Kingston, B. Parrein, F. Autrusseau, “Multiresolution Mojette transform", Chapter 6 in The Mojette transform : Theory and Applications, ISTE-Wiley, ISBN : 9781848210806, February 2009.
- [7] A. Kingston, F. Autrusseau, “Lossless compression", Chapter 9 in The Mojette transform : Theory and Applications, ISTE-Wiley, ISBN : 9781848210806, February 2009.
- [8] A. Kingston, F. Autrusseau, E. Grall, T. Hamon, B. Parrein, “Mojette based security”, Chapter 10 in The Mojette transform : Theory and Applications, ISTE-Wiley, ISBN : 9781848210806, February 2009.

Revues internationales

- [9] K. Flégeau, C. Toquet, G. Rethore, C. D’Arros, L. Messenger, B. Halgand, D. Dupont, F. Autrusseau, J. Lesoeur, J. Veziers, P. Bordat, A. Bresin, J. Guicheux, V. Delplace, H. Gautier, P. Weiss, “In Situ Forming, Silanized Hyaluronic Acid Hydrogels with Fine Control Over Mechanical Properties and In Vivo Degradation for Tissue Engineering Applications”, accepted in in Advanced Healthcare Materials, Wiley, Aug. 2020.

- [10] H. Hofbauer and F. Autrusseau and A. Uhl, "To Recognize or not to Recognize — A Database of Encrypted Images with Subjective Recognition Ground Truth", accepted for publication in Elsevier Information Sciences, Nov. 2020.
- [11] V. L'Allinec and S. Chatel and M. Karakachoff and E. Bourcereau and Z. Lamoureux and A. Gaignard and F. Autrusseau and S. Jouan and A.C. Vion and G. Loirand and H. Desal and O. Naggara and R. Redon and M. Edjlali and R. Bourcier, "Unruptured Cerebral Aneurysm, prediction of evolution : The UCAN Project", in Neurosurgery, Oxford University Press, Feb. 2020.
- [12] A. Nouri, F. Autrusseau, R. Bourcier, A. Gaignard, V. L'Allinec, C. Menguy, J. Veziers, H. Desal, G. Loirand and R. Redon, "Characterization of 3D Bifurcations in Micro-scan and MRA-TOF Images of Cerebral Vasculature for Prediction of Intra-Cranial Aneurysms ", in Elsevier Computerized Medical Imaging and Graphics, CMIG, Aug. 2020.
- [13] M. Urvoy, D. Goudia, F. Autrusseau, " Perceptual DFT watermarking with improved detection and robustness to geometrical distortions", in IEEE Transactions on Information Forensics and Security, IEEE-TIFS, 9(7), pp. 1108-1119, ISSN 1556-6013, Jul. 2014.
- [14] T. Stuetz, F. Autrusseau, A. Uhl, " Non-Blind Structure-Preserving Substitution Watermarking of H.264/CAVLC Inter-Frames", in IEEE Transactions on Multimedia, IEEE-TMM, 16(5), pp. 1337-1349, ISSN 1520-9210, Aug. 2014.
- [15] F. Zhang, W. Jiang, F. Autrusseau, W. Lin, " Exploring V1 by modeling the perceptual quality of images", in Journal of Vision, 14(1), Jan. 2014.
- [16] W. Chen, Z. Shahid, T. Stuetz, F. Autrusseau, P. Le Callet, " Robust drift-free bit-rate preserving H.264 watermarking", in Springer, Multimedia Systems, Special Issue on Digital Media Security and Right Management, 20(2), pp 179- 193, ISSN 0942-4962, 2014.
- [17] M. Urvoy and F. Autrusseau, "Print and scan robust image watermarking algorithm", Journal of South China University of Technology, JSCUT, pp 120-125, Vol 7, 2013.
- [18] J. Dong, L. Su , Y. Zhang , F. Autrusseau, Z. Yang, " Illumination direction estimation of three-dimensional surface texture based on active basis and Mojette transform", in SPIE, Journal of Electronic Imaging, 2012, 21(1), 013023, Apr. 2012.
- [19] F. Autrusseau, L. Thibos, S. Shevell, " Chromatic and Wavefront Aberrations : L-, M- and S-cone Stimulation with Typical and Extreme Retinal Image Quality", Elsevier Vision Research, 51(21-22), pp 2282-294, Nov. 2011.
- [20] A. Sarkar, F. Autrusseau, F. Vienot, P. Le Callet, L. Blondé, " From CIE 2006 Physiological Model to Improved Age-Dependent and Average Colorimetric Observers", in Journal of the Optical Society of America A, JOSA-A, 28(10), pp. 2033-2048, 2011.

- [21] A. Kingston, F. Autrusseau, “ Lossless Image Compression via Predictive Coding of Discrete Radon Projections”, Elsevier Signal Processing : Image Communication, Vol 23(4) pp 313-324, May 2008.
- [22] F. Autrusseau, P. Le Callet, “ A robust image watermarking technique based on quantization noise visibility thresholds”, Elsevier Signal Processing, Volume 87(6), Pages 1363-1383, June 2007.
- [23] F. Autrusseau, S. K. Shevell, “ Temporal nulling of induction from spatial patterns modulated in time”, Cambridge University press, Visual Neuroscience, Volume 23(3-4), pp 479-482, May 2006.

Revues nationales

- [24] F. Autrusseau, JP. Guédon, “ Chiffrement Mojette d’images médicales”, ISI journal, Ingénierie des Systèmes d’Information, Health information systems special issue, Vol 8, No 1, pp 113-134, March 2003.
- [25] A. Saadane, F. Autrusseau, “ Tatouage perceptuel et adaptatif d’images fixes”, revue Traitement du Signal, Volume 18, No 4, pp. 235-247, Watermarking special issue, 2002.

Conférences internationales (proceedings)

- [26] J. Guillou, F. Autrusseau, R. Bourcier, “Brain Vasculature Segmentation based on Human Perception Criteria”, in SPIE Medical Imaging 2020, Houston, TX USA, 15 - 20 Feb. 2020.
- [27] A. Nouri, F. Autrusseau, R. Bourcier, A. Gaignard, V. L’Allinec, C. Menguy, J. Veziers, H. Desal, G. Loirand, R. Redon, “3D bifurcations characterization for intra-cranial aneurysms prediction”, in SPIE Medical Imaging 2019, San Diego, CA USA, 16 - 21 Feb. 2019.
- [28] H. Hofbauer, F. Autrusseau, A. Uhl, “To See or Not To See : Determining the Recognition Threshold of Encrypted Images”, 7th European Workshop on Visual Information Processing, EUVIP, Tampere, Finland, 26-28 Nov. 2018.
- [29] P. Michel, V. Prezelin, P. Blery, Y. Amouriq, JP. Guédon, F. Autrusseau, “Automatic Generation of the Dental Scheme based on 2D radiographs”, in SPIE Medical Imaging 2018, Houston, TX, USA, 10 - 15 Feb. 2018.
- [30] L. Mahé, F. Autrusseau, H. Desal, JP. Guédon, H. Der Sarkissian, Y. Le Teurnier, S. Davila, “Quality metrics can help the expert during neurological clinical trials”, in Proc. SPIE Medical Imaging 2016, San Diego, CA, USA, 27 Feb. - 3 Mar. 2016.
- [31] F. Autrusseau, Y. Amouriq, JP. Guédon, “ Linking bone microarchitecture to projections texture analysis”, in Proc. SPIE Medical Imaging 2016, San Diego, CA, USA, 27 Feb. - 3 Mar. 2016.

- [32] E. Crauste, F. Autrusseau, JP. Guédon, P. Pilet , Y. Amouriq, P. Weiss, B. Giumelli, “Bone vascularization and bone microarchitecture characterizations according to the microCT resolution”, in Proc. SPIE Medical Imaging 2015, Orlando, FL, USA, 21-26 Feb. 2015.
- [33] E. Freuchet, B. Recur, JP. Guédon, A. Kingston, F. Autrusseau, Y. Amouriq, “Building a bone micro-CT images atlas for micro-architecture recognition”, in Proc. SPIE Medical Imaging 2015, Orlando, FL, USA, 21-26 Feb. 2015.
- [34] M. Urvoy, F. Autrusseau, “Application of Grubbs’ test for outliers do the detection of watermarks”, in the 2nd ACM Workshop on Information Hiding & Multimedia Security, IH&MMSec, pp. 49-60, Salzburg, Austria, June 11-13, 2014.
- [35] P. Bléry, F. Autrusseau, E. Crauste, E. Freuchet, P. Weiss, Jp. Guédon, Y. Amouriq, “Bone Vascularization : A Way To Study Bone Microarchitecture?”, in SPIE Medical Imaging, San Diego, CA, USA, Feb. 15-20, 2014.
- [36] F. Autrusseau, D. Goudia, “Non linear hybrid watermarking for High Dynamic Range Images”, in IEEE International Conference on Image Processing, ICIP’2013, pp. 4527-4531, Melbourne, Australia, Sept 15-18, 2013.
- [37] JP. Guédon, F. Autrusseau, Y. Amouriq, P. Bléry, JM. Bouler, P. Weiss, FX. Barbarin, T. Dallet, V. Dallerit, “Exploring relationships between fractal dimension and trabecular bone characteristics”, in SPIE Medical Imaging 2012, Proc. SPIE 8317, 831717, San Diego, USA, 4-9 February 2012.
- [38] F. Autrusseau, S. David, V. Pankajakshan, P. Campisi, “A perceptually driven hybrid additive-multiplicative watermarking technique in the wavelet domain”, in Electronic Imaging : Media Watermarking, Security and Forensics XIII, vol. 7880, San Francisco, USA, 23-27 Jan. 2011.
- [39] T. Stuetz , V. Pankajakshan, F. Autrusseau, A. Uhl, H. Hofbauer, “Subjective and Objective Quality Assessment of Transparently Encrypted JPEG2000 Images”, in the 12th ACM Workshop on Multimedia and Security, ACM-MMSEC’2010, pp. 247-252, Rome, Italy, Sept. 9-10 2010.
- [40] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, P. Morvan, J. Stauder, “ How Modern Displays Push Conventional Colorimetry to its Limit”, Colour in art, science, design, conservation, research, printmaking, digital technologies, textiles (CREATE), Gjovik, Norway, 2010.
- [41] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, P. Morvan, J. Stauder, “ Toward Reducing Observer Metamerism Issue in Industrial Applications : Colorimetric Observer Categories and Observer Classification”, in the eighteenth Color Imaging Conference, Color Science and Engineering Systems, Technologies, and Applications, CIC18, San Antonio, Texas, Nov. 8-12, 2010.

- [42] V. Pankajakshan and F. Autrusseau, "A Multi-purpose Objective Quality Metric for Image Watermarking", in IEEE International Conference on Image Processing, ICIP'2010, pp. 2589-92, Hong Kong, Sept 26-29, 2010.
- [43] F. Autrusseau, S. David and Vinod Pankajakshan, "A Subjective Study of Visibility Thresholds for wavelet domain watermarking", in IEEE International Conference on Image Processing, ICIP'2010, pp. 201-204, Hong Kong, Sept. 26- 29, 2010.
- [44] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, J. Stauder, P. Morvan, "Modern displays : Why we see different colors, and what it means ?", 2nd European Workshop on Visual Information Processing, EUVIP'2010, Paris, France, July, 5th-7th, 2010.
- [45] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, P. Morvan, J. Stauder, "A color matching experiment using two displays : design considerations and pilot test results", in Proceedings of the Fifth European Conference on Color in Graphics, Imaging and Vision, CGIV'2010, Finland, 2010.
- [46] M. Carosi, V. Pankajakshan, F. Autrusseau, "Toward a simplified perceptual quality metric for watermarking applications", Proceedings of the SPIE Electronic Imaging, 7542, San Jose, CA, USA, January 2010.
- [47] C. Zhang, J. Dong, J. Li, and F. Autrusseau, "A New Information Hiding Method for Image Watermarking Based on Mojette Transform", in the Second International Symposium on Networking and Network Security, ISNNS, pages 124-128, Jinggangshan, China, 2-4 April 2010.
- [48] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, J. Stauder, P. Morvan, "Study of Observer Variability in Modern Display Colorimetry : An Analysis of CIE 2006 Model", 11th Congress of the International Colour Association (AIC), Sydney, Australia, 27 Sept - 2 Oct 2009.
- [49] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, J. Stauder, P. Morvan, "Study of Observer Variability in Modern Display Colorimetry : Comparison of CIE 2006 Model and 10 degree Standard Observer", 11th Congress of the International Colour Association (AIC) 2009, Sydney, Australia, 27 Sept - 2 Oct 2009.
- [50] C. Strauss, F. Pasteau, F. Autrusseau, M. Babel, L. Bedat, O. Deforges, "Subjective and Objective Quality Evaluation of LAR coded art images", IEEE Intl. Conf. on Multimedia & Expo, ICME 2009, pp. 674-677, New York, USA, June 28 - July 3 2009.
- [51] P. Jia, J. Dong, L. Qi, F. Autrusseau, "Directionality Measurement and Illumination Estimation of 3D Surface Textures by Using Mojette Transform", 19th International Conference on Pattern Recognition, ICPR'08, Tampa, FL, USA, 2008.
- [52] A. Kingston, B. Parrein, F. Autrusseau, "Redundant Image Representation via Multi-Scale Digital Radon Projection", IEEE International Conference on Image Processing, ICIP'2008, pp. 2920-24, San Diego, CA, USA, Oct12-15, 2008.
- [53] S. Tourancheau, F. Autrusseau, Z. M. Parvez Sazzad, Y. Horita, "Impact of subjective dataset on the performance of image quality metrics", IEEE International Conference on Image Processing, ICIP'2008, pp. 365-368, San Diego, CA, USA, Oct12-15, 2008.

- [54] M.Bouchakour, G.Jeannic, F. Autrusseau, “JND Mask Adaptation for Wavelet Domain Watermarking”, IEEE Intl. Conf. on Multimedia & Expo, ICME 2008, pp 201-204, Hannover, Germany, June 23-26 2008.
- [55] D. Pitzalis, G. Aitken, F. Autrusseau, M. Babel, F. Cayre, W. Puech, “TSAR : Secure Transfer of High Resolution Art Images”, Electronic Imaging and the Visual Arts, EVA’08, Firenze, Italy, 2008.
- [56] W.Chen, F. Autrusseau, P. Le Callet, “ A Error-Propagation-Free Perceptual Watermarking Algorithm for H.264/AVC encoded video”, 1st Workshop on Visual Signal Processing and Analysis (VISPA), Paris, France, Nov. 12-13 2008.
- [57] F. Autrusseau, P. Le Callet, A. Ninassi, “A study of content based watermarking using an advanced HVS model”, IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 485-488, Nov 26-28, Kaohsiung, Taiwan, 2007.
- [58] A. Kingston, S. Colosimo, P. Campisi, F. Autrusseau, “ Lossless Image Compression and Selective Encryption Using a Discrete Radon Transform” IEEE International Conference on Image Processing, ICIP’07, vol. 4, pp. 465-468, San Antonio, TX, USA, 2007.
- [59] E. Marini, F. Autrusseau, P. Le Callet, P. Campisi, “ Evaluation of standard watermarking techniques”, in SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX, San Jose , USA, Jan. 2007.
- [60] F. Autrusseau, B. Parrein, M. Servieres, “ Lossless Compression Based on a Discrete and Exact Radon Transform : A Preliminary Study”, IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP’06, (2), pp. 425-428, Toulouse, France, 2006.
- [61] A. Ninassi, P. Le Callet, F. Autrusseau, “ Pseudo No Reference image quality metric using perceptual data hiding”, SPIE Human Vision and Electronic Imaging, Volume 6057, pp. 146-157, San Jose, CA, USA, January 2006 .
- [62] P. Evenou, F. Autrusseau, T. Hamon, “ Secure Distributed Storage based on the Mojette transform”, in NOTERE 2006, Hermes-Lavoisier editors, pp. 161-170, Toulouse, France, 2006.
- [63] V. Ricordel, F. Autrusseau, W. Dupuy, D. Barba, “ 1D-mosaics grouping using lattice vector quantization for a video browsing application”, 4th international workshop on Content-based multimedia indexing, CBMI’05, Riga, Latvia, June 21-23 2005.
- [64] F. Autrusseau, S. K. Shevell, “ Temporal nulling of induction from spatial patterns modulated in time”, 18th symposium of the International Color Vision Society ICVS’05, Lyon, France, May 2005.
- [65] A. Manoury, F. Autrusseau, A. Saadane, “Optimisation selon un critère psychovisuel de la qualité d’insertion d’un schema de tatouage”, TAIMA’03, Traitement et Analyse de l’Information : Méthodes et Applications, 2003, Hammamet, Tunisie, Oct. 1st - 3rd 2003.
- [66] F. Autrusseau, JP. Guédon, “A joint multiple description-encryption image algorithm”, IEEE International Conference on Image Processing, ICIP’03, (3), pp. 269-272, Barcelona, Spain, 2003.

- [67] F. Autrusseau, JP. Guédon, “Perceptual image watermarking using a secure Mojette transmission scheme”, 4th Workshop COST 276, Transmitting processing and watermarking multimedia contents, Bordeaux, March 31st - April 1st 2003.
- [68] F. Autrusseau, JP. Guédon, Y. Bizais, “ Watermarking and cryptographic schemes for medical imaging”, SPIE Medical Imaging, Image processing, vol. 5032-105, pp. 958-965, San Diego CA, USA, 15-20 February 2003.
- [69] F. Autrusseau, JP. Guédon, “ Image Watermarking in the Fourier Domain Using the Mojette Transform”, 14th IEEE International Conference on Digital Signal Processing (DSP2002), vol. II, pp. 725-728, Santorini Greece, July 2002.
- [70] F. Autrusseau, JP. Guédon, “ Image watermarking for copyright protection and data hiding via the Mojette transform”, SPIE Electronic Imaging, Security and Watermarking of Multimedia Contents IV, vol. 4675, pp. 378-386, San Jose, CA, January 2002.
- [71] JP. Guédon, N. Normand, P. Verbert, B. Parrein, F. Autrusseau, “ Load-balancing and scalable multimedia distribution using the Mojette transform”, Internet Multimedia Management Systems II, ITCOM, vol. 4519, pp. 226-234, Denver USA, August 2001.
- [72] F. Autrusseau, A. Saadane, D. Barba, A Psychovisual approach for watermarking ”, SPIE Electronic Imaging, Security and Watermarking of Multimedia Contents III, vol. 4314, pp. 495-504, San Jose, CA, Jan. 2001.
- [73] F. Autrusseau, A. Saadane, D. Barba, “ Psychovisual selection of auspicious sites for watermarking”, SPIE, Image Processing, Image Quality, Image Capture Systems Conference, PICS’2000, pp. 326-329, Portland, OR, 2000.

Abstracts dans des conférences internationales

- [74] G. Frangi, G. Kerckhofs, J. Boulestreau, F. Autrusseau, J. Veziers, J. Guicheux, G. Penel, C. Chauveau, X. Prieur, L. Beck, S. Beck-Cormier, “Role of PiT2 in Bone and Adipose Tissues Inter-Communication?”, in 4th Intl. Meeting on Bone Marrow Adiposity, BMA 2018, Lille, France, Aug. 29-31 2018.
- [75] N. Bon, J. Boulestreau, F. Autrusseau, G. Penel, C. Chauveau, L. Beck, S. Beck-Cormier, “PiT2 deficiency results in skeletal phenotype associated with alteration of adipose tissues”, in 3rd International Meeting on Bone Marrow Adiposity, BMA 2017, Lausanne, Aug. 31 - Sept. 1st 2017.
- [76] A. Henry, P. Colombier, M. Dutilleul, F. Autrusseau, J. Guicheux, A. Camus, “Understanding the mechanisms of notochord maturation and transformation during intervertebral disc morphogenesis in the mouse”, Gordon Research Conference on Cartilage Biology & Pathology, Lucca, Italy, April 2-7, 2017.

- [77] F. Autrusseau, S. Shevell, “Chromatic and Wavefront Aberrations : Estimates of Typical and Individual Variation in Retinal Images”, in the Thirty-Fifth Annual Interdisciplinary Conference, Teton Village, WY, USA, January 24-29, 2010.
- [78] F. Autrusseau, S. K. Shevell, “ Temporal nulling of induction from spatial patterns modulated in time”, in European Conference on Visual Perception, ECVP’04, Budapest Hungary, 22-26 August 2004.

Bases de données subjectives

- [79] H. Hofbauer and F. Autrusseau and A. Uhl, “University of Salzburg Encryption and Quality Evaluation Database, USEEQ DB”, <http://wavelab.at/sources/USEEQ/>, 2020.
- [80] H. Hofbauer and F. Autrusseau and A. Uhl, “University of Salzburg Encryption Evaluation Database, USEE DB”, <http://wavelab.at/sources/USEE/>, 2019.
- [81] F. Autrusseau and M. Urvoy, “2AFC evaluation of Fourier watermarked images and documents”, <http://www.polytech.univ-nantes.fr/autrusseau-f/DFTWmking/>, 2014.
- [82] F. Autrusseau and D. Goudia, “2AFC evaluation of wavelet watermarking in the smooth http://www.polytech.univ-nantes.fr/autrusseau-f/Databases/2AFC_WaveletWmking/, 2013.
- [83] T. Stuetz and F. Autrusseau, “H.264 CAVLC Structure preserving watermarking”, available online : <ftp://ftp.ivc.polytech.univ-nantes.fr/IRCCyNIVCH264WatermarkingStructurePreserving/>, 2011.
- [84] F. Autrusseau, T. Stuetz and V. Pankajakshan, “Subjective quality assessment of selective encryption techniques”, available online : <http://www.polytech.univ-nantes.fr/autrusseau-f/Databases/SelectiveEncryption/>, 2010.
- [85] F. Autrusseau, M. Babel, “Subjective Quality assessment of LAR coded art images”, available online : <http://www.polytech.univ-nantes.fr/autrusseau-f/Databases/LAR/>, 2009.
- [86] F. Autrusseau, P. Bas “Subjective Quality assessment of the Broken Arrows watermarking technique”, available online : <http://www.polytech.univ-nantes.fr/autrusseau-f/Databases/BrokenArrows/>, 2009.
- [87] F. Autrusseau, P. Meerwald, “DT-CWT versus DWT watermark embedding”, available online : <http://www.polytech.univ-nantes.fr/autrusseau-f/Databases/MeerwaldDB/>, 2009.

- [88] F. Autrusseau, “Fourier sub-band embedding”, available online : <http://www.polytech.univ-nantes.fr/autrusseau-f/Databases/FourierSB/>, 2008.
- [89] E. Marini, F. Autrusseau and P. Le Callet, “Evaluation of standard watermarking techniques”, available online : <http://www.polytech.univ-nantes.fr/autrusseau-f/Databases/Enrico/>, 2007.
- [90] P. Le Callet, F. Autrusseau, “Subjective quality assessment IRCCyN/IVC database”, available online : http://ivc.univ-nantes.fr/en/databases/Subjective_Database/, 2005.

Rapports et publications nationales

- [91] G. Frangi, G. Kerckhofs, J. Boulestreau, F. Autrusseau, J. Veziers, J. Guicheux, G. Penel, C. Chauveau, X. Prieur, L. Beck and S. Beck-Cormier, “Altered Bone Marrow adiposity in PIT-2 deficient mice”, Journées Françaises de Biologie des Tissus Minéralisés (JFBTM), 5 au 7 Juin 2019.
- [92] G. Frangi, J. Boulestreau, F. Autrusseau, J. Veziers, J. Guicheux, G. Penel, C. Chauveau, L. Beck and S. Beck-Cormier, “PIT2 Deficiency Results in Skeletal Phenotype Associated with Alteration of Bone Marrow Adipose Tissue”, Journées Françaises de Biologie des Tissus Minéralisés (JFBTM), 21 au 23 Mars 2018.
- [93] A. Maillat, F. Rivard, F. Autrusseau, R. Lehn, M. Gelgon, P. Bruneau, “Safim@ge”, 5eme Workshop Interdisciplinaire sur la Sécurité Globale, Troyes, France, 25-26 Janvier 2011.
- [94] A. Maillat, F. Rivard, F. Autrusseau, R. Lehn, M. Gelgon, P. Bruneau, “Safim@ge” 2009, WISG’09 - Workshop Interdisciplinaire sur la Sécurité Globale, Jan 27-28, 2009, Troyes, France.
- [95] A. Maillat, F. Rivard, R. Lehn, M. Gelgon, F. Autrusseau, “Safim@ge : projet de recherche”, WISG’08 - Workshop Interdisciplinaire sur la Sécurité Globale, Jan 29-30, 2008, Troyes, France
- [96] A. Maillat, F. Rivard, F. Autrusseau, R. Lehn, M. Gelgon, P. Bruneau, “Safim@ge”, 2008 NEM Summit, Networked & Electronic Media, Oct 13-15, 2008, Saint Malo, France
- [97] A. Saadane, F. Autrusseau, “ Aquamarquage des documents audiovisuels pour leur transmission, diffusion, circulation, distribution en toute sécurité sur les réseaux hétérogènes de telecommunication. Analyse et specifications”, RNRT AQUAMARS, Septembre 1999, Report D 2.1 (47 pages).

- [98] A. Saadane, F. Autrusseau, “ Aquamarquage des documents audiovisuels pour leur transmission, diffusion, circulation, distribution en toute sécurité sur les réseaux hétérogènes de telecommunication. Modélisation psycho-visuelle pour le tatouage des images, RNRT AQUAMARS, October 1999, Report D 2.2.1 (23 pages).

Education

- [99] F. Autrusseau, “Modélisation psychovisuelle pour le tatouage des images”, Ph. D. Thesis, Electronics, Ecole polytechnique de l’Université de Nantes, (244 pages), November 7 2002.
- [100] F. Autrusseau, “Modélisation psychovisuelle pour l’aquamarquage des images”, MSc. thesis, Graduate studies dissertation, IRESTE, Nantes (in French), Sept. 1999.

Introduction

Contents

2.1	Positionnement de mon travail	21
2.1.1	Les thématiques de recherche	21
2.1.2	Résumé succinct des contributions	24
2.2	Contributions personnelles	25
2.2.1	La perception visuelle appliquée à l'image	25
2.2.2	La perception des couleurs	27
2.2.3	La transformation Mojette	28
2.2.4	L'imagerie médicale et biologique	29

2.1 Positionnement de mon travail

Dans cette section, je vais introduire mes différentes thématiques de recherche, je présenterai succinctement mes principales contributions au cours de ces dernières années. Nous verrons que j'ai eu l'opportunité de contribuer sur quatre axes thématiques. Par la suite, quatre chapitres distincts seront dédiés à la présentation dans le détail de ces différents champs de recherche.

Au cours de mes années d'activité dans la recherche, j'ai formé et encadré plusieurs étudiants. C'est une partie de mon travail que j'affectionne tout particulièrement. Tout au long de ces années, j'ai obtenu divers financements pour conduire mes recherches, j'ai pu recruter plusieurs personnes, que j'ai parfois co-encadré, et parfois encadré seul.

2.1.1 Les thématiques de recherche

Mes travaux s'articulent autour du traitement et de l'analyse des images numériques. Depuis mes débuts dans la recherche, lors de mon stage de D.E.A. (actuellement Master 2) à l'IRESTE à Nantes en 1998-1999 en passant par ma thèse de Doctorat soutenue en 2002 à Polytech'Nantes, mon post-doctorat à l'Université de Chicago en 2003-2004, mon poste d'Ingénieur de Recherche à Polytech'Nantes entre 2004 et 2016 ou encore ma mise à disposition en santé au CHU de Nantes (UFR d'Odontologie) de 2016 à 2020, mes différentes activités ont principalement gravité autour de la perception des images. J'ai aussi eu l'opportunité de prospecter

d'autres champs thématiques, tels que la compression des images, le chiffrement sélectif ou encore l'analyse de textures. Parmi les différentes thématiques de recherche sur lesquelles j'ai pu me pencher il en est une qui est restée dominante et continue encore à ce jour à me taquiner, c'est la perception visuelle.

Lors de mon arrivée dans le laboratoire SEI (Systèmes Electroniques et Informatiques), devenu Institut de Recherche en Communication et Cybernétique de Nantes (IRCCyN) par la suite, et à présent Laboratoire des Sciences du Numérique de Nantes (LS2N) la modélisation psychovisuelle portée par Dominique Barba prenait une place très importante dans l'équipe Image et Vidéo-Communication (IVC). Outre les travaux de Dominique Barba sur la perception humaine, une toute récente thématique venait aussi d'émerger dans l'équipe IVC, la transformation Mojette, portée par Jeanpierre Guédon. C'est à la toute fin des années 1990 que la thématique de la dissimulation de l'information a fait son entrée dans les thématiques de recherche du laboratoire grâce au projet ANR *Aquamars*. C'est donc par mon stage de DEA que la protection des droits d'auteurs a commencé à occuper quelques esprits au sein du laboratoire. Le tatouage des images (appelé watermarking en Anglais) est une discipline du traitement des images qui a pour objectif d'insérer des données cachées dans des contenus multimédias (le plus souvent des images ou vidéos numériques). L'information insérée (watermark en Anglais, ce qui signifie filigrane) peut se présenter sous différentes formes, ça peut tout aussi bien être un bruit ou un signal intelligible décodable. Dans ce manuscrit je pourrai me référer au domaine du tatouage, de la protection des droits d'auteurs (ou de copyright) ou encore au watermarking, tous ces termes ont la même signification. La dissimulation de l'information est un terme plus générique qui peut englober d'autres applications comme nous le verrons pas la suite.

Mon stage de DEA entrait dans le cadre du projet Aquamars [Saadane 1999a, Saadane 1999b] et avait pour titre "Modélisation Psychovisuelle pour l'aquamarquage des images" [Atrousseau 1999], dans la continuité de ce stage, j'ai poursuivi en thèse, toujours sur ce même projet, mais tandis que le stage de DEA portait uniquement sur les critères de vision, la thèse de doctorat englobait la modélisation psychovisuelle et la transformation Mojette [Atrousseau 2002c].

Comme nous le verrons par la suite, je me suis donc attaché à faire cohabiter cette nouvelle thématique du tatouage des images avec celles déjà bien présentes dans l'équipe, les modèles perceptuels et la transformation Mojette. J'ai donc obtenu ma thèse de doctorat en 2002 sur le sujet "Modélisation psychovisuelle pour le tatouage des images", suite à ma thèse, j'ai effectué un stage post-doctoral au sein du *Visual Science Laboratories* de l'Université de Chicago en 2003-2004, sous la direction du Professeur Steve Shevell.

Lors de ce post-doctorat, je me suis penché sur les sciences visuelles et cognitives, je me suis de fait éloigné un peu de la perception des images pour m'attacher à comprendre la perception des couleurs. J'ai notamment contribué à deux principaux travaux. J'ai tout d'abord cherché à comprendre le fonctionnement des champs récepteurs neuronaux responsable de la discrimination des couleurs dans le cortex cérébral. Ces champs neuronaux, assimilables à des différences de Gaussiennes 2D

peuvent entraîner des illusions chromatiques. J’ai étudié leur comportement dans le cas d’une réponse temporelle [Autrusseau 2006b]. Par ailleurs, le second projet qui a occupé une grande partie de mon post-doc portait sur la modélisation optique de l’œil humain et notamment les aberrations chromatiques et de front d’onde. Grâce à une collaboration avec le Pr. Larry Thibos du département d’optométrie de l’Université de l’Indiana, nous avons pu modéliser tout le processus optique de l’œil, ainsi que l’absorption de la lumière par les cônes sur la rétine (récepteurs permettant la vision des couleurs) [Autrusseau 2011b].

Suite à cette expérience de post-doctorat à l’étranger, je suis revenu à l’École Polytechnique de l’Université de Nantes sur le poste d’Ingénieur de Recherche que j’occupe encore à ce jour. Pendant douze années (2004 à 2016) j’ai porté la thématique de la sécurité dans l’équipe IVC. Bien que le watermarking (tatouage des images) eût été une thématique très porteuse dès son apparition dans le début des années 1990, au fil des ans, l’intérêt pour la recherche sur la dissimulation de l’information s’est maintenu sur les applications de la stéganographie, mais s’est étioilé petit à petit concernant les problématiques de protection des droits d’auteurs. Effectivement, la recherche initialement fortement fondamentale sur la protection des copyrights numériques est maintenant largement transférée vers l’industrie privée. La société Digimarc aux États Unis commercialise de nombreuses suites logicielles permettant l’insertion de copyrights. La recherche dans cette discipline reste maintenant plus marginale.

Tout au long de mon parcours, à plusieurs reprises, je me suis intéressé à l’imagerie médicale, au début via des solutions de chiffrement des images au format “Digital imaging and communications in medicine” (DICOM), puis petit à petit, je me suis tourné vers la recherche en analyse d’images médicales. J’ai tout d’abord participé à des projets portant sur l’analyse de texture [Dong 2012] principalement pour des applications liées à la caractérisation osseuse [Bléry 2014, Guédon 2012, Crauste 2015]. Le domaine de l’imagerie médicale et biologique a donc assez tôt suscité mon intérêt, mais la recherche portant sur la dissimulation de l’information dans le médical est restée assez marginale.

C’est donc en 2016 que j’ai souhaité réorienter mes thématiques de recherches et obtenu une mise à disposition vers le Laboratoire d’Ingénierie Ostéo-Articulaire et Dentaire (LIOAD) (devenu depuis Janvier 2017 : Regenerative Medicine and Skeleton (RMeS)). C’est donc au RMeS que j’ai pu travailler sur divers outils d’analyse des images médicales et biologiques. J’ai intégré la plateforme “Service Commun de Microscopie électronique, Microcaractérisation et Morpho-histologie-imagerie fonctionnelle (SC3M)”, dont l’un des axes porte sur l’analyse des images biologiques et médicales. Au sein du laboratoire RMeS, sur la plateforme SC3M, j’ai eu la chance de pouvoir monter plusieurs projets. L’un de ces projets avait pour objectif la génération automatique de schémas dentaires à partir de radiographies panoramiques, ce projet, financé par l’Appel à Projet Interdisciplinaire de l’Université de Nantes a été conduit entre 2017 et 2018 et a pu aboutir sur la création d’une entreprise. Natalia De La Fournière qui avait été recrutée en stage de M2 Signaux et Images en Biologie et Médecine (SIBM) sur ce projet, a créé la société Artefakt-AI en 2019

sur cette même problématique de reconnaissance des dents. Le second projet sur lequel j'ai pu contribuer était en collaboration avec des neuro-radiologues de l'Hôpital Nord à Nantes, affiliés à l'Institut du Thorax : le Docteur Romain Bourcier et le Professeur Hubert Desal. Ma contribution dans ce projet, aidé par Anass Nouri, recruté en post-doctorat, portait sur l'analyse d'images cérébrales (micro-scanner et Imagerie par Résonance Magnétique (IRM)) dans le but de prédire la formation des anévrismes intra-crâniens. Finalement, un troisième projet, en collaboration avec les services de rhumatologie et d'odontologie du CHU de Nantes porte sur l'analyse de la densité osseuse sur des images de "Cone Beam Computed Tomography" (CBCT) en vue de prédire le risque ostéoporotique. C'est le projet *Tomosteo*. Lors de cette collaboration avec les Professeurs Yves Maugars (rhumatologue) et Assem Soueidan (chirurgien dentiste), j'ai pu bénéficier de bourses de stage de M2 (recrutement de Clément Morin) début 2020.

2.1.2 Résumé succinct des contributions

Mes principales contributions dans les différentes thématiques de recherche que j'ai pu aborder pourraient se résumer comme suit :

- J'ai cherché à connecter deux communautés qui malgré l'intérêt commun qu'elles devraient avoir, ne parviennent cependant pas à s'associer. La communauté de la dissimulation des données et la communauté de la perception humaine.
- En lien avec cette problématique du tatouage, j'ai pu proposer une nouvelle méthode de détection des watermarks basée sur l'inter-corrélation. L'exploitation de la méthode statistique de détection d'outliers (le test de Grubbs) s'applique parfaitement à la détection des watermarks. De plus, elle permet d'éviter une phase expérimentale très coûteuse qui nécessite des tests extensifs avec un grand nombre d'images afin de déterminer expérimentalement le seuil de détection.
- Encore une fois, sur la tentative de réconcilier les deux communautés (tatouage et vision humaine), je me suis attaché à tester les métriques de la littérature et proposer des métriques de qualité qui soient adaptées à la thématique de la dissimulation de l'information.
- L'étude de la perception des couleurs m'a amené à proposer un modèle optique de l'oeil. Nous avons ainsi pu construire un observateur standard (du point de vue de la diffraction optique), nous permettant de mieux comprendre les aberrations chromatiques.
- Les divers travaux que j'ai pu conduire sur la transformation Mojette m'ont amenés à proposer une méthode de compression sans perte basée sur la Mojette, qui jusqu'alors était principalement exploitée pour la redondance qu'elle apportait, de plus, en lien avec cette technique de compression, j'ai pu développer diverses approches de chiffrement sélectif exploitant l'instabilité de la rétro-projection Mojette.
- Finalement, j'ai récemment contribué sur un projet en santé portant sur la

détection des anévrismes intra-crâniens, nous avons ainsi pu proposer une méthode de caractérisation des bifurcations que forment les artères vasculaires cérébrales. Cette caractérisation nous a permis de prédire le risque de formation d'un anévrisme sur une bifurcation en fonction de sa configuration géométrique.

Voici pour un bref résumé des principales contributions de mes travaux. Dans la section suivante, je vais présenter dans le détail chacune des thématiques que j'ai pu aborder.

2.2 Contributions personnelles

Je présente donc dans cette section de façon un peu plus détaillée quelques uns de mes différents travaux, je donnerai une description exhaustive des nouveautés apportées par les approches proposées. Je détaille brièvement le cadre dans lequel ces travaux se sont inscrits, les projets de recherche, et les encadrements et collaborations qui y sont associés. Finalement, je présenterai les conclusions de chaque projet.

2.2.1 La perception visuelle appliquée à l'image

Depuis que l'exploitation des images numériques s'est démocratisée, des chercheurs ont eu pour objectif de comprendre la perception humaine desdites images. Tout d'abord, les images numériques occupent une place significative sur les supports de stockage, il était donc souhaitable de réduire, dans la mesure du possible leur taille, mais sans toutefois dégrader visuellement le contenu. Une fois l'avènement des réseaux informatiques, et notamment d'Internet, il était une fois de plus nécessaire de réduire les images et vidéos afin de réduire leur impact sur la bande passante. Des nombreux travaux ont donc porté sur l'étude de la perception des signaux numériques. Il a alors été nécessaire de comprendre le fonctionnement du Système Visuel Humain (SVH) afin de mieux pouvoir le modéliser et appliquer toute sorte de traitements sur les contenus multimedia.

Une partie de mes travaux depuis le tout début de ma thèse de doctorat en 1999 a porté sur la compréhension du SVH et l'application de ses modèles pour diverses applications, telles que la dissimulation de l'information, la constitution de métriques de qualité, ou encore la perception des couleurs. La dissimulation de l'information (*Data Hiding* en Anglais) est un terme générique qui englobe plusieurs champs de recherche. Nous trouvons principalement des travaux portant sur la protection des droits d'auteurs des contenus multimedia, la stéganographie, qui est l'art de dissimuler de l'information dans un support, généralement dans le but de transmettre des informations secrètes. La Stéganographie est principalement utilisée pour des applications militaires, la robustesse de l'information insérée importe peu, mais l'expéditeur du message doit s'assurer que même si le message est détecté il ne pourra pas être lu, des techniques de chiffrement sont alors utilisées. S'agissant de la problématique de la protection des droits d'auteurs, le message inséré prouve

les droits de propriété du contenu, il est donc nécessaire d'empêcher le piratage des images/vidéos protégées. Le copyright doit donc être robuste face à tout type d'attaques malveillantes ou même non-intentionnelles. Il est important de comprendre ici qu'un compromis doit absolument être trouvé entre l'invisibilité et la robustesse. Effectivement, plus la marque est insérée de façon robuste, plus l'invisibilité risque d'en pâtir. Pendant de nombreuses années, j'ai cherché à étudier conjointement la dissimulation de l'information et la modélisation du SVH. De trop nombreux travaux de la littérature présentent des méthodes de tatouage des images sans fournir aucune preuve sur l'invisibilité de l'information insérée. Il s'avère que très souvent les copyrights supposément invisibles peuvent être perçus à l'œil nu. Il me paraissait donc crucial de fournir à la communauté du *Data Hiding* des outils permettant de garantir l'invisibilité. J'ai principalement orienté mes travaux suivant deux axes. Dans un premier temps, j'ai cherché à proposer des méthodes de dissimulation de l'information qui optimisent le compromis entre l'invisibilité et la robustesse. Dans le cadre du projet *e-Map*, j'ai ainsi pu proposer diverses méthodes opérant soit dans le domaine de Fourier [Autrusseau 2007a, Urvoy 2014b], soit dans les ondelettes [Bouchakour 2008, Autrusseau 2010a, Autrusseau 2011a]. Avec Dalila Goudia (embauchée via un contrat de maturation et sur un poste d'ATER), nous avons aussi cherché à optimiser le tatouage des images High Dynamic Range (HDR). Nous avons proposé un tatouage adaptatif opérant dans la Discrete Wavelet Transform (DWT). Ces travaux étaient la suite et plus précisément l'adaptation aux images HDR, d'une méthode développée avec Vinod Pankajakshan et Patrizio Campisi qui avait pour objectif de déterminer précisément le seuil de visibilité d'une méthode de tatouage dans les ondelettes. La seconde thématique que j'ai cherchée à développer au sein de l'équipe IVC autour de la perception humaine est l'implémentation de métriques de qualité. Les métriques de qualité sont des outils logiciels dont l'objectif est d'estimer la qualité perçue des images et vidéos traitées. Nous verrons en détail cette problématique dans le chapitre 3. Les métriques de qualité ont tout d'abord vu le jour dans le but de valider la qualité des images et vidéos ayant subi des algorithmes de compression. Parmi les métriques les plus largement reconnues et utilisées, nous pouvons citer le Visible Difference Predictor (VDP) de Scott Daly [Daly 1993]. Concernant mes travaux sur les métriques de qualité, j'ai avant tout cherché à déterminer quelles métriques de la littérature étaient les plus appropriées pour les problématiques de dissimulation de l'information [Le Callet 2008]. Cette étude a en partie été conduite en collaboration avec Patrizio Campisi de l'Université de Rome. La plus grande majorité des travaux portant sur la protection des droits d'auteurs utilise deux métriques très répandues : le Peak Signal to Noise Ratio (PSNR) et Structural Similarity Index (SSIM) [Wang 2004]. Or, il s'avère que j'ai pu découvrir par mes travaux qu'elles comptent parmi les pires métriques lorsque les distortions se trouvent être proches du seuil de visibilité. Parfois, le PSNR et SSIM peuvent donner une idée de la perception des distortions sur certaines bases de données subjectives. Souvent, cela se produit sur des bases ayant été mal calibrées. Cependant, dans la plupart des cas, et en tout cas plus particulièrement sur des bases ayant rapport à la dissimulation de l'information, ces deux métriques échouent pour la prédiction

de la qualité, et ceci car les distortions sont théoriquement très proches du seuil de visibilité. Ces métriques sont bien trop simplistes et ne considèrent absolument aucun critère avancé de la vision humaine. De ce fait, en observant les lacunes des différentes métriques de la littérature, avec l’aide notamment de Vinod Pankajshankhan, recruté en post-doctorat au cours des projets *e-Map* et *StreamMaster* et de Maurizio Carosi stagiaire de M2 Erasmus, nous avons proposé différentes métriques particulièrement adaptées à la problématique du tatouage des images [Carosi 2010, Pankajakshan 2010] et du chiffrement sélectif [Stütz 2010]. Bien évidemment, ces différents travaux portant sur la perception humaine ont tous été validés par des tests subjectifs [Atrousseau 2013d, Atrousseau 2013b, Atrousseau 2009b, Marini 2007a, Atrousseau 2008, Stütz 2011, Le Callet 2005, Atrousseau 2009a, Atrousseau 2009c, Atrousseau 2009c, Atrousseau 2010b]

Très récemment, j’ai aussi eu l’opportunité de travailler sur une thématique très peu étudiée : la reconnaissance perceptuelle des images. J’entends ici par “reconnaissance”, l’identification du contenu. Effectivement, comme nous le verrons, il existe pléthore de métriques de qualité permettant d’estimer le jugement humain sur des images dégradées, cependant, à ma connaissance, il n’existe pas à ce jour d’outil capable d’estimer avec précision si une image sera reconnue par un être humain. C’est la notion de “Recognizability” que nous avons introduite dans de récents travaux avec mes collègues de l’Université de Salzbourg. Cette étude rentrait dans le cadre du projet *ViSecMetrics* financé par le *FWF* en Autriche et porté par Andreas Uhl. Pour étudier l’identification perceptuelle des images, il est nécessaire de mettre en place un nouveau protocole de tests subjectifs, les protocoles existants n’étant pas destinés à ce type de tâches. Nous avons donc proposé dans [Hofbauer 2018] un protocole adapté à ce nouveau type de tâche. Des travaux en cours portent sur l’évaluation des métriques dans ce contexte, mais nous proposons également des alternatives plus fiables permettant de juger du pouvoir de reconnaissance [Hofbauer 2020a].

Nous verrons dans le Chapitre 3 les différents travaux que j’ai pus conduire. J’expliquerai dans le détail les trois problématiques basées sur la perception que j’ai pues étudier.

2.2.2 La perception des couleurs

Au cours de mes travaux de post-doctorat et à travers l’encadrement de la thèse CIFRE d’Abhijit Sarkar (Thèse CIFRE en collaboration avec Technicolor R&D à Rennes), j’ai eu l’opportunité de travailler sur la perception des couleurs. Lors de mon post-doc, j’ai abordé le sujet du point de vue de la recherche fondamentale, nous avons cherché à comprendre les mécanismes de perception des couleurs en étudiant des champs récepteurs corticaux, ou encore en analysant la décomposition de la lumière à travers un système optique et l’impact de cette diffraction sur la perception des couleurs. Lors de l’encadrement de la thèse de doctorat d’Abhijit [Sarkar 2011b], nous avons travaillé sur la variabilité inter-observateurs dans une problématique de visualisation des couleurs. Ces travaux de recherche étaient, de par leur nature (thèse CIFRE Technicolor), plus orientés vers des applications de la perception de

la couleur sur des écrans modernes. La compréhension de la perception des couleurs est un champ de recherche captivant, dans lequel nous ne cessons de découvrir de nouvelles propriétés du SVH.

Mes travaux au Visual Science Laboratories (VSL) de l'Université de Chicago ont porté sur deux volets distincts, d'une part, j'ai étudié des champs récepteurs cérébraux antagonistes permettant de générer un métamérisme sous certaines conditions de visualisation. En particulier, nous avons cherché à comprendre quelle modulation des couleurs était nécessaire pour obtenir une compensation de l'induction chromatique lors d'une variation temporelle des stimuli chromatiques. Nous avons montré que l'annulation temporelle de l'effet d'induction chromatique pouvait être obtenue via un phénomène d'assimilation ou de contraste en fonction de l'antagonisme des champs récepteurs, soit en fonction de l'ordre de contiguïté des chrominances induisant le métamérisme [Autrusseau 2006b].

La seconde partie de mes travaux de post-doc a porté sur la modélisation optique de l'oeil humain, en particulier, nous nous sommes attachés à fournir un modèle d'observateur chromatique standard (du point de vue des imperfections optiques de l'oeil). Cette étude nous a permis de comprendre que la Longitudinal Chromatic Aberrations (LCA) ne pouvait en aucun cas être modélisée seule, et qu'elle devait absolument s'ajouter aux aberrations de front d'onde ayant un impact beaucoup plus significatif sur la perception des contrastes chromatiques [Autrusseau 2011b].

Lors de la thèse d'Abhijit Sarkar, nous avons étudié le métamérisme observateur. Nous avons donc cherché à analyser la variabilité individuelle dans la perception des couleurs. Nous avons pu catégoriser plusieurs observateurs notamment en fonction de leur tranche d'âge. Au cours de ces travaux, nous avons cherché à obtenir des correspondances de couleurs sur deux écrans aux propriétés (densités spectrales de puissance) bien différentes (un écran LCD et un écran CRT). Cette étude s'est faite en deux temps, nous avons tout d'abord testé extensivement les modèles fournis par la Commission Internationale de l'Éclairage (CIE) puis, par la suite nous avons conduit un test subjectif nous permettant d'établir les clusters d'observateurs.

2.2.3 La transformation Mojette

La transformation Mojette est née de l'imagination de Jeanpierre Guédon au cours des années 1990 [Guédon 1995]. Cette transformation peut être vue comme une version discrète de la transformée de Radon. Elle opère donc une ou des projections de l'image selon des directions données. Chacune des projections est obtenue par simple sommation des pixels traversés. L'inversion de la transformation Mojette a vu le jour en 1996 [Normand 1996a], et depuis, d'autres algorithmes plus rapides et plus robustes ont été proposés [Normand 2006a]. Pendant les quinze premières années de la Mojette, celle-ci a principalement été exploitée pour la redondance qu'elle pouvait apporter aux données. Elle est pour cette même raison actuellement exploitée par la société ROZO FS [Evenou 2019] dans des applications de cloud storage. La Transformation Mojette est extrêmement modulable, il est possible d'ajouter une multitude de projections d'un même support, chaque projection contient une partie

de l'information, et il est aussi possible de déterminer aisément à partir de combien de projections nous pouvons reconstruire le support. Il est donc aisé d'apporter une redondance modulable aux données [Parrein 2003].

La plupart du temps, des projections Mojette sont fortement décorrélées, c'est d'ailleurs ce qui sera recherché pour des applications de stockage ou de transmission sécurisé(e), il est toutefois possible de générer des projections corrélées entre elles. Pour ce faire, il suffit de choisir des projections ayant des angles très proches. Ce n'est qu'en 2006 qu'une première méthode de compression sans perte basée sur la Mojette a pu être proposée [Autrusseau 2006a]. Cette méthode, qui tirait profit des similarités intra- et inter-projections, fut développée dans le cadre du projet *TSAR*, elle était particulièrement adaptée aux types d'images traitées dans ce projet. Les images du musée étant de très grande dimensions (pouvant typiquement aller jusqu'à une résolution de 20000×20000 pixels). Plus la résolution de l'image est importante, plus il sera aisé de produire des projections d'angles très proches, et donc très fortement corrélées. Ces travaux ont été poursuivis grâce à l'arrivée d'Andrew Kingston au laboratoire IRCCyN, recruté sur un financement de bourse post-doctorale par la Région Pays de la Loire. Cette collaboration a permis d'améliorer significativement les performances de la compression Mojette [Kingston 2009a, Kingston 2008a], et a aussi permis de proposer une version multi-résolution de la transformation Mojette [Kingston 2009c, Kingston 2008b]. Au cours de ces travaux, les projections ne sont plus représentées sous leur forme 1D, mais elles sont "repliées" sur un support 2D, permettant ainsi d'appliquer des algorithmes de compression classiques 2D (tels que LOCO, CALIC ou Glicbawls). La compression Mojette faisait partie des contributions de l'IRCCyN dans le projet *TSAR*. Outre la compression, j'ai aussi eu l'occasion à travers différents projets de proposer des méthodes de chiffrement basés sur la transformation Mojette [Kingston 2009b]. Les premiers travaux, conduits au cours de ma thèse étaient destinés à protéger des images médicales [Autrusseau 2003a, Autrusseau 2003d, Autrusseau 2003b], par la suite, avec l'avènement de la compression Mojette, nous avons pu proposer des méthodes de crypto-compression et de chiffrement sélectif [Kingston 2007]. Ces travaux ont pu se faire encore une fois via l'encadrement d'Andrew Kingston mais aussi grâce à une collaboration avec l'Université de Rome et le co-encadrement de Simone Colosimo avec Patrizio Campisi. Ces travaux ont fait partie du projet *MILES*.

Je présenterai dans le chapitre 5 les différentes contributions que j'ai pu apporter à la thématique de la transformation Mojette.

2.2.4 L'imagerie médicale et biologique

J'ai commencé à m'intéresser à l'imagerie médicale lors de ma thèse de doctorat. Comme nous venons de le voir, mes premiers travaux ont principalement porté sur le chiffrement et la transmission sécurisée des images médicales. Par la suite, grâce à une mise à disposition dans un laboratoire en santé (le LIOAD), j'ai pu me pencher sur diverses problématiques de l'imagerie médicale et biologique.

Tout d'abord, par le biais d'un projet financé par l'Université de Nantes (Appel

à Projet Interdisciplinaire, 2017), j’ai collaboré avec Jeanpierre Guédon au LS2N et Yves Amouriq (au LIOAD) nous avons pu recruter Natalia De La Fournière en stage de M2 (formation SIBM). L’objectif de ce projet était de fournir une solution logicielle permettant de reconnaître les dents et les différents traitements d’un patient afin de générer de façon automatique le schéma dentaire [Michel 2018]. Le schéma dentaire (ou odontogramme) est une représentation schématique de la denture d’un patient, ainsi que des différents traitements qu’il a pu subir. Tout praticien doit établir le schéma dentaire de tous ses patients. Ce dernier permet de simplifier la prise en main de chaque dossier patient et peut aussi permettre l’identification de corps post-mortem.

Très récemment, j’ai eu l’occasion de monter un projet collaboratif avec des collègues du CHU de Nantes et de l’UFR d’odontologie. Le Pr. Yves Maugars en rhumatologie et le Pr. Assem Soueidan en chirurgie dentaire avaient pour projet de dépister l’ostéoporose via des acquisitions de CBCT. Classiquement, l’ostéoporose est dépistée grâce à la “Dual-energy X-ray Absorptiometry” (DXA), cependant, de tels scanners sont beaucoup plus coûteux, et moins accessibles que les CBCT. Des études ont montré que la position du Foramen Mentonnier (trou transperçant la mandibule) pouvait aider à prédire le risque ostéoporotique [Koh 2011]. Le Foramen Mentonnier est un interstice qui traverse la mandibule de part en part, il sert notamment à faire traverser le nerf mentonnier. L’épaisseur du foramen mentonnier, ou son éloignement du bas de la mandibule seraient des indicateurs de risque ostéoporotique. Ce projet (nommé *Tomosteo*) comporte plusieurs objectifs, d’une part, une étude clinique sera menée sur des patientes (âgées de 60 ans et plus) étant suivies au Centre de Soins Dentaires du CHU de Nantes et pour lesquelles une DXA a été prescrite.

Au cours de mon affectation dans le laboratoire RMeS, j’ai aussi eu la chance de pouvoir contribuer au projet ANR ICAN (Intra-Cranial ANeurysms : From familial forms to pathophysiological mechanisms) [Bourcier 2017a] et par la suite au projet UCAN [L’Allinec 2020]. Au cours du projet I-CAN, il a été montré qu’un facteur génétique important contribuait au risque de développement des anévrismes intra-crâniens [Bourcier 2017b]. Cependant, ce projet ne s’arrêtait pas à l’étude génétique des Anévrismes Intra Crânien (AIC), il importait aussi aux neuro-radiologues (le Docteur Romain Bourcier et le Professeur Hubert Desal) de comprendre l’intégralité des facteurs qui pouvaient conduire à la formation d’un anévrisme. Outre le risque génétique, il est à présent largement accepté que certains facteurs environnementaux peuvent avoir un impact sur la probabilité d’occurrence des anévrismes, par exemple, la consommation d’alcool, ou le tabagisme accroissent le risque, tout comme une alimentation trop salée, le manque d’exercice physique ou encore la prise de certains médicaments. Le docteur Romain Bourcier s’est aussi intéressé à la géométrie des bifurcations de l’arbre vasculaire cérébral. En effet, la majorité des anévrismes intra-crâniens surviennent sur une portion bien définie de l’arbre vasculaire cérébral : le polygone de Willis, et sur cet ensemble d’artères principales, le risque est prédominant au niveau des bifurcations. Effectivement, la répartition du flux sanguin dans les bifurcations n’est pas sans risque. La géométrie desdites bifur-

cations peut entraîner un impact trop important du sang contre la paroi, si la paroi est fragilisée, l'anévrisme peut se former. La présence d'un anévrisme en soi n'est pas problématique, n'entraîne aucun symptôme et n'engendre pas de risques particuliers, cependant, la rupture d'un anévrisme comporte de nombreux risques, elle peut conduire à de graves séquelles, voire même à la mort du patient. Au cours de ce projet, j'ai pu encadrer Anass Nouri pour un stage post-doctoral d'une année, nous avons conduit des études visant à proposer une caractérisation complète de l'arbre vasculaire cérébral [Nouri 2020]. Bien évidemment, nos analyses devaient se concentrer sur les risques liés à la configuration géométrique de l'arbre vasculaire uniquement. Il nous fallait donc écarter tous les risques liés aux facteurs environnementaux et génétiques. Pour ce faire, nous avons conduit une étude intra-patients. Nous avons donc mis en place une base de données d'images (d'IRM "Time Of Flight" : Angiographie par temps de vol. (TOF)) de 25 patients porteurs d'anévrismes sacculaires sur le polygone de Willis. Nous avons procédé à la caractérisation des bifurcations porteuses de l'anévrisme et aussi de la bifurcation miroir sur le polygone.

Nous avons ainsi été en mesure de déterminer quels facteurs géométriques semblaient être les plus propices à entraîner la formation d'un anévrisme. Selon nos résultats, il semblerait que les angles que forment les artères filles, la distance géodésique et la tortuosité soient les principales caractéristiques géométriques augmentant le facteur de risque.

Nous verrons dans le détails ces différentes études portant sur l'imagerie médicale dans le Chapitre 6.

Tout au long de ce chapitre, nous avons pu voir le positionnement de mon travail de recherche, le contexte des différentes études à travers des projets financés et des encadrements d'étudiants et de personnel contractuel. Au cours des quatre chapitres suivants, je vais présenter plus en détail le contenu scientifique de ces diverses études.

La perception visuelle

Contents

3.1 La dissimulation de l'information	33
3.1.1 Application à la protection des droits d'auteurs des images numériques	33
3.1.2 Protection des vidéos numériques	41
3.2 L'évaluation de la qualité	44
3.2.1 Métriques de qualité pour le tatouage	45
3.2.2 La reconnaissance des images	51

Mes travaux sur la perception visuelle se sont principalement orientés selon 3 axes distincts : La protection des droits d'auteurs, l'évaluation de la qualité et la perception des couleurs. Je présente ici les principaux travaux que j'ai pus conduire au cours de ces deux dernières décennies sur la modélisation de la vision humaine appliquée à des problématiques de sécurité des contenus multimédia.

3.1 La dissimulation de l'information

Une partie significative de mes travaux de thèse a porté sur la modélisation de la perception. L'objectif principal était alors de comprendre comment fonctionne le Système Visuel Humain, de pouvoir le modéliser et surtout de comprendre comment ajouter de l'information dans les images sans en modifier la perception.

3.1.1 Application à la protection des droits d'auteurs des images numériques

Les travaux sur la protection des contenus numériques ont été conduits principalement lors de ma thèse de doctorat, entre 1999 et 2002, le sujet de la thèse était "Modélisation psychovisuelle pour le tatouage des images". Le tatouage consiste à insérer de façon invisible des informations de copyright dans les documents multimedia. Effectivement, si le copyright est visible, alors il sera aisément possible de le supprimer pour n'importe quel faussaire ou hacker disposant de compétences basiques en manipulation des contenus multimedia. Il est donc crucial de rendre l'information insérée non seulement invisible pour l'oeil humain, mais aussi à la fois le plus robuste possible, afin que des traitements ne puissent pas supprimer le copyright sans dégrader visuellement le support.

Outre ces travaux initiaux ayant été conduits lors de ma thèse de doctorat, à l'issue de mon post-doctorat, j'ai intégré à nouveau le laboratoire IRCCyN, et plus particulièrement l'équipe IVC sur le poste d'Ingénieur de Recherche que j'occupe encore à ce jour. Pendant la décennie qui a suivi mon retour de post-doctorat, j'ai donc pu porter cette thématique de la sécurité des contenus numériques (entre 2004 et 2014) dans l'équipe IVC. S'en sont suivis plusieurs encadrements et co-encadrements sur cette thématique. Parmi les avancées significatives auxquelles j'ai pu participer, l'un des travaux les plus représentatifs de nos études est l'article [Urvoy 2014b] qui présente une méthode de tatouage perceptuel très robuste face à de nombreuses distorsions. Ces travaux avaient été initiés par un brevet déposé en 2013 [Atrousseau 2013a]. Les travaux autour de cette méthode de tatouage avaient pour principe de modifier la phase du spectre de Fourier de l'image tout en suivant une pondération perceptuelle exploitant la Contrast Sensitivity Function (CSF) et utilisant une fonction psychométrique pour permettre une modulation optimale de la force du tatouage. Ces travaux nous ont permis d'étudier certains aspects connexes, mais néanmoins très pertinents et nécessaires à la construction d'une technique de tatouage des images. Ceci m'a donc conduit à exploiter des outils d'analyse statistique permettant de vérifier la présence des informations cachées dans les images. Ainsi, le test de Grubbs de détection des outliers a pu nous servir à détecter un pic de corrélation [Urvoy 2014a].

Prenons le temps d'explorer dans le détail quelques uns de ces travaux. Dans l'article [Atrousseau 2007a], nous nous sommes intéressés à un modèle psychovisuel développé précédemment au laboratoire par Dominique Barba, ce dernier avait initialement été implémenté dans le but d'optimiser les algorithmes de compression des images. Dans ces travaux, datant de 2007, j'avais eu pour projet de détourner cette décomposition en canaux perceptuels afin de générer un masque de perception communément appelé "Just Noticeable Difference" (JND) dans la littérature. Un tel masque, défini dans le domaine spatial nous permet de garantir l'invisibilité de toute information insérée dans l'image. Cependant, afin de garantir une meilleure robustesse au copyright inséré, il est préférable que ce dernier soit inséré dans un domaine transformé. Dans notre cas, la marque était modulée autour d'une fréquence porteuse dans le domaine de Fourier, confinée à l'intérieur d'une sous-bande perceptuelle et insérée par simple multiplication de la partie réelle du spectre (ou alors spatialement après une transformée de Fourier inverse de la marque). Toute méthode de tatouage multimedia est composée de deux étapes distinctes, une étape d'insertion de la marque dans le contenu multimedia, et une étape de détection ou d'extraction de ladite marque. Dans notre cas la détection se faisait par simple extraction des coefficients de Fourier ayant potentiellement subi l'insertion puis par un calcul d'inter-corrélation normalisée avec la marque à chercher. Cette inter-corrélation est alors comparée avec un seuil de détection, si un pic de corrélation est supérieur à ce seuil, la marque est considérée comme présente dans l'image.

La différence de contraste $\Delta C_{i,j}$ qu'il est possible d'appliquer dans une sous-bande (i, j) est donnée par l'équation 3.1.

$$\Delta C_{i,j} = \frac{E_{i,j}}{L_0} \left(\frac{\Delta f_i}{f_{0,i}} \right)^{\lambda_{i,j}} \quad (3.1)$$

ou $E_{i,j}$ représente la puissance de la sous-bande (i, j) , L_0 est la luminance de l'écran, Δf_i représente la i^{eme} bande radiale, $f_{0,i}$ est la fréquence centrale de la sous-bande i , et $\lambda_{i,j}$ est une constante dépendant de la sous-bande (i, j) .

Cette différence de contraste perceptuel nous permet de définir le masque JND comme suit (équation 3.2) :

$$\Delta L_{i,j}(m, n) = \Delta C_{i,j} \times \bar{L}_{i,j}(m, n) \quad (3.2)$$

$\bar{L}_{i,j}(m, n)$ étant la luminance moyenne locale (la luminance formée par toutes les fréquences inférieures à la bande d'insertion).

Finalement, si l'insertion doit avoir lieu dans le domaine spatial, le coefficient de pondération spatiale ($K_{i,j}$) peut être défini comme suit :

$$K_{i,j} = \operatorname{argmin}_{m,n} \left(\left| \frac{\Delta L_{i,j}(m, n)}{W_S(m, n)} \right| \right) \quad (3.3)$$

ou $W_S(m, n)$ est la représentation spatiale de la marque en tout point (m, n) de l'image.

La Décomposition en Canaux Perceptuels (DCP) est représentée sur la Figure 3.1. Nous avons conduit une étude afin de déterminer quelle bande de fréquence était la plus à même de porter le watermark tout en optimisant conjointement la robustesse et l'invisibilité. Le patch de bruit servant de copyright a ainsi été modulé sur chacune des trois bandes de fréquences radiales, l'invisibilité étant assurée pour chaque insertion, car la force est modulée par le masque perceptuel, nous étudions alors la robustesse pour chaque bande de fréquence. La suite logicielle Stirmark [Petitcolas 1998, Petitcolas 2000] était largement utilisée à l'époque pour tester la validité des méthodes de tatouage, elle opérait 89 distorsions différentes sur les images à tester, ces distorsions étaient en grande partie constituées de méthodes de compression (JPG) et de distorsions géométriques. Nous représentons sur la Figure 3.2 les valeurs des pics de corrélations en fonction des 89 attaques de Stirmark et ce, dans les trois bandes de fréquences de la décomposition perceptuelle. Nous pouvons observer que la bande radiale de fréquences moyennes offre le meilleur compromis entre l'invisibilité et la robustesse.

De plus, cette même méthode de protection de copyright a été exploitée à nouveau mais dans un tout autre contexte. Nous avons détourné l'utilisation initiale de la méthode de tatouage pour proposer une métrique de qualité sans référence [Ninassi 2006]. Alexandre Ninassi en stage de M2 au sein de l'équipe IVC a travaillé sur l'exploitation du tatouage pour prédire la qualité des images. Cette métrique pseudo-sans-référence nous permettait d'estimer la qualité perçue en se basant sur la perte de qualité de la marque insérée. Par ailleurs, avec l'aide de Guillaume Jeannic et Mohamed Bouchakour (respectivement doctorant à IVC et en stage de

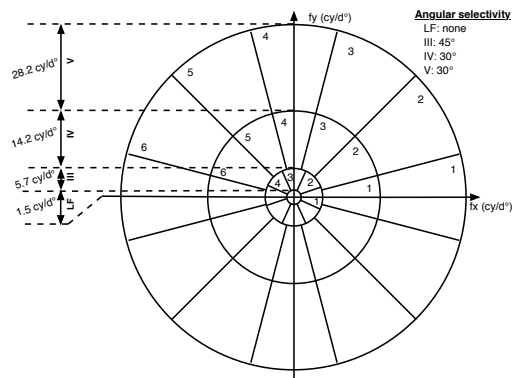


FIGURE 3.1 – Décomposition du spectre de Fourier en canaux perceptuels.

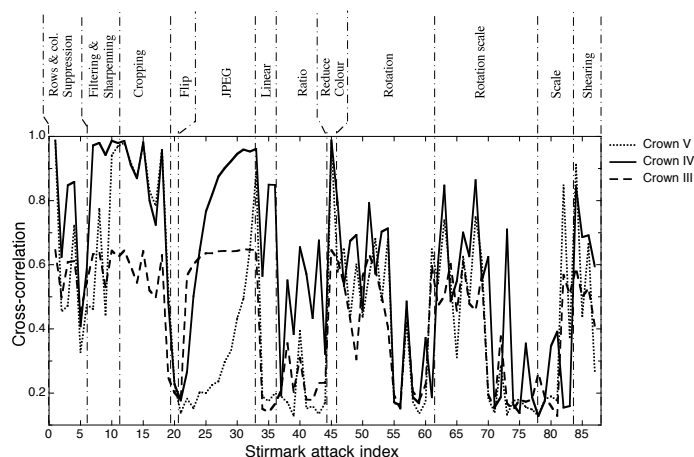


FIGURE 3.2 – Robustesse de la méthode dans les 3 bandes de fréquences face aux 89 attaques de Stirmark.

mémoire CNAM) une approche similaire a vu le jour dans le domaine des ondelettes [Bouchakour 2008]. Ces différentes approches ont ainsi pu être exploitées dans le projet *ANR-TSAR* [Pitzalis 2008] que j’ai pu conduire entre 2005 et 2008.

Quelques années plus tard, dans le cadre du projet de Recherche *StreamMaster* j’ai pu diriger les travaux de deux post-doctorants, Mathieu Urvoy et Dalila Goudia. Matthieu avait été recruté sur ce projet collaboratif, tandis que le poste de Dalila était rémunéré via un contrat de maturation porté par la SATT “Ouest Valorisation” qui faisait suite au dépôt du brevet [Autrusseau 2013a].

Nos travaux dans ce cadre se sont inspirés de l’étude que nous venons de voir.

Cependant, nous sommes allés plus loin et avons pu améliorer significativement les performances (tant du point de vue de l'invisibilité de la marque que du point de vue de sa robustesse). Dans ces travaux [Urvoy 2014b, Urvoy 2013], la DCP a été abandonnée et remplacée par une modélisation de la CSF ou Fonction de Sensibilité aux Contrastes. Nous avons alors retenu la CSF de Barten [Barten 1999].

Nous avons défini le seuil de contraste local $\mathcal{C}_{\text{local}}^\tau$ comme suit :

$$\mathcal{C}_{\text{local}}^\tau(f, \theta) = \frac{1}{\text{CSF}(f, \theta)} \quad (3.4)$$

Nous exploitons alors la fonction Psychométrique définie par Daly [Daly 1993] :

$$\Psi(\mathcal{C}_{\text{local}}^{\tau*}) = 1 - e^{-(\mathcal{C}_{\text{local}}^{\tau*})^\beta}, \quad (3.5)$$

ou $\mathcal{C}_{\text{local}}^{\tau*} = \mathcal{C}_{\text{local}} / \mathcal{C}_{\text{local}}^\tau$ représente le ratio entre le contraste local normalisé et sa valeur seuil donnée dans l'équation 3.4. Et β étant la pente de la fonction psychométrique.

Une fois le seuil de visibilité défini pour un seul signal sinusoïdal, nous devons considérer l'information insérée comme un patch convexe formé d'une multitude de signaux sinusoïdaux. Pour ce faire, nous avons utilisé un modèle de sommation de probabilités.

$$\Psi_{\text{mark}} = 1 - \prod_{k=0}^{N-1} (1 - \Psi(\mathcal{C}_{\text{local}}^{\tau*}(f_k, \theta_k))) \quad (3.6)$$

Ψ_{mark} étant la probabilité que la marque soit perçue par un observateur. f_k et θ_k sont respectivement les fréquences visuelles et orientations.

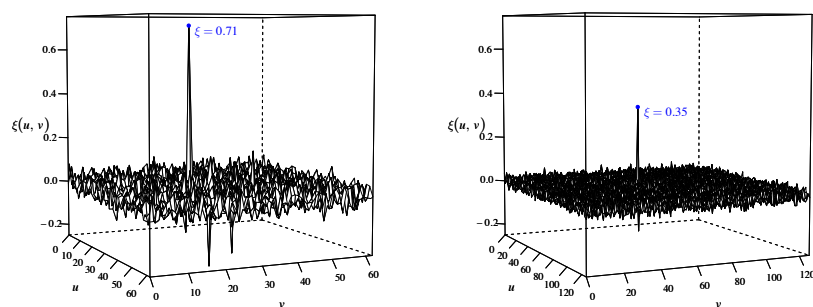
Le contraste normalisé, qui équivaut à l'amplitude optimale max du signal peut ainsi être obtenue par inversion de la fonction psychométrique (Eq. 3.5) :

$$A_{\text{peak}}(f_k, \theta_k) = \mathcal{C}_{\text{local}}^\tau(f_k, \theta_k) \cdot \left[-\ln \left((1 - \Psi_{\text{mark}})^{1/N} \right) \right]^{1/\beta} \quad (3.7)$$

Ici, la composition de la marque insérée est différente ; alors que dans les travaux antérieurs, un bruit aléatoire servait de copyright, dans le cadre du projet "Stream-Master" nous souhaitons insérer une information binaire. L'insertion dépend à la fois de l'amplitude et de la phase du spectre de Fourier. La modification de l'amplitude permet l'ajustement de la force du watermark, alors que le watermark binaire est inséré sur la phase du spectre.

Comme nous avons pu le voir précédemment [Autrusseau 2007a], la détection se fait par le calcul de l'inter-corrélation normalisée. En pratique, une méthode de Template Matching est appliquée [Lewis 1995].

Cette méthode de tatouage s'est avérée être très efficace, nous avons même pu la tester avec succès face à des distorsions de *Print and Scan* (Impression et numérisation). Ce type d'attaque est la plus dévastatrice sur des systèmes de tatouage des images, car elle combine une multitude de distorsions (géométriques pour la



(a) Watermark de taille 32×32 (b) Watermark de taille 64×64

FIGURE 3.3 – Matrices de corrélation pour deux tailles distinctes de patches insérés.

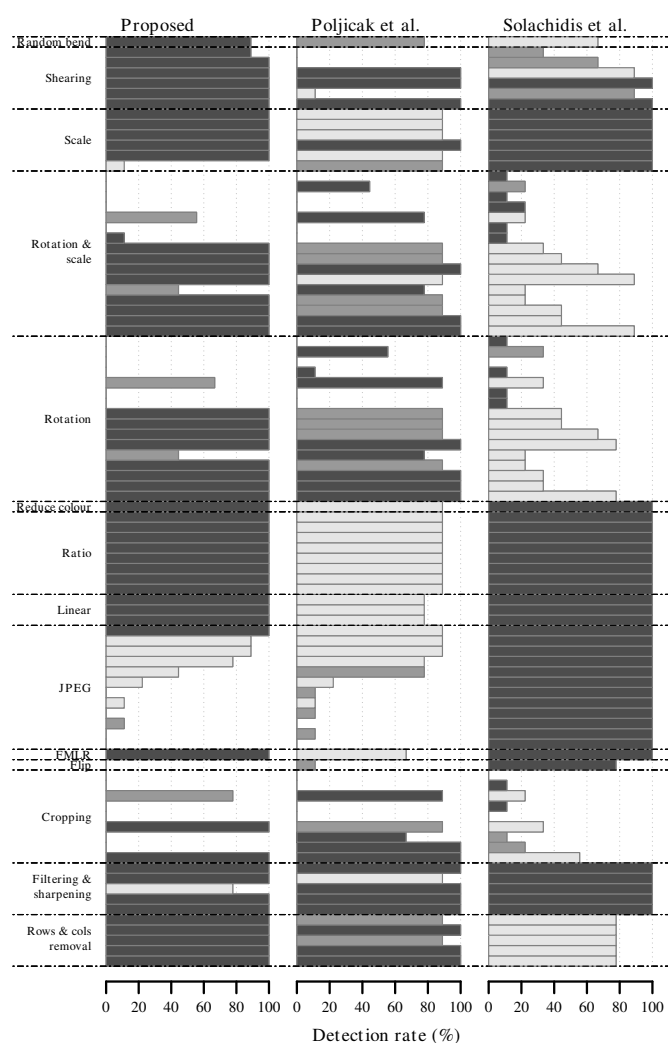


FIGURE 3.4 – Robustesse face à Stirmark au seuil de visibilité.

plupart). La Figure 3.4 nous montre les performances de robustesse de la méthode proposée face à deux méthodes de la littérature ([Solachidis 2001, Poljicak 2011])

Comme expliqué précédemment, la détection se fait par *Template Matching*, cependant, une fois la corrélation effectuée, il est nécessaire de la comparer à un seuil de détection. Classiquement, un test d'hypothèse est utilisé. Deux hypothèses peuvent être faites, soit $\mathcal{H}0$, l'image testée ne contient pas la marque recherchée, ou selon l'autre hypothèse $\mathcal{H}1$, la marque est bien présente. Si la méthode de tatouage a été efficacement mise en oeuvre, ces deux distributions doivent être suffisamment éloignées, ce qui permet alors de définir un seuil de détection entre les deux distributions, et permettra alors de minimiser à la fois les faux positifs et les faux négatifs. La Figure 3.5 nous montre les distributions $\mathcal{H}0$ et $\mathcal{H}1$ pour les méthodes [Urvoy 2014b], [Solachidis 2001] et [Poljicak 2011].

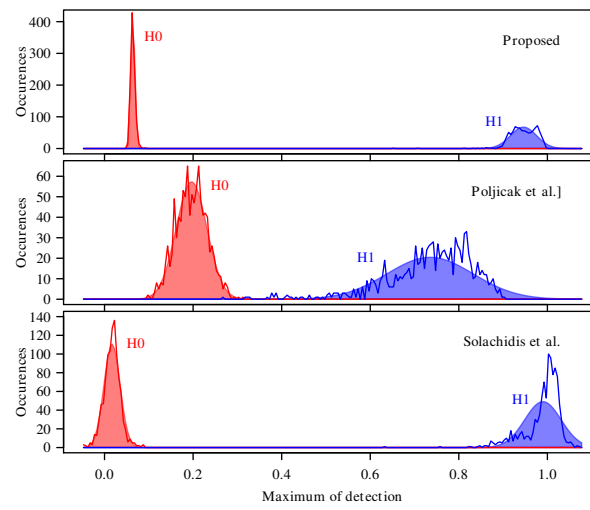


FIGURE 3.5 – Distributions $\mathcal{H}0$ et $\mathcal{H}1$ pour les trois méthodes testées.

Suite à ces travaux, j'ai souhaité me pencher sur l'optimisation du seuil de détection. Effectivement, comme nous venons de le voir, dans ces diverses études, le seuil de détection était défini de façon expérimentale, une grande quantité d'images était utilisée pour insérer le copyright, le détecteur était testé à la fois sur les images tatouées, et sur les images non tatouées, ou encore sur des images tatouées avec un autre identifiant. Nous avons proposé avec Matthieu Urvoy une méthode de détection basée sur le test statistique de Grubbs [Grubbs 1969]. Ce projet, publié dans [Urvoy 2014a] est très novateur. Le principe de base est l'analyse statistique de la carte de corrélation afin d'y détecter un outlier. Effectivement cette carte de corrélation est composée de valeurs réelles suivant le plus souvent une distribution Gaussienne, or, le test de Grubbs se base lui aussi sur un test d'hypothèses :

$\mathcal{H}0$: les données testées ne comportent pas d'outlier, et

$\mathcal{H}1$: il y a précisément un outlier sans les données.

Le test statistique de Grubbs s'écrit :

$$G = \frac{\max_{i=1,\dots,N} |Y_i - \bar{Y}|}{\sigma} \quad (3.8)$$

Y sont les données à tester, $\bar{Y}$ en est la moyenne, et σ l'écart type. Cette définition est valable pour faire face à une distribution unilatérale, cependant, dans le cas du watermarking, un test bilatéral est souhaitable. Effectivement, une corrélation atteignant -1 signifie que la marque est présente dans l'image mais juste inversée, ce qui peut tout à fait se produire dans un scénario de tatouage. Une distortion géométrique, une inversion des luminances peut produire une inversion du bruit inséré. Fort heureusement, il existe une version bilatérale du test de Grubbs :

$$G > \frac{N-1}{\sqrt{N}} \sqrt{\frac{t_{\alpha/(2N), N-2}^2}{N-2 + t_{\alpha/(2N), N-2}^2}} \quad (3.9)$$

ou $t_{\alpha/(2N), N-2}$ est la valeur critique de la loi de Student avec $N-2$ degrés de liberté.

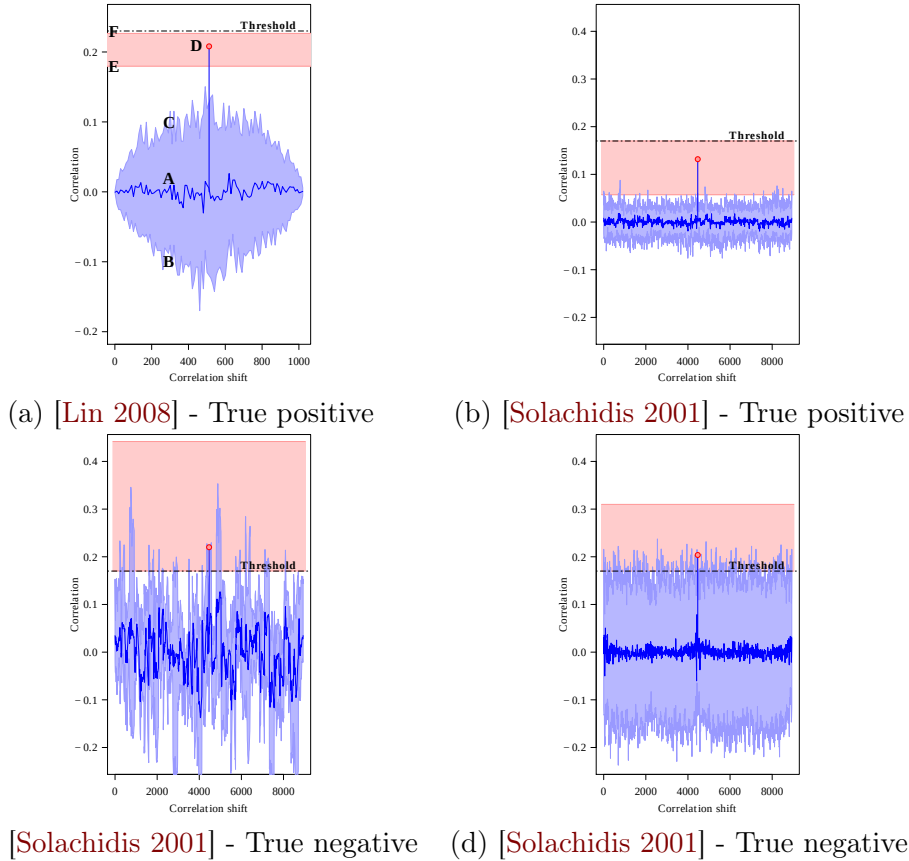


FIGURE 3.6 – Illustration du test de Grubbs, lors duquel une valeur de l'échantillon est vue comme un outlier dans une distribution Gaussienne. Nous présentons des exemples de Corrélations pour diverses méthodes de la littérature.

La Figure 3.6 présente quelques exemples de détection basées sur le test de Grubbs. Le seuil horizontal en pointillé représente la valeur définie dans les articles correspondants. Dans les sous-figures (a) et (b) Grubbs détecte la marque correctement, alors que les travaux respectifs de [Lin 2008] et [Solachidis 2001] l'estimaient perdue. Pour les deux dernières sous-figures, (c) et (d), les auteurs de [Solachidis 2001] détectaient dans leurs travaux un watermark (corrélations au-dessus du seuil) alors que Grubbs ne le détecte pas (à juste titre).

3.1.2 Protection des vidéos numériques

La participation de l'équipe IVC dans le projet *ANR-Safim@ge* [Maillet 2008a, Maillet 2008b, Maillet 2009, Maillet 2011] a été pour moi l'opportunité de me pencher sur la protection des droits d'auteurs des vidéos numériques. A travers divers encadrements, et notamment Wei Chen en stage de M2 en 2007 et un peu plus tard les post-doctorats de Thomas Stütz et Zafar Shahid en 2013-2014, j'ai pu aborder la problématique du tatouage vidéo par deux angles différents. Au cours de l'étude portant sur le projet *Safim@ge* nous avons proposé une méthode d'insertion dans le flux codé h.264 [Chen 2008, Chen 2014]. L'objectif de ces travaux était de s'assurer qu'une modification du flux de données compressés h.264 n'engendre pas d'étalement de l'erreur sur les blocs avoisinants.

Comme nous pouvons l'observer sur la Figure 3.7, le décodage d'un bloc se base sur les valeurs des blocs précédemment décodés (bloc supérieur et bloc de gauche, pour un décodage dans un ordre "raster scan"). L'objectif est alors de modifier les coefficients des blocs tout en maintenant ces valeurs de bord inchangées.

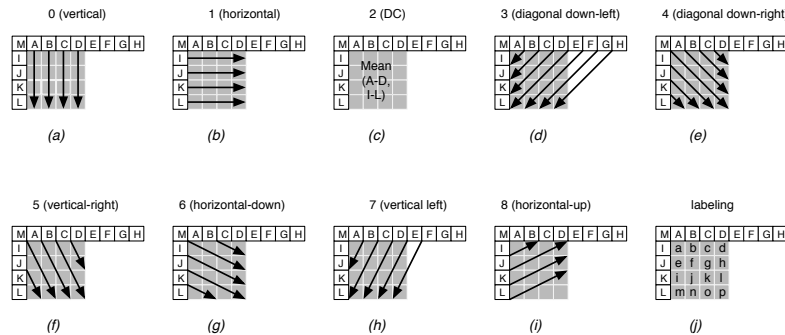


FIGURE 3.7 – Les neuf modes de prédiction *intra* pour le décodage de la vidéo.

Une modification maîtrisée des blocs DCT était ainsi opérée, de façon à ce que les coefficients utilisés pour la prédiction lors du codage ne soient pas impactés. De plus, encore une fois, dans un souci de garantir l'invisibilité de la marque insérée, nous avons exploité les matrices de quantification visuelles proposées par Watson [Watson 1993] afin de moduler la force du watermark (sa capacité). L'insertion ici se faisait par substitution des coefficients.

Le choix du seuil de détection se fait ici de façon expérimentale. Une grande quantité de contenus sont donc tatoués, nous distinguons deux hypothèses :

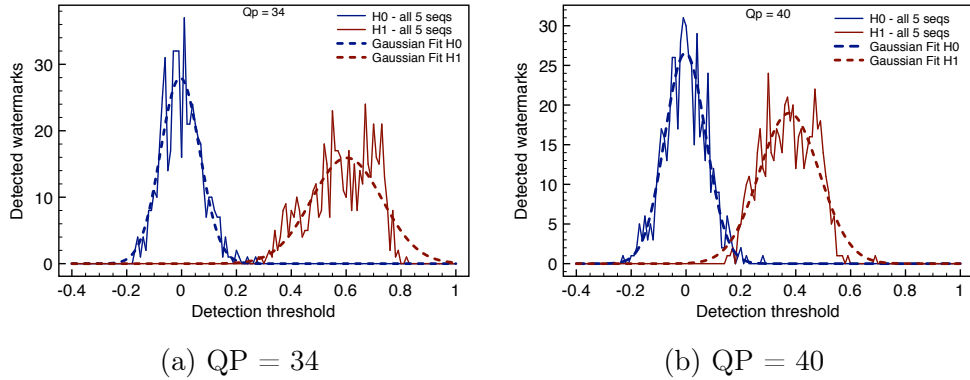


FIGURE 3.8 – Analyse des hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$ pour deux facteurs de quantifications.

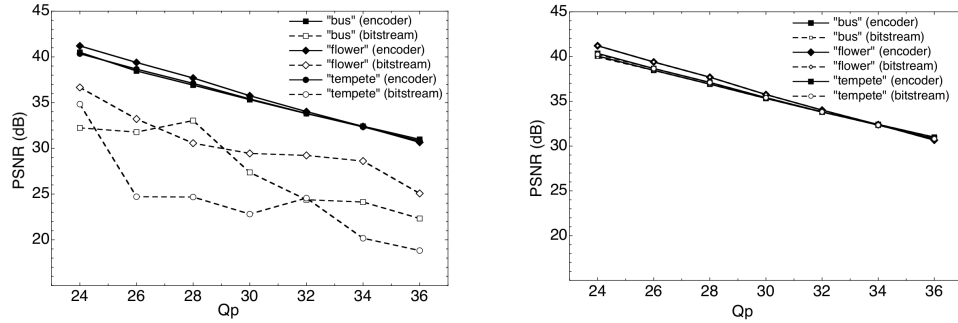
$\mathcal{H}_0$: Le contenu ne comporte pas de watermark, et
 $\mathcal{H}_1$: Le contenu comporte une information de copyright.

Une fois cette base constituée, le détecteur est testé sur tous les contenus, et la distribution des corrélations avec et sans watermark sont supposées être distinctes, un seuil de détection doit alors se trouver entre ces deux distributions. La Figure 3.8 nous montre comment sont distribuées les corrélations sur les deux hypothèses vues précédemment pour deux taux de quantification différents. Ces distributions dépendent de nombreux facteurs, la méthode de tatouage elle-même (et la méthode de détection associée) influe de façon significative sur ces deux distributions, mais aussi, la quantité d'informations insérées, la virulence des attaques, etc. Le plus souvent, les coefficients de corrélations du détecteurs suivent une distribution Gaussienne, comme nous pouvons l'observer sur la Figure 3.8. Il est souvent nécessaire de prendre en compte des attaques sur les contenus tatoués, ce qui va mécaniquement étaler et réduire la distribution $\mathcal{H}_1$, car les corrélations seront inévitablement réduites après attaques.

Cette méthode de sélection du seuil de détection a été très largement exploitée par la communauté du *watermarking*, cependant, elle requiert la constitution d'une base d'images assez volumineuse et rien ne garantit que les distorsions aient été correctement choisies et permettent de réduire les fausses alarmes. C'est pourquoi nous avons proposé en 2014 la méthode de détection par outlier de Grubbs que nous avons vu dans la section précédente.

La Figure 3.9 donne une idée de l'impact des méthodes de tatouage sur les séquences. Deux méthodes d'insertion sont testées, tatouage du bitstream ou tatouage intégré dans l'encodeur.

Au cours de ces travaux, nous avons donc pu insérer des informations de copyright dans les vidéos h.264, sans augmenter le bit-rate mais tout en garantissant à la fois l'invisibilité des informations et une forte robustesse. Contrairement aux travaux de Noorkami et al. [Noorkami 2007] l'insertion dans le bitstream n'entraîne pas d'étalement de la distortion à travers les frames, et ce, grâce au maintien des



(a) Algorithme [Noorkami 2007]

(b) Méthode proposée dans [Chen 2014]

FIGURE 3.9 – Comparaison des PSNR sur la méthode de [Noorkami 2007] et la méthode proposée dans [Chen 2014].

coefficients de bord, comme nous avons pu le voir précédemment.

Quelques années plus tard, toujours sur cette même problématique de tatouage psychovisuel des vidéos numériques, j'ai encadré Thomas Stütz en stage post-doctoral dans l'équipe IVC. En collaboration avec Andreas Uhl de l'Université de Salzbourg (Autriche), nous avons travaillé sur une méthode de tatouage substitutive adaptée au tatouage des vidéos H.264 CAVLC [Stütz 2014]. Dans ces travaux, nous proposons une méthode de tatouage par substitution qui opérait une modification des différences de vecteurs de mouvements Motion Vector Difference (MVD). La longueur du bitstream n'était donc pas impactée par l'insertion de l'information. Dans cette méthode, uniquement les inter-frames dites "no-reference" étaient modifiées. Nous avons particulièrement travaillé sur la maîtrise de l'impact visuel du tatouage. Un test de qualité subjectif avait ainsi été conduit [Stütz 2011], ce dernier avait montré que l'insertion de la marque n'entraînait pas de perte de qualité du contenu plus significative qu'un simple codage H264 uniquement (cf. Figure 3.10).

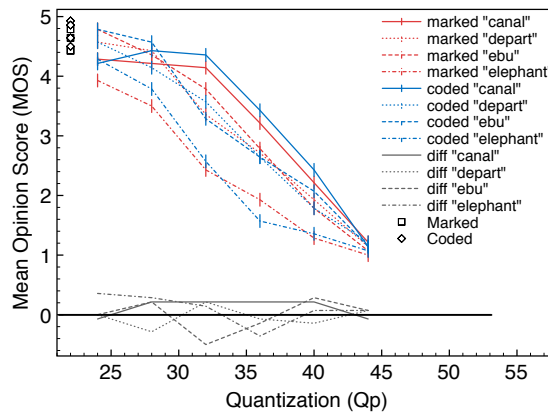


FIGURE 3.10 – Qualité Subjective des séquences tatouées pour différents pas de quantifications.

De plus, nous avons aussi pu tester plusieurs métriques de qualité, et notamment le PSNR et Visual Information Fidelity (VIF), qui ont, elles aussi, montré une très faible différence de qualité prédite entre les séquences tatouées et les séquences codées. Cependant, nous pouvons noter que le PSNR, SSIM, ainsi que la métrique “Carosi, Pankajakshan & Atrusseau (CPA)” [Carosi 2010, Pankajakshan 2010] n’ont pas été en mesure d’évaluer correctement la qualité d’une séquence animée (“*Elephants Dream*” générée par ordinateur). Seule la métrique VIF en a été capable, comme en atteste la Figure 3.11.

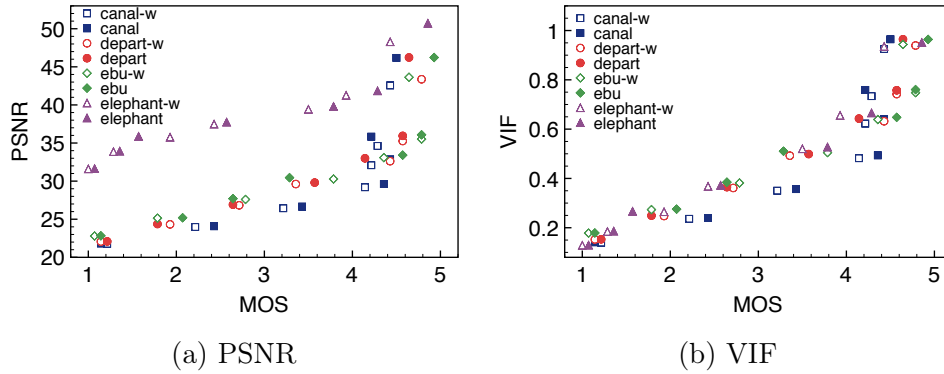


FIGURE 3.11 – PSNR et VIF sur les différentes séquences de test tatouées avec notre méthode.

Outre le fait que l’invisibilité ait été validée à travers un test subjectif, notre méthode s’est de plus, avérée être plus robuste que la méthode de référence proposée à l’époque par Zou et Bloom [Zou 2009a, Zou 2009b].

Dans cette section nous avons eu un aperçu des travaux que j’ai pu conduire sur la mise en oeuvre de méthodes de tatouage exploitant des critères perceptuels. Voyons à présent une partie de mes travaux portant sur les métriques de qualité dans un contexte de sécurité des images.

3.2 L’évaluation de la qualité

Comme nous avons pu le voir dans l’introduction de ce manuscrit, pour plusieurs applications (*watermarking*, stéganographie, etc.) il est crucial que l’information soit insérée de façon imperceptible. Il est donc primordial de s’assurer que les contenus modifiés soient de très bonne qualité. Pour s’en assurer, deux possibilités s’offrent à nous. Il est possible de conduire des tests subjectifs, lors desquels on sollicite des observateurs humains qui vont juger de la qualité des images modifiées. L’alternative est d’utiliser des métriques de qualité qui vont prédire la qualité telle qu’elle sera perçue par des observateurs humains. Je présente dans cette section une partie de mes travaux sur les métriques objectives d’évaluation de qualité pour deux applications distinctes : le tatouage d’une part et la reconnaissance des images, ou plus spécifiquement l’identification du contenu d’autre part.

3.2.1 Métriques de qualité pour le tatouage

Au cours de ces 10 années qui ont fait suite à mon post-doctorat (entre 2004 et 2014), outre la problématique du tatouage perceptuel, je me suis aussi penché sur différents travaux portant sur l'évaluation de la qualité des images. Ces travaux sur les Métriques Objectives de Qualité étaient à l'époque rattachées aux travaux menés au sein du laboratoire sur la protection des droits d'auteurs. En effet, nous avons cherché à construire une métrique la mieux adaptée à la problématique du tatouage numérique. De nombreuses méthodes de tatouage d'images de la littérature revendiquent une parfaite invisibilité de la marque insérée par l'utilisation de métriques de qualité. Ces métriques sont des outils logiciels ayant pour objectif d'estimer la qualité perçue des images. On les trouve sous différentes formes, les métriques dites "Full Reference" nécessitent d'avoir accès à l'image originale de bonne qualité, ainsi que l'image dégradée à évaluer. Les métriques "Reduced Reference" ont quant à elles besoin d'une certaine quantité d'informations représentatives de l'image originale, ainsi que l'image dégradée à évaluer. Finalement, les métriques "No Reference" nécessitent uniquement d'avoir accès à l'image dégradée. Ces trois types de métriques fournissent donc en sortie une évaluation de la note de qualité. Cette note prédite peut être normalisée entre 0 et 1 ou dans l'intervalle $[-1, 1]$, mais il est aussi possible que les notes soient sur un intervalle $[0, 100]$ voire même dans $[0, +\infty]$. Dans de très nombreux travaux (une vaste majorité des études sur le tatouage), les auteurs évaluent la qualité perceptuelle des images tatouées par l'utilisation de deux métriques simplistes, le PSNR et SSIM. Or, il s'avère que nous avons montré à plusieurs reprises dans nos travaux [Carosi 2010, Pankajakshan 2010] que ces deux métriques sont de loin les moins bien adaptées à une problématique de camouflage de l'information. Une étude exhaustive a été conduite dans [Le Callet 2008] montrant la complexité des méthodes d'évaluation de la qualité. Dans ce chapitre d'ouvrage, nous avons présenté les différentes possibilités qui s'offraient aux chercheurs en tatouage des images pour valider l'invisibilité des marques insérées. Une étude similaire en collaboration avec les collègues de l'Université de Salzbourg a porté sur l'étude des métriques de qualité pour évaluer la dégradation perceptuelle engendrée par des algorithmes de chiffrement sélectif [Stütz 2010].

Pour toute problématique d'évaluation perceptuelle, la vérité terrain ne peut être acquise que par des observateurs humains, utilisateurs finaux des images et seuls à même de déterminer la qualité des images ou l'invisibilité des traitements. Il est donc crucial de conduire des tests subjectifs standardisés, dans des conditions de visualisation contrôlées. Ces tests subjectifs doivent en effet répondre à un certain cahier des charges défini par des organismes internationaux de standardisation. Au cours des mes divers projets, j'ai pu être amené à conduire de nombreux tests subjectifs [Marini 2007a, Atrousseau 2008, Atrousseau 2009c, Atrousseau 2009a, Atrousseau 2009b, Atrousseau 2010b, Stütz 2011, Atrousseau 2013b, Atrousseau 2013d, Hofbauer 2019]. J'ai implémenté plusieurs de ces protocoles de tests, et en ai mis certains à disposition dans une

version en-ligne¹.

Par opposition aux métriques telles que le PSNR ou SSIM j'ai souhaité proposer une métrique qui soit la plus simple possible, mais tout en considérant quelques propriétés de base du SVH. Il est notamment primordial de prendre en compte la sensibilité au contraste de notre système visuel, or ni le PSNR, ni SSIM ne la prennent en considération. Nous avons évoqué précédemment la CSF dans ce chapitre, voyons à présent ce qu'elle représente. Il nous faut tout d'abord comprendre ce qu'est un contraste, et comment nous le percevons. Le contraste est un changement de luminosité ou de couleur d'un stimulus.



FIGURE 3.12 – Variation du contraste selon l'axe vertical, et variation de la fréquence spatiale sur l'axe horizontal.

La Figure 3.12 nous montre une variation progressive à la fois du contraste (selon l'axe vertical) et des fréquences spatiales (selon l'axe horizontal). La limite de perception de ces contrastes est précisément la CSF. Cette fonction nous permet donc de distinguer les contrastes qui sont visibles de ceux qui ne le sont pas.

La Figure 3.13 nous montre la Fonction de Sensibilité au Contraste dans ses versions 1D et 2D. Notre système visuel est particulièrement sensible aux fréquences spatiales, nous ne percevons pas identiquement les hautes et les basses fréquences, les zones uniformes et les zones texturées. Ce filtre 2D est à appliquer sur le spectre de Fourier d'une image et permet ainsi de moduler les sinusoides composant l'image en fonction de leur probabilités de perception.

Lors l'un projet impliquant Maurizio Carosi (alors en stage de M2), Vinod Panakajakshan (post-doctorant à IVC) et moi même, nous avons proposé une métrique dénommée CPA [Carosi 2010] qui se réduisait tout simplement à filtrer par la CSF les spectres de Fourier de l'image originale d'une part et de l'image test d'autre part. Une transformation de Fourier inverse était alors appliquée sur ces spectres modifiés, puis nous calculions la différence en tout point entre les deux images ainsi obtenues. Finalement, une sommation pondérée (pooling) était appliquée sur la différence des images. Une sommation de Minkowski avait été testée avec différents

1. <http://websubjexp.polytech.univ-nantes.fr>

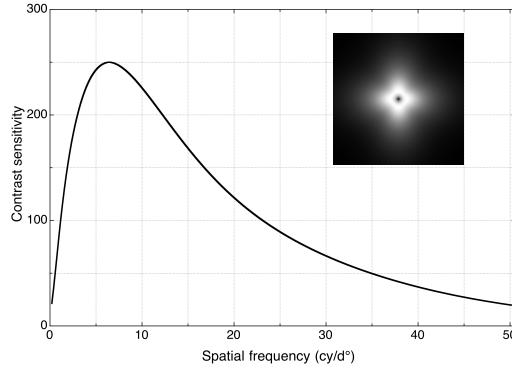


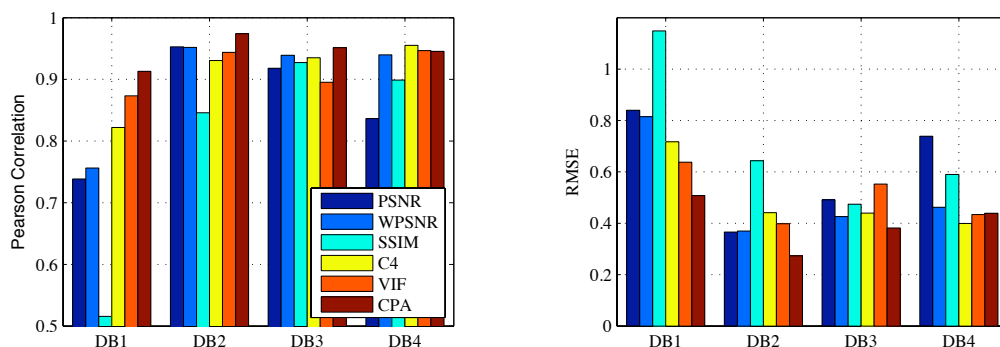
FIGURE 3.13 – Fonction de Sensibilité au Contraste en version 1D (et 2D dans l’encart).

paramètres P et R comme indiqué dans l’équation 3.10. Cette stratégie de pooling est classiquement utilisée pour regrouper les distorsions des images en une seule et unique note prédictive de qualité [Wang 2006].

$$Mink(P, R) = \sqrt[R]{\frac{1}{N} \sum_{i=1}^N |v_i|^P} \quad (3.10)$$

Bien que relativement simple, cette métrique de qualité a démontré des performances très intéressantes face à tout type de bases de données subjectives et notamment contre des distorsions de type “Data hiding”. Nous avons testé la métrique CPA sur 4 bases de données subjectives composées comme suit : **DB1** était composée d’images tatouées dans le domaine de Fourier [Atrousseau 2013d], **DB2** [Atrousseau 2009b] était composée d’images tatouées par l’algorithme [Furon 2008] et établie en collaboration avec le GIPSA-lab au cours du projet *TSAR*. La base **DB3**, établie en collaboration avec Peter Meerwald de l’Université de Salzbourg, comportait quand à elle des images tatouées dans 2 domaines distincts : Discrete Wavelet Transform et Dual-Tree Complex Wavelet Transform [Atrousseau 2009c]. Finalement, la base **DB4** comprenait des algorithmes de compression uniquement, (JPG, J2K & LAR) [Atrousseau 2009a] (Collaboration avec l’IETR dans le cadre du projet *TSAR*). Les résultats peuvent être observés sur la figure 3.14. Typiquement, lorsqu’une métrique est implémentée, la meilleure façon de la tester est de comparer les scores prédits par les scores fournis par des observateurs pendant des tests subjectifs. Les scores des observateurs représentent la vérité terrain, nous les appelons Mean Opinion Score (**MOS**), tandis que les scores fournis par les métriques (prédits) sont les Mean Opinion Score (predicted) (**MOSp**). Une fois les **MOS** et **MOSp** récoltés, nous sommes en mesure de les comparer entre eux, nous utilisons typiquement différentes mesures de corrélation (Pearson et Spearman) la Root Mean Square Error (**RMSE**) ainsi qu’un indice d’Outlier Ratio.

Dans la Figure 3.14 nous comparons les performances de *CPA* avec cinq autres métriques largement utilisées dans la littérature, dont le *PSNR* et *SSIM* mentionnés précédemment, mais aussi le *Weighted PSNR* [Voloshynovskiy 2000] (métrique elle aussi créée spécifiquement pour le tatouage des images) et deux métriques plus complexes qui considèrent des critères de perception avancés [Sheikh 2005, Carnec 2008]. Bien évidemment, il est souhaitable d’obtenir la meilleure corrélation et la plus faible *RMSE*. Nous constatons que *CPA* offre des performances très intéressantes, et se montre même être la meilleure métrique pour 3 des 4 bases de données testées (celles en lien avec le tatouage).



(a) Pearson (Corr) pour les 4 bases testées

(b) RMSE sur les 4 bases testées

FIGURE 3.14 – Performances de la métrique *CPA* sur 4 bases de données subjectives.

Dans des travaux ultérieurs [Pankajakshan 2010], nous avons complexifié légèrement la métrique en y ajoutant une décomposition par blocs de l’image et une fonction d’élévation de seuil (c’est la métrique dénommée *CPA2*).

Ensuite, par le biais d’une collaboration avec l’Université de Salzbourg, de travaux menés avec Thomas Stütz (qui deviendra plus tard post-doctorant sous ma direction) et Andreas Uhl, nous avons pu tester ces même métriques dans un contexte de chiffrement sélectif des images [Stütz 2010]. Encore une fois, *CPA* a démontré des performances très intéressantes. Au cours de ces travaux, les collègues de l’Université de Salzbourg ont proposé trois méthodes de chiffrement sélectif du format de compression d’images JPEG2000. Notre objectif était de déterminer si des métriques de qualité pouvaient être à même de prédire les notes fournies par des observateurs. Un protocole de tests Double Stimulus Impairment Scale (*DSIS*) a été mis en place. Les observateurs ont jugé de la qualité des images et 8 métriques de qualité ont été testées [Atrusseau 2010b].

Nous pouvons observer les tracés des courbes des scores subjectifs en fonction des scores prédits sur la Figure 3.15. Une métrique efficace se reconnaît par une distribution linéaire et étroite des points sur un tel tracé. À première vue, *CPA* semble offrir de bonnes performances face à une telle problématique. Nous pouvons noter que sur cette Figure, la distribution de *CPA* est inversée par rapport aux autres métriques, sa pente est négative, ceci s’explique simplement par le fait que

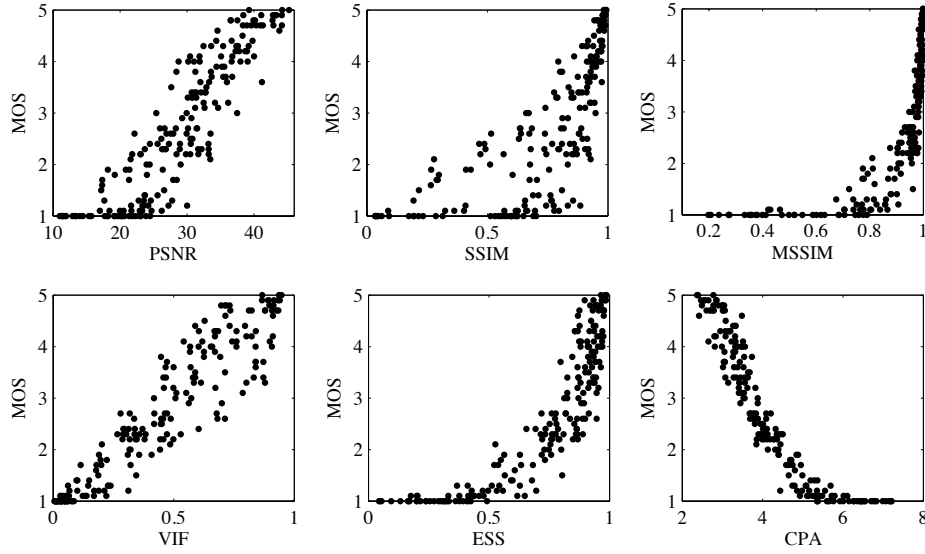


FIGURE 3.15 – Tracés des courbes MOS en fonction de MOSp pour six métriques sue la base [Autrusseau 2010b].

les autres métriques testées donnent une estimation des similarités entre les images, alors que **CPA** quantifie les différences.

Le tableau 3.1 présente les performances des 8 métriques. Bien que très simple dans son design, il s'avère que **CPA** obtient les plus faibles **RMSE** et Outlier Ratio, tout en obtenant les meilleures corrélations. Même les métriques **ESS** et **LSS** spécifiquement conçues pour évaluer des distorsions engendrées par des algorithmes de chiffrement [Mao 2006] ne parviennent pas à égaler les performances de **CPA**.

Metric	RMSE	Outlier Ratio	Pearson Corr	Spearman Corr
PSNR	0.555	0.460	0.909	0.912
SSIM	0.730	0.640	0.840	0.869
MSSIM	0.634	0.485	0.901	0.952
VSNR	0.540	0.445	0.914	0.918
VIF	0.459	0.345	0.939	0.940
CPA	0.372	0.300	0.961	0.970
ESS	0.53	0.400	0.912	0.93
LSS	0.54	0.400	0.914	0.927

TABLE 3.1 – Comparaison des performance de 8 métriques (chiffrement sélectif).

Dans cette étude, ce qui nous intéressait principalement n'était pas la corrélation des **MOS** avec les **MOSp**, mais plutôt de connaître le pouvoir de discrimination des métriques. C'est à dire que si nous fixons un seuil de détection sur les métriques, nous souhaitons idéalement être capable de diviser les données en deux clusters de

qualité cohérents. Nous pouvons observer sur la Figure 3.16 que pour la métrique **SSIM** lorsqu'un seuil $MOSP_T$ est fixé sur les valeurs prédites par **SSIM**, le nuage de points se divise en 4 quadrants : les Faux Positifs (FP), les Faux Négatifs (FN), les Vrais Positifs (VP) et les Vrais Négatifs (VN). Bien évidemment, notre objectif est de minimiser les Faux Négatifs et les Faux Positifs, tout en maximisant les Vrais Positifs et Négatifs.

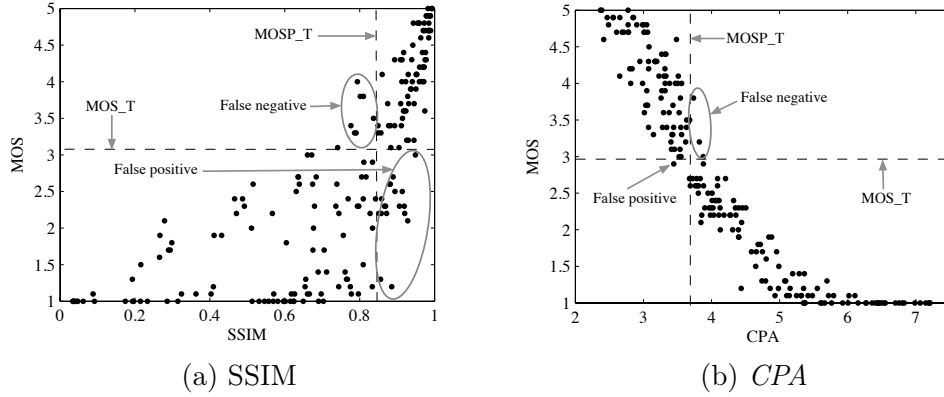


FIGURE 3.16 – Comparaison de la fidélité de deux métriques (SSIM & CPA) pour estimer la qualité du chiffrement sélectif.

La métrique qui sera à même de discriminer efficacement les points parmi ces 4 quadrants sera la plus à même de prédire la perte de qualité engendrée par le chiffrement sélectif. Nous avons donc établi une mesure de discrimination :

$$E(T) = \sum_{i \in (\mathcal{I}_{FP}^T \cup \mathcal{I}_{FN}^T)} |MOS(i) - MOS_T|^2 \quad (3.11)$$

Plus cette mesure de $E(T)$ est faible meilleure est la discrimination, c'est à dire que les 4 clusters (FP, FN, VP, VN) sont optimisés.

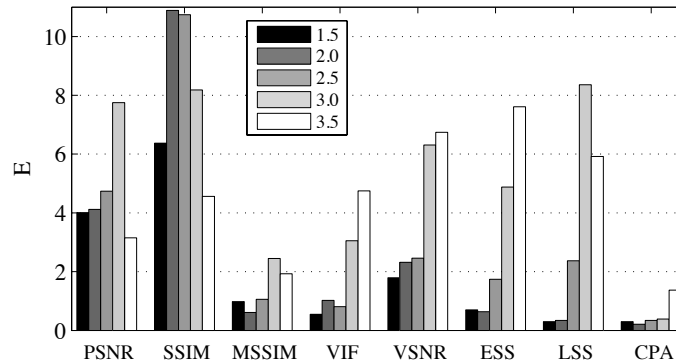


FIGURE 3.17 – Pouvoir de discrimination des différentes métriques dans les 4 clusters (FP, FN, VP, VN).

Nous constatons sur la Figure 3.17 que CPA offre ici les meilleures performances. Dans cette figure, nous avons fixé différents seuils de détection (MOS_T) et évalué pour chacun de ces seuils la fiabilité des métriques. Encore une fois, les deux métriques les plus utilisées dans la communauté de la sécurité multimedia (le PSNR et SSIM) présentent (et de loin) les pires performances.

3.2.2 La reconnaissance des images

Très récemment, par le biais d'un projet collaboratif entre le laboratoire RMeS et l'Université de Salzbourg en Autriche, j'ai pu me pencher sur une problématique encore très peu (voire même pas du tout) étudiée : la reconnaissance subjective des images chiffrées. Ces travaux sont assez proches de ce que nous venons de voir dans la section précédente, mais ici nous allons plus loin, nous ne cherchons pas simplement à évaluer la qualité des images, mais plutôt à déterminer si elles peuvent être reconnues par les observateurs. L'université de Salzbourg et plus particulièrement le laboratoire "WaveLab" d'Andreas Uhl possède une expertise mondialement reconnue dans le chiffrement sélectif des contenus multimedia. Andreas Uhl et son collègue Heinz Hofbauer ont fait appel à mes services pour conduire une étude approfondie sur la perception des distorsions, et plus particulièrement l'étude de la reconnaissance visuelle des contenus très fortement dégradés. Cette étude a été menée entre 2017 et 2020. Nous avons pu proposer à la communauté du chiffrement sélectif diverses recommandations afin de conduire correctement une étude de reconnaissance perceptuelle des images [Hofbauer 2018]. Dans le cadre de ces travaux, nous avons proposé un nouveau protocole de tests subjectifs adapté à l'évaluation de la reconnaissance des images chiffrées. Encore une fois, j'ai cherché ici à réconcilier deux communautés qui échangent assez peu : la communauté de la recherche sur la vision humaine et la recherche en cryptographie et sécurité des contenus.

En ce qui concerne la première communauté, celle de l'étude de la vision humaine, une réglementation assez stricte est déjà mise en place, il existe deux groupes distincts qui travaillent spécifiquement sur des recommandations pour la conduite de tests subjectifs et la bonne utilisation des métriques de qualité. Le Video Quality Experts Group (VQEG) et l'International Telecommunication Union (ITU) sont deux organisations qui ont diffusé nombre de rapports, spécifiant très précisément sous quelles conditions les tests doivent être mis en place [ITU-R BT.2022 2012, ITU-R-BT.500-11 2004, Cermak 2009, VQEG_Contributors 2008, VQEG_Contributors 2010, VQEG_Contributors 2011, IEC1999 1999].

Effectivement, il est souhaitable de conduire les tests subjectifs dans des environnements normalisés, permettant de respecter un maximum de contraintes techniques, de façon à pouvoir être en mesure de les répéter à l'identique dans un autre laboratoire. Tout facteur extérieur pouvant influencer la notation subjective doit être maîtrisé (bruit, inconfort de température, éclairage variable, *etc.*).

Les recommandations ITU ou VQEG définissent notamment les protocoles de visualisation, parmi ces divers protocoles, nous pouvons principalement trouver le

DSIS, le “Pair Comparison”, le Two Alternative Forced Choice (2AFC), le Double Stimulus Continuous Quality Scale (DSCQS), ou encore le Absolute Category Rating (ACR), et bien d’autres encore. A travers mes différents projets, j’ai été amené à programmer certains de ces protocoles pour une visualisation en ligne².

La disposition des images, leurs dimensions, le temps d’affichage, la durée maximale des tests, l’échelle de notation sont autant de recommandations fournies afin de standardiser la conduite des tests. Dans ces rapports se trouvent aussi des informations sur la bonne utilisation des métriques, les outils permettant d’évaluer les performances des métriques et l’interprétation des résultats. Cependant, toutes ces recommandations, bien que très complètes, ont été proposées pour des artéfacts relativement proches du seuil de visibilité, elles ne sont pas du tout adaptées à des images de très pauvre qualité. Si nous appliquons un des protocoles standards de tests, comme ceux mentionnés ci-dessus, nous ne serons pas en mesure d’évaluer l’aptitude des utilisateurs à reconnaître les images. Il nous faut alors conduire une étude comme celle que nous avons vu précédemment [Stütz 2010] et fixer des seuils de reconnaissance sur une échelle de qualité, ce qui nous force alors à conduire une interprétation inhabituelle des résultats.

Dans l’étude [Hofbauer 2018], nous étudions trois protocoles possibles³ pour répondre aux besoins d’une étude de “*recognizability*” nous pouvions ainsi proposer un nouveau protocole de test, assimilable au protocole 2AFC, Cette étude était préliminaire à un travail beaucoup plus vaste dans lequel nous avons analysé en détail la “*recognizability*” ou identification du contenu [Hofbauer 2020a].

Le protocole retenu lors de ce projet est donc à choix forcé, 6 images sont affichées sur l’écran, trois dans la partie supérieure, et trois dans la partie inférieure. Les trois images du haut sont des contenus originaux, tandis que les trois images du bas sont dégradées (par différents algorithmes de chiffrement sélectif). Il n’y a qu’un seul contenu qui soit présent à la fois en haut et en bas, les observateurs ont pour consigne de sélectionner la paire d’images correspondante. Une sélection est obligatoire à chaque affichage. La Figure 3.18 représente un exemple de configuration du protocole à choix forcé.

Ce même test a été conduit à la fois dans les locaux de l’Université de Nantes et de l’Université de Salzbourg. Le test conduit à Nantes s’est fait dans des conditions de visualisation standardisées : dans une salle obscure, avec une paroi de couleur uniforme derrière l’écran de visualisation, la luminosité ambiante de la pièce était maîtrisée, les observateurs ont fait un test d’acuité et de vision des couleurs avant de commencer le test. De plus, les observateurs étaient isolés dans la pièce. En ce qui concerne la partie du test conduite à Salzbourg, l’écran de visualisation était placé dans une salle de travail, l’illumination n’était pas maîtrisée et dépendait donc des conditions météorologiques, l’acuité des observateurs n’était pas testée et la salle, étant occupée par plusieurs personnes n’était pas plongée dans une atmosphère

2. <http://websubjexp.polytech.univ-nantes.fr>

3. <http://websubjexp.polytech.univ-nantes.fr/FCexp/index.html>,
<http://websubjexp.polytech.univ-nantes.fr/FCexp2/index.html> et
<http://websubjexp.polytech.univ-nantes.fr/FCexp3/index.html>



FIGURE 3.18 – Configuration du protocole de test à choix forcé, l’observateur a pour objectif de trouver la paire d’images.

silencieuse.

Outre le setup subjectif normalisé ou dans un environnement de bureau, nous avons aussi proposé une version “application web” du test⁴.

La base de données générée lors de ces travaux était constituée de 6 distorsions différentes, le volume d’image généré étant assez important, nous avons divisé la base en 2 ensembles distincts (Set 1 & Set 2). Pour chacun de ces 2 ensembles d’images, 45 observateurs ont été enrôlés à Nantes et 30 observateurs à Salzbourg pour les tests subjectifs de reconnaissance des images.

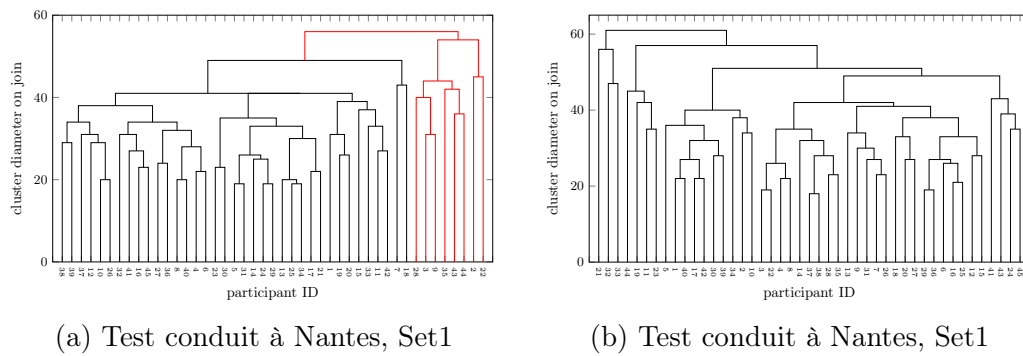


FIGURE 3.19 – Clustering par dendrogrammes sur les observateurs des 2 tests conduits à Nantes.

Une fois les tests subjectifs conduits, il est alors nécessaire de s’assurer que les évaluations fournies par les observateurs soient toutes cohérentes, nous avons proposé dans des travaux antérieurs [Urvoy 2014b] d’utiliser les dendrogrammes comme méthode de clustering des observateurs. Dans ces travaux sur la reconnaissance des images chiffrées, nous utilisons encore une fois cette méthode pour trier les observateur en fonction de la cohérence de leurs observations. La Figure 3.19 nous montre deux exemples de dendrogrammes, dans celui de gauche, les votes de 8 observateurs

4. <http://websubjexp.polytech.univ-nantes.fr/FCexp3SubSet/index.html>

(affichés en rouge) sont jugés incohérents par rapport au reste des évaluations.

Nous disposons donc d'un ensemble $I_i \in I, i = 1, \dots, \#I$ d'images évaluées par $O_j \in O$ pour $j = 1, \dots, \#O$ observateurs. Les votes de chaque observateur sont représentés par un vecteur binaire (image reconnue ou pas : match des images ou pas) : $O_j = (o_j^1, \dots, o_j^{\#I}), j = 1, \dots, \#O$ avec $o_j^i \in \{0, 1\}$. Un score de reconnaissance est défini pour chaque image $S(I_i)$:

$$S(I_i) = \frac{1}{\#O} \sum_{j=1}^{\#O} \#O o_j^i. \quad (3.12)$$

La probabilité de trouver fortuitement une correspondance s'élève à $p_{\text{Match2}}^r = \frac{1}{3} \frac{1}{3} = 0.1\bar{1}$, effectivement, nous avons une chance sur 3 dans la partie supérieure, et aussi une chance sur 3 dans la partie inférieure de l'écran. Nous souhaitons définir un seuil T tel qu'une image I avec $S(I) > T$ sera considérée comme non reconnue, nous souhaitons en outre que moins de 5% des images non reconnues soient classifiées par erreur comme reconnues.

La probabilité pour une image non reconnue U d'avoir V décisions reconnues, et un score $S(U) = \frac{\#O-V}{\#O}$ est alors le nombre de vecteurs possibles de la forme (c_V) fois la probabilité d'un tel vecteur (p_V) , avec

$$c_V = \binom{\#O}{V}, \quad (3.13)$$

$$p_V = p_*^r V (1 - p_*^r)^{\#O-V}. \quad (3.14)$$

La probabilité de n'obtenir pas plus de V décisions "reconnues", soit $S(U) \geq \frac{\#O-V}{\#O}$ peut ainsi être calculée

$$P(V) = \sum_{i=0}^V c_i p_i = \sum_{i=0}^V \binom{\#O}{i} p_*^r i (1 - p_*^r)^{\#O-i}. \quad (3.15)$$

Ceci nous permet donc de calculer un seuil T et V_T tel que $P(V_T) > 0.95$. Ceci signifie que la probabilité qu'une image non reconnue par les observateurs soit considérée reconnue à tort est inférieure à 5%.

Dans un premier temps, nous avons testé 8 métriques de qualité de la littérature sur les bases de données subjectives obtenues. Pour rappel, il y avait 2 emplacements de test : Nantes et Salzbourg, et pour chaque ville, il y avait 2 sous-ensembles d'images. Chaque test comportait 45 observateurs à Nantes et 30 à Salzbourg. Nos tests ont montré que pour une telle tâche de reconnaissance, l'environnement de visualisation importait peu, il n'est donc pas nécessaire de conduire un tel tests dans des conditions de visualisation standardisées telles que celles recommandées par les standards ITU ou VQEG. Ceci est tout à fait logique, car les distorsions sont ici bien éloignées du seuil de visibilité, le processus de reconnaissance des images se fait à un tout autre niveau que l'évaluation de la qualité ou le tracking d'un seuil de visibilité.

Comme nous nous y attendions, les métriques de qualité ne sont pas en mesure de prédire si une image sera reconnue par des observateurs humains. Effectivement,

comme nous l'avons vu précédemment, les métriques de qualité sont implémentées de façon à évaluer les distorsions dans une gamme de qualité qui, sans être très proches du seuil de visibilité, restent néanmoins globalement relativement élevées, elles sont totalement inaptes à juger de la reconnaissance des images.

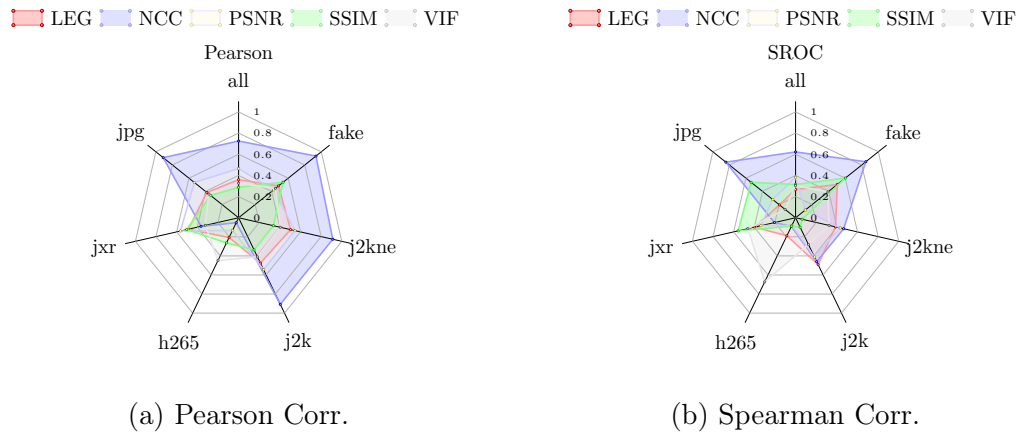


FIGURE 3.20 – Corrélations (Pearson & Spearman)

Au cours de ces travaux, j'ai donc suggéré d'utiliser des outils plus adéquats, tels que le "Template Matching" [Lewis 1995] ou encore du "Feature Matching", tels que Scale-Invariant Feature Transform (SIFT) [Lowe 1999] ou Speeded-Up Robust Features (SURF) [Bay 2006]. Il s'est avéré que l'utilisation du "Template Matching" face à ce type de tâche de reconnaissance était beaucoup plus efficace que les métriques de qualité. La Figure 3.20 nous montre que la Normalized Cross-Correlation (NCC) (ou Template Matching) présente effectivement des performances très intéressantes. Nous notons cependant que face à deux types de distorsions (*jxr* et *h265*), aucune des métriques n'a été en mesure de prédire convenablement la reconnaissance des images. Des travaux sont en cours avec cette même équipe de l'Université de Salzbourg pour tenter de lier l'identification du contenu (*recognizability*) à l'évaluation de la qualité [Hofbauer 2020b].

Au cours de cette étude, nous nous sommes aussi posé 2 autres questions très importantes : Quel est le nombre d'observateurs minimal pour un tel genre de test, et quelle est la taille minimale de la base de données subjectives. Les tests étant parfois assez fastidieux, il est parfois nécessaire de réduire la configuration du test subjectif. Il est alors soit possible de réduire le nombre d'observateurs, mais tout en conservant des évaluations qui soient exploitables, ou alors, il peut être souhaitable de n'exploiter qu'un nombre restreint d'images.

Pour répondre à la première question, nous avons réduit progressivement le nombre d'observateurs et avons étudié l'impact de cette réduction sur la fiabilité (via la Corrélation et la RMSE). Les résultats sont visibles sur la Figure 3.21. Nous estimons qu'une quinzaine d'observateurs suffisent à prédire la reconnaissance des images.

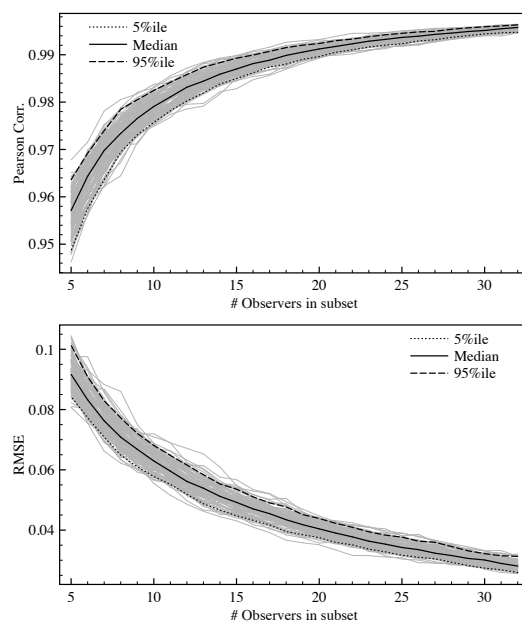


FIGURE 3.21 – Corrélation de Pearson et RMSE sur l'intégralité de la base de données subjectives lorsqu'un nombre variable d'observateurs est considéré.

Quant à la seconde question posée, concernant la taille minimale de la base, nous avons procédé de la même façon que précédemment, nous avons sélectionné plusieurs sous-ensembles de la base initiale et avons testé les différentes métriques sur ces bases réduites. Les résultats peuvent être observés sur la Figure 3.22. Au vu de ces résultats, il semblerait que le dimensionnement de la base pour une problématique de reconnaissance des contenus soit relativement peu important.

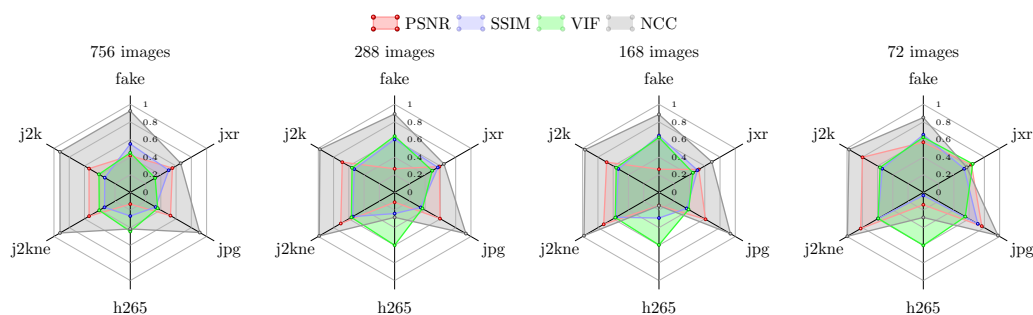


FIGURE 3.22 – Corrélations de Pearson pour différentes tailles de bases de données lorsque 15 observateurs ont été retenus.

Nous avons vu dans ce chapitre mes différentes tentatives pour réconcilier deux thématiques de recherche distinctes : la perception humaine et la sécurité de l'in-

formation. Comme expliqué précédemment, ces deux communautés scientifiques éprouvent souvent quelques difficultés pour se mélanger. À travers plusieurs travaux portant sur l'évaluation objective et subjective de la qualité, ou encore sur la modélisation du SVH, j'ai tenté d'étudier conjointement ces deux thèmes, comme nous venons de le voir.

À travers ces différents travaux, de 2004 à 2016, je portais donc la thématique de recherche "protection des droits d'auteurs des images et documents numériques" au sein du laboratoire IRCCyN. La spécificité de cette thématique dans l'équipe était de conjuguer le tatouage des images avec des critères de vision avancés. Dans le début des années 2010, la protection des copyrights est devenue une thématique assez mature et largement commercialisée, la recherche fondamentale n'était plus que marginale. J'ai donc réorienté mes thématiques de recherche vers d'autres domaines, dont principalement l'imagerie médicale comme nous le verrons par la suite dans le Chapitre 6.

Le prochain chapitre est encore dédié à la présentation de mes travaux portant sur la perception mais cette fois ci, nous nous pencherons plus particulièrement sur la perception des couleurs.

La perception des couleurs

Contents

4.1	Adaptation temporelle des champs récepteurs +s/-s	59
4.2	Modélisation optique de l'oeil	62
4.3	La variabilité des observateurs	66

Mes travaux sur la perception des couleurs ont été conduits selon deux approches différentes. J'ai tout d'abord conduit des recherches fondamentales au cours de mon post-doctorat au sein de l'Université de Chicago, dans le [VSL](#). Par la suite, via l'encadrement de la thèse CIFRE d'Abhijit Sarkar, nous avons conduit des travaux de recherche appliquée et adaptée à l'industrie de la colorimétrie.

Lors de mon post-doctorat, j'ai pu travailler aux côtés du Professeur Steve Shevell sur l'étude de l'absorption des photorécepteurs réiniens [[Autrusseau 2006b](#)]. Nous avons notamment cherché à comprendre quels mécanismes neuronaux entraient en jeu lors d'une variation temporelle d'un stimulus engendrant une induction chromatique. Ensuite, toujours au [VSL](#) avec Steve Shevell, nous avons travaillé sur les aberrations chromatiques de front d'onde. Cette étude m'a amené à développer un modèle optique de l'oeil [[Autrusseau 2011b](#)], nous nous intéressons aux imperfections optiques et leur impact sur la perception des couleurs. Une collaboration très intéressante a vu le jour avec le Professeur Larry Thibos du collège d'optométrie à l'Université de l'Indiana. Son modèle optique a donc été extensivement testé et nous a permis de modéliser un observateur standard d'un point de vue optique.

À mon retour dans le laboratoire [IRCCyN](#), fort de ces compétences nouvelles en perception des couleurs, j'ai alors entrepris de co-encadrer une thèse CIFRE (en collaboration avec Technicolor à Rennes) sur le métamérisme observateur. Ce projet avait pour application l'étude de la science des couleurs dans l'industrie de l'entertainment multimedia. L'encadrement de la thèse d'Abhijit Sarkar nous a permis d'approfondir ce sujet encore assez peu étudié [[Sarkar 2009b](#), [Sarkar 2009a](#), [Sarkar 2010d](#), [Sarkar 2010c](#), [Sarkar 2010b](#), [Sarkar 2010a](#), [Sarkar 2011a](#)].

4.1 Adaptation temporelle des champs récepteurs +s/-s

Au cours de mon stage post-doctoral au [VSL](#), j'ai pu étudier la réponse des champs récepteurs antagonistes +s/-s (absorbant majoritairement le spectre de lumière bleue), et notamment la réaction de ces champs neuronaux face à des stimuli variant dans le temps. Nous pouvons observer sur la [Figure 4.1](#) deux configurations

différentes de cercles concentriques de couleurs respectives vert-violet parmi lesquels est inséré un anneau de couleur orange. L'anneau orange se trouve être soit contigu avec un cercle vert (à gauche), soit avec un cercle violet (à droite). La perception visuelle du cercle de test peut varier (en fonction de la distance de visualisation, il peut être nécessaire de s'éloigner de la figure pour que l'illusion prenne forme) pour typiquement paraître orange dans la configuration représentée à gauche et plutôt rose dans la configuration de droite. Ce phénomène s'appelle métamérisme, c'est la capacité qu'a notre cerveau de percevoir des couleurs différemment en fonction de leur entourage ou de leur illumination. Effectivement, des couleurs métamères sont des couleurs composées d'un spectre lumineux différent, mais que la vision humaine n'est pas en mesure de différencier.

La génération de ce métamérisme par la manipulation de la chromaticité du voisinage, comme nous le faisons dans ces travaux s'appelle l'induction chromatique. Mes travaux ont fait suite aux travaux de Patrick Monnier à l'Université de Chicago [Monnier 2004]. Cette illusion est due à la présence des champs récepteurs corticaux ressemblant à une différence de Gaussiennes 2D. Les couleurs choisies ici (vert et violet) sont choisies de façon à être en opposition sur l'axe chromatique des cônes absorbant le bleu (cônes S), la fréquence d'affichage des anneaux concentriques est déterminée de façon à ce que 3 anneaux concentriques englobent la largeur des champs récepteurs. Lors du test, les stimuli étaient générés spécifiquement dans un espace chromatique perceptuel, celui de MacLeod & Boynton [MacLeod 1979] qui permet de maîtriser très précisément l'excitation des trois types de cônes rétinien.

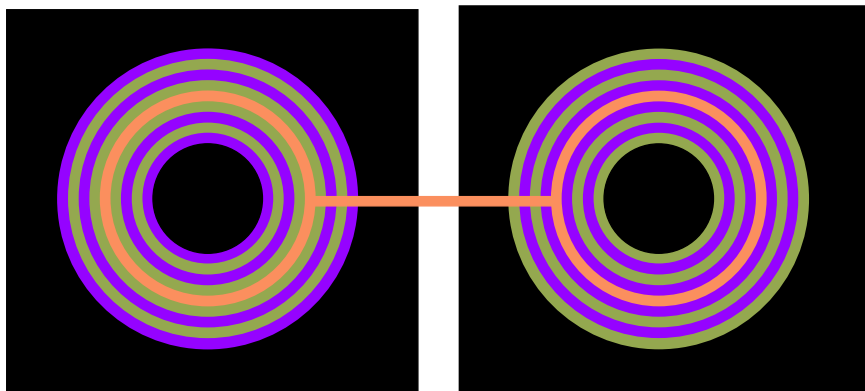


FIGURE 4.1 – Perception altérée d'un anneau orange / rose intercalé entre des anneaux vert-amanche / violet.

Au cours de ces travaux, un test subjectif a été conduit. Lors de ce test, des observateurs humains étaient installés dans une salle de test subjective (salle totalement noire) et devaient grâce à un logiciel adéquat, moduler l'amplitude des chromaticités de anneaux concentriques vert et violet de façon à annuler l'induction chromatique, et de ce fait, observer très précisément la même couleur pour les deux anneaux de test (orange/rose).

La Figure 4.2 nous montre pour un observateur du test la modulation d'am-

plitude nécessaire à compenser l'induction chromatique, nous pouvons ainsi observer que dans le cas d'une configuration d'anneaux concentriques vert-violet (Figure 4.2(a)), une amplitude négative est nécessaire, tandis que dans le cas d'une configuration violet-vert ((Figure 4.2(b)) l'amplitude de compensation est positive. Nos résultats ont donc montré que la modulation chromatique temporelle contigüe a nécessité une inversion de phase pour atteindre la compensation de l'induction chromatique (phénomène d'assimilation) dans le même temps, l'induction chromatique induite par les anneaux non-contigus ont nécessité une modulation en phase avec l'anneau de test pour compenser l'induction chromatique (phénomène de contraste induit).

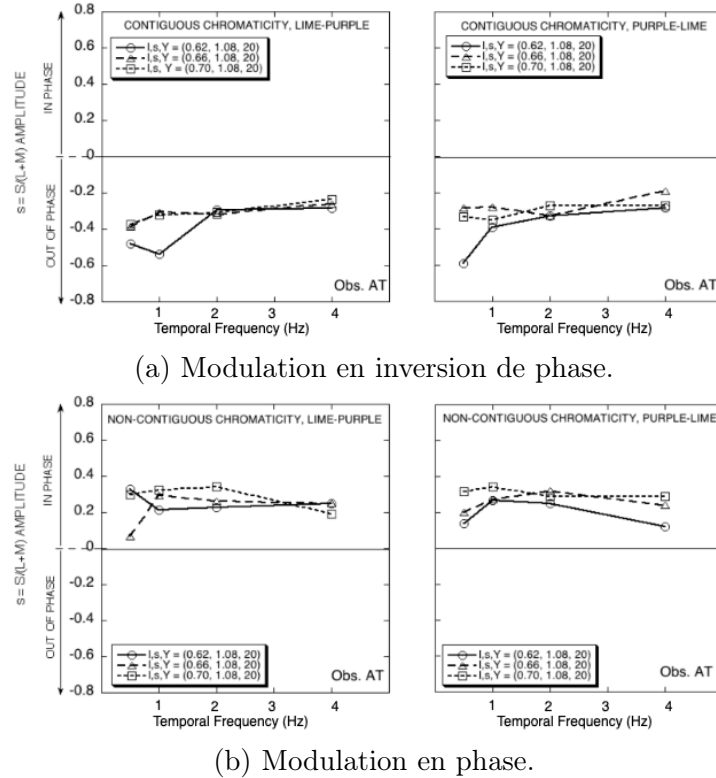
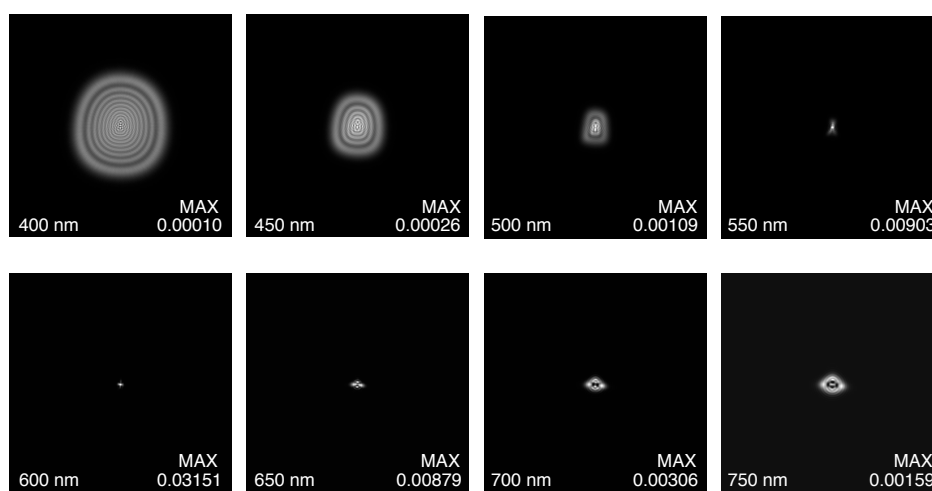


FIGURE 4.2 – Aperçu de l'inversion de phase que l'on peut rencontrer (phénomène d'assimilation) (a) et à l'inverse l'induction chromatique est induite en phase (contraste) en (b).

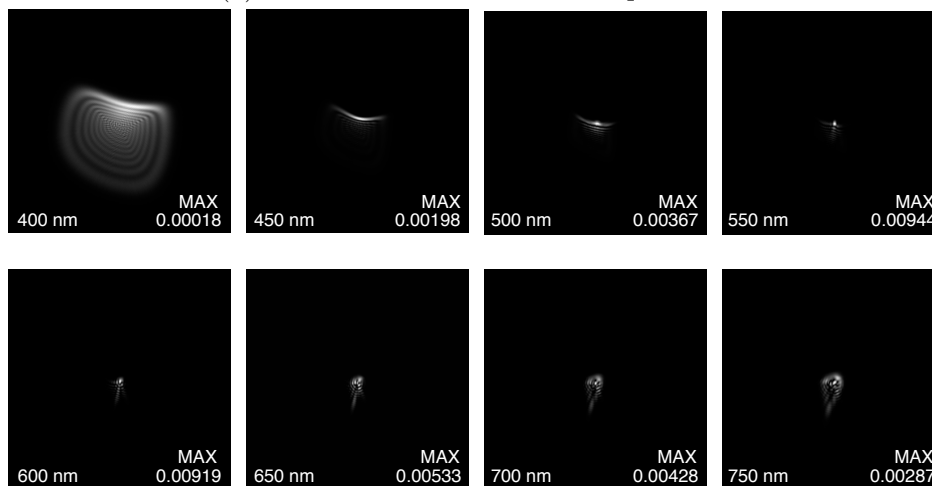
Par la suite, au cours de ce post-doctorat, j'ai pu me pencher sur des problèmes différents de la perception des couleurs, qui cette fois impliquaient des déformations optiques de l'oeil et leur impact sur la fonction d'étalement du point, la Point Spread Function (PSF), voyons, à présent, le détail de ces travaux dans la prochaine section.

4.2 Modélisation optique de l'oeil

La fonction d'étalement du point (ou **PSF**) est la réponse de tout système optique face à un stimulus lumineux ponctuel. La **PSF** peut en quelque sorte caractériser le système optique, une **PSF** étalée signifie généralement que l'optique du système est imparfaite. Chaque oeil humain possède bien évidemment sa propre **PSF**. Outre l'impact des imperfections optiques sur l'acuité de l'observateur, elles jouent aussi un rôle important sur la perception des couleurs. Effectivement, un flux lumineux chromatique traverse un système optique avec différentes distances focales en fonction de la longueur d'onde, ce phénomène s'appelle **LCA**.



(a) Modélisation de la LCA uniquement.



(b) Prise en compte de la LCA et des aberrations de front d'onde.

FIGURE 4.3 – Comparaison entre l'aberration chromatique seule (a) et la prise en compte des aberrations de front d'ondes (b).

De nombreux modèles de la littérature ne considèrent que la **LCA** comme

aberrations chromatiques ayant une influence sur la perception des couleurs [Marimont 1994], or, dans nos travaux [Autrusseau 2011b], nous avons étudié l'impact des aberrations de front d'onde sur la PSF et donc sur la perception chromatique. Les aberrations de front d'onde sont des aberrations de plus haut niveau qui peuvent être vues comme une modélisation des imperfections de la surface de la lentille, chaque personne dispose d'aberrations de front d'onde différentes, et bien évidemment, chacune de ces aberrations a un impact différent sur la PSF. Au cours de nos travaux, nous avons proposé un modèle optique de l'oeil qui comprend à la fois la LCA, les aberrations de front d'onde et l'absorption des photorécepteurs rétiniens (les trois types de cônes, L, M et S). Les trois types de cônes qui tapissent notre rétine sont les photorécepteurs qui permettent d'absorber les longueurs d'onde correspondant grossièrement au rouge, vert et bleu. Nous avons tout d'abord adapté le modèle optique de Larry Thibos [Thibos 1999], qui était destiné à modéliser les imperfections optiques pour un signal ne comportant qu'une composante de luminance (pas de composantes chromatiques), puis grâce à un modèle statistique d'observateurs [Thibos 2002], nous avons pu générer cent PSF correspondantes à des observateurs différents (plus précisément cent yeux différents via cent aberrations de front d'onde différentes). Les aberrations de front d'onde sont générées à partir de polynômes de Zernike, nous avons donc calculé les PSF pour chacun et en avons déduit un "oeil standard". L'objectif ici étant d'estimer la différence entre un modèle comportant uniquement la LCA et un modèle comportant à la fois la LCA et les aberrations de front d'onde. La Figure 4.3 nous montre les PSF entre 400 et 750 nm (dans le spectre visible tous les 50 nm) lorsque la LCA est considérée uniquement (figure 4.3(a)), et quand les aberrations de front d'onde sont prises en compte dans le modèle (figure 4.3(b)).

Sur cette Figure, il paraît évident que les aberrations de front d'ondes distordent sensiblement la PSF et nous pouvons notamment observer une dissymétrie importante comparée au scénario modélisant uniquement la LCA. Hormis la PSF un autre outil est largement utilisé pour évaluer les systèmes optiques, c'est la Modulation Transfer Function (MTF). Cette dernière nous fournit la réponse d'un système aux fréquences spatiales. Plus la PSF est étalée, plus elle risque de distordre des stimuli fréquentiels complexes. La MTF représente la perte de contraste engendrée par divers stimuli de fréquences variables.

Des stimuli comme ceux représentés sur la Figure 4.4 ont du être générés pour tester la réponse du système optique face à différentes fréquences spatiales, et ainsi étudier la perte d'amplitude du signal engendrée par ces imperfections.

Nous verrons plus tard dans la Figure 4.7 comment un stimulus disposant d'une certaine fréquence peut se retrouver avec une amplitude réduite suite à la diffraction optique. Cet écart (cette perte) d'amplitude est justement représenté par la MTF.

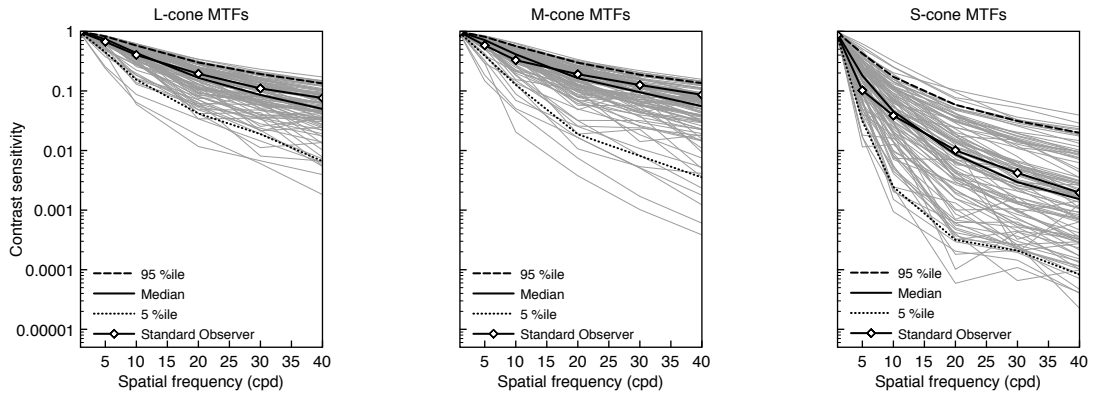
Notre objectif dans ces travaux était de trouver la configuration optique moyenne dans une large population. Cette MTF médiane a pu être obtenue en extrayant l'observateur médian parmi les cent modèles, c.f. Figure 4.5. Nous avons donc généré différents stimuli sinusoïdaux 2D, de fréquences variant entre 1 et 40 cycles par degré d'angle visuel, puis notre modèle optique a ainsi pu être testé sur ces diffé-



(a) Stimulus de faible fréquence spatiale. (b) Stimulus de haute fréquence spatiale.

FIGURE 4.4 – Génération de deux stimuli de diverses fréquences spatiales pour calculer la *MTF*.

rents signaux afin d'en étudier la réponse (la perte de contraste). L'observateur se rapprochant le plus de la médiane peut ainsi être considéré comme un observateur "standard" d'un point de vue optique. Comme nous le voyons sur la Figure 4.5, plus la fréquence spatiale des stimuli augmente, plus la perte de contraste est significative, c'est tout à fait normal, la *PSF*, trop large, va s'étendre et "baver" sur les fréquences voisines du stimulus.



(a) Absorption des cônes L (b) Absorption des cônes M (c) Absorption des cônes S

FIGURE 4.5 – Absorption des trois types de cônes (L, M & S) pour 100 modèles optiques d'observateurs.

Finalement, nous pouvons observer sur la Figure 4.6 la différence de perception lorsque le modèle optique est appliqué sur la lettre "E". Nous pouvons apprécier la différence de contraste lorsque les aberrations de front d'onde sont omises (en haut), et lorsqu'elles sont considérées (en bas). Encore une fois, comme nous avons pu l'observer dans la Figure 4.3, la différence est significative, et nous constatons que les différences sont tangibles sur les trois types de cônes.

Finalement, nous pouvons observer sur la Figure 4.7 la différence d'absorption sur les cônes *S* lorsque la *LCA* est modélisée seule, et lorsque les aberrations de front

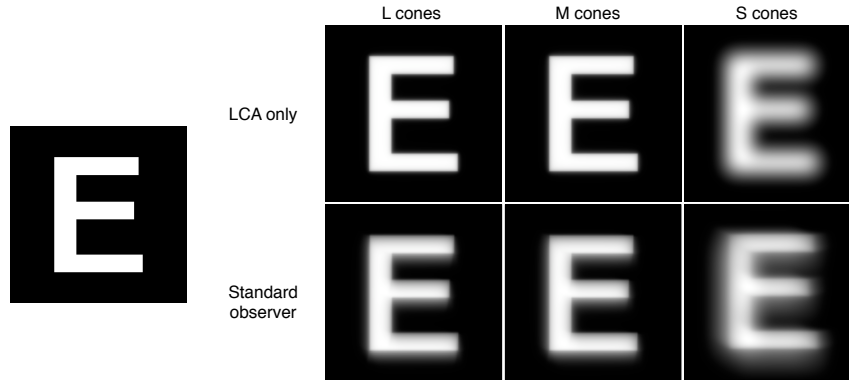


FIGURE 4.6 – Différence de perception sur la lettre “E” , lorsqu’uniquement la LCA est considérée, et quand les aberrations de front d’onde s’y ajoutent.

d’onde de l’observateur standard s’y ajoutent. Nous pouvons aussi voir la différence avec le modèle proposé dans travaux de [Marimont 1994]. Le stimulus de test est une image2D (nous ne représentons ici que le profil 1D, soit une ligne du stimulus) qui prend la forme de bandes blanches sur un fond noir. Encore une fois, l’asymétrie apportée par les aberrations de front d’onde apparait de façon évidente sur cette Figure.

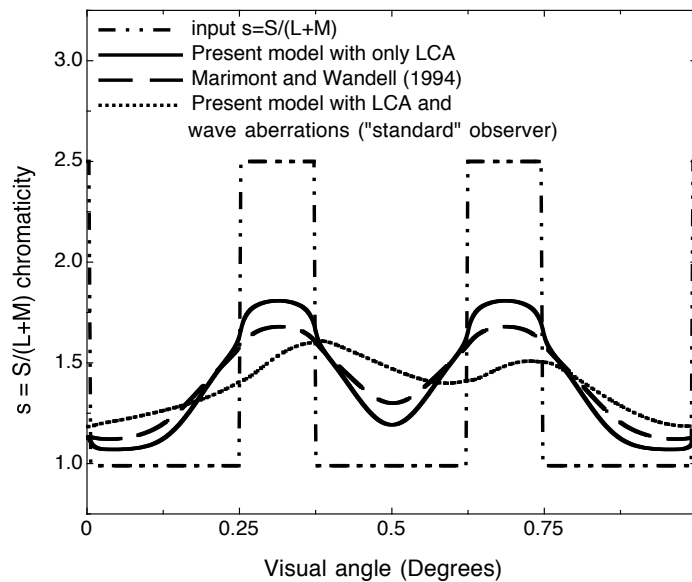


FIGURE 4.7 – Comparaison du profil 1D sur une image stimulus de bandes verticales noires et blanches. On distingue les profils avec et sans aberrations de front d’onde et les travaux de [Marimont 1994].

4.3 La variabilité des observateurs

Lorsque deux stimuli colorés produisent la même réponse visuelle, nous parlons de correspondance de couleur. Deux stimuli ayant des distributions spectrales de puissances différentes peuvent donner lieu à des réponses de cônes identiques, résultant en une correspondance des couleurs. Cependant si cette correspondance est établie pour un observateur elle peut donner, et donne souvent, une discordance ou un désaccord pour un autre observateur, car ce second observateur a des fonctions colorimétriques, les Color Matching Functions (CMF), différentes du premier. Ce phénomène est communément dénommé métamérisme observateur.

Plusieurs études passées ont donné une vue précise sur la variabilité observateur pour la mise en correspondance des couleurs, et sur ses ramifications en science de la couleur et en colorimétrie appliquée. Alors que notre connaissance des causes physiologiques sous-jacentes à la variabilité individuelle de la perception de la couleur s’est enrichie considérablement ces vingt dernières années, il manque toujours une solution pratique pour prendre en compte la variabilité dans les applications de la colorimétrie. Contrainte par un unique modèle d’observateur moyen, la colorimétrie ne sait prédire comment des mises en correspondance individuelles peuvent différer d’une mise en correspondance moyenne. Les conséquences ne sont pas anodines pour certaines applications industrielles où la perception de la couleur est critique.

A titre d’exemple, on peut citer le processus d’ajustement de couleurs (appelé étalonnage couleur ou “color grading”) réalisé dans la phase de postproduction du cinéma ou de la télévision où le contenu filmé brut issu du tournage est modifié pour obtenir l’effet couleur recherché. Le Coloriste doit collaborer avec le Directeur de la Photographie pour ajuster les couleurs du contenu original afin d’assurer cohérence et homogénéité au long des différentes scènes, tout en conservant l’intention artistique souhaitée par le réalisateur et par le directeur de la photographie. Toutefois, si le Coloriste et le Directeur de la Photographie ont des caractéristiques de vision colorée différentes, ils perçoivent les couleurs différemment, et des couleurs qui semblent similaires à l’un seront perçues différentes pour l’autre. Alors que le savoir-faire du Coloriste permet de combler l’écart, la colorimétrie conventionnelle ne saura pas rendre compte de ces différences de vision des couleurs.

L’objectif général du travail de thèse d’Abhijit Sarkar (entre 2008 et 2011) était de proposer un cadre fonctionnel et un processus de traitement couleur prenant en compte la variabilité individuelle, et de fournir une solution appliquée pour l’industrie.

Lors de ce travail, nous supposons que les observateurs humains présentant une vision normale des couleurs pouvaient être classés en un nombre restreint de catégories correspondant à leur vision des couleurs. Ces catégories d’observateurs, caractérisées par leurs fonctions colorimétriques (CMFs), peuvent être identifiées par une analyse statistique exploitant les données d’un large ensemble d’observateurs. Sur la base d’une telle catégorisation de l’ensemble d’une population d’observateurs, on pourra définir plusieurs modèles d’observateurs utilisables en colorimétrie appliquée.

Il était de plus souhaitable que ces catégories puissent être correctement identi-

fiées. En d'autres termes, nous nous basions sur l'hypothèse qu'il existe une façon simple et pratique de déterminer de façon expérimentale, dans quelle catégorie un observateur donné peut être classifié. Les dispositifs affectés par le métamérisme et la variabilité sont nos meilleurs alliés pour aboutir à une solution à ces problèmes. Ces dispositifs sont par exemple les nouveaux écrans, à couleurs étendues, qui ont des primaires spectralement étroites.

Nous pouvons cependant redouter que les caractéristiques spectrales des fonctions colorimétriques propres à une catégorie d'observateurs ne satisferont pas exactement les observateurs individuels assignés à cette catégorie. Cependant, si la catégorie est correctement identifiée, les résultats colorimétriques obtenus globalement seront vraisemblablement plus précis que les résultats fournis par une autre catégorie ou par un observateur colorimétrique de référence. Ainsi, l'utilisation de catégories d'observateurs colorimétriques pourrait résoudre le problème du métamérisme observateur.

Alors que l'utilisation de catégories d'observateurs colorimétrique peut sembler inadéquate pour des applications où plusieurs observateurs regardant simultanément les couleurs d'un écran, cela peut s'avérer utile dans certaines conditions. Par exemple, lorsque tous les observateurs concernés sont dans un groupe d'âge élevé, il sera plus approprié d'utiliser les catégories sélectionnées préférentiellement par les groupes d'observateurs d'âge élevé.

En effet, ce travail a montré que certaines catégories sont choisies préférentiellement par des observateurs d'âge élevé. Cependant il existe un bénéfice plus direct pour les applications où l'on désire une reproduction précise des couleurs pour un observateur individuel.

Avec Abhijit, nous avons donc étudié le phénomène de métamérisme observateur, et évalué comment la variabilité observateur peut modifier la mise en correspondance des couleurs. Nous avons fait différentes tentatives de quantification du métamérisme observateur dont le "CIE Standard Deviate Observer" sur les déviations par rapport à l'observateur de référence, puis passé en revue les expérimentations classiques de mise en correspondance des couleurs utilisant des stimuli monochromatiques.

Un des aspects fondamentaux de la colorimétrie appliquée est la trichromacie de notre système visuel, Cette dernière nous permet de représenter n'importe quelle couleur en la décomposant sous forme de trois composantes de couleurs primaires (dits "tristimulus"). Calculer les valeurs des tristimulus pour un objet nécessite l'utilisation de la réflectance spectrale de la couleur dudit objet, la distribution spectrale de puissance de l'illumination de la scène, ainsi que les caractéristiques spectrales d'un observateur dit colorimétrique. Pour la communauté de l'imagerie couleur, il est particulièrement intéressant de déterminer la meilleure représentation des observateurs, les fonctions colorimétriques (où CMF) dérivées du modèle d'observateur physiologique proposé par la CIE en 2006, où l'observateur colorimétrique standard (dénommé CIE 10°). Au cours de la thèse d'Abhijit, cette problématique a été explorée via une analyse théorique dans un contexte de colorimétrie des moniteurs.

Dans sa méthode pour définir un observateur fondamental, le comité technique de la CIE (CIE TC 1-36) a ignoré la variabilité individuelle des observateurs [CIE 2006,

[Stockman 1993]. Certaines études [MacLeod 1983] se sont penchées sur la variation individuelle des CMFs, en analysant les données collectées par Stiles et Burch pour un champ de 10° , elles ont examinées les différences entre les CMF de l'observateur colorimétrique standard de la CIE 1931, la révision de cet ensemble proposée par Judd et l'ensemble de CMF 2° collectées par Stiles et Burch [Smith 1983]. Une comparaison a été effectuée entre la variabilité inter et intra-observateurs des CMFs expérimentales [Viénot 1980].

Nous avons exploité le modèle CIEPO06 afin d'examiner via une analyse théorique les effets de l'âge sur la CMF des observateurs et sur des équivalences de couleurs individuelles sur des moniteurs.

Nous avons comparé les effets de différents facteurs en termes de perception des couleurs sur deux moniteurs ayant des caractéristiques spectrales différentes. La distribution spectrale de puissance des couleurs primaires de ces moniteurs est représentée en Fig. 4.8. WG-LCD est représentatif d'un moniteur moderne à large gamut avec des primaires ayant un spectre présentant de fortes variations et Ref-CRT est représentatif d'un moniteur CRT typique et de diffusion HDTV standard.

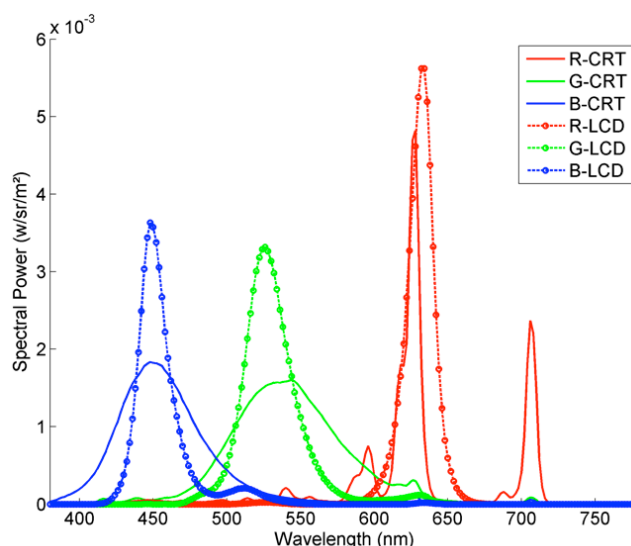


FIGURE 4.8 – Densité spectrale de puissance des deux moniteurs utilisés dans notre analyse.

Afin de simuler les effets induits par divers facteurs physiologiques lors de la visualisation de stimuli chromatiques sur différents moniteurs, nous avons généré sept stimuli de tests qui pouvaient se trouver dans l'intersection des deux gamuts des moniteurs de l'expérience (LCD et CRT). Les chromaticités de ces stimuli ont été choisies de façon à couvrir au mieux le gamut dans l'espace de coordonnées CIE 1976 (u', v').

Nous avons donc conduit une simulation des shifts chromatiques de sept stimuli par simple modification des courbes fondamentales d'absorption des cônes dans un diagramme chromatique (u', v'). La Figure 4.9 nous montre ces shifts de chromatici-

tés (u' , v') des sept stimuli de tests. Ceci représente les effets des courbes d'absorption fondamentales de cônes modifiées sur la perception des couleurs sur CRT (symboles verts) et LCD (symboles rouges). Les carrés représentent une augmentation de 25% de la densité optique du média oculaire (Fig. 4.9(a)), du pigment maculaire (Fig. 4.9(b)) des photopigments des cônes (Fig. 4.9(c)) et finalement un décalage de la sensibilité maximum des longueurs d'ondes des cônes de 4nm (Fig. 4.9(d)). Les triangles, quant à eux, représentent une réduction de 25%.

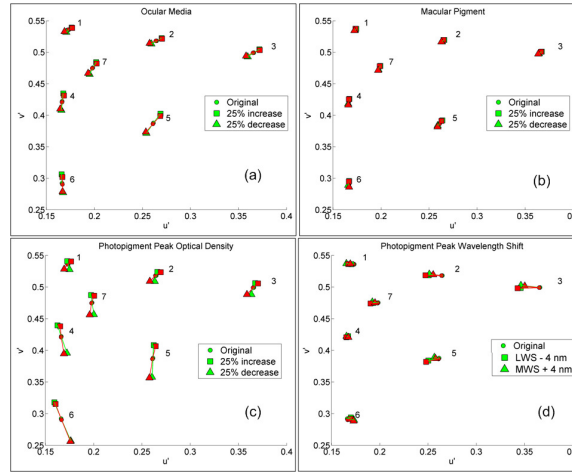


FIGURE 4.9 – Simulation des shifts chromatiques de sept stimuli par simple modification des courbes fondamentales d'absorption des cônes dans un diagramme chromatique (u' , v') (voir texte pour les détails).

Dans un second volet de la thèse d'Abhijit Sarkar, nous avons conduit une expérience subjective lors de laquelle les observateurs devaient atteindre un matching (une mise en correspondance) des couleurs sur les deux moniteurs de tests. Ici, neuf chromaticités différentes ont été choisies, de façon à faire varier l'absorption des cônes L ou M.

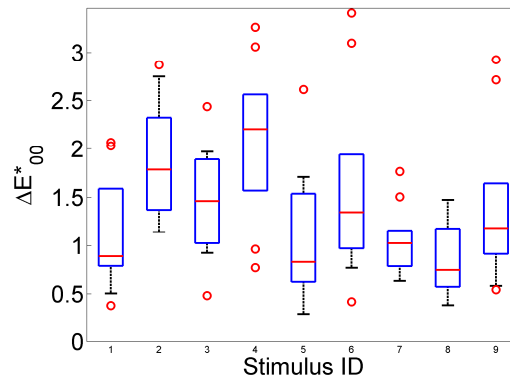


FIGURE 4.10 – Box plots représentant la variabilité inter-observateurs pour l'expérience de "color matching".

Les deux moniteurs étaient placés à 90° l'un par rapport à l'autre, le LCD était positionné en face des observateurs, et le CRT était décalé sur la gauche, un miroir placé entre les deux écrans permettait de composer un stimulus visuel comportant une double chromaticité : affichage CRT dans la partie gauche du champ visuel et LCD dans la partie droite. Le système était installé de façon à ce que la lumière diffusée par chacun des écrans ne puisse pas parasiter l'autre écran. Le champ visuel était calibré pour occuper 10 degrés d'angle visuel. Cette même expérience a été conduite avec deux couleurs de fond différentes, lors du premier test le background était exclusivement noir, tandis que dans le second test, le background était uniformément blanc. Grâce à une molette, chacun des 10 observateurs participant à cette étude devaient faire *matcher* les deux couleurs. La Figure 4.10 représente l'écart chromatique ΔE^* entre les deux couleurs ainsi obtenues. Pour chacun des 9 chromaticités choisies, nous traçons les écarts chromatiques à travers les 10 observateurs ayant pris part à l'étude. Nous pouvons observer une variabilité significative entre les différents observateurs. De plus, cette variabilité semble dépendre de la chromaticité des stimuli. Lors de cette étude, les différences se sont avérées assez minimes entre les écarts de couleurs en fonction de la couleur du fond.

Comme nous venons de le voir dans ce Chapitre, mes travaux portant sur la perception des couleurs se sont faits par deux approches différentes, une étude plus fondamentale a été conduite avec Steve Shevell tandis que nos travaux avec Abhijit Sarkar et ses collègues de Technicolor était plus appliqués.

Dans le chapitre suivant, nous abordons une thématique complètement différente : la Transformation Mojette. Nous verrons comment cette dernière a été exploitée pour s'intégrer dans différents projets ayant trait à la compression ou à la sécurité des données.

La transformation Mojette

Contents

5.1	Transformation Mojette	71
5.1.1	Transformation Directe	71
5.1.2	Inversion de la Transformation Mojette	72
5.2	Compression Mojette	76
5.3	Sécurité basée Mojette	83
5.3.1	Applications au tatouage	84
5.3.2	Applications au chiffrement	86

5.1 Transformation Mojette

La “Mojette” est une transformation par projection de l’image. Elle peut être vue comme une transformée de Radon discrète [Deans 1983]. Elle permet de calculer des projections en une dimension par simple sommation des pixels traversés dans une image 2D. Elle peut tout aussi bien projeter un volume 3D sur des plans 2D. Je présente dans cette section le principe de base de la transformation, ainsi que son inversion.

5.1.1 Transformation Directe

La transformation Mojette Directe est définie sous plusieurs formes, la Transformée Mojette Dirac peut être résumée par l’équation 5.1. Pour tout angle discret défini par le vecteur (p, q) , la Mojette calcule les bins formés par la somme des pixels traversés. La transformation Mojette, ainsi que ses différentes applications est largement détaillée dans cet ouvrage : [Guédon 2013].

$$\mathcal{M}_{p_i, q_i} \{f(k, l)\} = \text{proj}_{p_i, q_i}(b) = \sum_{k=-\infty}^{+\infty} f(k, l) \Delta(b + kq_i - lp_i), \quad (5.1)$$

La Figure 5.1(a) nous montre la sommation de deux pixels, de valeurs respectives 3 et 5 vers un bin de valeur 8 pour un angle de projection $(p, q) = (2, 1)$. Les angles de projections, θ_i , sont donc définis par un ensemble de vecteurs (p_i, q_i) tels que $\theta_i = \tan^{-1}(q_i/p_i)$. Les vecteurs (p_i, q_i) sont choisis de façon à ce que p et q soient premiers entre eux, c’est à dire $\text{pgcd}(p_i, q_i) = 1$. Chaque valeur d’une projection est

dénommée “bin”, comme en tomographie. Comme nous pouvons l’observer sur la Figure 5.1(b), les projections n’ont pas toutes la même longueur (le même nombre de bins). Le nombre de bins, B_i , de chaque projection dépend de la direction choisie pour le vecteur (p_i, q_i) , ainsi, pour une image de dimensions $P \times Q$ le nombre de bins sera :

$$B_i = (Q - 1)|p_i| + (P - 1)q_i + 1. \quad (5.2)$$

La complexité algorithmique de la transformation Mojette pour une image $P \times Q$ sur I projections est $O(PQI)$.

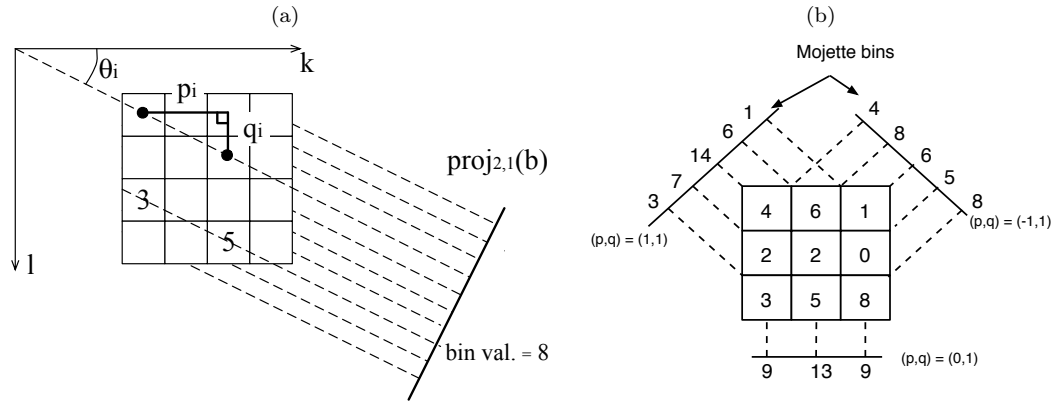


FIGURE 5.1 – (a) Angle θ_i formé par (p_i, q_i) , projection des pixels sur le bin b de la proj $proj_{p_i, q_i}(b)$, (b) Exemple de transformation Mojette sur une grille 3×3 suivant les directions $\{(1, 0), (1, 1), (-1, 1)\}$.

Outre la transformation Mojette dite Dirac, nous pouvons noter qu’il est aussi possible de proposer une Mojette “Spline”. Différents modèles de pixels peuvent donc être utilisés. Nous pouvons observer dans la Figure 5.2 les projections d’un pixel selon le modèle *Spline 0*. Le chapitre 3 du livre [Guédon 2013] explique en détail le fonctionnement de la Mojette *Spline*. Je ne rentrerai pas dans ces explications ici, car tout au long de mes travaux sur la Mojette, à l’exception d’une étude portant sur l’analyse de textures [Dong 2012] j’ai quasi-exclusivement utilisé le modèle Dirac. Nous pouvons cependant noter que sur ce modèle Dirac, il est possible d’effectuer une transformation “modulo”. Chaque bin peut être réduit en modulo (typiquement 256) tout en conservant les propriétés de rétro-projections.

La figure 5.3 nous montre un cas concret de projections sur une image de textures (exemple tiré de [Dong 2012]).

Voyons à présent l’inversion de la transformation Mojette et ses conditions.

5.1.2 Inversion de la Transformation Mojette

Les travaux de Myron Katz ont été très importants pour fixer les conditions de restructibilité [Katz 1978]. Il a été montré qu’en respectant le critère donné en équation 5.3, il était possible de reconstruire un ensemble de pixels de taille $P \times Q$.

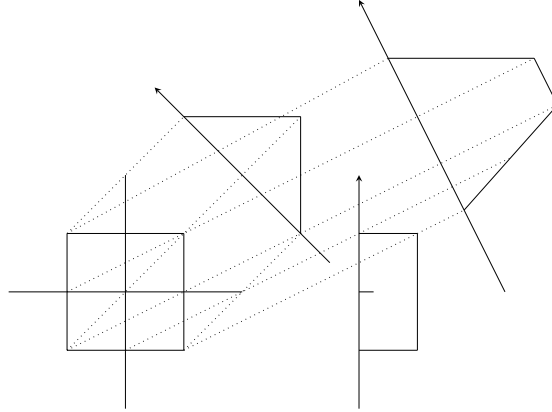


FIGURE 5.2 – Projection d’un pixel via une Transformation Mojette *Spline 0* selon les angles $(1, 0)$, $(2, 1)$ et $(1, 1)$.

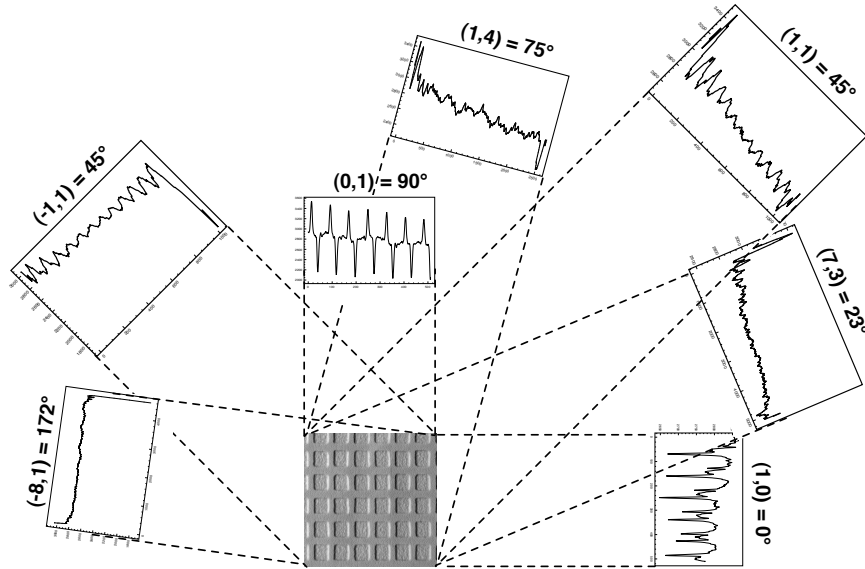


FIGURE 5.3 – Quelques exemples de projections Mojette “Haar” sur une image de texture.

La transformation Mojette est donc inversible sous certaines conditions. La première étude sur le sujet de l’inversion de la Mojette a été publiée en 1996 [Normand 1996b].

$$P \leq \sum_{i=1}^I |p_i| \quad \text{or} \quad Q \leq \sum_{i=1}^I q_i, \quad (5.3)$$

Ici, P et Q représentent les dimensions de l’image à reconstruire, tandis que (p_i, q_i) représente le set de projections à notre disposition.

La transformation Mojette inverse telle que proposée dans [Normand 1996b] peut être représentée comme sur la Figure 5.4. Cette méthode de reconstruction

est itérative, l'algorithme recherche les correspondances 1 bin - 1 pixel, puis rétro-projette la valeur du bin dans la position du pixel à reconstruire. S'ensuit une mise à jour de la projection et des correspondances pixel-bin. Ce processus est répété itérativement jusqu'à la reconstruction complète du support. Une seconde méthode, plus rapide et plus stable a été proposée quelques années plus tard dans [Normand 2006b].

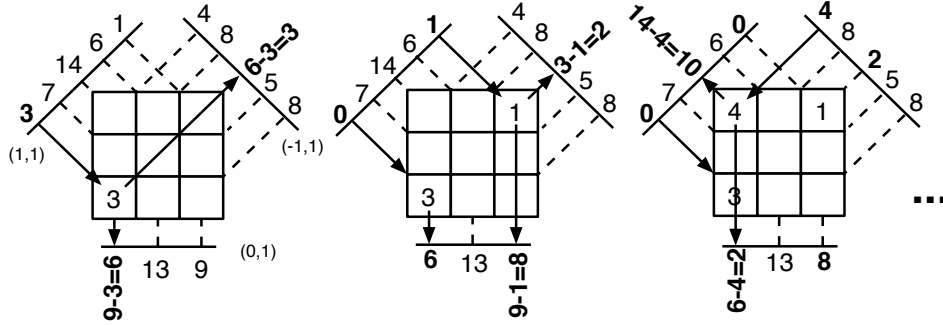


FIGURE 5.4 – Trois premières étapes d'une reconstruction Mojette. 5.1.

La transformation Mojette Directe peut être décrite par l'algorithme suivant :

Algorithme 1 Algorithme de la Mojette Directe

```

pour  $lig = 1$  à  $Q$  faire
  pour  $col = 1$  à  $P$  faire
     $m \leftarrow p \times lig - q \times col$  // calcul du bin  $m$  correspondant
     $proj+ \leftarrow image(lig, col)$  // incrémentation du bin

```

L'algorithme de la Mojette inverse est plus complexe, il requière plusieurs étapes :

Algorithme 2 Algorithme de la Mojette Inverse

```

pour  $compteur = 1$  à  $P \times Q$  faire
   $(i, m) \leftarrow Phase\_1$  // chercher correspondance univoque
   $(lig, col) \leftarrow Phase\_2(i, m)$  // chercher élément d'image correspondant
   $Phase\_3$  // mise à jour sur chaque projection incrémentation du bin
  //l'image est reconstruite

```

Comme nous pouvons le voir, il comporte 3 phases distinctes que nous allons à présent détailler :

Tout d'abord la recherche des bins à correspondance univoque (bins n'étant issus que de la projection d'un unique pixel).

Algorithme 3 *Phase_1*

```

// chercher correspondance univoque
pastrouvé ← vrai, pasfini ← vrai
tant que pastrouvé et pasfini faire
    on parcourt les projections  $proj(i, m)$ 
// Ici, on a trouvé  $(i, m)$  ou bien le support n'est pas restructurable

```

Lors de la seconde phase, on “rétro-projette” le bin à correspondance univoque dans l'image à reconstruire.

Algorithme 4 *Phase_2*

```

// reconstruisons le pixel correspondant : il n'existe qu'un seul élément de image
dont la valeur est  $-1$  sur la ligne  $m = p_i \times lig + q_i \times col$ 
 $col \leftarrow 1$ , pastrouvé ← vrai
tant que pastrouvé faire
    si  $(m - q_i \times col)$  est divisible par  $p_i$  (donne  $lig$ ) et  $(image(lig, col) = -1)$ 
    alors pastrouvé ← faux
    sinon  $col \leftarrow +1$ 
fin si

```

La dernière phase permet, pour chaque bin “rétro-projeté”, de mettre à jour le bin correspondant au pixel dans toutes les autres projections.

Algorithme 5 *Phase_3*

```

0: // mise à jour des projections
pour  $i \leftarrow 1$  à  $I$  faire
     $m \leftarrow p_i \times lig - q_i \times col$  // calcul du bin  $m$  correspondant
     $proj(i, m) \leftarrow image(lig, col)$  // décrémentation du bin
     $projélém(i, m) \leftarrow 1$  // décrémentation du # éléments

```

L'inversion de la transformation Mojette par cette méthode est très sensible au bruit. Une modification, même minime des projections entraîne inévitablement une propagation importante de l'erreur lors de la reconstruction. Il y a cependant des alternatives pour reconstruire des projections bruitées, notamment, la “Filtered Back-Projection” (FBP) (se référer au Chapitre 8 de [Guédon 2013] pour les détails) ou l'utilisation du théorème de la tranche centrale.

Effectivement, la transformation Mojette étant une version discrète de la transformation de Radon, le théorème de la tranche centrale peut là aussi s'appliquer [Chandra 2014]. Rappelons ce théorème, qui opère dans le domaine continu. Soit la transformée de Fourier :

$$F(u, v) = \int \int f(x, y) \exp(-2\pi i (ux + vy)) dx dy \quad (5.4)$$

pour la ligne $v = 0$ du spectre de Fourier, nous avons :

$$\begin{aligned}
F(u, 0) &= \int \left[\int f(x, y) dy \right] \exp(-2\pi i (ux)) dx \\
&= \mathcal{F}(proj_{\theta=0}(t))
\end{aligned} \tag{5.5}$$

où $proj_{\theta}(t)$ représente une projection de l'image suivant un angle θ . Pour une ligne passant par l'origine du spectre de Fourier avec un angle θ nous avons :

$$\begin{aligned}
F(r \cos \theta, r \sin \theta) &= \int \int f(x, y) \exp(-2\pi i (xr \cos \theta + yr \sin \theta)) dx dy \\
&= \int \left[\int f(t \cos \theta - s \sin \theta, s \cos \theta + t \sin \theta) ds \right] \exp(-2\pi i r t) dt \\
&= \mathcal{F}(p_{\theta}(t))
\end{aligned} \tag{5.6}$$

avec le changement de variable suivant :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} t \\ s \end{pmatrix} \tag{5.7}$$

Une ligne du spectre de Fourier d'orientation θ correspond donc à la Transformée de Fourier de la projection de l'image selon cet angle.

Pour conclure cette introduction sur la transformation Mojette. Cet outil a été utilisé depuis le milieu des années 1990 pour diverses applications. Son principal intérêt étant la redondance modulable, elle a surtout été exploitée pour des problématiques de stockage ou de transfert de données numériques. Elle dispose de certaines caractéristiques très intéressantes, mais a montré une instabilité notable lors de la rétro-projection de bins erronés et n'a pas été exploitée pour la compression des données image. Il m'a donc paru intéressant de tenter d'exploiter la Mojette pour proposer une méthode de compression sans pertes des données, nous verrons cet aspect dans le détail dans la prochaine section. De plus, il est aisé de tirer profit de l'instabilité de la rétro-projection pour crypter les données, ceci fera l'objet de la section 5.3.2.

5.2 Compression Mojette

Alors que la transformation Mojette a vu le jour au cours des années 1990 [Guédon 1995], et a été exploitée pour une grande diversité d'applications, il a fallu attendre 2006 pour voir apparaître la première méthode de compression sans perte basée sur la Mojette [Atrusseau 2006a]. Et c'est en 2009 que la Mojette est devenue Multi-résolution [Kingston 2009c].

5.2.0.1 Compression et redondance

Alors que la Transformation Mojette a été largement exploitée pour sa capacité à ajouter une redondance modulable au support initial, en 2006, nous avons proposé d'exploiter les similarités que peuvent présenter des projections d'angles très

proches [Atrousseau 2006a]. Ces similarités inter-projections nous ont donc permis de réduire l'entropie via un codage de la différence entre les projections. De plus, nous avons aussi pu constater qu'une certaine périodicité pouvait être présente au sein même d'une projection. Les similarités inter-projections peuvent être observées sur la Figure 5.5, où sont représentées 3 projections d'angles assez proches sur l'image "watch" : (1,128), (1,129) et (1,130). La Figure 5.6 quant à elle représente une portion de ces projections (zoom sur les bins dans l'intervalle [11000, 11400]). La périodicité intra-projection y est aussi représentée et nous pouvons ainsi observer la diminution sensible de l'amplitude sur la projection de différence.

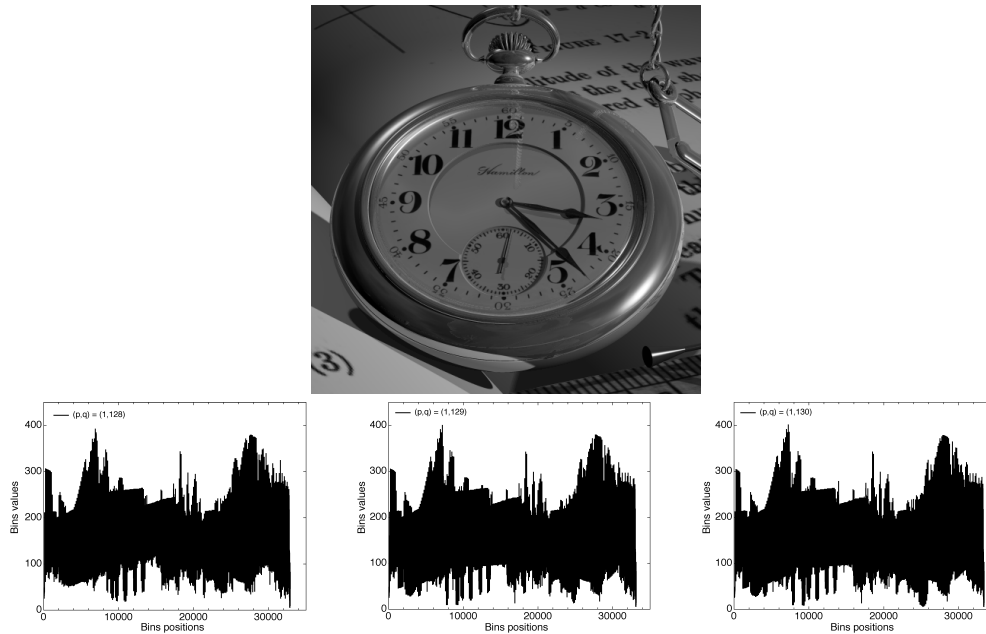


FIGURE 5.5 – Projections Mojette sur l'image "watch" selon des orientations (1,128), (1,129) et (1,130).

La Figure 5.7 nous montre les résultats de l'entropie sur 11 images de test. Ces valeurs d'entropie sont à comparer avec des méthodes de compression sans perte. Effectivement, l'instabilité de la transformation Mojette lors de la rétro-projection ne nous autorise pas à modifier les projections sans dégrader fortement l'image qui sera reconstruite. Cette méthode de compression est donc sans perte. Les taux de compression à ce stade sont encore perfectibles, mais il faut tout de même noter que jusqu'ici la Transformation Mojette n'avait pas encore été exploitée dans une optique de compression des données. Suite à cette première étape très encourageante, afin de poursuivre ces travaux et de tenter d'améliorer les performances de notre méthode, j'ai pu recruter Andrew Kingston, diplômé de l'Université de Melbourne en Australie pour effectuer un stage post-doctoral d'une année via un financement de la Région Pays de la Loire. Andrew a donc rejoint l'équipe IVC en 2006-2007. Au cours de cette année de post-doc, il a non seulement pu améliorer les taux

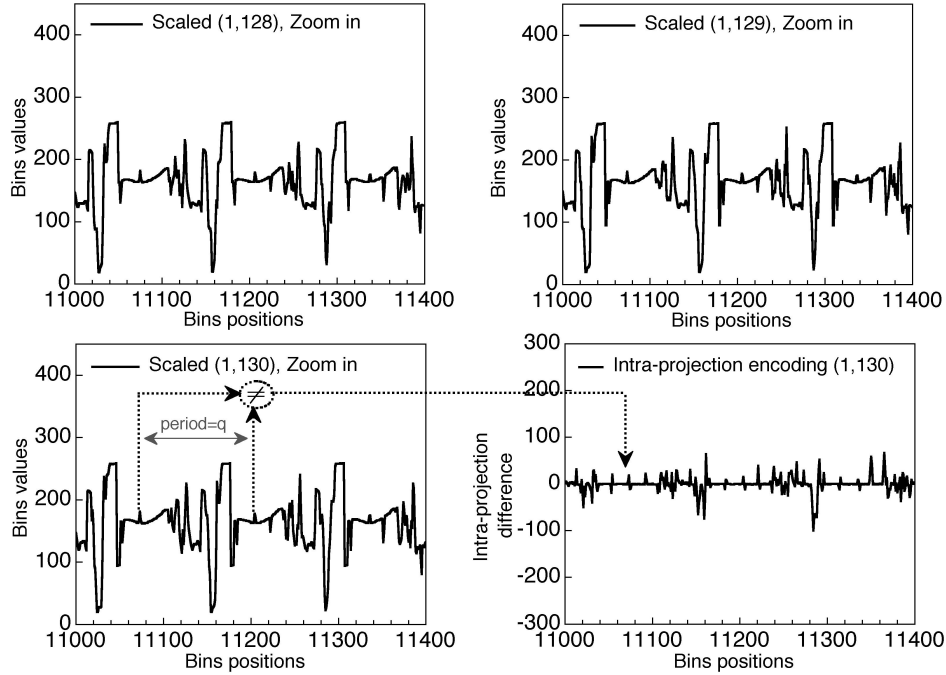
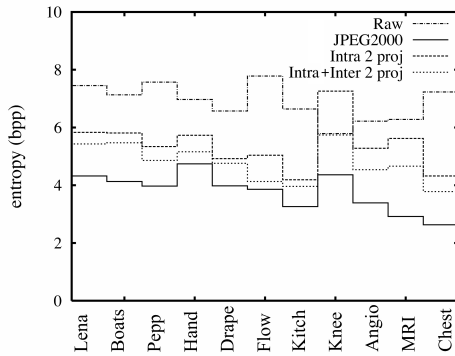
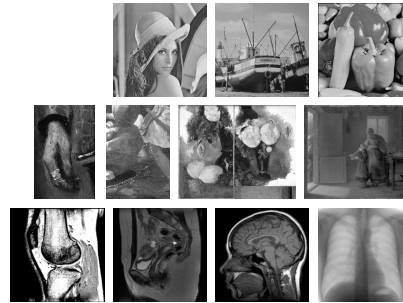


FIGURE 5.6 – Zoom sur les projections (1,128), (1,129), (1,130), différence intra-projection sur (1,130). Range des bins : [11000, 11400].



(a)



(b)

FIGURE 5.7 – (a) Entropie pour la compression Intra- et Inter-projections comparée avec JPEG2000 pour les 11 images de test montrées en (b).

de compression de la compression Mojette sans perte, mais il a aussi proposé une méthode de transformation Mojette Multi-Résolution, que nous verrons par la suite dans la section 5.2.0.2, ainsi qu'une méthode de chiffrement sélectif (section 5.3.2). Nous verrons par la suite ces différents travaux, pour l'instant attachons nous à rentrer dans le détail de la méthode de compression sans perte.

Andrew s'est dit qu'en appliquant des algorithmes de compression 2D sur les

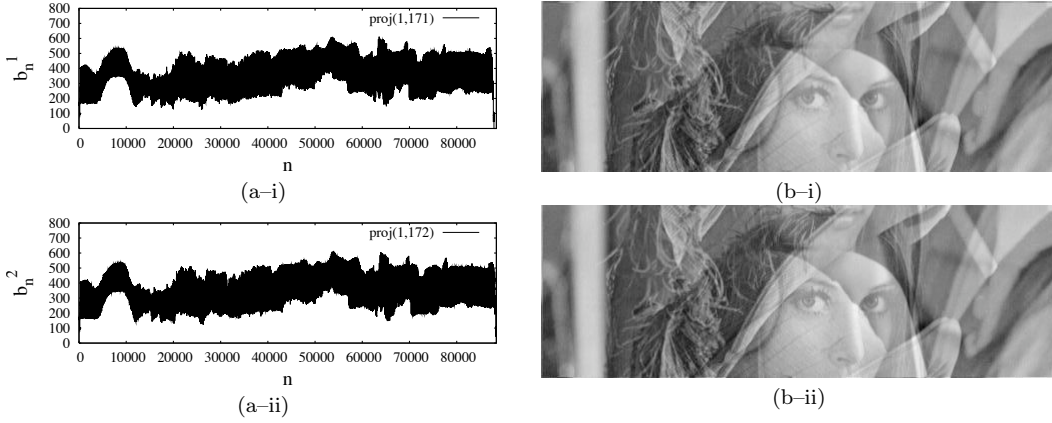


FIGURE 5.8 – (a) Projections 1D de l’image “Lena” et (b) ces mêmes projections affichées sous forme 2D.

projections (au lieu d’un simple DPCM 1D), nous pouvons augmenter les performances de codage. Il nous faut alors des projections 2D. Il est effectivement possible de “mapper” les projections sur un support 2D. L’équation 5.8 nous montre comment ce changement de disposition peut être effectué en projetant un pixel de valeur $I(k, l)$ sur le bin de la projection $\text{proj}_{p_i, q_i}(n_k, n_l)$:

$$\begin{aligned} \text{if } |p_i| \geq q_i & \begin{cases} n_k = \left\lfloor \frac{k}{|p_i|} \right\rfloor \\ n_l = l - \frac{(k - n_k)}{(k - n_k)} q_i \end{cases} , \\ \text{otherwise} & \begin{cases} n_l = \left\lfloor \frac{l}{|q_i|} \right\rfloor \\ n_k = k - \frac{(l - n_l)}{q_i} p_i \end{cases} , \end{aligned} \quad (5.8)$$

Le bin correspondant dans la projection 1D $\text{proj}_{p_i, q_i}(b)$ peut être retrouvé comme suit :

$$b = n_l p_i - n_k q_i \quad (5.9)$$

Une fois les projections représentées sous leur forme 2D, en ce qui concerne la compression intra-projection, il est alors possible d’appliquer des méthodes de compression sans pertes, qui vont pouvoir bénéficier de la redondance spatiale en deux dimensions. Pour ce faire, nous nous sommes tournés vers les méthodes suivantes de la littérature [Weinberger 1996, Wu 1997, Meyer 2001].

Il est important de noter que malgré la compression qu’il est possible d’appliquer sur les diverses projections, il est encore possible d’ajouter de la redondance aux données. Effectivement, il reste possible de rajouter des projections bien que le critère de Katz soit respecté. La Figure 5.10 représente l’évolution de l’entropie pour trois scénarios distinct :

1. Pas de projection de redondance, toutes les projections sont nécessaires à la reconstruction.

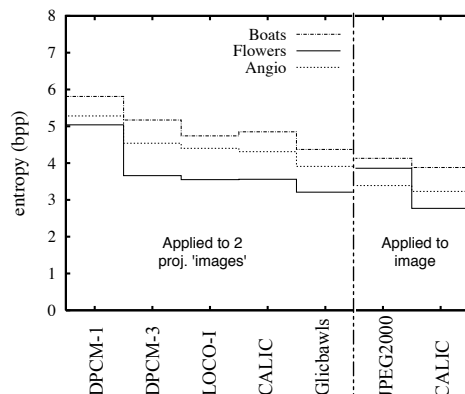


FIGURE 5.9 – Résultats en termes d’entropie pour la compression intra-projections entre deux projections. Nous utilisons les méthodes de compression : DPCM-1, DPCM-3, LOCO, CALIC, et Glicbawls. Nous comparons ces résultats à un codage JPEG2000 et CALIC appliqués directement sur les images

2. Une projection redondante est utilisée, N projections parmi les $N + 1$ seront suffisantes pour reconstruire l’image.
3. Deux projections redondantes sont présentes, N projections parmi les $N + 2$ seront suffisantes pour reconstruire l’image.

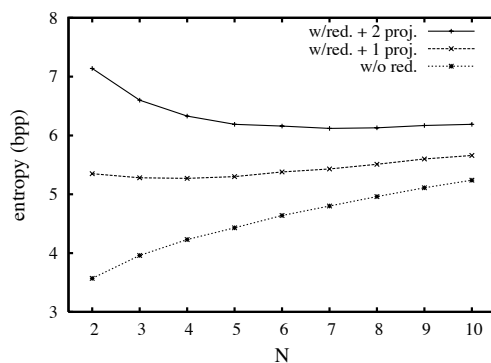


FIGURE 5.10 – Entropie tracée en fonction du nombre de projections, N , nécessaires pour la reconstruction. Les trois courbes représentent trois scénarios différents : 1) “sans projection redondante”, 2) “une projection supplémentaire” et finalement 3) “deux projections supplémentaires”.

En ce qui concerne la compression inter-projections, nous avons pu nous appuyer sur les travaux [Wu 2000] pour optimiser les performances de la méthode de base [Autrusseau 2006a]. Les travaux dans [Wu 2000] avaient pour objectif d’étendre la méthode de compression CALIC aux images multi-spectrales.

La Figure 5.11 présente les résultats de la méthode proposée. Nous comparons les entropies dans le cas des codages intra- et inter-projections avec l'entropie sur l'image d'origine, les performances de l'étude préliminaire ainsi que le taux de compression appliqué par JPEG2000. Nous pouvons constater que les taux de compression offerts par la transformation Mojette sont assez proches des taux atteints par JPEG2000 (parfois même meilleurs). De plus, nous devons garder à l'esprit que la Mojette offre non seulement des taux de compression compétitifs, mais qu'en sus, il est possible de moduler la redondance.

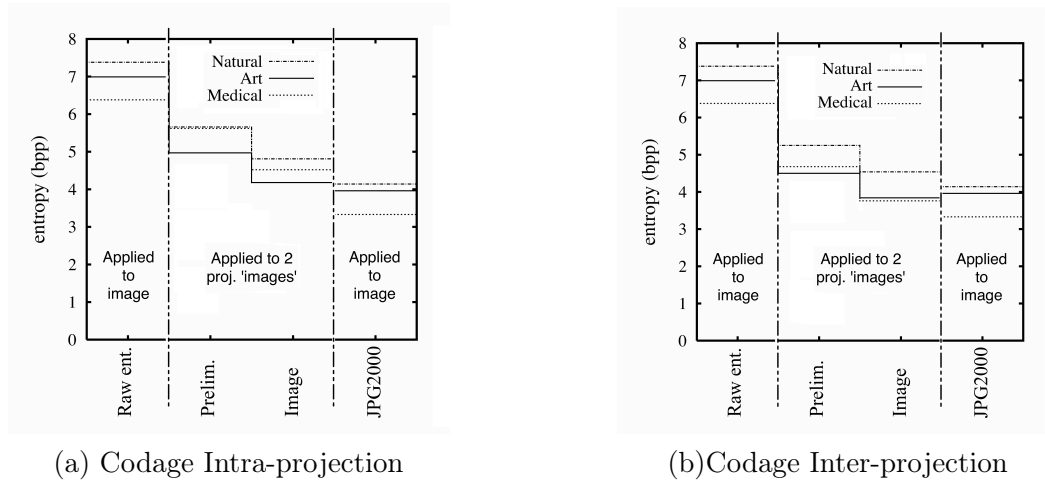


FIGURE 5.11 – Taux de compression moyen (sur 11 images de test) pour le scénario de codage intra-projections (a) et inter-projections (b).

5.2.0.2 La Mojette Multi-echelle

De nouveau avec Andrew Kingston, au cours de son post-doc en 2006-2007, outre ces travaux sur la compression sans perte basée sur la transformation Mojette (que nous venons de voir dans la section précédente), nous nous sommes aussi intéressés à la possibilité de fournir une représentation multi-résolution de la Mojette. Ces travaux ont fait l'objet d'un chapitre d'ouvrage [Kingston 2009c].

Au cours de son post-doctorat, Andrew a montré que si un sous-échantillonnage des projections était possible, nous pouvions alors associer la transformation Mojette avec la transformée en ondelettes. Par exemple, la projection $(15, 2)$ et $(21, 2)$ peut être sous-échantillonnée par les vecteurs $(1, 2)$, $(3, 1)$ ou encore $(3, 2)$. La Figure 5.12 nous montre comment une projection Mojette peut subir ce sous-échantillonnage.

Si nous notons la transformation Mojette $\mathcal{M}$ et la transformée en ondelettes $\mathcal{D}$, nous avons montré que $\mathcal{M}(\mathcal{W}(f))$ était équivalent à $\mathcal{M}(\mathcal{W})(\mathcal{M}(f))$. En d'autres termes, les projections Mojette de la DWT appliqué à l'image sont équivalentes à la projection Mojette de la DWT appliqué à la projection Mojette.

La Figure 5.13 nous montre la scalabilité des projections Mojette (on applique la DWT sur des projections Mojette représentées en 2 dimensions). Sur la Figure 5.14

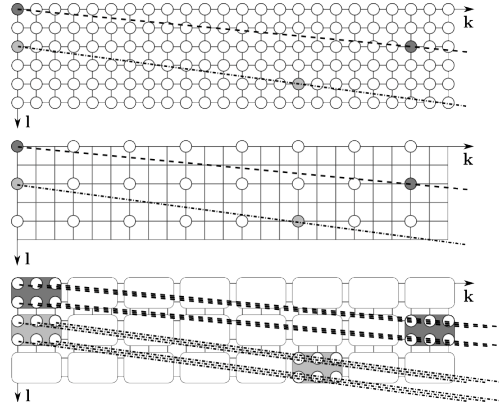


FIGURE 5.12 – Les lignes de directions des vecteurs $(15, 2)$ (points) et $(21, 2)$ (tirets) peuvent subir un sous-échantillonnage $(3, 2)$.

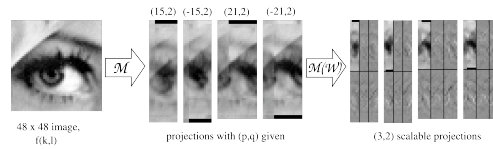


FIGURE 5.13 – Représentation de $\mathcal{M}(\mathcal{W})$, on observe la scalabilité des projections Mojette.

nous pouvons observer la propriété de redimensionnement de l'image et des projections Mojette.

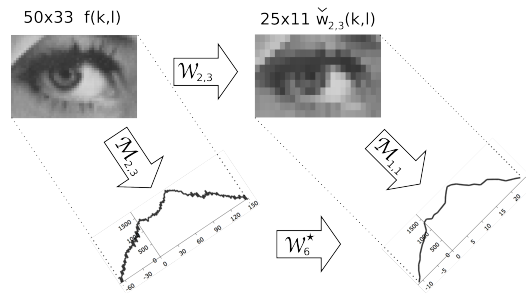


FIGURE 5.14 – Propriété de redimensionnement sur une image au format 48×48 pixels, $\mathcal{W}$ représente l'opérateur de scaling.

Enfin, la Figure 5.15 présente deux chemins possibles de reconstruction pour les projections Mojette multi-échelle. La première possibilité est donc tout simplement d'appliquer une reconstruction Ondelette suivie d'une reconstruction Mojette soit $\mathcal{M}(\mathcal{W})^{-1}$. La seconde possibilité, plus utile en pratique, est de décomposer

les hautes et basses fréquences des Ondelettes, une transformation Mojette inverse, appliquée séparément sur les hautes et basses fréquences, soit : $\mathcal{M}^{-1}(\mathcal{M}(f)_L)$ et $\mathcal{M}^{-1}(\mathcal{M}(f)_H)$.

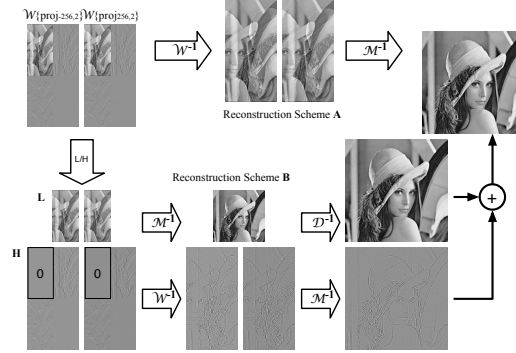


FIGURE 5.15 – Deux possibilités différentes de reconstruire une image à partir de ses projections scalables.

Il est important de noter qu'un sur-échantillonnage sera nécessaire sur l'image de basses fréquences ainsi reconstruite, et finalement, une simple addition des basses et hautes fréquences ainsi reconstruites nous conduira à l'image originale.

C'est donc ainsi qu'en 2008 la Mojette est devenue multi-résolution. Cette nouvelle propriété de la transformée peut ainsi ouvrir de nombreuses perspectives. Par exemple, il deviendrait à présent plus aisé de proposer une méthode de tatouage basée sur la Mojette. Effectivement, maintenant qu'une décomposition fréquentielle est possible, l'insertion d'un bruit (watermark) peut se faire tout en préservant la qualité de l'image reconstruite.

5.3 Sécurité basée Mojette

L'instabilité de la reconstruction Mojette dont nous avons déjà discuté précédemment dans la section 5.1.2 fait tout naturellement de cette transformation une bonne candidate pour proposer des algorithmes de chiffrement. Une simple modification des bins des projections peut aisément entraîner une reconstruction erronée de l'image, la perte de qualité est suffisamment significative pour permettre d'en déduire une méthode de chiffrement. Il est toutefois plus difficile de proposer une méthode de tatouage basée sur la Mojette. L'éclatement du support 2D de l'image vers plusieurs projections 1D nous complique la tâche, car encore une fois, si les modifications ont lieu sur les projections 1D, la rétro-projection risque fort de dégrader sensiblement la qualité de l'image sans que nous ne puissions maîtriser ces distortions. Nous voyons dans cette section ces deux aspects de la sécurité, abordés au cours de ma thèse de doctorat et ensuite les travaux poursuivis avec l'aide d'Andrew Kingston.

5.3.1 Applications au tatouage

Comme je l'expliquais ci-dessus, j'ai pu proposer quelques méthodes de tatouage au cours de ma thèse [Atrousseau 2002a, Atrousseau 2002b, Atrousseau 2003c]. Etant donné que la modification des projections elles mêmes est un peu périlleuse, la méthode la plus efficace est d'insérer des fantômes Mojette. Un fantôme Mojette est une combinaison (une convolution) d'éléments structurants à deux pixels qui va permettre de maintenir inchangées les projections qui le constituent.

La Figure 5.16 nous montre comment un fantôme est constitué et comment il peut ainsi générer des projections nulles.

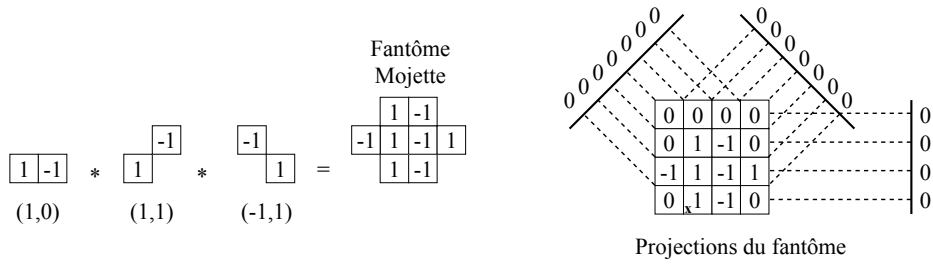


FIGURE 5.16 – Convolution de 3 éléments structurants à 2 pixels $((1,0)$, $(1,1)$ et $(-1,1)$) pour composer un fantôme et projections Mojette du fantôme le long de ces orientations.

Un tel fantôme peut alors être totalement invisible (d'où son nom) sur les orientations qui ont été utilisées lors de sa construction. La clef de détection peut alors résider dans les projections sur lesquelles le fantôme est détectable. Dans mes travaux, j'avais utilisé des fantômes Mojette comme watermark dans le domaine de Fourier de l'image. Les fantômes étaient modulés sur une porteuse spatiale dans une sous-bande perceptuelle du modèle de vision (cf. modèle de perception visuelle dans le Chapitre 3). Ceci nous permettant donc de moduler l'amplitude du tatouage tout en garantissant l'invisibilité de la marque lors du retour de l'image dans le domaine spatial. La Figure 5.17 nous présente les différentes images issues de cette méthode d'insertion. La détection se faisait alors directement sur les projections du spectre tatoué. Effectivement, si le spectre modifié était projeté sur un angle qui n'a pas été utilisé pour former les fantômes, alors, une répétition devrait survenir sur les projections, correspondant à la somme des éléments structurants à deux pixels.

La Figure 5.18 nous présente quelques résultats, et nous pouvons notamment observer que le fantôme apparaît bien sur les projections Mojette, même après attaques (ici un filtrage Gaussien).

Ces travaux ont donc fait partie de ma thèse, je n'ai pas effectué d'encadrement sur ce sujet, cependant, j'ai encadré plusieurs personnes sur la thématique du chiffrement des images. Simone Colosimo, qui a fait son stage de Master (échange Erasmus en collaboration avec Patrizio Campisi, Professeur à l'Université de Rome)

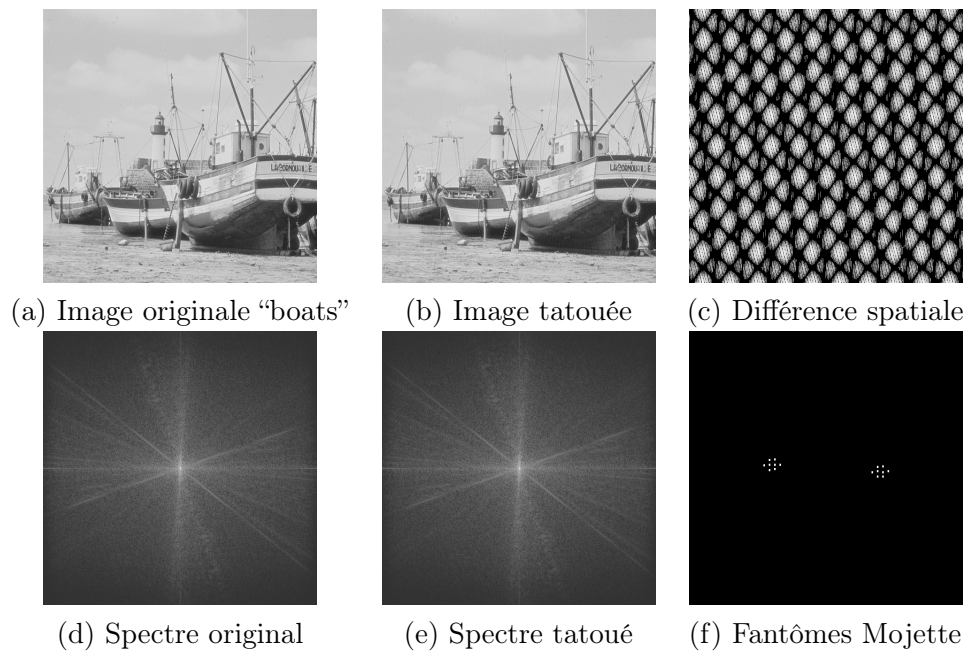


FIGURE 5.17 – Image originale et sa version tatouée, ainsi que le watermark dans sa version spatiale (images du haut) et dans le domaine de Fourier (images du bas).

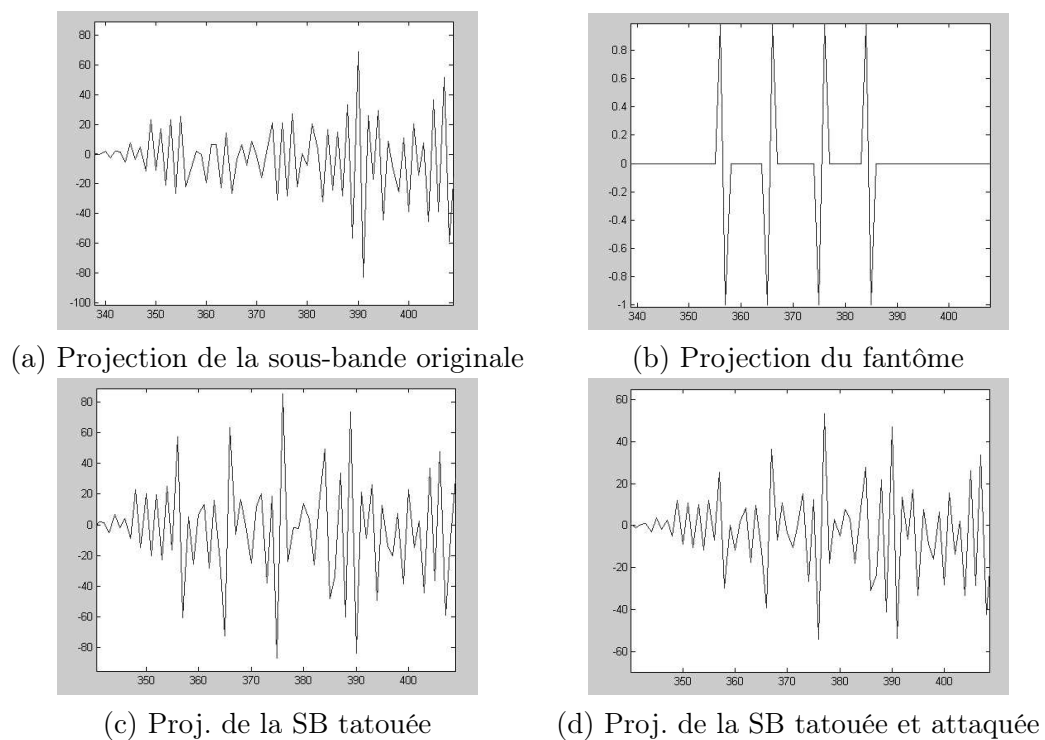


FIGURE 5.18 – Détection du fantôme sur les projections Mojette.

et Andrew Kingston ont travaillé sur le chiffrement Mojette dans le cadre du projet *TSAR* [Kingston 2007].

5.3.2 Applications au chiffrement

Une petite partie de ma thèse a porté sur le chiffrement Mojette. Nous avons montré que d'infimes modifications des projections pouvait entrainer un brouillage sévère de l'image reconstruite par rétro-projection. Au cours de ma thèse, j'avais proposé une méthode simple de chiffrement qui consistait à rétro-projeter les bins des projections en suivant un chemin pré-défini. Effectivement, à chaque étape de la rétro-projection, nous disposons d'un ensemble de bins à correspondance unitaire (1 bin correspond à 1 pixel) il est alors possible de choisir dans cet ensemble celui qui sera rétro-projeté, nous pouvons alors modifier tous les autres bins. La Figure 5.19 illustre ce principe.

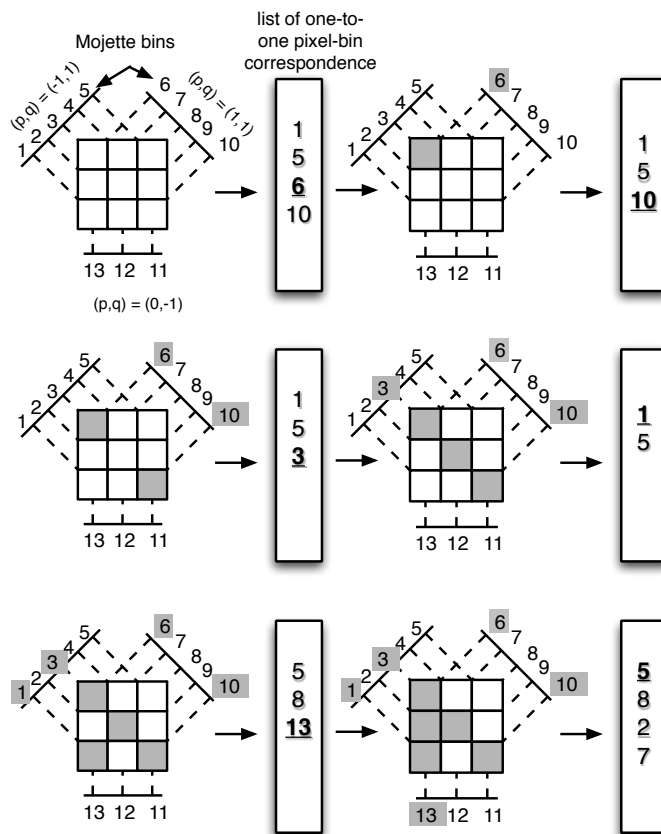


FIGURE 5.19 – Rétro-projection en suivant un chemin spécifique de l'ensemble des bins à correspondance unitaire.

La complexité (nombre de combinaisons possibles) peut sembler restreinte, cependant, comme nous allons le voir dans le Tableau 5.1, une faible redondance peut

#bins	19	20	22	23	24	25	30	34
C_n^k	969	4845	74613	245157	$7.3.10^{15}$	$2.04.10^6$	$1.45.10^9$	$2.20.10^9$
Directions	(1,5)	(1,1) (1,3)	(1,6)	(1,2) (1,3)	(1,0) (1,1) (1,3)	(1,7)	(1,1) (1,2) (1,3)	(1,0) (1,1) (1,2) (1,3)
Redondance	0.187	0.250	0.375	0.437	0.500	0.562	0.875	1.125

TABLE 5.1 – C_n^k et redondance pour différents sets de projections

être suffisante pour obtenir une complexité très importante. Les chances d'un pirate de rétro-projeter les bins non modifiés sont majorées par l'équation C_n^k que nous rappelons ici (Eq. 5.10). Le Tableau 5.1 nous montre le comportement des C_n^k pour une image de taille 4×4 en fonction de différentes directions de projections, soit différentes redondances.

$$C_n^k = \frac{n!}{k! \times (n-k)!} \quad (5.10)$$

Tandis que ces travaux avaient fait l'objet d'une partie de ma thèse, par la suite, dans le cadre du projet *TSAR*, j'ai dû me pencher à nouveau sur ces problématiques de chiffrement. Cette fois-ci, étant donné que la compression Mojette venait de voir le jour, il était alors intéressant de combiner la compression sans pertes avec le chiffrement sélectif.

Comme nous l'avons vu précédemment (section 5.2) la compression Mojette peut se faire de deux façons, il est possible de coder l'information intra-projection (tirer profit de la périodicité interne aux projections dues aux pixels voisins se projetant sur les bins) ou alors, il est possible de coder l'information inter-projections, i.e. deux projections d'angles proches sont similaires, et donc la différence entre les deux sera faible et d'entropie réduite. Il est bien évidemment possible de combiner le codage intra- et inter-projections. Dans un tel scénario, nous aurons à notre disposition une projection de base et une projection de différences, comme nous pouvons le voir sur la Figure 5.20. La projection de base est donc indispensable pour la bonne reconstruction des données. Nous avons donc proposé un algorithme de chiffrement sélectif tirant profit de cette forte dépendance entre les projections.

Dans le cadre du projet *TSAR*, il était souhaitable de fournir une méthode de chiffrement sélectif rapide opérant sur des images de très grandes tailles (images du laboratoire du musée du Louvre). Nous avons donc proposé un algorithme de chiffrement opérant une cascade de projections intra- et inter-projections, à chaque boucle de la cascade, les projections se trouvent donc être de dimensions réduites, après plusieurs boucles (dont le nombre est à définir, nous verrons des suggestions dans les résultats) l'ultime projection de base peut être chiffrée. Dans nos travaux, cette projection de base était chiffrée avec des algorithmes standards de la littérature de la cryptographie, tels que AES ou 3DES, cependant, si la projection de base est

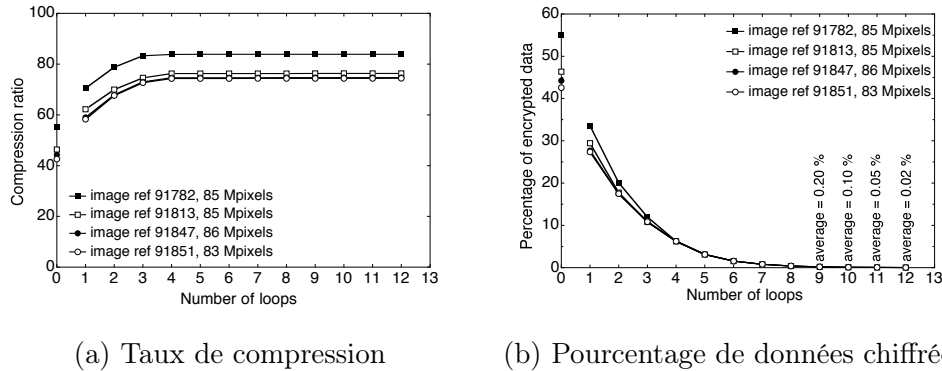


FIGURE 5.22 – Résultats fournis par la méthode de chiffrement par cascade de projection. Lorsque le nombre de boucles de la cascade augmente, il est nécessaire de chiffrer moins de données, mais au prix d'un taux de compression accru.

Ces travaux sur le chiffrement sélectif basé sur la Mojette (2006-2007) étaient antérieurs à l'étude portant sur la Mojette Multi-résolution (2008-2009), nous n'avons malheureusement pas eu l'opportunité de chercher à combiner les deux méthodes. Le chiffrement sélectif peut se voir sous deux angles différents, d'une part, il peut être souhaitable de ne chiffrer qu'une portion des données tout en garantissant un brouillage (visuel) complet des données (ce que nous venons de présenter dans ce chapitre), ou d'autre part, il est parfois souhaitable de proposer un chiffrement progressif des données, une qualité visuelle modulable (nous avons vu cette problématique dans le Chapitre 3 (plus précisément dans la section 3.2.2)). Il aurait effectivement été très intéressant de pouvoir proposer une méthode de chiffrement progressif (en terme de qualité visuelle).

Dans le prochain Chapitre, nous allons aborder une nouvelle thématique, qui occupe la majeure partie de mon temps depuis ces trois dernières années : l'analyse des images médicales et biologiques.

Imagerie Médicale et Biologique

Contents

6.1	Détection des dents	91
6.1.1	Approche “Computer Vision”	92
6.1.2	Approche “Deep Learning”	93
6.2	Détection de l’ostéoporose	95
6.2.1	Caractérisation osseuse	96
6.2.2	Reconnaissance du foramen mentonnier	97
6.3	Caractérisation de l’arbre vasculaire cérébral	98
6.3.1	Caractérisation des bifurcations	99
6.3.2	Estimation du risque de formation des anévrismes	102

Mon activité sur les thématiques de recherche de l’imagerie médicale et biologique s’est accrue lors de ma nouvelle affectation à l’UFR d’odontologie. J’ai intégré le laboratoire [LIOAD](#) en Novembre 2016 (devenu le [RMeS](#) en Janvier 2017). C’est un laboratoire de recherche spécialisé dans le squelette, différentes approches sont abordées. Le laboratoire est composé de deux équipes. Au sein de l’équipe Regenerative Medicine of Bone Tissues ([REGOS](#)), les thématiques de recherche se tournent vers l’étude du vieillissement osseux, la régénération osseuse et les biomatériaux tandis que l’équipe Skeletal Physiopathology and Joint Regenerative Medicine ([STEP](#)) se concentre plus particulièrement sur la physiologie du squelette, les cellules souches pour la régénération osseuse et les hydrogels.

Entre 2017 et 2019, j’ai eu l’opportunité de contribuer à différents projets, j’en détaille trois dans ce chapitre. Nous verrons tout d’abord un projet portant sur la détection des dents et la génération automatique du schéma dentaire, puis je détaillerai le projet *Tomosteo* dont l’objectif est de prédire le risque ostéoporotique en se basant sur des acquisitions de [CBCT](#). Finalement, je présenterai le projet *I-CAN*, un projet ambitieux de caractérisation de l’arbre vasculaire cérébral dans le but de prédire le risque d’occurrence des anévrismes intra-crâniens en se basant exclusivement sur l’imagerie.

6.1 Détection des dents

Le projet de Génération automatique du schéma dentaire avait pour objectif de reconnaître les dents et traitements sur des radiographies panoramiques. Il existe certaines études portant sur la segmentation automatique des dents, mais la plupart

opèrent sur des radiographies intra-orales (ou radios périapicales), ce type de clichés est beaucoup plus simple à traiter qu’une radio panoramique, dans lesquelles la résolution est moins bonne, la courbure de la mâchoire entraîne un chevauchement des dents, etc.

Grâce tout d’abord à l’encadrement d’un binôme d’étudiants de Polytech’Nantes, puis ensuite avec l’encadrement d’une stagiaire de M2 [SIBM](#) (Natalia De La Fournière) nous avons pu proposer différentes approches pour la problématique de détection des dents dans des radiographies panoramiques. Nous avons abordé ce problème via deux approches différentes que je détaille dans les prochaines sous-sections.

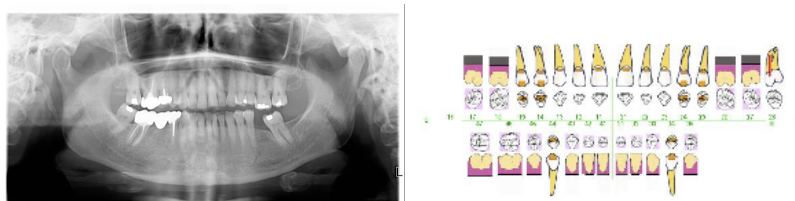


FIGURE 6.1 – Un exemple de radiographie panoramique et de schéma dentaire.

6.1.1 Approche “Computer Vision”

Au cours de son stage de M2, Natalia a pu exploiter de très nombreuses pistes pour proposer une méthode de détection des dents efficaces. Elle s’est tout d’abord penchée sur des méthodes classiques de l’analyse des images, elle a étudié les profils 1D de l’image, comme nous pouvons le voir sur la Figure 6.2.

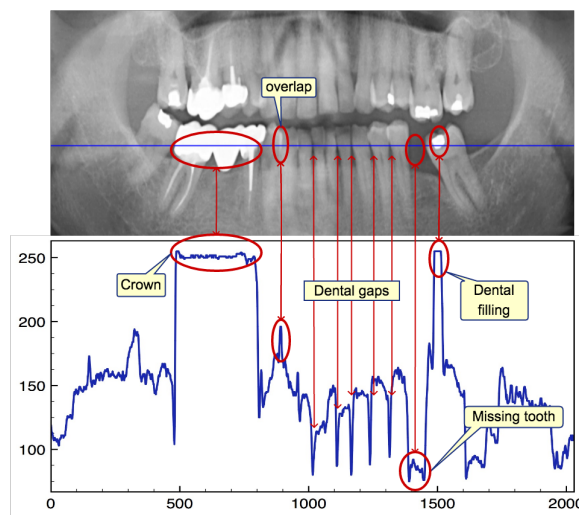


FIGURE 6.2 – Une radiographie panoramique avec une représentation d’un profil 1D de l’image (en bleu).

Natalia a testé une multitude d’outils de la littérature, tout d’abord via une

étude fréquentielle, le périodogramme du signal 1D a été calculé et analysé, puis une approche plus orientée vers l'analyse des images a été envisagée, nous avons testé le watershed, les snakes (contours actifs) et bien d'autres méthodes encore. Aucune méthode de la littérature ne s'est montrée efficace. Les travaux similaires de détection des dents portent surtout sur des radiographies périapicales ayant une bien meilleure qualité et résolution, très peu d'études ont été conduites sur des radiographies panoramiques. Finalement, une simple décomposition fréquentielle des profils 1D pouvait nous permettre d'estimer la fréquence de répétition du signal, et de fait, la largeur des dents.

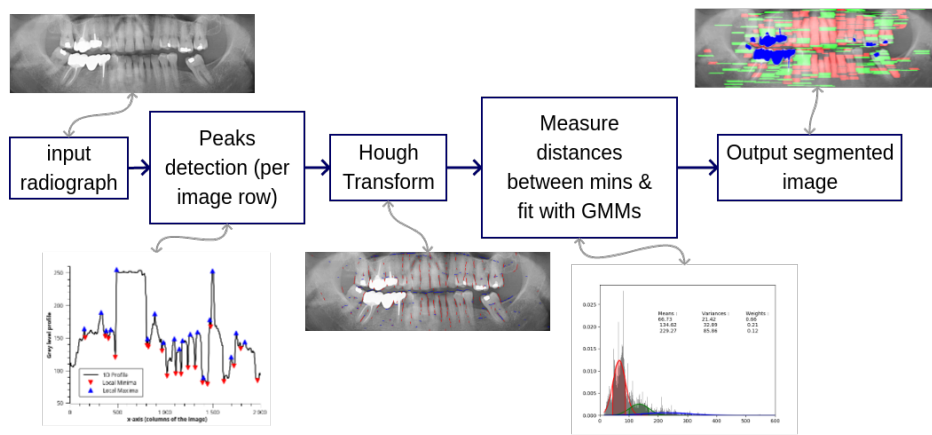


FIGURE 6.3 – Schéma bloc représentant des différentes étapes du processus de détection des dents.

Le processus complet de détection des dents est représenté sur la Figure 6.3. Une détection de maxima et minima locaux, combinée à une transformée de Hough nous permet de détecter des structures linéaires verticales pouvant correspondre aux interstices inter-dentaires. Nous utilisons ensuite des Gaussian Mixture Model (GMM) afin de procéder à un clustering des distances entre les segments détectés par la transformée de Hough. Ces GMM nous ont alors permis de classifier les dents, et de distinguer des structures plus larges, telles que des bridges dentaires par exemple. Une couleur différente est alors allouée à chacun des GMM, en fonction de l'espacement, comme nous pouvons l'observer sur la Figure 6.4, où les dents sont globalement correctement détectées. Cependant, les résultats sont perfectibles, nous nous sommes alors tournés vers les méthodes d'apprentissage.

6.1.2 Approche “Deep Learning”

Au cours de son stage de Master, Natalia de la Fournière s'est donc aussi intéressée aux méthodes d'apprentissage, et a souhaité tester une technique de segmentation des dents par réseau de neurones.

Natalia a mis en oeuvre une méthode de deep learning qui s'est avérée assez effi-

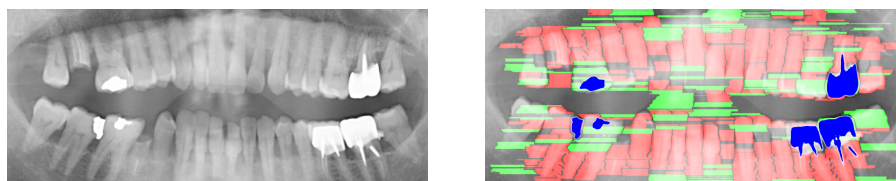


FIGURE 6.4 – La plupart des dents sont détectées par l’un des 2 GMM

cace. Pour ce faire, nous avons dû faire appel à des étudiants et internes en chirurgie dentaire (en projet de TER : Travaux d’Études et de Recherche) pour segmenter manuellement les dents sur des radiographies panoramiques [Michel 2018]. Les étudiants de l’UFR d’odontologie ont donc segmenté manuellement les dents de 26 radiographies panoramiques, ces segmentations nous ont alors servi d’images labellisées pour alimenter un réseau de neurones, les résultats ont montré des performances très intéressantes sur la reconnaissance des dents.

Cette expérience a été conduite avec un réseau de neurones U-Net, une phase d’apprentissage (*training*) composée de 26 images, avec les masques correspondants a été lancée, suivie par une phase de test composée de 10 images. La Figure 6.5 montre un exemple d’image de denture avec le masque binaire correspondant.



FIGURE 6.5 – Radiographie panoramique avec le détourage manuel des dents.

La Figure 6.6 présente quelques résultats obtenus par le réseau de neurones sur une denture acquise à la radiographie panoramique.

Bien que le réseau de neurones ait montré des résultats très encourageants pour la détection des dents, nous n’avons malheureusement pas eu le temps de nous pencher sur la détection des traitements effectués aux patients (traitements des racines, amalgames, couronnes, bridges, etc.). Cependant, à l’issue de son stage, Natalia a créé une entreprise ayant le même objectif que celui de son stage de M2. L’entreprise, créée en 2019 se nomme *Artefakt-AI*, nous avons d’ores et déjà eu l’occasion de répondre à des appels à projets afin de recruter du personnel (post-doctorant) sur les approches basées Deep-Learning.

Artefakt-AI est déjà très active dans le domaine de l’intelligence artificielle au service de l’analyse des radiographies dentaires.

Nous avons répondu conjointement (Natalia et moi même) à un appel à projets “Innovez” de l’I-Site NExT en 2019 pour conduire le projet “Odon” sur une durée d’une année, au cours de laquelle il est prévu de recruter un post-doctorant. Le projet “Odon” a été accepté en Février 2020, nous sommes en attente du financement.

Artefakt-AI est membre de deux pôles de compétitivité : Atlanpole Biotherapies

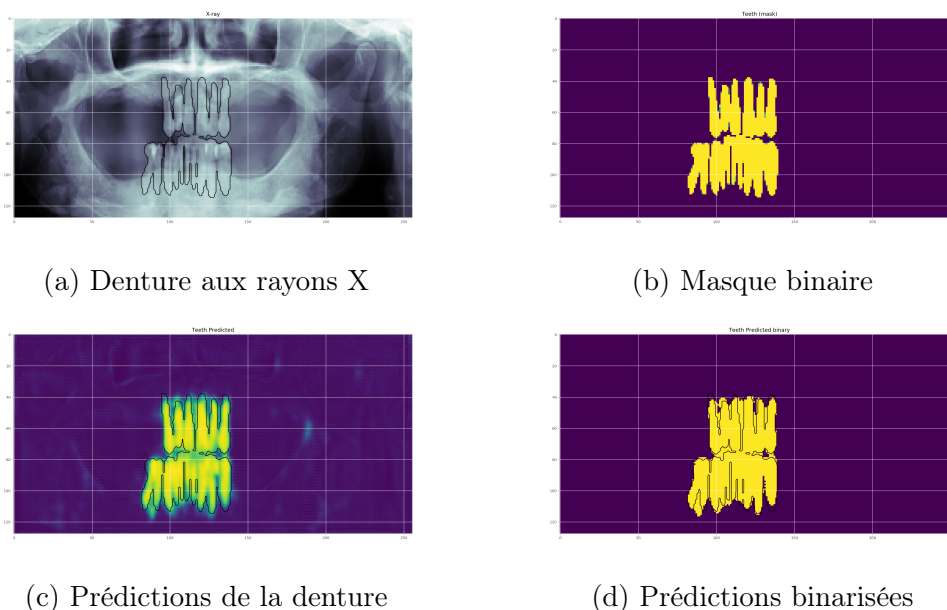


FIGURE 6.6 – Résultats du réseau de neurones U-Net sur la segmentation de la denture

et Images & Réseaux pour ses activités à l'interface entre le numérique et la santé en médecine personnalisée. L'entreprise *Artefakt-AI* et le projet "Odon" sont également accompagnés par l'incubateur Atlanpole depuis juillet 2019 avec le programme d'accompagnement sur 5 ans et ont rejoint l'incubateur national privé *1Kubator* depuis Mars 2019.

6.2 Détection de l'ostéoporose

Trois millions de femmes ménopausées en France sont atteintes d'ostéoporose. Cette maladie est responsable chaque année de 130 000 fractures. Concernant principalement des personnes âgées, ces accidents sont particulièrement graves et handicapants. De manière probabiliste, en Europe, environ 1/3 des femmes en post-ménopause et 1/8 des hommes après 50 ans auront une fracture d'origine ostéoporotique. L'ostéoporose est donc identifiée comme un problème de santé publique majeur.

Notre étude vise principalement à obtenir un meilleur dépistage de l'ostéoporose en substituant ou multipliant les méthodes d'acquisition, en particulier le CBCT en regard de sa résolutivité et son informativité, et en permettant de faire rentrer dans la chaîne de dépistage des professionnels de santé comme les chirurgiens-dentistes, les chirurgiens maxillo-faciaux et ORL.

Pendant le projet Tomosteo, deux pistes ont été privilégiées pour détecter l'ostéoporose via l'analyse des images. Nous avons d'une part cherché à caractériser l'os, afin de pouvoir estimer sa solidité, et d'autre part, proposé diverses mesures sur le

Foramen Mentonnier qui pourraient être un indicateur de risque d'après certaines études.

6.2.1 Caractérisation osseuse

L'estimation du risque ostéoporotique peut se faire via une étude de la densité osseuse avec des outils d'analyse d'images. Le CBCT dispose d'une résolution moindre que certains scanners, mais peut toutefois permettre d'obtenir une estimation relativement fiable de la densité. Classiquement, la densité osseuse peut être mesurée grâce à la dimension fractale.

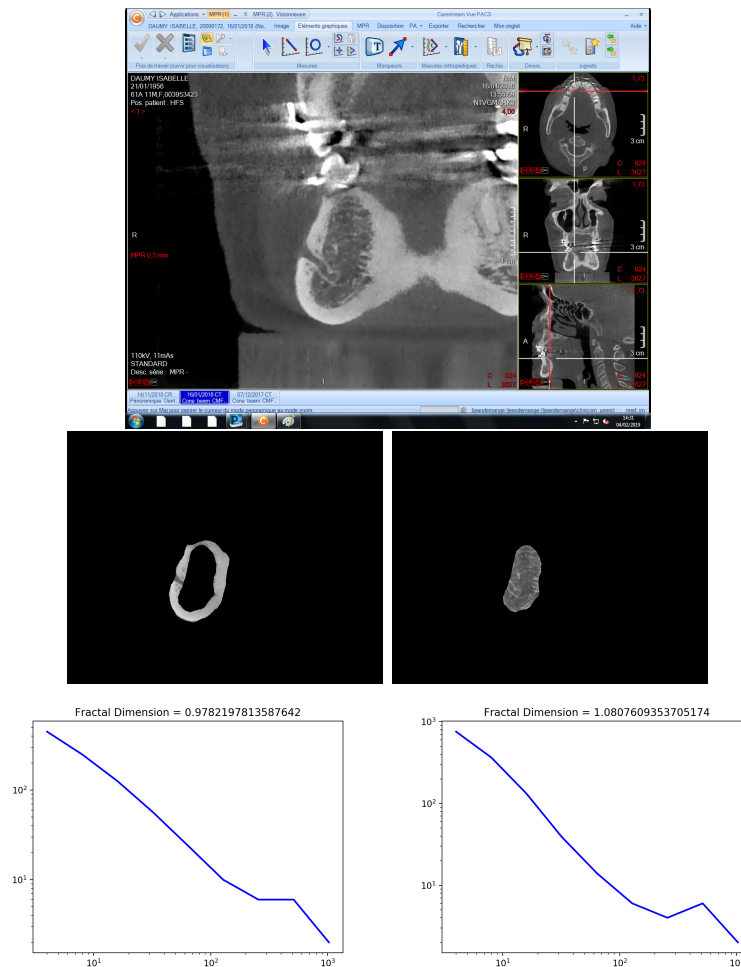


FIGURE 6.7 – Dimension fractale dans l'os cortical (à gauche) et dans l'os trabéculaire (à droite).

La Figure 6.7 montre une portion de la mandibule, avec une segmentation manuelle de l'os cortical (représenté à gauche) et de l'os trabéculaire (à droite), les courbes représentent la dimension fractale selon la méthode de Minkowski-Bouligand [Falconer 2003]. Lors de cette étude, qui vient de débuter, nous comptons

non seulement mener une comparaison des dimensions fractales entre patients sains et patients ostéoporotiques, mais bien évidemment, nous nous tournerons aussi vers d'autres méthodes d'analyse des textures. Notre objectif final étant d'exploiter plusieurs mesures de caractérisation des textures pour alimenter un réseau de neurones qui nous permettra de prédire le risque.

6.2.2 Reconnaissance du foramen mentonnier

Comme nous avons pu le voir brièvement dans l'introduction, le Foramen Mentonnier est un trou qui transperce la mandibule, il a pour fonction de faire passer le nerf mentonnier de part et d'autre de la mandibule. Certaines études [Koh 2011] ont montré qu'une caractérisation précise du Foramen Mentonnier pouvait aider à prédire le risque ostéoporotique chez les patients. La Figure 6.8 représente sur la coupe frontale (en bas à droite) le Foramen Mentonnier dans un cercle rouge.

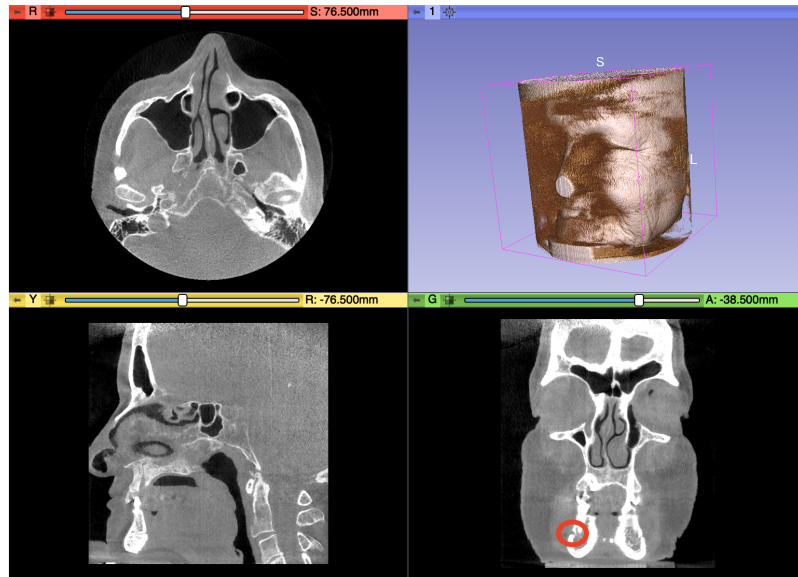


FIGURE 6.8 – Le Foramen Mentonnier est marqué par le cercle rouge dans la coupe frontale (image en bas à droite).

La Figure 6.9 présente quatre images de CBCT de quatre patientes différentes, nous pouvons faire plusieurs observations. Tout d'abord, le Foramen Mentonnier peut non seulement présenter un espacement variable, mais aussi une distance variable au bas de la mandibule. De plus, l'épaisseur de l'os cortical semble elle aussi varier.

D'un point de vue strictement analyse d'images, lorsqu'il y aura suspicion d'ostéoporose sur l'imagerie DXA, nous chercherons à déterminer sur les scans 3D du CBCT si le Foramen Mentonnier présente des caractéristiques particulières en comparaison avec des patients non atteints d'ostéoporose. De plus, au cours de cette étude, nous comptons aussi proposer une méthode de caractérisation des textures

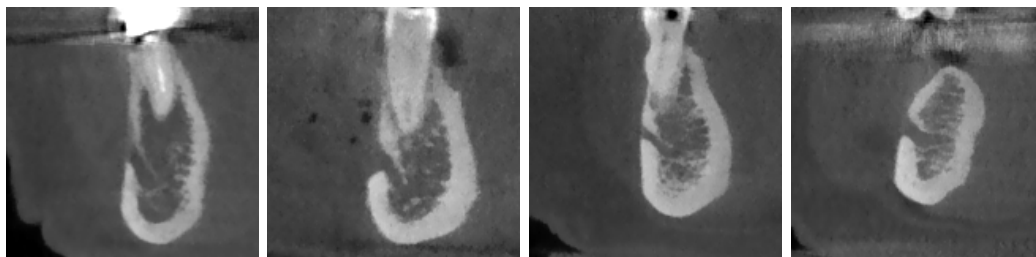


FIGURE 6.9 – Différents Foramens Mentonniers sur des acquisitions au CBCT.

nous permettant d’estimer la Densité Minérale Osseuse. L’os est principalement composé de deux parties distinctes, l’enveloppe extérieure est appelée “os cortical”, c’est la partie la plus dure, l’os cortical est un matériaux plein, tandis que la partie intérieure, l’os “trabéculaire” est plus friable, elle est constituée de trabécules, qui sont des structures alvéolaires, cette partie de l’os est plus creuse (c’est l’os spongieux). L’espacement des trabécules, leur épaisseur, leur nombre contribuent à la solidité de l’os. Au cours du projet *Tomosteo* nous chercherons à mesurer automatiquement l’épaisseur de l’os cortical et à évaluer la solidité osseuse grâce à une analyse des textures nous permettant de quantifier la proportion d’os spongieux. Dans le début de l’année 2020, un stagiaire de M2, Clément Morin a été recruté (via un financement du cluster “FAME” de l’I-Site NExT) pour proposer ces méthodes de caractérisation de l’os. Ce projet vient juste de débuter, mais nos premiers résultats sont plutôt encourageants.

6.3 Caractérisation de l’arbre vasculaire cérébral

Comme nous avons pu le voir dans l’introduction de ce manuscrit, lors de mon implication dans le projet ANR *I-CAN*, j’ai pu me pencher sur la caractérisation de l’arbre vasculaire cérébral en vue de prédire le risque de formation des anévrismes intra-crâniens. Notre objectif dans ce projet était de collecter plusieurs mesures représentant la géométrie des bifurcations. La Figure 6.10 donne une représentation graphique des différentes mesures que nous pouvons effectuer sur l’arbre vasculaire cérébral. Voici comment nous pouvons résumer la problématique : Si le flux sanguin arrive avec une vitesse importante sur une bifurcation un anévrisme risque de se former sur la paroi, surtout si cette dernière est fragilisée (risque génétique par exemple). Le flux sanguin peut prendre de la vitesse si l’artère mère est rectiligne (moins de ralentissement par frottement sur les parois) si la distance géodésique depuis la bifurcation précédente est importante, si l’angle que forment les artères filles est important, et finalement, s’il existe une différence significative de diamètres entre l’artère mère et les artères filles. La Figure 6.10 représente donc ces différentes mesures.

Dans cette section, je présente les travaux effectués avec Anass Nouri (post-doctorant au [RMeS](#) en 2017-2018). Ces travaux ont été valorisés à travers un ar-

ticle de conférence [Nouri 2019], un brevet [Nouri 2018] et une revue internationale [Nouri 2020].

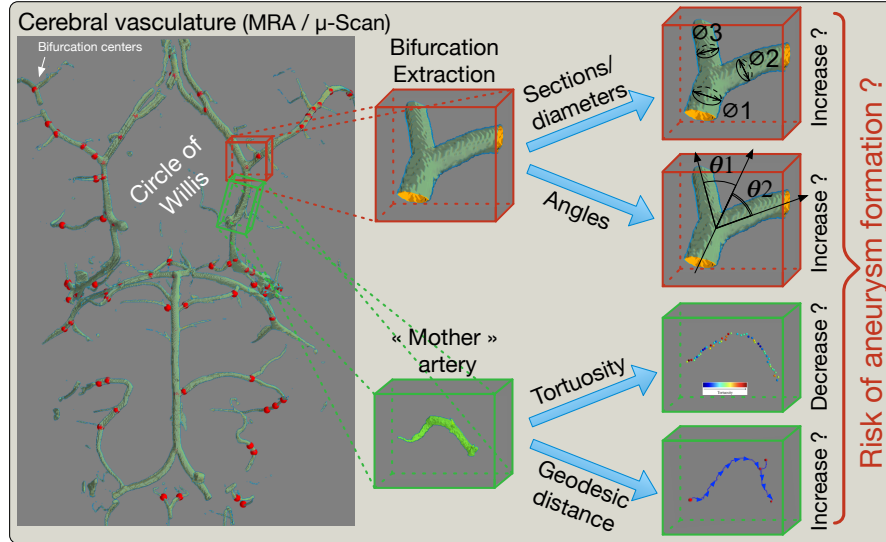


FIGURE 6.10 – Représentation graphique des mesures prises sur l'arbre vasculaire.

6.3.1 Caractérisation des bifurcations

Cette étude a porté sur deux études distinctes, l'une portait sur la caractérisation de l'arbre vasculaire de souris sur des acquisitions au micro-scanner, l'autre étude portait sur des **IRM TOF** sur des patients humains. La Figure 6.11 montre un arbre vasculaire sous trois représentations différentes. Une squelettisation 3D est appliquée sur l'arbre vasculaire segmenté, puis un graphe 3D est calculé sur ledit squelette. Cette représentation sous forme de graphe permet de déterminer les noeuds qui ne sont autre que les bifurcations (ou trifurcations) de l'arbre vasculaire.

Nous pouvons observer plus particulièrement cette représentation sous forme de graphe dans la Figure 6.12. Les sphères rouges représentent les noeuds du graphe.

Les bifurcations d'intérêt (celles qui sont le plus susceptibles de voir apparaître un anévrisme) se trouvent sur le polygone de willis, comme nous l'avons déjà vu dans l'introduction. La Figure 6.13 représente des bifurcations d'intérêt.

Voyons à présent les méthodes que nous avons mises en oeuvre afin de collecter les caractéristiques géométriques des bifurcations. Nous avons exploité la transformation Mojette (cf. Chapitre 5) pour déterminer les diamètres minimum et maximum des artères, comme le montre la Figure 6.14. Cette méthode très simple dans sa mise en oeuvre nous permettait néanmoins de collecter les diamètres avec une très bonne précision, et sans accroître significativement la complexité algorithmique.

Nous avons demandé à plusieurs experts (neuro-radiologues au CHU de Nantes) de mesurer manuellement les diamètres sur une base d'images, puis nous avons pu les comparer avec les mesures automatiques collectées par notre modèle. Le modèle

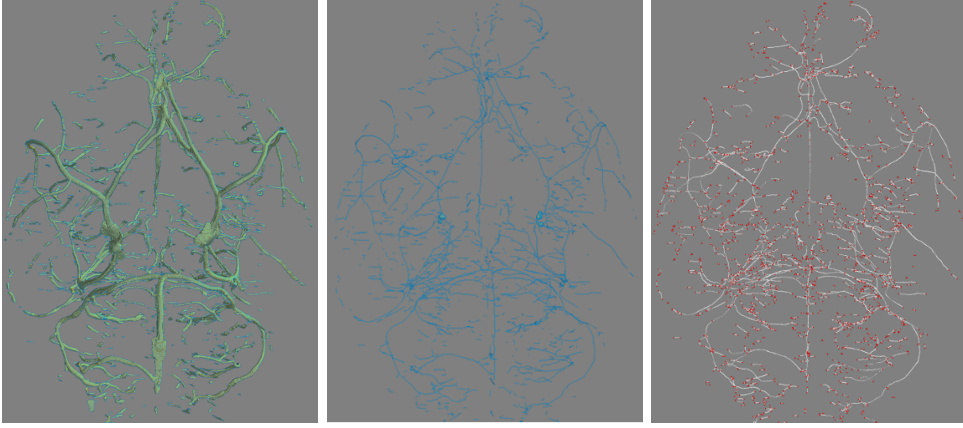


FIGURE 6.11 – (a) Arbre vasculaire de souris après acquisition au micro-scanner, (b) squelette de la vascularisation obtenue par morphologie mathématique, et (c) graphe de connexion entre les branches du squelette.

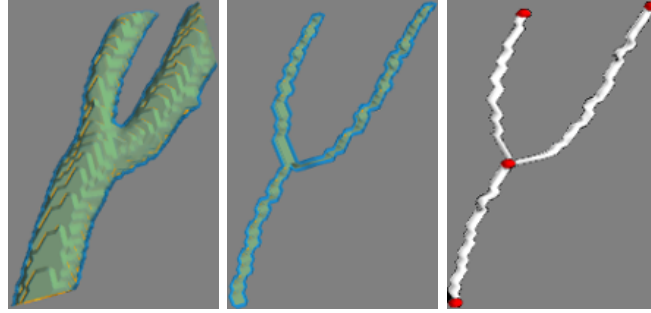


FIGURE 6.12 – (a) Bifurcation d'une TOF d'IRM sur patient humain, (b) squelette 3D et (c) graphe 3D.

s'avère être très proche des mesures des experts, comme en atteste la Figure 6.15.

Concernant la mesure des angles, l'équation 6.1 et la Figure 6.16 nous montre comment les angles ont pu être calculés en 3 dimensions.

$$\widehat{A}_1 = \text{atan2}(\| -\vec{CN} \bullet \vec{CL} \|_2, -\vec{CN} \cdot \vec{CL}) \quad (6.1)$$

Quant à la tortuosité, nous mesurons une courbure entre les voxels voisins le long du graphe grâce à la distance entre les vecteurs normaux au squelette. Une sommation de ces mesures de courbures à plusieurs résolutions a été calculée et fournit une mesure globale de tortuosité de l'artère considérée.

$$\mathcal{C}(v_i) = \frac{C_{left} + C_{right}}{2} \quad (6.2)$$

ou C_{left}^k est donné par :

$$C_{left} = \frac{(\vec{n}_{v_i} - \text{mean}(\vec{n}_{v_{left}})) \cdot (\vec{v}_i - \text{mean}(\vec{v}_{left}))}{\|\vec{v}_i - \text{mean}(\vec{v}_{left})\|_2^2} \quad (6.3)$$

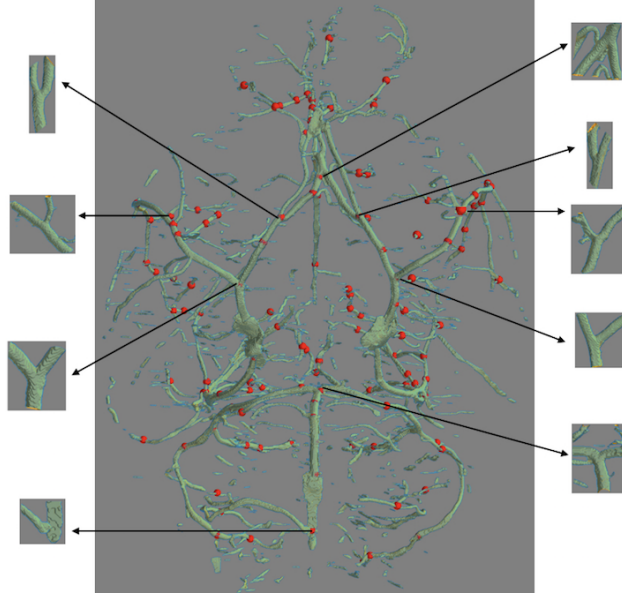
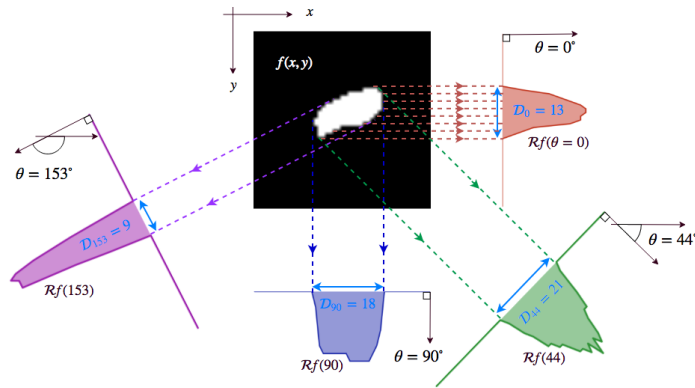


FIGURE 6.13 – Polygone de Willis sur un micro-scan de souris.

FIGURE 6.14 – Projections Mojet selon différents angles d'une cross-section 2D, obtention des diamètres min et max (resp. 9 et 21 bins (aux angles $\theta = 153^\circ$ et $\theta = 44^\circ$)).

$mean(\vec{n}_{left}) = \frac{\sum_{i=0}^{|\mathcal{N}_{left}|} \vec{n}_i}{|\mathcal{N}_{left}|}$ et $mean(\vec{v}_{left}) = \frac{\sum_{i=0}^{|\mathcal{N}_{left}|} \vec{v}_i}{|\mathcal{N}_{left}|}$. $|\mathcal{N}_{left}|$ se réfèrent au voisinage. La courbure du côté droit est définie de la même façon (C_{right}^k).

L'agrégation des courbures se fait via une sommation de Minkowski pondérée :

$$\mathcal{T} = \left(\frac{\sum_{i=0}^{|v|} \mathcal{C}(v_i)^p}{|v|} \right)^{1/r} \quad (6.4)$$

ou les indices p et r ont été déterminés expérimentalement.

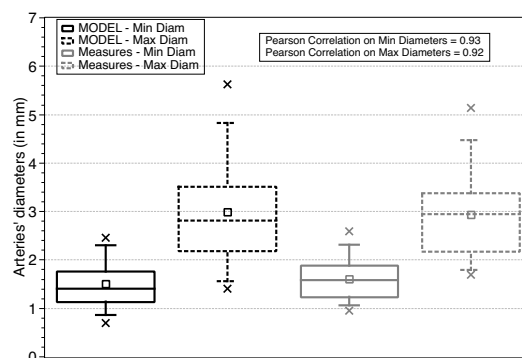


FIGURE 6.15 – Distributions des diamètres min et max d’après le modèle (en noir) et mesures manuelles par des experts (en gris).

La Figure 6.17 détaille ce processus. Encore une fois, à ce stade, il nous a fallu valider les mesures calculées par le modèle, nous avons donc mis en place un test subjectif de mesure des tortuosités, nous avons demandé à 20 observateurs de juger de la tortuosité de 27 artères sur une application web¹. Ceci nous a donc permis de valider les mesures obtenues par notre modèle.

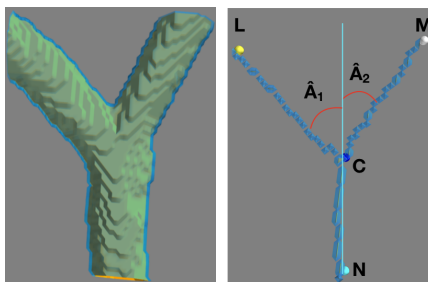


FIGURE 6.16 – (a) Bifurcation 3D, (b) Angles 3D formés par la bifurcation.

6.3.2 Estimation du risque de formation des anévrismes

Nous pouvons constater sur la Figure 6.18 qu’en moyenne, les bifurcation porteuses d’anévrismes (représentées en rouge) présentent des angles plus prononcés, une distance géodésique de l’artère mère plus importante, et une tortuosité plus faible que les bifurcations sans anévrismes.

Nos résultats expérimentaux n’ont pas pu montrer que la section des artères, ou la différence de section entre l’artère mère et les artères filles puissent avoir un impact quelconque sur le risque de formation des anévrismes. De plus, nous avons conduit une Analyse en Composante Principale (ACP), qui ne nous a pas non plus

1. http://websubjexp.polytech.univ-nantes.fr/Tortuosity/index_IsolatedArteries.php

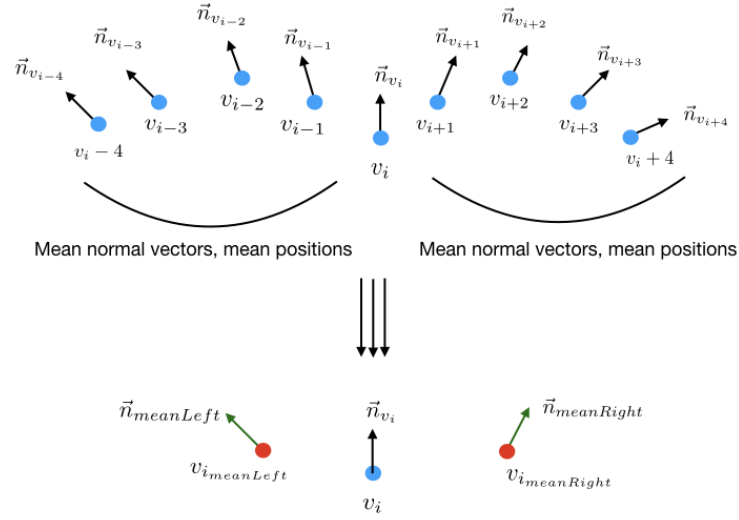


FIGURE 6.17 – Calcul de la tortuosité sur les artères.

permis de trouver de lien entre un regroupement quelconque des propriétés géométriques et le risque de survenue d'un anévrisme. À l'heure actuelle, nous sommes encore à la recherche de liens entre ces différentes propriétés géométriques et surtout nous tentons de comprendre si une combinaison entre les caractéristiques identifiées pourrait augmenter le risque.

Cette méthode de caractérisation des bifurcations de l'arbre vasculaire cérébral est fortement dépendante de la toute première étape du processus qui est la segmentation. Si la segmentation s'avérait être imparfaite, toutes les méthodes qui en découlent (squelettisation, graphe 3D, mesures automatiques de paramètres géométriques) risquent d'en pâtir. J'ai donc cherché via l'encadrement de Julia Guillou à optimiser l'étape de segmentation. Le défi était alors d'obtenir la meilleure segmentation possible du polygone de Willis tout en écartant dans la mesure du possible les veines et artères périphériques. Pour ce faire, nous avons exploité la CSF après l'avoir adaptée à la taille des artères d'intérêt [Guillou 2020]. Une fois de plus, INSERM-Transfert m'a aidé dans cette démarche et m'a permis de déposer un brevet sur cette problématique au tout début de l'année 2020 [Atrousseau 2020].

Une collaboration est toujours en cours avec Anass Nouri, qui est à présent Maître de Conférences à l'Université de Kénitra au Maroc. Nous avons pu proposer un projet au CNRST (Centre National pour la Recherche Scientifique et Technique) au Maroc, qui nous a permis de recruter deux doctorantes sur ce projet. Ibtissam Essadik et Sara Chater ont toutes deux commencé leur thèse de doctorat en Janvier 2020 à l'Université de Kénitra. Je co-encadre la thèse d'Ibtissam avec Anass et Raja Touahni et la thèse de Sara avec Anass et Youssef el Merabet. Ibtissam Essadik a pour objectif d'établir un atlas anatomique permettant de connecter automatiquement chacune des bifurcations détectées à des labels prédéterminés sur un polygone de Willis standard, tandis que Sara Chater travaille sur l'optimisation

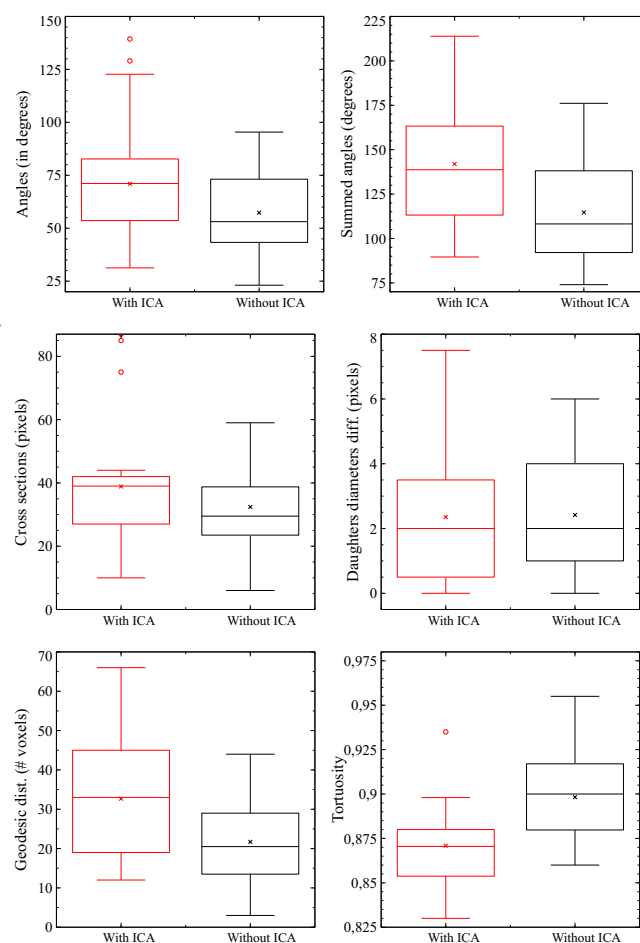


FIGURE 6.18 – (a) BoxPlot des Angles, (b) Somme des angles, (c) section des artères, (d) Différences de diamètres (mère/fille), (e) distances géodésiques, (f) tortuosité.

de la segmentation et la détection des anévrismes par l'utilisation de méthodes de Deep Learning. A compter du premier Septembre 2020, Nathan Lauzeral a rejoint notre projet en post-doctorat et travaillera lui aussi sur la segmentation basée Deep Learning de l'arbre vasculaire cérébral. Nous tenterons de générer des artères de souris dites "humanisées". L'apprentissage des réseaux de neurones se fera sur des artères de souris acquises au Micro-Scanner. Nous tenterons de modéliser au mieux le bruit propre à l'IRM (images TOF) que nous superposerons au Micro-scans. Une fois l'apprentissage effectué sur ces images de souris "augmentées", la phase de test se fera sur les TOFs humaines.

Ceci clos la présentation des mes différentes thématiques de recherche et des résultats obtenus. Nous allons à présent passer brièvement en revue les projets collaboratifs sur lesquels j'ai pu contribuer.

Collaborations et projets collaboratifs

Contents

7.1	Implication dans des Projets de recherche	105
7.2	Collaborations académiques et industrielles	111
7.2.1	Collaborations nationales	111
7.2.2	Collaborations internationales	112
7.3	Transfert de Technologie	113

7.1 Implication dans des Projets de recherche

Tout au long de mon parcours académique, j’ai eu la chance de participer à plusieurs projets collaboratifs. Pour la plupart de ces projets, j’ai activement contribué à la phase de montage des projets. J’ai donc pu répondre à de nombreux appels à projets, chercher des financements, établir les besoins en concertation avec différents partenaires, qu’ils soient académiques ou industriels, français ou étrangers, collègues chercheurs, ou uniquement utilisateurs du produit de la recherche développée.

Le tableau 7.1 recense les différents projets dans lesquels j’ai pu être impliqué depuis 2005. J’ai agi en tant que porteur pour le projet ANR “TSAR”, pour tous les autres projets, j’étais le responsable scientifique du partenaire Polytech/IRCCyN ou Polytech/RMeS. Cela signifie donc que pour chacun de ces projets, je conduisais les travaux de recherche et développement de mon laboratoire. Je procédais au recrutement et à l’encadrement des personnels contractuels (Masters, Ingénieurs ou post-docs). Je contribuais à l’avancement scientifique, seul ou en collaboration avec le personnel recruté. Pour chaque projet, j’ai développé des méthodes de traitement d’images issues de la recherche amont.

A plusieurs reprises, des logiciels développés suite à mes activités de recherche ont été fournis à nos partenaires. Pour chaque projet, j’ai rédigé les rapports d’avancement ainsi que le rapport final du projet. De plus, pour chacun des projets que j’ai pu conduire, j’ai publié nos résultats dans des revues et conférences scientifiques. Je contribue actuellement sur un projet financé par un partenaire autrichien au cours duquel je suis en charge de conduire des tests subjectifs sur des images/vidéos chiffrées. J’ai aussi obtenu début 2020 un financement de 60 K€ sur un Projet

Financier	Projet	Années	Partenaires	Sujet(s)
ANR (ARA)	TSAR	2005-08	IRCCyN, IETR, LIRMM, GIPSA, C2RMF	Watermarking, chiffrement, compression
ANR (CSOSG)	Safim@ge	2007-10	IRCCyN, LINA , Alcatel, IST	Watermarking Vidéo, modèles de perception
Région PdL	Bourse post-doctorale	2007-08	N/A	Compression Mojette & multi-résolution
Région PdL	Miles	2006-09	LINA, IREENA, LERIA, LIUM, LISA, IMN, LAUM	Compression & chiffrement sélectif
OSEO	e-Map	2010-12	LINA, Ec. de Design, IM'INFO, Human Connect	Tatouage perceptuel, masques psychovisuels
OSEO	StreamMaster	2012-14	EMN, Ec. de Design, IM'INFO, DataSysCom	Tatouage perceptuel Robuste aux dist. géométriques
FWF	ViSecMetrics	2017-20	Univ. of Salzburg / WaveLab	Perception, reconnaissance des images chiffrées
Univ. Nantes	Interdisciplinaire	2017-18	RMeS, Polytech'Nantes / LS2N	Reconnaissance des formes, détection de dents
ANR	Post-doc (ICAN)	2018	RMeS, Institut du Thorax	Caractérisation de bifurcations sur IRM
NExT "FAME"	TOMOSTEO	2020	RMeS, CHU de Nantes (rhumato & odonto)	Reconnaissance des formes, morpho math
NExT "Innovez"	Odon	2020	RMeS, Artefakt-AI/ Deep Meaning	Reconnaissance des formes, Deep Learning
PHRC	Post-doc (U-CAN)	2020	RMeS, CHU Nantes Institut du Thorax	Segmentation vascu cérébrale, Deep Learning

TABLE 7.1 – Contributions à des projets nationaux et internationaux.

Hospitalier de Recherche Clinique (**PHRC**) pour mener à bien la détection et la caractérisation des anévrismes intra-crâniens avec mes partenaires du CHU de Nantes.

Ces projets m'ont permis de me pencher sur des problématiques scientifiques assez variées, telles que la protection des copyrights, la cryptographie, la biométrie, la compression sans pertes, la perception des couleurs, l'évaluation de la qualité des images/vidéos, l'imagerie médicale, l'analyse statistique des données biologiques. Ma production scientifique reflète assez bien la diversité de mes travaux.

Je détaille ici brièvement les principaux projets dans lesquels j'ai été impliqué, les différents encadrements que j'ai pu dispenser, ainsi que mes contributions.

Le projet TSAR

Le projet *TSAR* faisait partie de l'appel à projet "Axes de Recherche Amont" de l'ANR, il s'est étalé sur une période allant de 2005 à 2008. J'étais le porteur de ce projet. Il avait pour objectif de sécuriser la base de données numérique du Centre de Recherche et de Restauration des Musées de France (C2RMF). Ce der-

nier, basé au Musée du Louvre à Paris a numérisé une grande partie des oeuvres des musées de France. De nombreuses peintures et sculptures sont photographiées, scannées (dans la lumière visible, mais aussi en Infra-Rouge et Ultra-Violet). Cette immense base de donnée est donc stockée au C2RMF et devait à l'époque du projet *TSAR* être sécurisée et compressée (sans pertes). C'est dans ce cadre que j'ai pu conduire mes travaux portant sur la compression Mojette et le chiffrement sélectif [Autrusseau 2006a, Kingston 2007] (cf. Chapitre 5). Certains de mes travaux portant sur les métriques de qualité ont aussi été conduits pendant ce projet [Marini 2007b, Ninassi 2006].

Dans ce projet, la contribution du LIRMM portait sur les méthodes de chiffrement, le GIPSA-Lab avait pour objectif de proposer des méthodes de *watermarking*, et l'IETR développait une méthode de compression sans pertes (le LAR).

Le projet Safim@ge

Le projet *Safim@ge* a été accepté dans le cadre de l'appel à projets "Concepts, Systèmes et Outils pour la Sécurité Globale" de l'ANR. Il était porté par Alcatel-Lucent entre 2007 et 2010. Ce projet a été organisé autour de plusieurs axes :

- Développement d'une sonde permettant l'analyse en temps réel du contenu des flux IP de l'internet sur des liens ayant un débit jusqu'à 10 GB/s
- Application d'interception légale du trafic Internet et plus généralement tout trafic IP)
- Machine Virtuelle Java (JVM) utilisant directement l'ensemble des accélérateurs hardware des « Network Service Processors » de la plateforme pour développer rapidement de nouvelles applications et services
- Ontologie de suivi de session portée sur la JVM pour prise en compte rapide de nouveaux protocoles et critères de décision
- Tatouage de vidéos
- Classification statistique des flux multimédias.

Les partenaires étaient donc Alcatel-Lucent qui portait le projet, le LINA, pour les aspects de la classification de données Multimedia et l'ontologie. La société IS2T était en charge des travaux sur la JVM, et l'IRCCyN (ma contribution) avait pour objectif de proposer une méthode de tatouage de contenus vidéos et de fournir des métriques de qualité associées. Les travaux que j'ai pus conduire dans ce projet ont été présentés dans le Chapitre 3, et notamment, les publications [Chen 2014, Maillet 2011, Maillet 2008a, Carosi 2010, Pankajakshan 2010].

Le projet Miles

Le projet Miles était financé par la Région Pays de la Loire et s'est étalé sur la période allant de 2006 à 2009. De nombreux laboratoires Ligériens étaient impliqués (IRCCyN, LINA, IREENA, LERIA, LUIM, LISA, IMN-PCM, LGMPA, LAUM, LSO, UCO2M). Concernant la contribution de l'IRCCyN, nos tâches dans ce projet portaient principalement sur le développement de méthodes de compression sans pertes et de chiffrement sélectif. Ce projet a donc été pour nous l'occasion

de poursuivre des travaux initiés lors du projet *TSAR*. Avec Andrew Kingston, nous avons donc pu parfaire les méthodes initiées précédemment, ce qui a abouti sur les articles [Kingston 2008b, Kingston 2009a, Kingston 2009c, Kingston 2009b, Kingston 2008a]. Le financement de la bourse de post-doc d'Andrew avait été obtenu via un appel à projet de la région Pays de la Loire et d'OSEO. Grâce à une collaboration de longue date avec l'Université de Rome (et en particulier avec Patrizio Campisi, Professeur à Università Roma Tre), j'ai pu accueillir Simone Colosimo en stage de Master Erasmus, qui a pu nous aider sur le chiffrement sélectif Mo-jette [Kingston 2007].

Les projets e-Map et StreamMaster

Je regroupe ici volontairement ces deux projets car le consortium était sensiblement le même pour les deux projets, *StreamMaster* peut être vu comme une suite logique d'*e-Map*. Voici donc la composition du consortium : LINA-GRIM, Ecole de Design, IRCCyN-IVC, IM'INFO, Human Connect et DataSysCom. DataSysCom était porteur sur ces deux projets, le LINA travaillait sur du clustering, l'école de design sur des IHM, quant à l'IRCCyN, nous devions proposer des méthodes de tatouage des images. Tout d'abord, dans *e-Map* nous avions pour objectif de protéger des cartes visuelles (cartes de *Mind mapping*), alors que dans *StreamMaster*, nous devions protéger des documents administratifs (bulletins de salaires) dans un contexte de dématérialisation de documents administratifs officiels. Ces projets ont abouti à diverses publications, avec notamment un brevet [Autrusseau 2013a] et une publication dans la revue IEEE TIFS [Urvoy 2014b]. Les travaux sur le test de Grubbs rentraient eux aussi dans ce contexte [Urvoy 2014a]. Au cours du projet *e-Map*, j'ai recruté Vinod Pankajakshan en tant que post-doctorant et Sylvain David sur un poste d'ingénieur d'étude. Via le projet *StreamMaster* deux recrutements ont eu lieu, Matthieu Urvoy en post-doctorat, et Dalila Goudia sur un poste d'ingénieur maturation (via la SATT Ouest Valorisation). Ces différentes études ont été détaillées dans le Chapitre 3.

Le projet ViSecMetrics

Le projet *ViSecMetrics* était porté par Andreas Uhl de l'Université de Salzbourg. Ce projet traite de la sécurité des images et vidéos par des méthodes de chiffrement sélectif. Contrairement à mes travaux antérieurs portant sur la mise en oeuvre de méthodes de chiffrement sélectif, dans ce projet, ma contribution était toute autre. Effectivement, Andreas m'a contacté pour mes compétences en modélisation perceptuelle, mon objectif était ici de déterminer la *recognizability*, ou l'identification du contenu. Au cours de ce projet l'Université de Salzbourg proposait des méthodes de chiffrement sélectif, et de mon côté, je devais conduire des tests subjectifs sur la base d'image ainsi établie et tester différentes métriques objectives de qualité. Ce projet nous a permis de découvrir que parmi les nombreuses métriques de qualité actuellement disponibles dans la littérature, aucune n'était en mesure de prédire

efficacement si une image donnée sera reconnue par des observateurs humains. J’ai ainsi pu proposer une métrique structurelle qui permet de mieux évaluer la perception dans un tel scénario. Le *Template Matching* [Lewis 1995] semble beaucoup plus adapté à ce type de tâches, ainsi que, dans une moindre mesure, les descripteurs du type SIFT [Lowe 1999] ou SURF [Bay 2006]. Nous avons vu dans le détail la description de ces travaux dans le Chapitre 3.

Le projet Odontogramme

J’ai monté le projet *Odontogramme* en arrivant dans le laboratoire RMeS fin 2016, le projet s’est étalé sur les années 2017-2018. C’était un projet Interdisciplinaire financé par l’Université de Nantes monté en collaboration avec les Professeurs Jeanpierre Guéron à Polytech’Nantes et Yves Amouriq, PU-PH en Chirurgie Dentaire et moi même. J’ai pu recruter une stagiaire de Master 2 (SIBM), Natalia De La Fournière pour travailler sur ce projet. L’objectif était de proposer un algorithme de détection des dents sur des radiographies panoramiques. Dans le cahier des charges formulé par le Pr. Yves Amouriq, il nous était demandé de générer automatiquement le schéma dentaire d’un patient en fonction de sa radiographie panoramique. Cela signifie, qu’outre la détection des dents, il est aussi souhaitable de détecter les différents soins que le patient a pu subir. Natalia s’est attaquée au problème par deux approches différentes, une approche était orientée vers l’analyse des images, avec une partie de ses travaux portant sur le clustering de données, et une autre partie de son stage portait sur l’exploitation d’un algorithme de Deep Learning (U-Net). Ces travaux ont été résumés au cours du Chapitre 6.

L’objectif initial étant très ambitieux, lors du stage de Natalia, nous n’avons pu nous pencher que sur une seule partie du problème : la détection des dents. Cependant, suite à ce projet, Natalia a souhaité poursuivre l’aventure et a créé une entreprise dédiée à cette même problématique de génération automatique de l’odontogramme. Avec cette entreprise, j’ai pu proposer un projet en réponse à l’appel à projet “Innovez” de l’I-Site NExT.

Nous accueillerons Clément Morin pour son stage Ingénieur entre juillet et décembre 2020 dans le cadre de ce projet *Odon*. Ce dernier sera financé par la société *Artefakt-AI* et je l’encadrerai au sein du laboratoire RMeS. De plus, si le projet est accepté, nous pourrions recruter un post-doctorant sur une période d’une année pour travailler sur l’utilisation de méthodes de Deep Learning pour la génération du schéma dentaire.

Le projet I-CAN

Contrairement aux autres projets présentés ici, je n’ai pas contribué au montage du projet *I-CAN* et n’en ai pas non plus été partenaire, je l’ai rejoint en cours de route, mais j’ai pu bénéficier d’une bourse pour recruter et encadrer un post-doctorant. Je me suis fortement impliqué dans ce projet, et bien que n’ayant pas été à l’initiative du montage du projet initial, j’ai par la suite activement cherché des sources de financement connexes pour poursuivre les travaux initiés lors

d'*I-CAN*. Ce vaste projet était porté par l'équipe de neuro-radiologie du CHU de Nantes (Hôpital Nord Laënnec), le Dr. Romain Bourcier (Neuro-radiologue) et le Pr. Hubert Desal (PU-PH, chef du service de neuro-radiologie). Le projet *I-CAN* s'intitule précisément "Anévrismes intra-crâniens : des formes familiales aux mécanismes physiopathologiques" / "Understanding the pathophysiology of Intracranial Aneurysm". L'objectif est donc d'étudier la formation des anévrismes intra-crâniens. Le consortium étant principalement composé de médecins et de biologistes, une grande partie du projet portait sur l'étude des prédispositions génétiques pouvant entraîner la formation des anévrismes intra-crâniens. Romain Bourcier m'a contacté via la plateforme [SC3M](#) pour apporter une expertise sur l'analyse des images. L'objectif était alors de proposer des outils permettant de caractériser l'arbre vasculaire cérébral, il nous fallait notamment détecter automatiquement les bifurcations dans l'arbre vasculaire, et collecter plusieurs mesures sur les artères formant ces bifurcations (diamètres, angles, etc.). J'ai donc recruté Anass Nouri sur un post-doctorat, il a rejoint l'aventure de la détection des anévrismes en 2017 et 2018. Nous avons vu dans le détail nos contributions sur le projet *I-CAN* dans le Chapitre 6. Suite au post-doctorat d'Anass, dans une optique d'amélioration des performances de la segmentation, j'ai encadré une stagiaire de Master 2 de l'Université de Paris Est Créteil. Julia Guillou a effectué son stage au laboratoire [RMeS](#) entre Mars à Août 2019 sur l'exploitation de critères perceptuels pour la segmentation des arbres vasculaires cérébraux. Julia était docteur en médecine, avait une formation de neuro-pédiatre, et avait suivi les cours théoriques du Master "Sciences pour l'Ingénieur" spécialité "Signaux et Images en Médecine" avant d'effectuer son stage de recherche.

L'aventure du projet *I-CAN* s'est poursuivie après le projet ANR initial, un [PHRC](#) a été accepté début 2020. Dans le cadre de ce [PHRC](#), je dispose de 60 K€ me permettant de recruter un post-doctorant pour travailler sur la détection et la caractérisation des anévrismes (c'est le projet U-CAN). De plus, deux doctorantes travaillent actuellement à l'amélioration de nos précédents résultats et cherchent à établir un atlas anatomique et une méthode de détection des anévrismes. Un post-doctorant a rejoint l'aventure en Septembre 2020.

Le projet Tomosteo

Trois parties sont impliquées dans le projet *Tomosteo*, le service de rhumatologie, représenté dans le projet par le Pr. Yves Maugars (PU-PH, chef du service de rhumatologie du CHU de Nantes), le département de parodontologie, avec le Pr. Assem Soueidan, (PU-PH, UFR d'Odontologie, chef de service de parodontologie) et moi-même concernant la partie sur l'analyse des images numériques. L'objectif de ce projet est de dépister l'ostéoporose à partir de volumes acquis au [CBCT](#), classiquement, l'ostéoporose est diagnostiquée via des scans de [DXA](#), mais le coût moindre des [CBCT](#) et leur plus grande disponibilité peut rendre très attractive la possibilité d'un dépistage par ce biais. Pendant ce projet, j'ai obtenu deux bourses de stage de Master 2 (via le cluster "FAME" de l'I-Site NExT), l'une de ces bourses était destinée à une étudiante de l'UFR d'odontologie (Marianne Lucazeau), encadrée par

le Pr. Soueidan, et l'autre (Clément Morin) était en rapport direct avec l'analyse des images de CBCT. Avec Clément, nous avons pu proposer diverses méthodes de mesures automatiques des caractéristiques géométriques du Foramen Mentonnier, et nous avons de plus exploité différentes méthodes d'analyse des textures nous permettant d'estimer la solidité de l'os cortical et de l'os trabéculaire.

7.2 Collaborations académiques et industrielles

Au cours de ces nombreuses années d'affectation dans différents laboratoires de recherche de l'Université de Nantes, j'ai eu la chance de tisser des liens avec des collègues chercheurs dans diverses thématiques. Pour la majorité des collaborations auxquelles j'ai pu prendre part, mes partenaires étaient affiliés dans des laboratoires de recherche publics, mais via certains projets ou diverses rencontres, j'ai aussi pu collaborer avec des équipes de recherche privées.

Ma participation dans des différents projets de recherche listés dans la section précédente m'a permis de nouer de nombreux liens avec différents collègues sur différentes thématiques de recherche de par le monde. Je vais ici dénombrer quelques unes de ces collaborations.

7.2.1 Collaborations nationales

Comme nous avons pu le voir dans la section précédente, grâce à plusieurs projets de recherche nationaux, j'ai pu collaborer avec certains laboratoires de recherche Français, parmi lesquels l'IETR à Rennes, le Gipsa-lab à Grenoble, le LIRMM à Montpellier, le LINA à Nantes, et bien d'autres encore.

Ma collaboration avec l'IETR s'est faite avec Olivier Déforges et Marie Babel. J'ai conduit une évaluation de la qualité de l'algorithme de compression LAR [Autrusseau 2009a, Strauss 2009]. Un test subjectif a été conduit, et diverses métriques de qualité objectives ont été testées. Avec mes collègues du Gipsa-Lab (Patrick Bas et François Cayre) j'ai proposé des métriques de qualité adaptées au tatouage des images. L'utilisation de la métrique *Komparator* qui avait démontré certaines qualités dans ce domaine était suggérée dans le contexte du concours BOWS 2 (organisé par Patrick Bas).

Ces collaborations se sont faites dans le cadre du projet *TSAR* comme nous avons pu le voir dans la section précédente [Pitzalis 2008].

Avant de rejoindre le LIOAD j'ai eu plusieurs opportunités de collaborer avec Pauline Bléry et Yves Amouriq sur la caractérisation osseuse [Guédon 2012, Autrusseau 2016] ou la segmentation de la vascularisation osseuse [Bléry 2014, Crauste 2015]. Ces collaborations se sont bien évidemment poursuivies après mon intégration au sein du laboratoire RMeS [Michel 2018].

De façon similaire, j'ai collaboré avec le Pr. Hubert Desal via l'encadrement de la thèse de Laure Mahé en 2016 sur un projet portant sur la détection des lésions vasculaires cérébrales [Mahé 2016]. Ce n'est que quelques années plus tard que cette

collaboration a repris son cours sur les projets relatifs à l'étude des anévrismes intra-craniens [Nouri 2019, Guillou 2020, Nouri 2020] avec le Dr. Romain Bourcier.

Sur le plan national, mes collaborations avec des équipes de recherche privées se sont surtout faites par le biais des projets de recherche, avec notamment la thèse CIFRE d'Abhijit Sarkar chez Technicolor à Cesson Sévigné. Laurent Blondé et Jurgen Stauder étaient les encadrants d'Abhijit chez Technicolor. Nous avons travaillé sur la perception des couleurs, comme nous avons pu le voir dans le Chapitre 4. Via le projet *Safim@ge*, j'ai pu travailler avec Alcatel-Lucent à Orvault (Alain Maillet et Laurent Bureau) sur le tatouage des vidéos [Maillet 2011, Maillet 2009]. Plus récemment, des contacts avaient été pris avec Gwenaël Doërr (Technicolor à Cesson Sévigné), nous envisagions de conduire des tests subjectifs sur les méthodes de tatouage vidéo de Technicolor, mais la collaboration n'a malheureusement pas pu se poursuivre suite à mon changement d'affectation, et au même moment, Gwenaël quittait Technicolor pour fonder la société *ContentArmor*.

Grace à mon premier brevet, déposé en 2013 [Autrusseau 2013a], et surtout à travers l'accompagnement de la SATT "Ouest Valorisation" (et en particulier Julien Roca) j'ai pu prendre contact avec diverses entreprises, telles que *Advanced Track & Trace* ou encore *Agelia*, mais c'est l'entreprise *DataSysCom*, partenaire sur les projets *e-Map* et *StreamMaster* qui a finalement acquis des licences du brevet dans le cadre des travaux que nous conduisions ensemble. De plus, un contrat de collaboration a été signé avec *DataSysCom* pour que je conduise une étude de faisabilité de la détection des watermarks via des tablettes.

7.2.2 Collaborations internationales

Outre ces diverses collaborations avec des acteurs français de la recherche, j'ai aussi eu la chance de conduire plusieurs travaux avec des laboratoires académiques à l'étranger. J'ai par exemple collaboré à plusieurs reprises avec le Pr. Junyu Dong de la *Ocean University of China* à Qingdao en Chine sur différents outils exploitant la transformation Mojette. Nous avons ainsi pu proposer une méthode de tatouage basée sur la Mojette [Zhang 2010], ou encore une technique d'analyse de textures [Jia 2008, Dong 2012] encore une fois se basant sur la Mojette.

Certaines collaborations ont été très ponctuelles, comme par exemple avec Fan Zhang, de Lenovo à Hong Kong avec qui j'ai pu travailler sur un article [Zhang 2014] dans lequel nous proposons une métrique de qualité se basant sur les propriétés du cortex visuel V1. D'autres collaborations se sont inscrites dans la durée, comme par exemple avec l'Université de Salzbourg. C'est une collaboration qui a débuté en 2010 avec Thomas Stütz (à l'époque à l'Université de Salzbourg) [Stütz 2010] s'est poursuivie lors du post-doc de Thomas à Nantes sous ma direction en 2011 [Stütz 2011], puis s'est poursuivie par la suite [Stütz 2014]. J'ai pu rendre visite à Thomas à la *Fachhochschule Salzburg* (où il est à présent enseignant chercheur) en 2012 et 2014. Jusqu'en 2020, je suis resté en contact avec le Pr. Andreas Uhl (responsable du département "Computer Science" et directeur du laboratoire "WaveLab" à l'Université de Salzbourg) et Heinz Hofbauer, avec qui je contribue sur le projet *ViSecMetrics*

(cf. détail dans la section précédente).

De façon similaire, la collaboration avec Patrizio Campisi de l'Université de Rome a duré de nombreuses années. J'ai co-encadré avec lui différents stagiaires de Master Erasmus, tels que Enrico Marini [Marini 2007b], Maurizio Carosi [Carosi 2010], ou encore Simone Colosimo [Kingston 2007]. Nous avons conduit de nombreuses études sur les métriques de qualité dans une problématique de tatouage [Le Callet 2008], et je l'ai aussi associé à nos travaux avec Andrew Kingston (qui venait de l'Université de Monash à Melbourne et est à présent à l'Université de Canberra) portant sur le chiffrement sélectif.

Finalement, j'ai eu la chance de faire un post-doctorat à l'Université de Chicago, sous la direction du Pr. Steve Shevell, lors duquel j'ai pu collaborer avec le Pr. Thibos de l'Université de l'Indiana. Au retour de mon post-doctorat, j'ai pu poursuivre ma collaboration avec Steve jusqu'en 2011 [Autrusseau 2011b].

7.3 Transfert de Technologie

Certains de mes travaux ont abouti à des dépôts de brevets. Un premier brevet a été déposé en 2011 sur la dissimulation de l'information, et plus spécifiquement sur la protection des droits d'auteurs des images numériques [Autrusseau 2013a]. Les programmes ayant permis de proposer cette nouvelle méthode de tatouage ont quant à eux été déposés à l'Agence de Protection des Programmes (APP). Ce brevet a été exploité par la société DataSysCom qui était l'un de nos partenaires sur le projet *StreamMaster*. Grace à l'accompagnement de la SATT "Ouest Valorisation", une étude de marché a été conduite par Quadrium et a permis d'identifier nombre d'acteurs industriels sur le territoire Français, j'ai ainsi pu échanger avec de nombreuses entreprises (Novadis, Nexima, AlpVision, Civolution, Hologram Industries,...). Un "Non Disclosure Agreement" a été signé pour chacune de ces prises de contact. Un démonstrateur a été fourni pour montrer l'efficacité de notre méthode de tatouage. J'ai aussi pu bénéficier du support d'un poste d'ingénieur maturation en 2013-2014 pour tester extensivement les performances de la méthode de tatouage (poste de Dalila Goudia).

Un second brevet a été déposé en 2019 par la filiale de l'INSERM dédiée à la valorisation (INSERM-Transfert) sur la caractérisation des bifurcations de l'arbre vasculaire cérébral [Nouri 2018], encore une fois, le code qui y était rattaché a été déposé à l'APP. Nous sommes trois inventeurs sur ce brevet, Anass Nouri, Romain Bourcier et moi même. Le brevet est en cours d'extension. S'en suivra peut être un projet de maturation porté par INSERM-Transfert. Si tel était le cas, je pourrais encore une fois bénéficier d'un poste d'ingénieur maturation pour peaufiner la méthode de détection des bifurcations, et surtout tenter d'améliorer la rapidité des algorithmes implémentés.

J'ai très récemment (Février 2020) déposé un troisième brevet [Autrusseau 2020] (encore avec Anass Nouri et Romain Bourcier) portant sur l'exploitation de critères perceptuels pour optimiser la segmentation de l'arbre vasculaire cérébral. La procé-

de dépôt est à ce jour toujours en cours. Ce dernier est une fois encore porté par INSERM-Transfert.

Nous avons pu voir dans ce chapitre que mes travaux de recherche se sont faits en grande partie en collaboration avec diverses personnes, équipes de recherche, académiques ou industriels, Français ou étrangers. C'est, à mon sens la meilleure façon de faire de la recherche, en s'ouvrant aux autres, en collaborant, en s'organisant autour de projets construits, avec des objectifs clairement définis et des problématiques et des besoins bien identifiées.

Conclusions et perspectives de travaux

Contents

8.1 La Perception visuelle	115
8.1.1 Principales contributions	116
8.1.2 Perspectives	117
8.2 La Mojette	117
8.2.1 Principales contributions	118
8.2.2 Perspectives	118
8.3 L'imagerie médicale et biologique	118
8.3.1 Principales contributions	119
8.3.2 Perspectives	119
8.4 Perspectives générales	119

Comme nous avons pu le voir tout au long de ce manuscrit, mes travaux se sont principalement articulés autour de trois axes. Une partie de mes recherches a porté sur la compréhension et la modélisation de la perception visuelle, et ce, pour diverses applications. J'ai aussi beaucoup travaillé sur la transformation Mojette, une version discrète de la transformation de Radon. Encore une fois, plusieurs applications en ont découlées. Finalement, mes travaux les plus récents ont porté sur l'analyse d'images médicales et biologiques. Ces travaux ont majoritairement été conduits dans le cadre de projets de recherche financés par des organismes tels que l'ANR, l'OSEO, la Région Pays de la Loire ou encore NExT. Comme nous l'avons vu, ceci a impliqué de nombreuses collaborations, tant au niveau national qu'international, avec des laboratoires académiques ou des équipes d'industriels. Pour chacun de ces projets, j'ai pu encadrer des étudiants et/ou du personnel contractuel.

Dans ce chapitre, je vais dresser quelques conclusions sur mes diverses études. Je tacherai de faire ressortir les points forts de ces contributions, et je mènerai aussi une réflexion sur les différentes perspectives possibles de ces travaux.

8.1 La Perception visuelle

Comme nous avons pu le voir dans les chapitres 3 et 4, au cours de ma thèse de doctorat à l'IRCCyN, mon post-doctorat à l'Université de Chicago, et finalement

mon poste d'ingénieur de recherche à l'Université de Nantes, j'ai eu l'occasion de mener divers travaux sur l'étude et la modélisation de la perception humaine. Ces travaux ont eu différents objectifs, je me suis tout d'abord penché sur l'exploitation de modèles psychovisuels pour garantir l'invisibilité de méthodes de dissimulation de l'information. En effet, la meilleure façon d'optimiser la robustesse des techniques de *watermarking* (protection de copyrights) est de placer la marque au plus près du seuil de visibilité. Dans le monde de la recherche, la communauté du *Data Hiding* et la communauté de la perception humaine sont malheureusement assez peu connectées. A travers nombre de mes travaux, j'ai tenté de réconcilier ces deux thématiques, et je l'ai fait via deux approches distinctes. Dans un premier temps, j'ai proposé des schémas de tatouage des images [Autrusseau 2007a, Urvoy 2014b] et des vidéos [Chen 2014, Stütz 2014] qui comportaient une modélisation du Système Visuel Humain, ceci m'a permis de garantir non seulement l'invisibilité des modifications apportées, mais aussi d'assurer un compromis optimal entre l'invisibilité et la robustesse. Ces travaux m'ont par la suite entraîné à me pencher sur le problème de la sélection du seuil de détection (test de Grubbs) afin d'optimiser les méthodes de tatouage développées. Dans un second temps, j'ai testé des métriques de qualité de la littérature afin de déterminer celles qui se prêtaient le mieux à une problématique donnée [Le Callet 2008]. Constatant que les métriques de l'état de l'art étaient pour la plupart conçues pour évaluer des distorsions engendrées par des algorithmes de compression ou de filtrage, mais absolument pas pour faire face à un contexte de camouflage de l'information, j'ai donc proposé des métriques de qualité adaptées à des distorsions proches du seuil de visibilité [Carosi 2010, Pankajakshan 2010], ou, au contraire pour répondre à un besoin d'identification du contenu [Hofbauer 2020a].

Outre ces travaux relatifs à des problématiques de sécurité des contenus, j'ai aussi cherché à comprendre et à modéliser la vision des couleurs, avec notamment la définition d'un observateur standard du point de vue des aberrations chromatiques (optiques) [Autrusseau 2011b], et l'étude du métamérisme observateur avec la mise en oeuvre de cluster d'observateurs partageant une vision des couleurs similaire [Sarkar 2011a].

8.1.1 Principales contributions

Parmi mes contributions les plus marquantes dans ce domaine, je citerais principalement mon rôle de charnière entre l'univers du *watermarking* et celui de la perception humaine. Les métriques de qualité adaptées au tatouage que nous avons proposées se sont avérées être parmi les plus efficaces de la littérature. La méthode de tatouage perceptuel publiée dans [Urvoy 2014b] fait aussi partie des méthodes les plus efficaces qui soient aptes à résister à des distorsions géométriques sévères. La méthode de sélection du seuil de détection par le test de Grubbs est aussi particulièrement novatrice et a démontré une efficacité redoutable tout en garantissant une mise en oeuvre très simplifiée. Quant aux modèles observateurs de perception des couleurs, il n'en existait pas de comparable à l'époque et il est crucial de disposer de tels modèles tant dans un scénario plus théorique et fondamental (mes travaux

avec Steve Shevell) qu'appliqué (encadrement de la thèse d'Abhijit à Technicolor).

8.1.2 Perspectives

Parmi les perspectives possibles de ces travaux, celle qui me vient tout naturellement à l'esprit est l'exploitation des critères de vision humaine dans des applications de segmentation et de reconnaissance des formes pour l'imagerie médicale. Comme nous l'avons vu, j'ai d'ores et déjà commencé à exploiter des modèles de vision humaine pour la segmentation [Guillou 2020, Autrusseau 2020], mais il reste beaucoup à faire dans ce domaine. Je compte bien profiter du financement obtenu sur le PHRC *WE-CAN* pour recruter un post-doctorant avec lequel j'explorerai cette thématique qui me paraît très prometteuse. Nous en profiterons aussi pour exploiter des méthodes d'apprentissage profond conjointement avec des critères de vision afin d'optimiser la détection et la reconnaissance d'*objets* (arbre vasculaire entre autres, mais aussi des cellules par exemple).

8.2 La Mojette

Bien qu'étant, par sa construction même, un outil très simple (des projections par simples additions) la Mojette permet toutefois de répondre à différents besoins et peut s'adapter à de nombreuses applications. Comme nous avons pu le voir, dans mes travaux, la Mojette a servi pour répondre à des problématiques de tatouage, chiffrement, compression, et même pour de l'analyse de textures [Dong 2012]. Initialement, mes travaux sur la transformation Mojette s'inscrivaient dans un contexte de protection des droits d'auteurs des images numériques (sujet de ma thèse). Ensuite, au fil du temps, j'ai constaté que la Mojette disposait de propriétés très intéressantes pour ce qui touchait à la redondance de l'information, mais que malheureusement, personne n'avait encore réussi à compresser l'information avec cette transformée. Je me suis donc fixé le défi de proposer une méthode de compression Mojette. En exploitant la similarité des projections Mojette d'angles proches, ainsi que la périodicité des bins intra-projections, il a donc été possible de compresser (sans pertes) les données Mojette [Autrusseau 2006a]. Avec Andrew Kingston nous avons pu optimiser sensiblement les performances de la compression grâce à l'utilisation d'une représentation 2D (et non plus 1D) des projections [Kingston 2008a, Kingston 2009a]. S'en sont suivies plusieurs études portant sur le chiffrement sélectif basé sur la Mojette. L'instabilité de la rétro-projection est alors un atout pour crypter les données. De plus, le codage inter-projection nécessitant une projection de référence permettant de décoder les différences, peut être exploité pour réduire la quantité d'information à chiffrer. Nous avons alors pu proposer une cascade de projections permettant de ne chiffrer qu'une quantité infime d'information tout en rendant le support impossible à reconstruire. Finalement, à nouveau avec Andrew, nous avons proposé une autre propriété qui manquait à la Mojette : une représentation Multi-résolution [Kingston 2009c].

8.2.1 Principales contributions

Mon apport le plus significatif sur la transformée Mojette aura probablement été de lui donner la capacité de compresser des données. La Mojette était principalement utilisée dans un contexte où l'apport de la redondance était cruciale (stockage/transmission sécurisée). Or, avant 2006 il était impossible de compresser avec cette transformation. Mes travaux l'ont permis, et avec Andrew nous avons pu accroître le taux de compression. Le chiffrement sélectif Mojette a aussi vu le jour au cours de ces mêmes années. L'instabilité de la rétro-projection qui était plutôt une faiblesse de la Mojette pour des problématiques de transmission sur des réseaux numériques est alors devenue une force dans un cadre de protection des données.

8.2.2 Perspectives

Bien que depuis quelques années, je me sois un peu éloigné de la Mojette (ne faisant plus de recherches sur la transformée), je n'exclue bien évidemment pas de l'utiliser pour répondre à des besoins pour lesquelles elle pourrait se prêter, par exemple, nous l'avons récemment exploitée pour mesurer des diamètres d'artères [Nouri 2020].

De plus, comme je l'ai brièvement mentionné dans la section 5.2, avec l'arrivée de la transformation Mojette multi-résolution, il serait à présent plus aisé d'imaginer une méthode de dissimulation de l'information. Effectivement, l'insertion d'un watermark dans les bandes de fréquences élevées pourrait se faire sans toutefois dégrader fortement la qualité du contenu. Cependant, comme je l'ai expliqué, mes thématiques de recherche ont évolué et se sont éloignées du watermarking, mais il n'est reste pas moins qu'il serait intéressant de tester une méthode de tatouage dans des fréquences Mojette.

8.3 L'imagerie médicale et biologique

De par mon parcours, je ne me suis investi qu'assez récemment dans la thématique de l'analyse des images médicales et biologiques. Effectivement, ma mise à disposition au LIOAD/RMeS date de fin 2016. Cependant, c'est un vaste champ de recherches qui est captivant. Les problématiques de recherche en analyse d'images médicales sont très variées, souvent très intéressantes, et il reste encore de nombreux défis à relever. J'ai eu la chance de prendre part à trois projets distincts, comme nous avons pu le voir dans le chapitre 6 de ce manuscrit. L'un de ces projets portait sur la détection des dents et des soins dentaires subis par un patient sur des radiographies panoramiques. Nous avons tenté de répondre à ce besoin via deux approches, l'une était basée sur une approche orientée analyse des images, et l'autre exploitait des outils d'apprentissage profond. Avec Natalia De La Fournière, nous avons ainsi pu détecter avec assez de précision les dents, et sommes actuellement impliqués dans un projet qui devrait nous permettre de poursuivre ces travaux. Nous souhaitons être rapidement en mesure de détecter les différents traitements que le patient a subi. Le second projet porte sur la détection de l'ostéoporose en se basant sur des images

de CBCT. Encore une fois, nous tentons de résoudre ce problème avec des outils d'apprentissage profond (réseau U-Net). Ces travaux sont encore en cours au moment de la rédaction de ce mémoire. Nos résultats préliminaires sont toutefois très encourageants et nous espérons obtenir des résultats significatifs dans les prochains mois. Finalement, le troisième et dernier projet dans lequel je me suis impliqué sur l'analyse des images médicales est le plus significatif. C'est le projet *I-CAN* et par la suite le projet *WE-CAN*. Mes travaux dans ce contexte ont porté sur la détection des anévrismes, la recherche et caractérisation des bifurcations dans l'arbre vasculaire cérébral. Avec l'aide d'Anass Nouri, nous avons obtenu des résultats très intéressants, nous avons même été en mesure de déterminer quelles caractéristiques géométriques des bifurcations semblaient être les plus prépondérantes dans le risque de formation des anévrismes.

8.3.1 Principales contributions

Les travaux portant sur la détection des dents et sur l'estimation du risque ostéoporotique étant encore assez récents, et à l'heure de la rédaction de ce manuscrit non encore aboutis, ma contribution principale dans la thématique de l'analyse des images médicales touche aux travaux sur la détection des anévrismes. Effectivement, comme expliqué précédemment, nous avons été en mesure de déterminer les caractéristiques géométriques des bifurcations qui seraient responsables de la formation des anévrismes. À notre connaissance, de tels résultats n'avaient pas encore été atteints. Bien que de nombreux aspects restent encore à étudier (plusieurs projets encore en cours sur ce sujet) nous avons d'ores et déjà obtenu des résultats significatifs. Nous avons bon espoir d'aller encore beaucoup plus loin avec les embauches en cours (deux doctorantes ayant débuté leur thèse début 2020 et un post-doctorant qui sera recruté d'ici peu).

8.3.2 Perspectives

Dans un futur proche, il me reste encore de nombreuses pistes à explorer sur l'analyse des images médicales, avec notamment les différents projets en cours, qu'il va falloir mener à bien. Petit à petit depuis quelques années, j'oriente mes recherches vers les méthodes d'apprentissage. Le deep learning a pris un essor incroyable au cours des dernières années, et notamment dans l'imagerie médicale. Je compte bien développer de plus en plus ces méthodes d'apprentissage profond dans les outils que je vais développer. De plus, comme indiqué précédemment, je souhaite aussi faire cohabiter la modélisation de la perception humaine et les méthodes d'apprentissage pour répondre à des besoins en imagerie médicale.

8.4 Perspectives générales

D'une manière plus globale, hormis les axes de recherche que j'ai pus emprunter jusqu'ici, mes thématiques de recherche vont progressivement être amenées à

s'orienter vers de nouvelles applications. À compter de Janvier 2020, j'ai rejoint le Laboratoire de Thermique et Énergie de Nantes ([LTeN](#)) (UMR 6607), j'ai ainsi pu commencer à m'imprégner des besoins en imagerie dans ces domaines nouveaux pour moi. Il y aura, à n'en pas douter, des collaborations captivantes à mettre en place. Mais il est encore un peu tôt pour en parler.

Sélection d'articles

Contents

9.1	Chromatic and wavefront aberrations: L-, M- and S-cone stimulation with typical and extreme retinal image quality	122
9.2	Lossless image compression via predictive coding of discrete Radon projections	136
9.3	Perceptual DFT watermarking with improved detection and robustness to geometrical distortions	149
9.4	Characterization of 3D Bifurcations in Micro-scan and MRA-TOF Images of Cerebral Vasculature for Prediction of Intra-Cranial Aneurysms	162

Dans cette section, je présente quatre articles représentatifs des différents travaux de recherche que j'ai pu conduire au cours de ces 20 dernières années. Le premier article [Autrusseau 2011b] porte sur la perception des couleurs et notamment sur la modélisation de l'optique de l'oeil dans le but de proposer un observateur standard et d'étudier l'impact des différentes aberrations chromatiques sur la vision des couleurs.

Dans le second article [Kingston 2008a], nous proposons une méthode de compression sans perte basée sur la transformation Mojette. L'un des principaux intérêts de cette méthode est la possibilité qu'elle offre de combiner à la fois redondance des données et taux de compression (sans perte). Effectivement, il est possible d'ajouter autant de projections que nécessaire, mais ces projections, si elles sont bien choisies peuvent conduire à des taux de compression intéressants.

Nous présentons dans l'article suivant [Urvoy 2014b] une nouvelle méthode de tatouage des images qui exploite des propriétés avancées du SVH afin de garantir le meilleur compromis possible entre l'invisibilité de la marque et sa robustesse. La méthode proposée a démontré des performances très intéressantes face à un type de distortion qui est souvent craint par la communauté du *watermarking* : les distortions géométriques et notamment la robustesse face au *print & scan*.

Finalement, dans le quatrième et dernier article [Nouri 2020], nous avons proposé une méthode de caractérisation des bifurcations d'un arbre vasculaire cérébral (acquisitions au micro-scanner pour des souris ou IRM-TOF pour des humains). Cette caractérisation complète des artères et des bifurcations nous a permis d'identifier les caractéristiques géométriques qui semblent avoir une influence majeure sur la formation des anévrismes.

9.1 Chromatic and wavefront aberrations : L-, M- and S-cone stimulation with typical and extreme retinal image quality



Contents lists available at SciVerse ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres



Chromatic and wavefront aberrations: L-, M- and S-cone stimulation with typical and extreme retinal image quality

Florent Autrusseau^{a,c}, Larry Thibos^d, Steven K. Shevell^{a,b,*}

^a Psychology, University of Chicago, 940 E. 57th Street, Chicago, IL 60637, USA

^b Ophthalmology & Visual Science and Institute for Mind & Biology, University of Chicago, 940 E. 57th Street, Chicago, IL 60637, USA

^c Ecole Polytechnique de l'Université de Nantes, IRCCyN, Rue Ch. Pauc, La Chantrerie, BP 50609, 44306 Nantes Cedex 03, France

^d School of Optometry, Indiana University, 800 E. Atwater Avenue, Bloomington, IN 47405, USA

ARTICLE INFO

Article history:

Received 23 May 2011

Received in revised form 11 August 2011

Available online 31 August 2011

Keywords:

Chromatic aberration

Wave aberrations

Optics

Retinal image

Spread light

Point spread function

ABSTRACT

The first physiological process influencing visual perception is the optics of the eye. The retinal image is affected by diffraction at the pupil and several kinds of optical imperfections. A model of the eye (Thibos & Bradley, 1999), which takes account of pupil aperture, chromatic aberration and wavefront aberrations, was used to determine wavelength-dependent point-spread functions, which can be convolved with any stimulus specified by its spectral distribution of light at each point. The resulting retinal spectral distribution of light was used to determine the spatial distribution of stimulation for each cone type (S, M and L). In addition, individual differences in retinal-image quality were assessed using a statistical model (Thibos, Bradley, & Hong, 2002) for population values of Zernike coefficients, which characterize imperfections of the eye's optics. The median and relatively extreme (5th and 95th percentile) modulation transfer functions (MTFs) for the S, M and L cones were determined for equal-energy-spectrum (EES) 'white' light. The typical MTF for S cones was more similar to the MTF for L and M cones after taking wavefront aberrations into account but even with aberrations the S-cone MTF typically was below the M- or L-cone MTF by a factor of at least 10 (one log unit). More generally, the model presented here provides a technique for estimating retinal image quality for the S, M and L cones for any stimulus presented to the eye. The model is applied to some informative examples.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

A single point of monochromatic light that enters the eye has a two-dimensional retinal image due to diffraction and imperfections of the eye's optics. The retinal image can be characterized by a point spread function (PSF), which varies with the wavelength of light. With broadband light, the retinal image is determined by decomposing the point of light into its spectral components; the PSF is applied separately at each wavelength (Barnden, 1974; Ravikumar, Thibos, & Bradley, 2008). The distribution of light on the retina is the superposition of the light distributions for each of the wavelengths. When an observer views a complete scene rather than a single point, each point in the scene is independently affected by the eye's optics; conceptually, the resulting retinal image at each wavelength is the superposition of the distribution of light from each point in the scene. Retinal image quality depends on both the PSF for each wavelength and the spatial and spectral distribution of the light in view.

Optical models of the eye have been sought for centuries for a variety of applications (Emsley, 1952; Smith, 1995). A relatively recent model (Thibos & Bradley, 1999) was used here to determine the spatial and spectral distribution of light on the retina; this distribution then was used to find the spatial distribution of light for each receptor cone class, L, M and S. This model has two advantages in comparison to Marimont and Wandell's (1994) well known model of retinal image quality. First, higher-order wavefront aberrations were considered explicitly here rather than as an implicit property of a wavelength-independent point spread function (Marimont & Wandell, 1994, p. 3116). Second, the Thibos and Bradley model-eye depends on specific parameters (Zernike coefficients) that characterize an individual eye, and these parameters have a known multivariate population distribution. The population distribution allowed estimates of individual differences in retinal image quality among people with normal corrected vision (Thibos, Bradley, & Hong, 2002). While a general comparison of the retinal image quality given by the Marimont and Wandell model versus the one used here is not possible because the models depend on different assumptions, results from the two models are compared in Section 5 using a typical eye from the population distribution given by Thibos, Bradley, and Hong (2002). The two models agree well in this special case (as discussed later).

* Corresponding author at: Institute for Mind & Biology, University of Chicago, 940 E. 57th Street, Chicago, IL 60637, USA.

E-mail address: shevell@uchicago.edu (S.K. Shevell).

The first part of this paper focuses on retinal image formation. The optical model is described and the calculated photoreceptor absorptions are explained. In the second part, the model is applied to broadband 'white' spectral stimuli to assess typical and extreme retinal contrast sensitivity in a normal population of human observers. The third part considers some specific cases to show how retinal image quality depends on particular features of a visual stimulus, and to demonstrate how the model may be applied to particular types of images.

2. Part 1: Retinal image model

The retinal image is determined by characterizing the eye's optics. Any optical system can be fully described mathematically over an isoplanatic area by its optical transfer function (Williams & Becklund, 1989) so calculating the retinal image involves determining the eye's optical transfer function (OTF) or the closely related point spread function (PSF). The PSF gives the retinal image of a monochromatic point source, taking account of the optics of the eye. The OTF is the Fourier transform of the intensity PSF.

A precise model of the eye includes various optical factors that affect the PSF at each wavelength. This section covers the main properties of the model eye used to determine the PSFs. Recall that the external stimulus pattern entering the eye is decomposed into multiple monochromatic stimulus patterns, and then the stimulus pattern at each wavelength is convolved with its wavelength-dependent PSF. This gives the retinal image $I(x, y, \lambda)$ at each wavelength λ . Superposition of these monochromatic retinal patterns yields the retinal spectral distribution of light for each spatial location (x, y) . These spectral distributions allow calculation of the spatial distribution $I_C(x, y)$ of cone excitation for each cone type C ($C = L, M$ or S) by applying the appropriate cone spectral sensitivity function $S_C(\lambda)$ as a weighting factor:

$$I_C(x, y) = \int S_C(\lambda) I(x, y, \lambda) d\lambda. \quad (1)$$

2.1. Optics of the human eye

The human eye has three main optical components that affect retinal image quality: the pupil, the cornea and the lens. The pupil diffracts light entering the eye, resulting in a PSF with a central point surrounded by concentric rings (Roorda, 2002; Williams & Hofer, 2004). The cornea accounts for most of the eye's refraction (about 43 diopters) while the lens, which refracts light after passing through the pupil, adds more than 20 diopters in a young adult (Roorda, 2002). The cornea and lens are the primary contributors to wave aberrations, which degrade retinal image quality. In the study here, all of these factors were combined into a reduced-eye optical model containing a pupil and a single refracting surface that is distorted from an optically perfect ellipse to exactly mimic the monochromatic aberrations of the whole eye (Ravikumar, Thibos, & Bradley, 2008). To account for individual variation, numerous such models were constructed using a statistical distribution of aberrations in a normal population of healthy adult eyes, as elaborated below. The monochromatic imaging capability of each model eye is summarized by its PSF.

Longitudinal chromatic aberration (LCA) is a consequence of refraction by a dispersive medium: light of different wavelengths is brought into focus at different distances. The focal distance increases with wavelength. LCA was included in the model by allowing the Zernike coefficient C_2^0 for defocus to vary with wavelength as prescribed by the Indiana Eye model of chromatic aberration (Thibos, Ye, Zhang, & Bradley, 1992). Transverse chromatic aberration was ignored because of its weak marginal influence on image

quality when LCA and wave aberrations are considered (Ravikumar, Thibos, & Bradley, 2008). Similarly, the slight effect of higher-order chromatic aberrations on image quality was not included (Nam, Rubinstein, & Thibos, 2010).

Retinal image quality depends on pupil size and the wavelength of light. The greatest loss of image quality from diffraction occurs with a small pupil and long wavelengths. On the other hand, the greatest loss from refractive elements occurs with a large pupil and short wavelengths. The sharpest retinal image, therefore, typically is at an intermediate pupil size near 3 mm, which balances the tradeoff between diffraction (worse at smaller pupil diameters) and the deleterious effects of wave aberrations (worse at larger diameters).

The best known wave aberrations caused by the eye's optics are defocus and astigmatism, which are ameliorated by standard corrective lenses. Higher-order wave aberrations (trefoil, coma, spherical, as well as other still higher-order aberrations) also reduce image quality (Packer & Williams, 2003). The imperfections captured by the higher-order wave aberrations include the irregularities in optical elements within the eye. The eye's lower- and higher-order wave aberrations can be modeled accurately using Zernike polynomials; the first 15 Zernike mode numbers were used here for the polynomials (Thibos & Bradley, 1999; Thibos, Hong, Bradley, & Cheng, 2002). Statistical sampling of Zernike aberration coefficients produced a random sample of 100 model eyes for analysis, each of which yielded monochromatic PSFs that were representative of human eyes (Thibos, 2009).

As mentioned above, each wavelength of light is affected differently by the optics of the eye so each wavelength has its own distinct PSF; we refer to the set of PSFs for all wavelengths in the visible spectrum as a hyperspectral PSF. Spectral sampling here was every 10 nm, a choice supported by the analysis of Ravikumar, Thibos, and Bradley (2008). A hyperspectral PSF captures two important aspects of retinal image quality. First, an object with a single broadband chromaticity, such as equal-energy-spectrum 'white', does not necessarily produce on the retina an image at only that chromaticity because some wavelengths are more strongly dispersed than others. Second, spectral mixtures that are visually indistinguishable in color as large homogenous patches (color metamers) may not match at other spatial frequencies because the wavelengths composing each metamer are unequally affected by optics (Marimont & Wandell, 1994; Poirson & Wandell, 1993).

An implication of the second point is that the influence of optics on the retinal image can be determined only from the full spectral distribution of light entering the eye. In general, a trichromatic description of the light – for example CIE X, Y, Z tristimulus values or the excitations of the three types of cones – is not sufficient. While some special cases can reduce the computational burden of multiple convolutions (for example, spectral homogeneity where every point in the image emits the same relative radiance spectrum, or when the full spectral distribution is uniquely determined by the trichromatic specification; Barnden, 1974; Ravikumar, Thibos, & Bradley, 2008), these conditions rarely occur in the natural world.

2.2. Cone quantal absorptions

Transduction of light at the photoreceptors establishes the neural responses that mediate vision. There are three classes of cone photoreceptor, labeled S, M and L, with peak sensitivity in the short-, middle- or long-wave part of the visible spectrum, respectively. The response of each cone type depends on the rate of quantal absorption. The relative spectral sensitivity of each type of cone is known (Smith & Pokorny, 1975) so the rate of quantal absorption for S, M or L cones can be calculated directly from the spectral energy distribution of light at each point on the retina. (The spacing

between adjacent cones in the retinal mosaic is ignored here.) This gives the spatial distribution of stimulation for each cone type.

The retinal spatial distribution of excitation for each cone type, therefore, requires the full spectral distribution of light at each point within the viewed scene. This implies, for example, that a scene on a color display with each pixel specified by its R ('red'), G ('green') and B ('blue') components must be transformed into a pixel-by-pixel spectral energy distribution in order to determine the retinal image. For example equal-energy-spectrum (EES) 'white' light, which by definition has all wavelengths in the visible spectrum at the identical energy, should not be used to calculate the retinal image of the *metamer* to EES on a color display because the display's weighted sum of the R, G and B components has a different physical spectral composition than true EES with all wavelengths at the same energy (compare dashed and solid lines, Fig. 1). Instead, calibration of the display is required to transform trichromatic coordinates of the display to the full spectral distribution of light at each point of the stimulus. This transformation is specific to each individual display; it differs among displays of the same manufacturer and model, and even for the same display over time.

3. Part 2: Typical and population extremes of retinal image quality

3.1. Point spread functions

The effects of pupil size and wave aberrations on the shape of the PSF can be determined separately. The PSFs shown in Fig. 2 ignore wavefront aberrations; only diffraction and longitudinal chromatic aberration (LCA) are included in the three PSFs, shown for wavelengths 400, 550 and 700 nm and with a 6 mm diameter pupil.

Zernike polynomials are used to model wavefront aberrations. A study conducted on 200 normal eyes was the basis for prior development of a multivariate statistical model for the population distribution of Zernike coefficients up to the 36th Zernike mode number. Fig. 3 shows the first 15 mean Zernike coefficients for a pupil diameter of 3.0, 4.5 or 6.0 mm from the work of Thibos, Bradley, and Hong (2002; some values in the figure were determined in that study though not explicitly presented in the paper). These results are consistent with other population studies (Castejon-Mochon, Lopez-Gil, Benito, & Artal, 2002; Porter, Guirao, Cox, & Williams, 2001; Salmon & van de Pol, 2006). In general, values of the Zernike coefficients vary with both the characteristics of a particular eye and with pupil size. As mentioned above, the slight variation expected in higher-order aberrations with wavelength

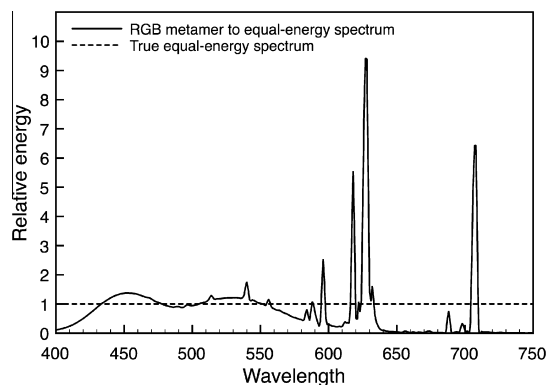


Fig. 1. Spectral power distribution of a true equal-energy-spectrum 'white' stimulus (dashed line) and from a typical color monitor displaying a light metameric to the equal-energy spectrum (solid line).

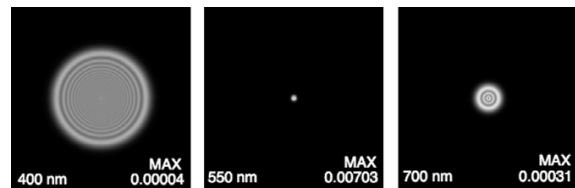


Fig. 2. Point spread functions due to only diffraction and longitudinal chromatic aberration (no wavefront aberrations) for monochromatic wavelengths 400, 550 and 700 nm. Focus wavelength is 570 nm; 6 mm diameter pupil. The height and width of each panel is 1° of visual angle. The sum of the volume under the PSF in each panel is 1.0 but, for visual clarity, the maximum in each plot is scaled to appear white; the actual maximum in each plot, which varies by more than 175:1, is shown in the bottom right of each panel.

and the variation of the prismatic terms responsible for transverse chromatic aberration were ignored (Nam, Rubinstein, & Thibos, 2010).

Point spread functions that include wave aberrations are shown in Fig. 4 for a 6 mm diameter pupil, which is close to the pupil size for 20–30 year-old observers at a luminance of about 100 cd/m² (Winn, Whitaker, Elliott, & Phillips, 1994). Analyses that follow also use a 6 mm diameter pupil. The mean Zernike coefficients from Fig. 3c were used for the PSFs in Fig. 4a but, as discussed below, these PSFs are not characteristic of a typical human eye. The PSFs in Fig. 4b are based on Zernike coefficients for a "standard observer" described later.

3.2. Modulation transfer functions and phase shifts

The modulation transfer function (MTF) is used to assess the eye's optical quality. Conceptually, the MTF can be determined for an EES light or any other spectral distribution by finding the contrast in the retinal image for sine waves at various spatial frequencies.

L-, M- and S-cone MTFs were determined here for 100 different independent random samples of Zernike coefficients taken from the population distribution given by Thibos, Bradley, and Hong (2002). The first 15 Zernike coefficients were used. The spectral distribution of light was EES, and luminance was varied sinusoidally in the horizontal direction (i.e., a vertically oriented sine wave). The wavelength in focus was fixed at 570 nm, which is the average wavelength that is optimally focused when observers report that white-light targets are best focused (Coe, 2009). With the PSFs determined for each visible wavelength, the retinal image at each wavelength was found by convolving the PSF for that wavelength with the visual stimulation at that wavelength. Repeating this for each wavelength in the stimulus and then taking the superposition of all wavelengths gave the spectral distribution of light at each point on the retina resulting from a particular visual stimulus (for example, an EES sine-wave grating varied in luminance at 10 cycles per degree). Then, at each point on the retina, the amount of each wavelength was weighted by the Smith and Pokorny (1975) cone spectral sensitivity function, separately for the L, M and S cones. This gave the relative quantal absorption for each cone type at that point on the retina. The MTF for each cone type then was determined by finding the retinal contrast as a function of the spatial frequency of the stimulus.

The complete set of one hundred MTFs, determined separately for the L, M, and S cones, is shown in Fig. 5 (thin gray lines). The median (thick black line), 5th percentile (dotted line) and 95th percentile (dashed line) values are also shown. The figure reveals the large individual differences in the MTFs estimated to occur within a normal population. The S-cone median MTF is well below the L- and M-cone median curves, though the population variation for each cone's MTF is large.

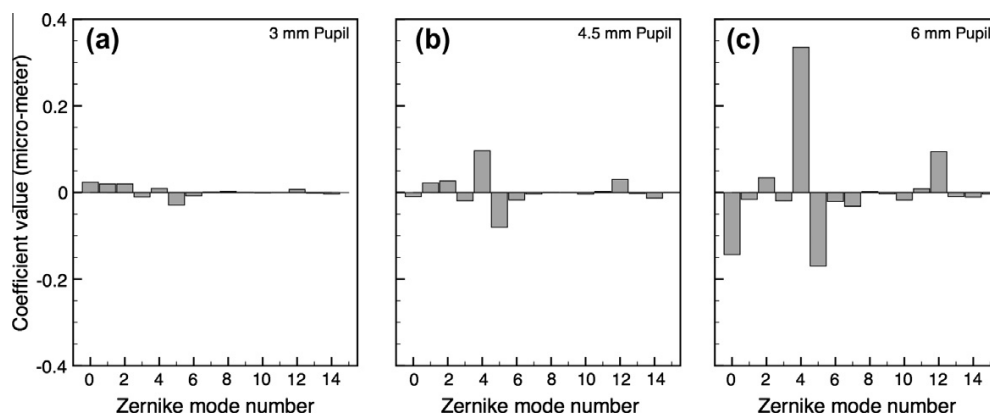


Fig. 3. Mean Zernike coefficients for the first 15 Zernike modes, for pupil diameter (a) 3 mm, (b) 4.5 mm or (c) 6 mm (from the model of Thibos, Bradley, & Hong, 2002).

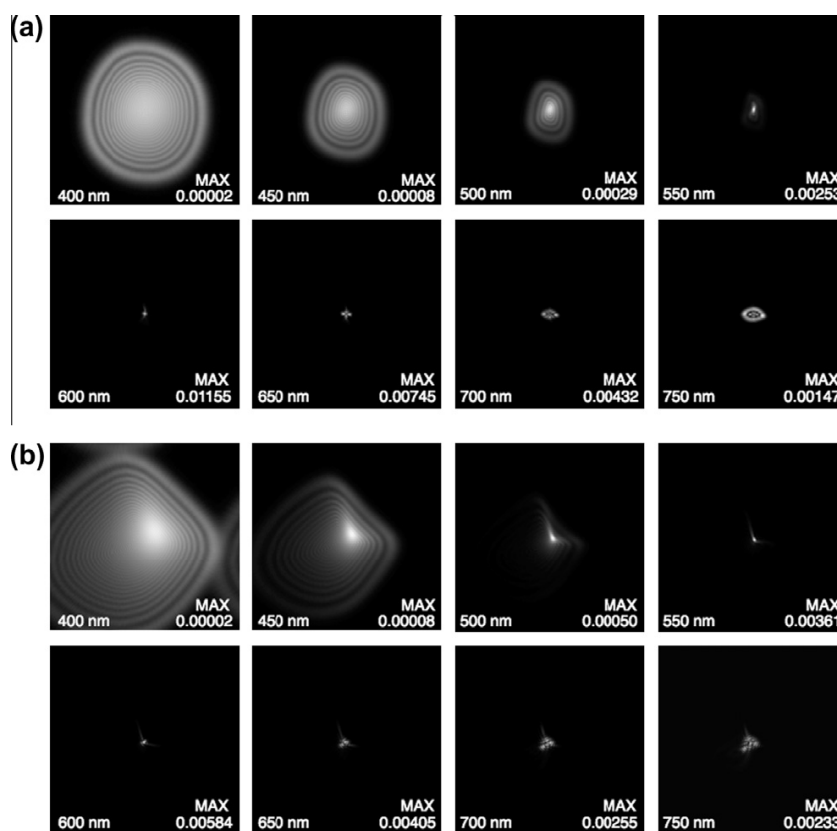


Fig. 4. As Fig. 2 but for point spread functions due to diffraction, longitudinal chromatic aberration, and wavefront aberrations, for monochromatic wavelengths 400, 450, 500, 550, 600, 650, 700 and 750 nm. The maximum of each plot is set to be white; the actual maximum, which varies by over 500:1, is shown in the bottom right of each panel. (a) Based on mean population Zernike coefficients from Fig. 3c. (b) Based on Zernike coefficients for a "standard observer" (see text and Table 1).

One sampled set of Zernike coefficients gave L-, M- and S-cone MTFs that were close to the median values (see lines connecting symbols, Fig. 5). This set of Zernike coefficients (Table 1) was used to define a "standard observer" for a 6 mm pupil; the MTFs from this set were consistent with an observer with about half the sample having better retinal cone contrast and half having worse contrast. While the standard observer's S-cone MTF is slightly better than the median (at worst the 65 percentile instead of the median's 50

percentile, at 10 cycles per degree), this set of sampled Zernike coefficients was selected for the standard observer because the set of coefficients gave values close to both the median MTFs (Fig. 5) and the median magnitudes of image displacement (computed as the ratio of absolute value of phase shift to spatial frequency), for all three types of cone. The absolute value of image displacements from the 100 random independent samples of Zernike coefficients is shown in Fig. 6. The median absolute values were close to the image

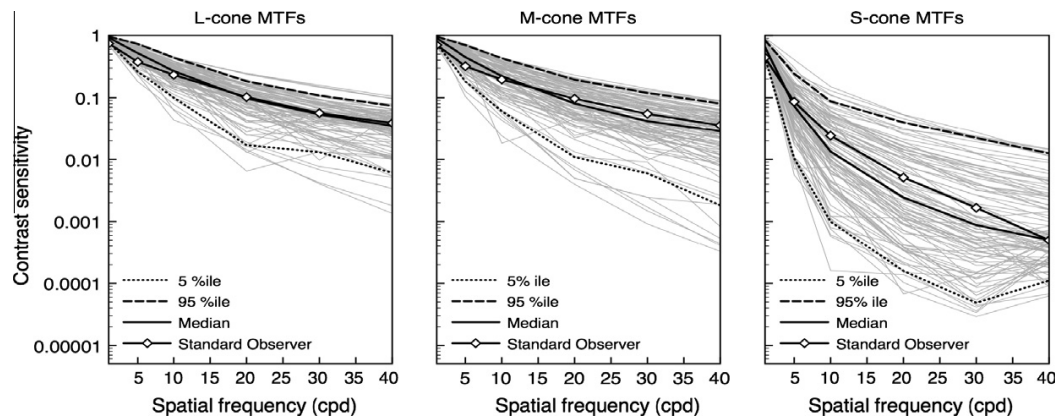


Fig. 5. Modulation transfer functions for 100 randomly sampled sets of Zernike coefficients for pupil diameter 6 mm (gray lines). The median contrast sensitivity among the 100 values at each spatial frequency is shown by the solid black line; the 5th and 95th percentile contrast sensitivity is shown by dotted and dashed lines, respectively. The solid black lines connecting the symbols are MTFs for the “standard observer” (see text). (Left) L-cone MTFs. (Middle) M-cone MTFs. (Right) S-cone MTFs.

Table 1
Zernike coefficients for “standard observer” with 6 mm pupil.

Zernike mode number	Coefficient value
0	0.3243
1	−0.5796
2	0.7083
3	−0.1115
4	0.8638
5	0.0148
6	−0.1116
7	−0.1644
8	0.3053
9	−0.0673
10	−0.0273
11	−0.0115
12	0.2394
13	0.0228
14	0.0784

displacements for the standard observer (compare solid lines, with and without symbols, Fig. 6). Of course, this particular set of Zernike coefficients is not unique in terms of giving MTFs and image displacements that are near the medians for all three cone types.

The significance of spatial phase shifts in the image plane should not be ignored when wave aberrations are introduced. A circularly symmetric PSF due only to diffraction and longitudinal chromatic aberration (i.e., no wave aberrations, as in Fig. 2) causes either a zero or 180° phase shift (i.e., phase reversal) in the stimulus. Wave aberrations, however, can shift phase in the image plane by arbitrary amounts. A PSF reflects both the MTF and phase shifts so a full description of the retinal image requires an accurate representation of both. While several other sampled sets of Zernike coefficients gave MTFs close to the median for the three cone types (Fig. 5), the standard observer was selected to closely represent both the median MTFs (Fig. 5) and median image displacements (Fig. 6).

Note the larger image displacements for S than for M and L cones in the population, for both the median and extreme (95th percentile) values (Fig. 6).

To assess the reliability of the median and extreme MTFs, the 100 sampled sets of Zernike coefficients were randomly split into two groups of 50 samples each. Then the median, 5th percentile and 95th percentile MTFs for the L-, M- and S-cone were compared for the two halves. The comparison of the two groups showed good agreement (Fig. 7).

A straightforward way to quantify the differences in MTFs among the L, M and S cones is to plot relative M/L and S/L contrast sensitivity for the median, 5th percentile and 95th percentile MTFs (Fig. 8). This shows that median contrast sensitivity is similar for the M and L cones. Median S-cone contrast sensitivity is a log unit or more lower than L, except at the lowest spatial frequencies.

3.3. Zernike coefficients: population means versus values for standard observer

There are substantial differences between the Zernike coefficients for the “standard observer” (Table 1, and gray bars in Fig. 9) and the mean Zernike coefficients for the population of human observers (diamonds, Fig. 9; replotted from Fig. 3c). For most purposes, calculations based on Zernike-coefficient population means are not representative of a typical observer’s image quality (Thibos et al., 2002). A way to visualize the significance of the difference between the two sets of Zernike coefficients is to compare the wavelength-dependent PSFs based on population-mean coefficients with the PSFs for the “standard observer” (Fig. 4a and b, respectively). The PSFs for the standard observer are less circularly symmetric, as expected for Zernike coefficients that are farther from zero.

Even more significant is the difference in the MTFs based on mean Zernike coefficients versus the Zernike coefficients for the standard observer. Consider again the median, 5th percentile and 95th percentile MTFs for the 100 randomly sampled sets of Zernike coefficients for pupil diameter 6 mm (Fig. 10, replotted from Fig. 5), and compare them to (i) MTFs for the standard observer and (ii) MTFs implied by the mean Zernike coefficients (open and solid symbols, respectively, in Fig. 10). The MTFs for the standard observer fall close to the medians, of course, because of the criteria used to define the standard observer. The L- and M-cone MTFs based on population-mean Zernike coefficients are near or above the 95th percentile MTFs; they show that mean Zernike coefficients give estimates of image quality appropriate for only a very small fraction of the population with the best optics.

4. Part 3: Estimating specific retinal images: issues and examples

The first two parts of the paper considered retinal image quality and how it affects stimulation of the L, M and S cones for typical

and relatively extreme normal human observers, assuming a 'white' stimulus with an EES spectral distribution. MTFs were derived using vertical sine-wave gratings. More generally, the model

developed above may be applied to any visual stimulus. In this part, retinal image quality is determined for other types of stimuli, with a focus on some informative cases.

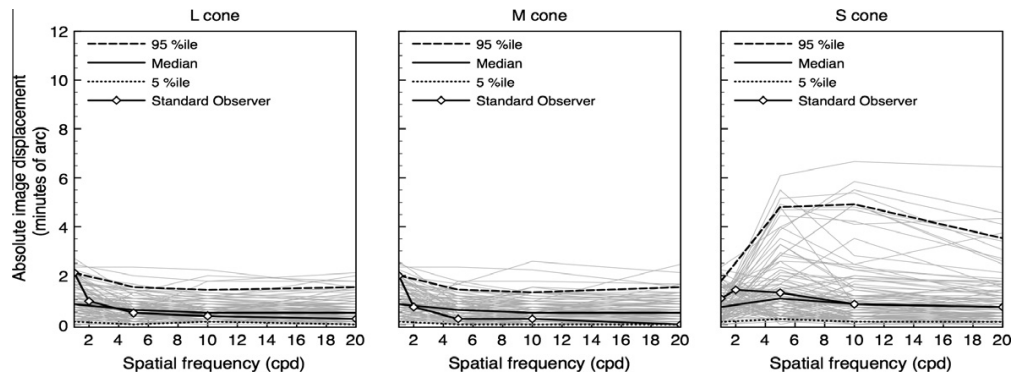


Fig. 6. Image displacement functions for 100 randomly sampled sets of Zernike coefficients (6 mm diameter pupil). Ordinate values are the ratio of $|\text{phase shift}|/\text{spatial frequency}$. The median, 5th percentile and 95th percentile shifts are indicated by solid, dotted and dashed black lines, respectively. Solid black lines connecting the symbols are image displacements for the standard observer whose Zernike coefficients are given in Table 1. (Left) L cone. (Middle) M cone. (Right) S cone.

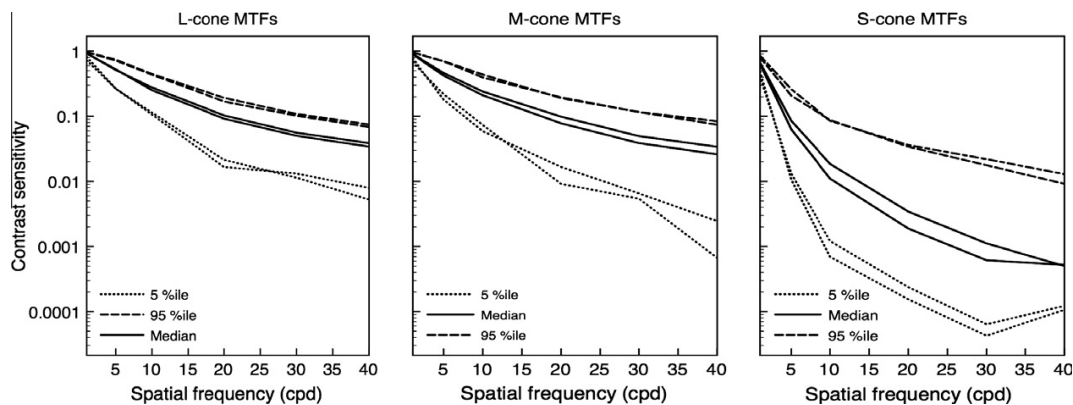


Fig. 7. Modulation transfer functions for split halves of the 100 randomly sampled sets of Zernike coefficients for pupil diameter 6 mm. Each split half had 50 sampled sets of Zernike coefficients. The median, 5th percentile and 95th percentile contrast sensitivity for each split half are shown as a function of spatial frequency by the solid, dotted and dashed lines, respectively. (Left) L-cone MTFs. (Middle) M-cone MTFs. (Right) S-cone MTFs.

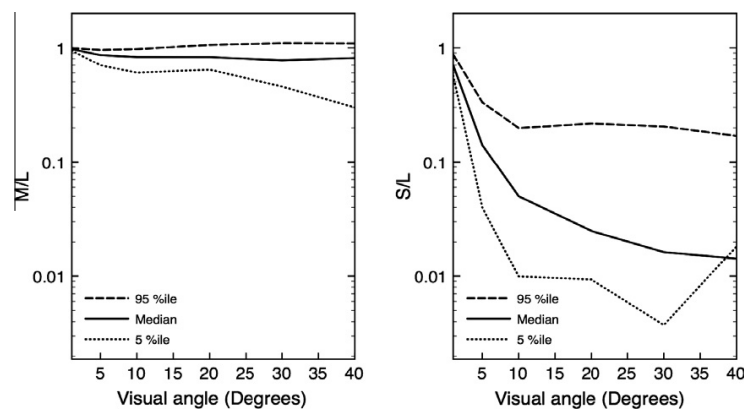


Fig. 8. The ratio of contrast sensitivity as a function of spatial frequency, based on 100 randomly sampled sets of Zernike coefficients, for (left) M cones relative to L cones and (right) S cones relative to L cones. The median, 5th percentile and 95th percentile are shown in each panel by the solid, dotted and dashed line, respectively. Pupil diameter is 6 mm.

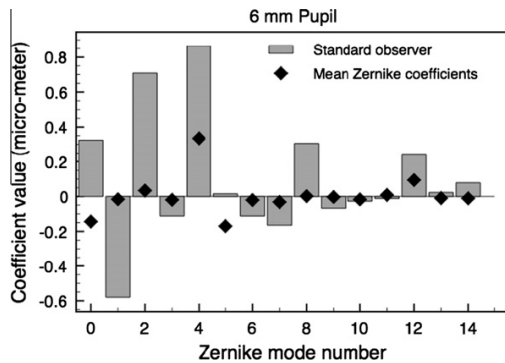


Fig. 9. The first 15 Zernike coefficients for the “standard observer” (6 mm diameter pupil, gray bars), whose MTFs and image displacements are similar to the median values from the 100 randomly sampled sets of Zernike coefficients. Mean Zernike coefficients are also shown (diamonds, replotted from Fig. 3c).

4.1. Equal-energy-spectrum ‘white’ and its CRT metamer

As discussed above, the eye’s optical imperfections can be quantified by wavelength-dependent PSFs (Fig. 4), the effects of which are determined by decomposing the visual stimulus into its monochromatic components, applying each wavelength-dependent PSF, and then combining each monochromatic spatial light distribution by superposition. Thus, the effect of the eye’s optics on the retinal spatial and spectral distribution of light depends on the spectral distribution of the stimulus.

The significance of the underlying spectral distribution of the stimulus can be missed by wrongly applying the color-vision principle of trichromacy, which holds that the color of any spectral distribution of light can be perfectly matched by a mixture of just three wavelengths (for example, 440 nm, 540 nm and 660 nm, which alone appear violet, yellowish-green and red, respectively). The basis for trichromacy is the three types of cone photoreceptors: L, M and S. Any given spectral distribution is perceived to be an exact match to an entirely different spectral distribution so long as both distributions have the identical rate of quantal absorption by the L, M and S cones; unequal spectral distributions that satisfy this matching criterion are called *metamers*. An example of two spectral distributions that appear identical is in Fig. 1. The flat energy distribution as a function of wavelength for EES ‘white’

(dashed line) results in the same stimulation of the L, M and S cones as the irregularly shaped spectral distribution (solid line), which is typical of a CRT display set to exactly match EES ‘white’.

Two metameric spectral distributions of light, however, may not have exactly the same appearance when presented as complex patterns of light. The reason is that the underlying wavelength distributions that deliver identical rates of L-, M- and S-cone quantal absorptions for large uniform patches (as used for color matching) may be unequally affected by the eye’s optics, because the effects of optics vary according to the wavelengths composing the two stimuli. For example, MTFs for the L and M cones for the standard observer are similar for true EES light and for its CRT metamer but contrast sensitivity for S cones is lower for true EES light compared to the metameric CRT chromaticity (compare open and filled symbols, Fig. 11). In general, retinal images of stimulus patterns composed of metameric lights may not be assumed to be equal.

4.2. Asymmetric retinal-image distortion and stimulus orientation

In the second part of the paper, MTFs were determined from luminance modulation of EES light varied in the horizontal direction (i.e., vertically oriented gratings). In the absence of wave aberrations, any orientation of the stimulus gives the same results because PSFs are circular (Fig. 2). PSFs that include effects of wave aberrations, however, are not circular in shape (Fig. 4) so the retinal image of even a uniform bar of light can (i) be asymmetric with respect to the center of the uniform bar and (ii) depend on stimulus orientation.

Asymmetric distortion from wave aberrations can be illustrated with a square-wave EES grating of two cycles per degree with 100% luminance contrast (solid lines, Fig. 12); in the figure, the one-dimensional square-wave is a profile of a two-dimensional vertical square-wave grating. Retinal-image profiles at monochromatic wavelengths between 400 and 750 nm are shown for a 6 mm pupil (Fig. 12). In the top panel, only diffraction and LCA are considered (no wave aberrations). The wavelength-dependent distortion of the square-wave stimulus is symmetric around the center of the stimulus peaks and troughs. By comparison, when wave aberrations for the standard observer are included (middle panel, Fig. 12), there is clear left–right asymmetry; moreover, the shape of the asymmetry is wavelength dependent. The asymmetry carries over to the spatial profile of stimulation for the L, M and S cones (bottom panel, Fig. 12).

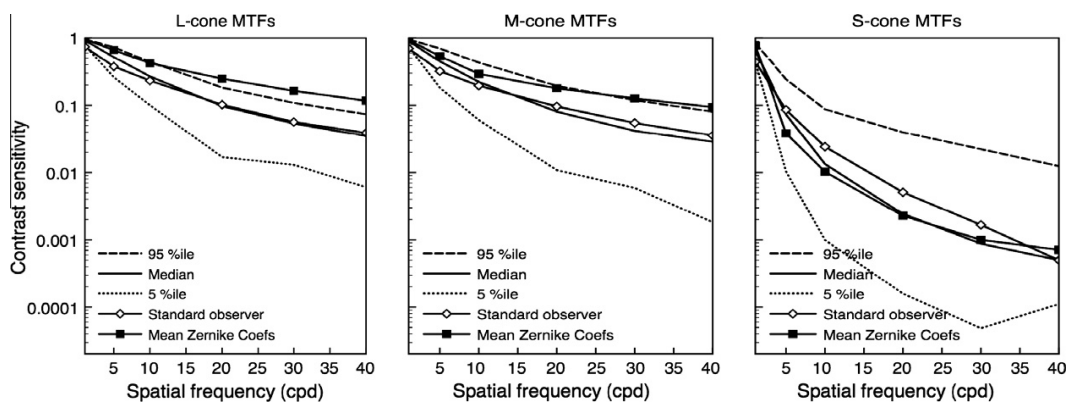


Fig. 10. Median, 5th and 95th percentile modulation transfer functions based on 100 randomly sampled sets of Zernike coefficients for pupil diameter of 6 mm (replotted from Fig. 5), compared to MTFs for the “standard observer” (solid lines connecting open symbols) and to MTFs based on the mean Zernike coefficients in Fig. 3c (solid lines connecting solid symbols). (Left) L-cone MTFs. (Middle) M-cone MTFs. (Right) S-cone MTFs.

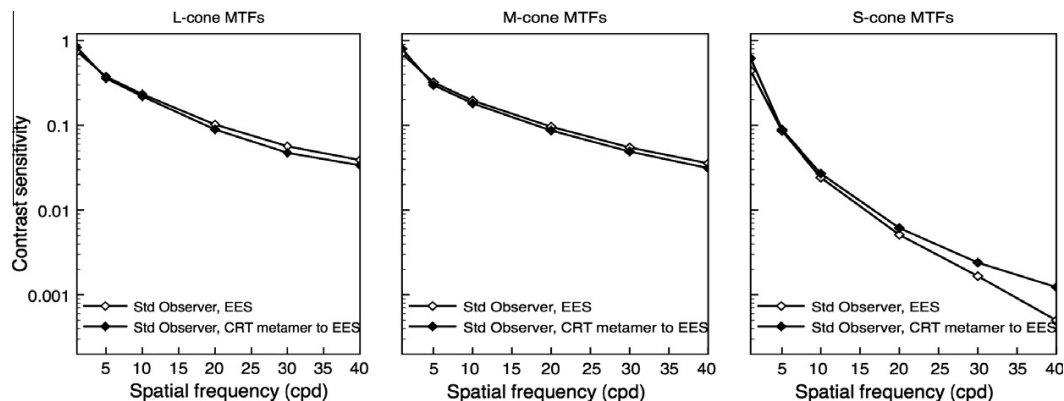


Fig. 11. Modulation transfer functions for the “standard observer” for vertical gratings composed of (i) EES ‘white’ light (open symbols) or (ii) the typical spectral distribution from a CRT display set to be metameric to EES ‘white’ (filled symbols). (Left) L-cone MTFs. (Middle) M-cone MTFs. (Right) S-cone MTFs.

The influence of stimulus orientation can be seen with the same EES, two cycle per degree square-wave grating. For the standard observer (6 mm pupil), the retinal-image profiles for the L, M and S cones are different for a vertically compared to horizontally oriented grating (see dotted and dashed lines, Fig. 13). In general, when wave aberrations are considered, the retinal image depends on the overall stimulus orientation as well as the spectral and spatial distribution of light.

4.3. The retinal image of an EES ‘white’ E

The loss in retinal image quality caused by the eye’s optics can be visualized by the retinal image of the letter “E” of width 0.5° visual angle (cf. Williams & Hofer, 2004).

An E presented to the eye as EES ‘white’ light is shown in the far left panel of Fig. 14. The retinal image of the E, assuming only diffraction and LCA (no wave aberrations), is in the three top panels on the right, separately for stimulation of the L, M and S cones. A modest drop in image quality is apparent for the L-cone and M-cone retinal images, while the S-cone retinal image is far worse than for L or M. Adding wave aberrations of the standard observer (bottom three panels at right) further degrades the L-cone and M-cone retinal images but not the S-cone image, which actually appears a bit better than the S-cone image without wave aberrations. Note, however, that the S-cone retinal image with wave aberrations (bottom right panel) is poorer than the L- and M-cone images that include wave aberrations (compare three bottom panels at right, Fig. 14). The effect of wave aberrations on the retinal image for each type of cone is taken up again in Section 5, where previous studies are considered.

4.4. Retinal image quality for typical R, G and B components of a color video display

The model developed here may be applied to a visual stimulus with any spectral distribution. Three spectra of practical interest are the R, G and B components of a color video display. The analysis below is for the R, G and B guns of a Sony GDM-F520 CRT, a display used in many vision laboratories. (The exact R, G and B spectra vary, of course, from one video display to another.)

For any known spectral distribution of light, the MTF can be determined for the L, M or S cones of the standard observer. The R, G and B spectra of the CRT were measured using a PhotoResearch PR-650 spectroradiometer. For each type of cone (L, M or

S), the R, G and B phosphor MTFs are plotted together in a single panel of Fig. 15. There are two main points. First, the MTFs for R and G are higher than for B, for every type of cone. Second, the MTFs for the R and G components are similar to each other for the L and M cones but not for the S cones, for which G has substantially higher contrast sensitivity than R. For S cones, in fact, contrast sensitivity for R is nearly as poor as for B.

5. Discussion

The approach here used a comprehensive model of the eye (Thibos & Bradley, 1999) that incorporates pupil aperture, longitudinal chromatic aberration and wavefront aberrations. The model gives the spatial and spectral distribution of light on the retina from any image presented to the eye. Such distributions were used to determine typical and extreme (5th and 95th percentile) MTFs for each type of cone (L, M and S), by incorporating known individual differences for the eye’s optics (Thibos, Bradley, & Hong, 2002).

A large number of random samples from the statistical distribution for individual differences also revealed median levels of contrast sensitivity, from which a “standard observer” was defined with retinal image quality near the central tendency of the normal human population. Typical image quality is better represented by this standard observer than by calculations based on population-mean Zernike coefficients because mean values tend to cancel out positive and negative coefficients and thus underestimate the loss of image quality, which is largely dependent on the coefficients’ magnitudes regardless of sign.

Other investigators have examined related questions about retinal image quality for spectrally broadband images. Results from the approach here are compared below to two well known studies.

5.1. A comparison to the Marimont and Wandell model

As mentioned in the Introduction, Marimont and Wandell (1994) consider a similar problem but with a model that has two important differences. First, the model here explicitly considers higher-order wavefront aberrations. Second, the model here can be used to estimate individual differences based on the statistical population distribution for characteristics of the eye’s optics (Thibos, Bradley, & Hong, 2002). Nonetheless, in special cases the two models should give similar results. In particular, when only diffraction and longitudinal chromatic aberration are included in the model here (no wave aberrations), the modeled retinal image

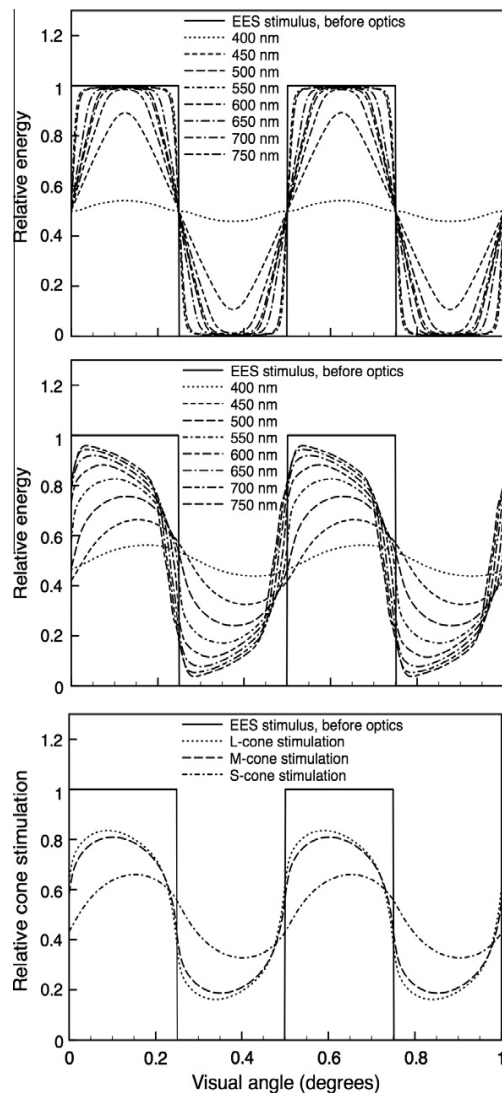


Fig. 12. Retinal-image profiles for an EES, two cycle per degree, 100% contrast square-wave grating. The solid line in each panel shows the square-wave stimulus. (Top and middle) Profiles for wavelengths 400, 450, 500, 550, 600, 650, 700 and 750 nm (6 mm pupil, 570 nm focus wavelength), with effects of only diffraction and longitudinal chromatic aberration (no wave aberrations, top) or with added effects of wave aberrations of the standard observer (middle). (Bottom) Retinal-image profiles for L-, M- and S-cone stimulation with wave aberrations of the standard observer.

should be somewhat better than the results from the Marimont and Wandell model, which incorporates an approximation for higher-order aberrations. Also, the full model here with wave aberrations of the standard observer (Table 1) should give a retinal image similar to but different in shape than the Marimont and Wandell model, because introducing explicit wave aberrations results in asymmetries in the PSFs (Fig. 4) and thus in the retinal light profile (e.g., middle panel of Fig. 12).

Retinal light profiles for an inhomogeneous chromatic image (Fig. 16) were determined from the model here, both without and with wave aberrations (dotted and dash-dot lines, respectively), and from the Marimont and Wandell model (dashed line). For this test case, the stimulus was composed of a square-wave

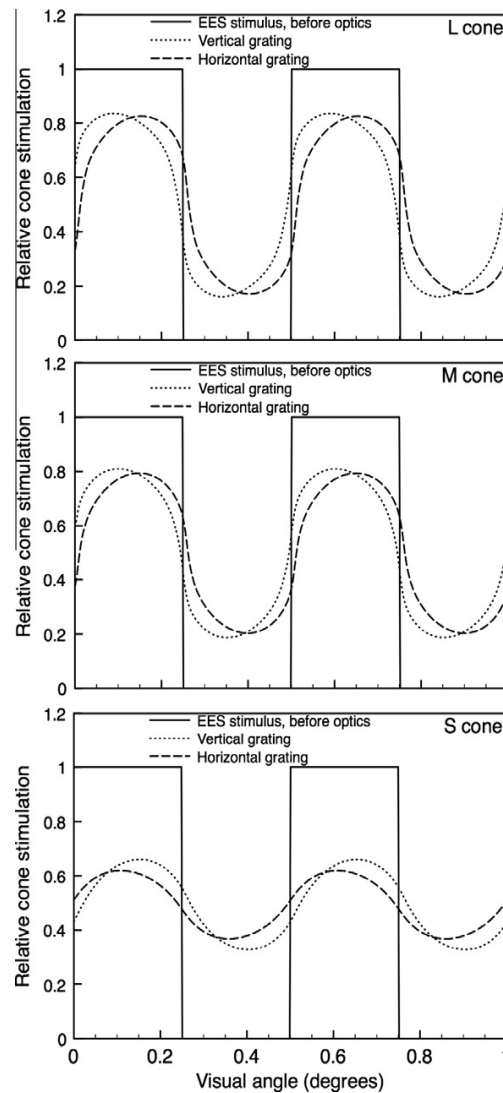


Fig. 13. Retinal-image profiles for an EES, two cycle per degree, 100% contrast square-wave grating (6 mm pupil, 570 nm focus wavelength) including effects of wave aberrations for the standard observer. The solid line in each panel shows the square-wave stimulus profile. The dotted and dashed lines show the profiles for a vertically or horizontally oriented grating, respectively. (Top) Relative L-cone stimulation; (Middle) relative M-cone stimulation; (Bottom) relative S-cone stimulation.

vertical stripe of width 7.3 min of arc, centered within a 1° wide surround. Both the stripe and surround had the same chromaticity (0.665) in the $l = L/(L + M)$ direction of the MacLeod and Boynton (1979) l, s cone-based coordinate system but differed in $s = S/(L + M)$: the stripe had an s chromaticity of 2.50 so when viewed alone would appear decidedly bluish, while the surround had an s chromaticity of 0.99 so was virtually metameric to EES 'white'. The stripe and surround were equal in luminance. The full spectral distributions of the chromaticities were assumed to be from the video display considered in Fig. 15.

In Fig. 16, the vertical axis is the light profile for the s chromaticity, derived from the retinal image for S, M and L and then applying the definition of $s = S/(L + M)$. The square wave (solid line) is

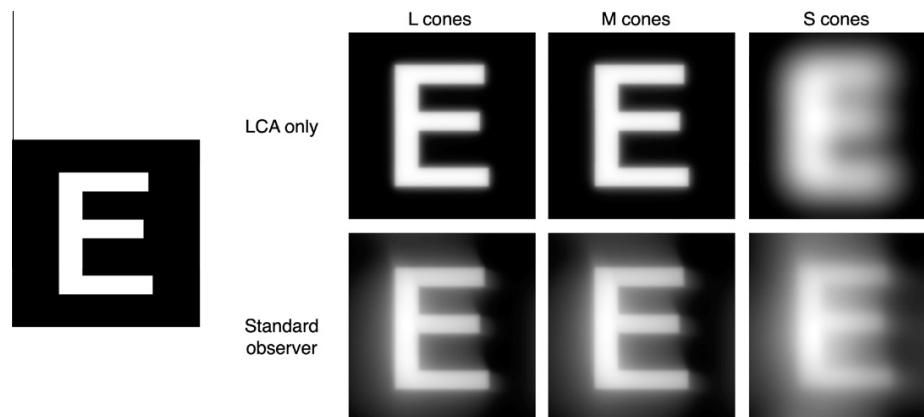


Fig. 14. Retinal images of an EES letter “E” subtending in width 0.5° visual angle. Each panel is 1° square and rendered relative to the maximum light level in the stimulus before entering the eye (shown at the far left). Pupil diameter is 6 mm. (Right, top row) The retinal image for the L, M and S cones assuming only diffraction and LCA (no wave aberrations). (Right, bottom row) The retinal image for the L, M and S cones with the added effects of wave aberrations of the standard observer.

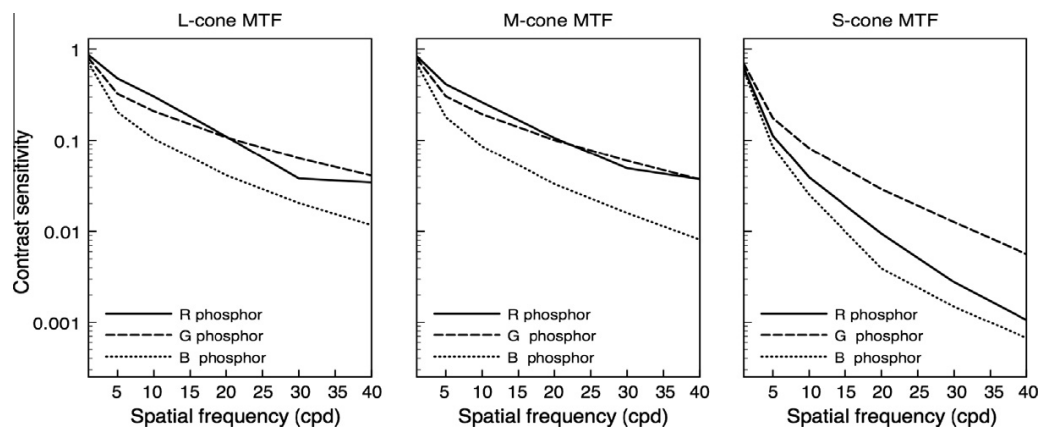


Fig. 15. Modulation transfer functions for the standard observer (6 mm diameter pupil) for vertical gratings composed of a typical video display’s R phosphor, G phosphor or B phosphor. The MTFs for the L, M and S cones are shown separately in the left, middle and right panels, respectively. Focus wavelength 570 nm.

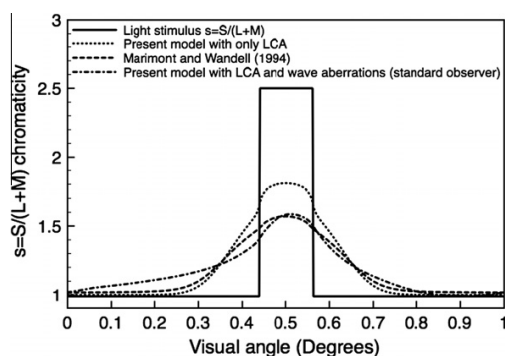


Fig. 16. Comparison between the retinal light profiles for $s = S/(L+M)$ from the model used here (focus wavelength 570 nm), without wave aberrations (dotted line) or with the wave aberrations of the standard observer (dash-dot line), and from the line spread function of Marimont and Wandell (1994) (dashed line). The corneal stimulus (solid line) is a vertical stripe of width 7.3 min within a 1° wide surround. The stripe and surround have different chromaticities (see text). Pupil diameter is 6 mm.

the image at the cornea. A technical point for calculation of s is selection of units for S, M and L; previous results for S-, M- and L-cone MTFs and light profiles were not dependent on units because plotted quantities were relative values. For Fig. 16, L and M were normalized so that $L + M$ gave luminance (as for the Smith and Pokorny (1975) cone fundamentals), and S was scaled so that $s = 1.0$ for the spectrally flat distribution of EES ‘white’.

As expected, the model here without wave aberrations specified a light profile for s that that was sharper (better retinal image quality) than either the Marimont and Wandell model or the model here that includes wave aberrations. The light profiles from the Marimont and Wandell model and from the model here with wave aberrations were similar but only the latter captured the asymmetry in the distribution of light implicit in the asymmetric PSFs (as seen in Fig. 4). Note that the light profile with higher-order wave aberrations (dash-dot line) depends on the orientation of the stripe in the stimulus. Overall, the comparison in this special case shows that the model used in this paper with the optics of the standard observer (Table 1) gives results in accord with those from the Marimont and Wandell model.

5.2. How do wave aberrations affect relative L-, M- and S-cone contrast sensitivity?

Longitudinal chromatic aberration (LCA) reduces the retinal image quality of broadband spectral light because the focal distance from refraction increases with wavelength. When light of some wavelength is in perfect focus, light at other wavelengths is blurred. MTFs that take account of both LCA and diffraction (but not wave aberrations) quantify the substantial reduction in retinal image quality that they cause for broadband EES 'white' light (left panel Fig. 17, lines without symbols).

Introducing wave aberrations also reduces image quality, compared to diffraction alone. The retinal image from diffraction, LCA and wave aberrations together – all of which, of course, affect natural viewing – might be expected to be worse than from only diffraction and LCA; somewhat surprisingly, however, adding the effects of wave aberrations to those of diffraction and LCA can improve image quality in some cases (McLellan, Marcos, Prieto, & Burns, 2002). The reason is that wave aberrations reduce the influence of wavelength on contrast sensitivity; although wave aberrations reduce contrast sensitivity at the wavelength of focus, they can increase sensitivity at other wavelengths that are badly defocused by LCA. The result is that wave aberrations in some cases "counteract retinal image blur from LCA (McLellan, Prieto, Marcos, & Burns, 2006, p. 3009)". For example, adding the standard observer's wave aberrations to diffraction and LCA reduces L- and M-cone contrast sensitivity but improves S-cone sensitivity (left panel Fig. 17, lines connecting symbols).

Comparing MTFs with versus without wave aberrations shows that introducing wave aberrations can make contrast sensitivity more similar for the three types of cone. For example at 20 cpd (see left panel, Fig. 17), with the wave aberrations of the standard observer the contrast sensitivity for L cones (and M cones) is about 20 times better than for S cones; without wave aberrations (only LCA and diffraction), L-cone (and M-cone) sensitivity is about 200 times better than S. These results confirm that wave aberrations can reduce the difference in contrast sensitivity among the three types of cone.

A related issue is whether wave aberrations fully (or nearly) eliminate the differences in contrast sensitivity for the L, M and S cones. McLellan et al. (2002) report that they do: the MTFs for all three cone types are nearly identical to each other, for each of their three observers (their Fig. 3b–d, p.175). To consider this question,

the random sample of 100 sets of Zernike coefficients was searched for sets with the S-cone MTF near or above the L- and M-cone MTFs. For example, one sample (#14; middle panel, Fig. 17) had close L-, M- and S-cone MTFs (cf. the standard observer's MTFs in left panel of Fig. 17); the difference between L- and M-cone contrast sensitivity was about the same as between L and S sensitivity. At 20 cpd for #14, the contrast sensitivity for L cones was less than twice that of S cones. This set of MTFs is similar to the set for an observer reported by McLellan et al. (2002; their Fig. 3b). Another sample at 20 cpd (#92; right panel, Fig. 17) also had L-cone sensitivity less than two times higher than S, and at 40 cpd had S sensitivity fractionally better than L or M sensitivity (reminiscent of the MTFs for the observer in Fig. 3c of McLellan et al. (2002)). In sum, a few of the 100 samples were consistent with the view that wave aberrations can (nearly) eliminate contrast sensitivity differences among the three types of cone.

Recall, however, that the MTFs in the middle and right panels of Fig. 17 were selected from the sample of 100 sets because contrast sensitivity was similar for the L, M and S cones. To consider whether wave aberrations typically eliminate contrast-sensitivity differences among the three types of cone, the MTF for the S cones was compared to the MTFs for L and M cones at 10 and at 20 cpd for all 100 randomly sampled sets of Zernike coefficients. Because L- and M-cone MTFs tended to be similar, average L and M contrast sensitivity was compared to S sensitivity (that is, Average (L-sensitivity, M-sensitivity) relative to S-sensitivity). This sensitivity ratio for the 100 samples is plotted in a histogram in Fig. 18 (left panel, 10 cpd; right panel, 20 cpd). A value of zero indicates equal contrast sensitivity for S cones compared to the average for L and M; a positive [negative] value indicates L and M cones had higher [lower] sensitivity than S. The horizontal axis in Fig. 18 is a log scale so, for example, a value of +2.0 indicates lower S contrast sensitivity by a factor of 100.

Overall, S-cone contrast sensitivity nearly always was lower than L and M sensitivity (in 99% of cases at 10 cpd, and 97% of cases at 20 cpd). In the vast majority of cases, S contrast sensitivity was at least 10 times lower ($>+1.0$ on the log scale) than L and M sensitivity (73% of cases at 10 cpd, 77% of cases at 20 cpd). While S-cone sensitivity that is lower by a factor of 10 is a much smaller sensitivity difference than a factor of more than 200 for retinal image quality based on only diffraction and LCA (see arrows, Fig. 18), the results show that wave aberrations very seldom bring L-, M- and S-cone contrast sensitivity to comparable levels. The simulated

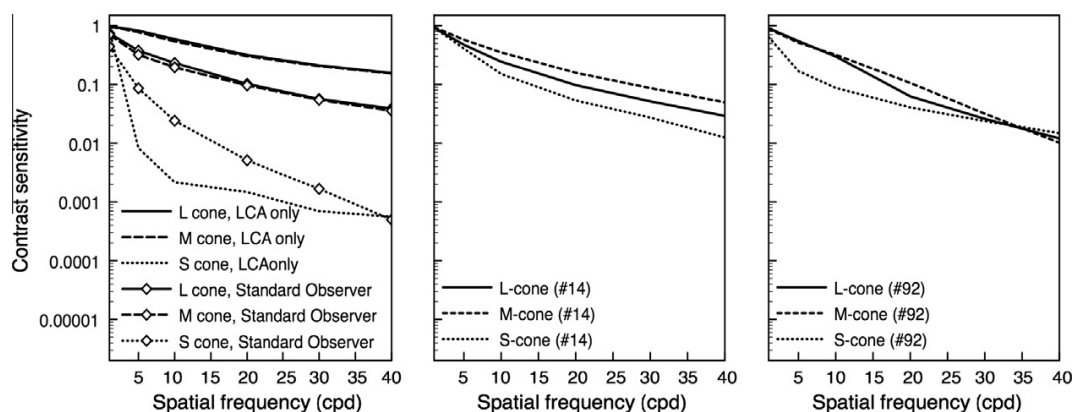


Fig. 17. (Left) The MTFs for each type of cone (L, M and S) with only diffraction and LCA (no wave aberrations, lines without symbols) and with the added wave aberrations of the standard observer (lines with symbols). (Middle and right) The L-, M- and S-cone MTFs for two (#14 and #92) of the 100 randomly sampled sets of Zernike coefficients, for which MTFs with wave aberrations were unusually similar for all three types of cone. For all panels, the spectral light distribution is EES, pupil diameter is 6 mm and focus wavelength is 570 nm.

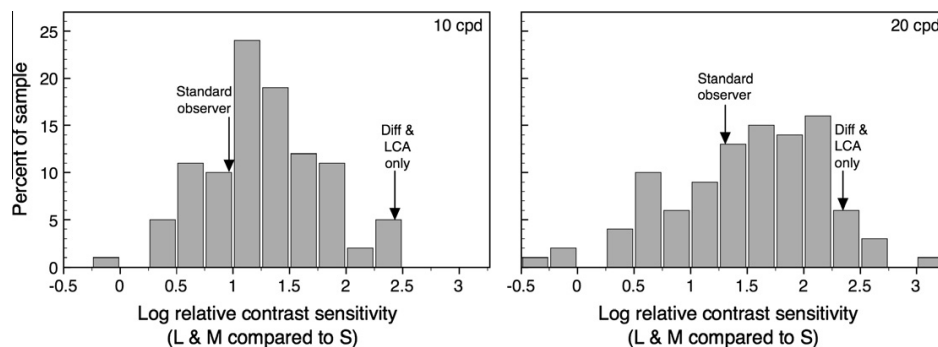


Fig. 18. Histogram of the contrast sensitivity ratio for the average for L and M cones compared to S cones, for the 100 randomly sampled sets of Zernike coefficients. Note the log scale on the horizontal axis. A positive value indicates lower S than L and M sensitivity. Pupil diameter is 6 mm and focus wavelength is 570 nm. (Left) 10 cpd. The value for the standard observer is 0.98, and for only diffraction and LCA (no wave aberrations) is 2.43. (Right) 20 cpd. The value for the standard observer is 1.30, and for only diffraction and LCA (no wave aberrations) is 2.33.

retinal image of the letter E (Fig. 14) is a demonstration of this point for the standard observer. The top three panels on the right of that figure include only diffraction and LCA; the image for S cones is far worse than for L and M. Wave aberrations, included in the images in the three bottom panels, further degrade the image for L and M cones but improve somewhat the image for S cones; nonetheless, even with wave aberrations the retinal image for S cones does not approach the quality of the image for L or M cones.

5.3. Other applications for the model of retinal image quality

The model used here can be employed for many other purposes. Psychophysical experiments aiming to infer neural processes often depend on knowledge of an accurate retinal (not corneal) spatial distribution of light. The model can be used to determine a typical retinal image for the standard observer, given any external stimulus; moreover, a population range of retinal images can be estimated for, say, the central 90% of the population by finding the extreme L-, M- and S-cone MTFs for the best and worst 5% of the population. This can be useful for excluding optical stimulus distortion as a viable account for experimental measurements, thereby implicating neural processes.

The pupil is largest, and retinal image quality often poorest, with dim illumination. A straightforward generalization of the model is to substitute the scotopic luminosity function $V(\lambda)$ for the cone spectral sensitivity function $S_C(\lambda)$ in Eq. (1). This gives the spatial distribution $I_R(x, y)$ for rod excitation, from which a rod MTF or other characteristics of rod stimulation can be determined.

The model's capability to incorporate population variation in retinal image quality can be applied to several practical issues, including design decisions for instruments, spectral illumination and chromatically complex surfaces. An estimate of the retinal image for any light stimulus, including any object under any illuminant, can be useful for example to determine the legibility of text for, say, 99% of the population by assuming retinal image quality for the worst 1% of normal human observers. This may be particularly valuable for examining nighttime visibility (for example, for signage), when the pupil is largest. More generally, standards can be developed for detection or discrimination that take account of losses in retinal image quality for the vast majority of the normal population (say 99%, again by using retinal image quality for the worst 1% of the normal population). A related application is assessment of image quality (digital or otherwise). The ability to incorpo-

rate normal variation in the eye's optics permits perceived image-quality assessments to take account of the substantial variation in the retinal image, which depends on both the particular image presented to the eye and individual differences within the population of human observers.

Acknowledgment

Supported by NIH Grants RO1 EY-05109 (L.T.) and RO1 EY-04802 (S.S.).

References

- Barnden, R. (1974). Calculation of axial polychromatic optical transfer function. *Optica Acta*, 21, 981–1003.
- Castejon-Mochon, J. F., Lopez-Gil, N., Benito, A., & Artal, P. (2002). Ocular wave-front aberration statistics in a normal young population. *Vision Research*, 42, 1611–1617.
- Coe, C. (2009). *Effect of monochromatic aberrations and wavelength on refraction and visual performance*. Doctoral dissertation, Indiana University (UMI Number: 3358910).
- Emsley, H. H. (1952). *Visual optics* (5th ed.). London: Hatton Press.
- MacLeod, D. I. A., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America*, 69, 1183–1185.
- Marimont, D. H., & Wandell, B. A. (1994). Matching color images: The effects of axial chromatic aberration. *Journal of the Optical Society of America A*, 11, 3113–3122.
- McLellan, J. S., Marcos, S., Prieto, P. M., & Burns, S. A. (2002). Imperfect optics may be the eye's defence against chromatic blur. *Nature*, 417, 174–176.
- McLellan, J. S., Prieto, P. M., Marcos, S., & Burns, S. A. (2006). Effects of interactions among wave aberrations on optical image quality. *Vision Research*, 46, 3009–3016.
- Nam, J., Rubinstein, J., & Thibos, L. (2010). Wavelength adjustment using an eye model from aberrometry data. *Journal of the Optical Society of America A*, 27, 1561–1574.
- Packer, O., & Williams, D. R. (2003). Light, the retinal image, and photoreceptors. In S. K. Shevell (Ed.), *The science of color* (2nd ed., pp. 41–102). Oxford: Elsevier.
- Poirson, A. B., & Wandell, B. A. (1993). Appearance of colored patterns: Pattern-color separability. *Journal of the Optical Society of America A*, 10, 2458–2470.
- Porter, J., Guirao, A., Cox, I. G., & Williams, D. R. (2001). Monochromatic aberrations of the human eye in a large population. *Journal of the Optical Society of America A*, 18, 1793–1803.
- Ravikumar, S., Thibos, L. N., & Bradley, A. (2008). Calculation of retinal image quality for polychromatic light. *Journal of the Optical Society of America A*, 25, 2395–2407.
- Roorda, A. (2002). Human visual system – Image formation. In J. P. Hornak (Ed.), *The encyclopedia of imaging science and technology* (Vol. 1, pp. 539–557). New York: John Wiley & Sons.
- Salmon, T. O., & van de Pol, C. (2006). Normal-eye Zernike coefficients and root-mean-square wavefront errors. *Journal of Cataract and Refractive Surgery*, 32, 2064–2074.
- Smith, G. (1995). Schematic eyes: History, description and applications. *Clinical and Experimental Optometry*, 78, 176–189.
- Smith, V. C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research*, 15, 161–171.

- Thibos, L. N. (2009). Retinal image quality for virtual eyes generated by a statistical model of ocular wavefront aberrations. *Ophthalmic and Physiological Optics*, 29, 288–291.
- Thibos, L. N., & Bradley, A. (1999). Modeling the refractive and neuro-sensor systems of the eye. In P. Mouroulis (Ed.), *Visual instrumentation: Optical design and engineering principles* (pp. 101–159). New York: McGraw-Hill.
- Thibos, L. N., Bradley, A., & Hong, X. (2002). A statistical model of the aberration structure of normal, well-corrected eyes. *Ophthalmic and Physiological Optics*, 22, 427–433.
- Thibos, L. N., Hong, X., Bradley, A., & Cheng, X. (2002). Statistical variation of aberration structure and image quality in a normal population of healthy eyes. *Journal of the Optical Society of America A*, 19, 2329–2348.
- Thibos, L. N., Ye, M., Zhang, X., & Bradley, A. (1992). The chromatic eye: A new reduced-eye model of ocular chromatic aberration in humans. *Applied Optics*, 31, 3594–3600.
- Williams, C. S., & Becklund, O. A. (1989). *Introduction to the optical transfer function*. New York: John Wiley & Sons.
- Williams, D., & Hofer, H. (2004). Formation and acquisition of the retinal image. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (Vol. 1, pp. 795–810). Cambridge, MA: MIT Press.
- Winn, B., Whitaker, D., Elliott, D. B., & Phillips, N. J. (1994). Factors affecting light-adapted pupil size in normal human subjects. *Investigative Ophthalmology & Visual Science*, 35, 1132–1137.

9.2 Lossless image compression via predictive coding of discrete Radon projections



Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image



Lossless image compression via predictive coding of discrete Radon projections

Andrew Kingston^{a,b}, Florent Autrusseau^{a,*}

^a IRCCyN lab, Polytech'Nantes, Rue Ch. Pauc, BP 50609, 44306 Nantes, France

^b Department of Applied Mathematics, Research School of Physical Sciences and Engineering, Australian National University, Canberra, ACT 0200, Australia

ARTICLE INFO

Article history:

Received 20 November 2007
Accepted 11 March 2008

Keywords:

Lossless image coding
Discrete Radon transform
Mojette
Redundancy

ABSTRACT

This paper investigates predictive coding methods to compress images represented in the Radon domain as a set of projections. Both the correlation within and between discrete Radon projections at similar angles can be exploited to achieve lossless compression. The discrete Radon projections investigated here are those used to define the Mojette transform first presented by Guédon et al. [Psychovisual image coding via an exact discrete Radon transform, in: T.W. Lance (Ed.), Proceedings of the Visual Communications AND Image Processing (VCIP), May 1995, Taipei, Taiwan, pp. 562–572]. This work is further to the preliminary investigation presented by Autrusseau et al. [Lossless compression based on a discrete and exact radon transform: a preliminary study, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. II, May 2006, Toulouse, France, pp. 425–428]. The 1D Mojette projections are re-arranged as two dimensional images, thus allowing the use of 2D image compression techniques onto the projections. Besides the compression capabilities, the Mojette transforms brings an interesting property: a tunable redundancy. As the Mojette transform is able to both compress and add redundancy, the proposed method can be viewed as a joint lossless source-channel coding technique for images. We present here the evolution of the compression ratio depending on the chosen redundancy.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

This work is motivated by the limitations of image processing tools in large multi-media databases. Numerous paintings belonging to French museums are stored in an image database within the 'Centre of Research and Restoration of French Museums' (C2RMF).¹ It is required that these high payload (up to 1000 Megapixel) images be losslessly compressed, stored securely (i.e., with some redundancy) and encrypted for transmission purposes. The French TSAR project (Secure Transfer of High Resolution Art Images²) aims to develop a method to securely

transfer images from the image database of artwork contained in the Louvre Museum. All this can also be said of medical image databases. It has been shown that the Mojette transform can be used for distributed storage [7] and encryption [10], if competitive lossless compression can also be achieved on the Mojette projection data, the majority of the above objectives can be achieved using only the Mojette transform. Although joint source-channel coding has been extensively studied in the literature [20], most of the interest has been focused on lossy coding [3] with only limited research being conducted on lossless joint source-channel coding for images [2]. Applications in geophysics and telemetry as well as the previously mentioned museum and medical image databases all require lossless compression [13]. We do not focus here on the properties of the Mojette transform compared to state of the art joint source-channel coding, but our goal is rather to improve

* Corresponding author.

E-mail address: Florent.Autrusseau@univ-nantes.fr (F. Autrusseau).

¹ <http://www.c2rmf.fr>

² <http://www.lirmm.fr/tsar/>

a previous study [1] in terms of lossless compression rate. We nevertheless want to point out the important link between the Mojette transform and maximum distance separable (MDS) codes used for joint source-channel coding (e.g. Reed–Solomon or BCH codes). The performances of the proposed algorithm regarding both compression and redundancy will be shown in Section 4.1.

The Mojette transform is an entirely discrete mapping (from a discrete image to discrete projections), which requires only the addition operation and is exactly invertible. It retains the major properties of the Radon transform such as the Fourier slice theorem and the related convolution property but also introduces new properties such as redundancy. It was first proposed by Guédon et al. in 1995 [5] in the context of psychovisual image coding. It has since been applied in many aspects of image processing such as image analysis, image watermarking, image encryption, and tomographic image reconstruction from projections. The unique properties of the transform have also made it a useful multiple description tool with applications in robust data transmission and distributed data storage. A summary of the evolution and applications of the Mojette transform to date can be found in [6].

Since the Mojette transform already has other advantages, the objective of this work is not to develop a superior image compression standard. Rather, we seek to extend the work by Autrusseau et al. in [1] on the compression of Mojette projection data to become comparable with results from existing techniques. The preliminary study introduced the idea of a compression scheme which exploits correlation within a projection (intra-projection coding) as well as compression scheme which exploits correlation between projections (inter-projection coding). Since the Mojette projections are principally used in a data transmission and storage context, developing compression techniques which are effective is important. Both types of coding must be investigated as, if redundant projections are required in the transform, only intra-projection coding may be possible. This paper investigates several methods to compress projections by adapting multi-spectrum image compression techniques to multi-projection data.

The paper is organised as follows: The Mojette transform, projection properties and inverse are presented in the next section. A summary of the preliminary study by Autrusseau et al. [1] follows in Section 3. Section 4 demonstrates how inter-band image compression techniques can be applied. This is followed by some concluding remarks and future research directions in Section 5.

2. A discrete Radon transform: The Mojette transform

2.1. Mojette projections

The Radon transform³ maps a continuous 2D function to a set of 1D continuous projections at all angles $\theta \in [0, \pi)$.

³ Ref. [4] provides an English translation of Radon's original 1917 paper.

A projection at angle, θ , is obtained as the linear integration of the function over all parallel lines with gradient $\tan \theta$. One of the most important properties of the Radon transform is that it is invertible. This implies that the internal structure of an object can be determined non-destructively from its projections (tomography). The Radon transform is utilised in areas ranging from medical tomography (CT, MRI, ultrasound) to astronomy and seismology. In recent years it has also been applied to many aspects of image analysis, image representation and image processing. Since the projection data and reconstructed image are both discrete, the implementation of the Radon transform and its inverse must be discretised. Many methods involve filtering and interpolating the discrete data; a numerically intensive procedure. There have also been several discrete Radon transforms proposed which naturally deal with discrete data, e.g. [11,9]. This paper is concerned with one particular discrete Radon transform known as the Mojette transform.

The Mojette transform is an exact, discrete form of the Radon transform defined for specific 'rational' projection angles. Like the classical Radon transform, the Mojette transform represents the image as a set of projections, however, in contrast, the Mojette transform has an exact inverse from a finite number of discrete projections (as few as 1 depending on the angle set). The rational projection angles, θ_i , are defined by a set of vectors (p_i, q_i) as $\theta_i = \tan^{-1}(q_i/p_i)$, as depicted in Fig. 1a for $(p_i, q_i) = (2, 1)$. These vectors must respect the condition that p_i and q_i are coprime (i.e., $\gcd(p_i, q_i) = 1$) and since $\tan$ is π -periodic q_i is restricted to be positive except for the case $(p_i, q_i) = (1, 0)$. The transform domain of an image is a set of projections where each element (called a 'bin' as in tomography) corresponds to the sum of the pixels centred on the line of projection as depicted in Fig. 1a. This is a linear transform defined for each projection angle by the operator:

$$\begin{aligned} \mathcal{M}_{p_i, q_i}\{f(k, l)\} &= \text{proj}_{p_i, q_i}(b) \\ &= \sum_{k=-\infty}^{+\infty} f(k, l) \Delta(b + kq_i - lp_i), \end{aligned} \quad (1)$$

where (k, l) defines the position of an image pixel and $\Delta(b)$ is the Kronecker delta function which is 1 when $b = 0$ and zero otherwise. Invertible projections can be obtained not only with addition but using any linear operation. Other practically useful operations include modulo 256 addition and bitwise XOR. The Mojette transform, $\mathcal{M}_I f(k, l)$, corresponds to the set of I projections as $\mathcal{M}_I f(k, l) = \{\text{proj}_{p_i, q_i}, i \in [1, \dots, I]\}$.

As depicted for the example images in Fig. 1a and b, each bin value equals the sum of the pixels crossed by the appropriate line

$$b = lp_i - kq_i. \quad (2)$$

The principle difference from the classical Radon transform is the sampling rate on each projection, which is no longer constant but depends on the chosen angle as $1/\sqrt{p_i^2 + q_i^2}$. This can be seen for the different projections in Fig. 1b which demonstrates the Mojette transform for

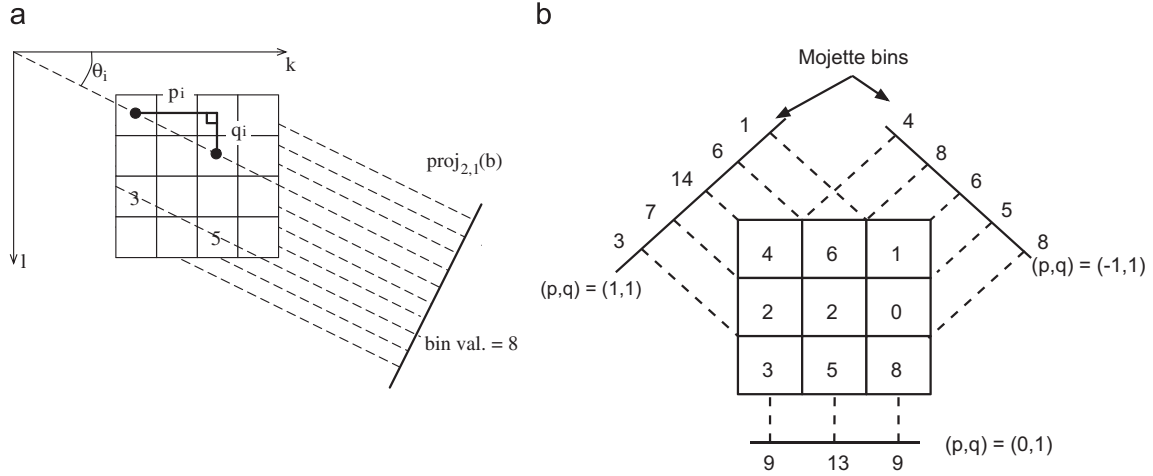


Fig. 1. (a) A depiction of (p_i, q_i) , the corresponding angle, θ_i , and the method of projection, i.e., summing pixel values centred on the line to give a bin value. (b) An invertible Mojette transform of a 3×3 example image using direction vectors $\{(1, 0), (1, 1), (-1, 1)\}$. Note the spacing between adjacent line sums varies with projection angle.

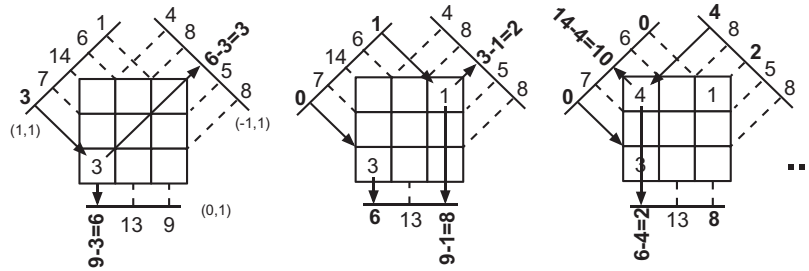


Fig. 2. Three first possible steps of the inverse Mojette transform of the projections obtained in Fig. 1.

the directions set $S = \{(1, 0), (-1, 1) \text{ and } (1, 1)\}$. The number of bins, B_i , for each projection depends on the chosen direction vector (p_i, q_i) , and for a $P \times Q$ image is found as

$$B_i = (Q - 1)|p_i| + (P - 1)q_i + 1. \quad (3)$$

The algorithmic complexity of the Mojette transform for a $P \times Q$ image with I projections is $O(PQI)$.

2.2. Conditions for reconstructability

Since the set of projection directions is selected arbitrarily, the original data cannot necessarily be recovered from the set of projections chosen. A criterion is required to determine if a set of projections is sufficient to uniquely reconstruct the data.

The first result on the conditions for the existence of a unique reconstruction from a given set of I projections came from Katz [8] in a very similar context. He showed that if the following criterion is satisfied, any rectangular $P \times Q$ data set can be uniquely reconstructed:

$$P \leq \sum_{i=1}^I |p_i| \quad \text{or} \quad Q \leq \sum_{i=1}^I q_i. \quad (4)$$

This result has been extended in an independent manner by Normand and Guédon [14] to apply to data with compact support of any shape.

2.3. Reconstruction from Mojette projections

The inverse Mojette transform is a fast and simple algorithm [14]. Searching for and updating 1–1 pixel-bin correspondence enables a simple iterative procedure to recover the image. The bin value is back-projected into the pixel and subtracted from the corresponding bins in all other projections. The number of pixels belonging to the corresponding bins is also decremented. The algorithmic complexity of the inverse Mojette transform for a $P \times Q$ image with I projections is $O(PQI)$ [14]. Fig. 2 shows one possibility of the first three steps of the inverse Mojette transform of the example projections given in Fig. 1.

This ‘accounting’ inverse was improved by Normand et al. in [15] where the need to search for 1–1 pixel-bin correspondence was removed. It was proven that when ordered by angle, $\theta_i = \tan^{-1}(q_i/p_i)$, each projection reconstructs the subsequent q_i rows of the image for the case where $\sum q_i = Q$ (or subsequent $|p_i|$ columns where

$\sum |p_i| = P$). This knowledge enables the periodic sequence of reconstructible pixels to be predetermined, removing the need for the accounting images in the reconstruction.

This result improves reconstruction time by a factor of 5 but is also useful for removing unwanted redundancy. Since the rows (or columns) of the image that are reconstructed by a given projection are known, any projection bins not containing pixels from these rows (or columns) can be removed. Thus for pure compression applications, the Mojette transform can be completely non-redundant mapping $P \times Q$ pixels to PQ bins.

3. A review of the preliminary study

Autrusseau et al. [1] noted that the Mojette projection data are highly correlated within a projection and also that a strong correlation exists between projections at similar angles. This can be seen in Fig. 3a(i)–(iii) for the projection set $\{(1, 171), (1, 172), (1, 173)\}$ with respective projection angles of 89.665° , 89.667° , and 89.669° . This implies that a form of differential coding should be an efficient compression technique.

The technique presented in the paper defines a simple compression technique based on two differential pulse code modulation (DPCM) schemes of order 1. One scheme is applied within a projection, defined as ‘intra-projection’ coding, and the other is applied between projections, defined as ‘inter-projection’ coding.

Let b_n^i denote the value of the n th bin of the i th projection, i.e., $\text{proj}_{p_i, q_i}(n)$. Assume this is the current bin value to be coded and let $\hat{b}_n^i$ be the prediction of b_n^i with the encoded prediction error defined as $e_n^i = b_n^i - \hat{b}_n^i$.

According to (2), horizontally adjacent pixels in the image are separated by q_i in the projection bins. Therefore the projection data are periodic with q_i and an appropriate prediction for intra-projection coding is

$$\hat{b}_n^i = b_{n-q_i}^i. \quad (5)$$

Fig. 3b(i) depicts the result of this coding applied to the section of $\text{proj}_{1,171}$ given in Fig. 3a(ii). A prediction for inter-projection coding is simply defined as

$$\hat{b}_n^i = \tilde{b}_n^{i+1}, \quad (6)$$

where $\tilde{b}_n^{i+1}$ is the bin in $\text{proj}_{p_{i+1}, q_{i+1}}$ that ‘best’ corresponds to b_n^i . This is more difficult to realise in practice since the projections are of different length and the most appropriate bin to utilise for prediction is not obvious. This is explored more fully in Section 4, here linear interpolation is used. Fig. 3b(ii) depicts the result of this coding applied to the section of $\text{proj}_{1,171}$ given in Fig. 3a(ii) using $\text{proj}_{1,172}$ as the reference projection. Both schemes used together produce a DPCM of order 3 as follows:

$$\hat{b}_n^i = b_{n-q_i}^i + \tilde{b}_n^{i+1} - \tilde{b}_{n-q_{i+1}}^{i+1}. \quad (7)$$

Fig. 3b(iii) shows the result of this coding applied to the section of $\text{proj}_{1,171}$ given in Fig. 3a(i) once again using $\text{proj}_{1,172}$ as the reference projection.

It is important to note that given I projections, the total number of bins according to (3) is $I + (Q - 1) \sum |p_i| + (P - 1) \sum q_i$ and Katz criterion, (4), must be satisfied for inversion. This implies projection sets that minimise redundancy are those of the form $\{(1, q_1), (1, q_2), \dots, (1, q_I)\}$ such that $\sum q_i$ is equal to or only slightly greater than Q (or $\{(p_1, 1), (p_2, 1), \dots, (p_I, 1)\}$ such that

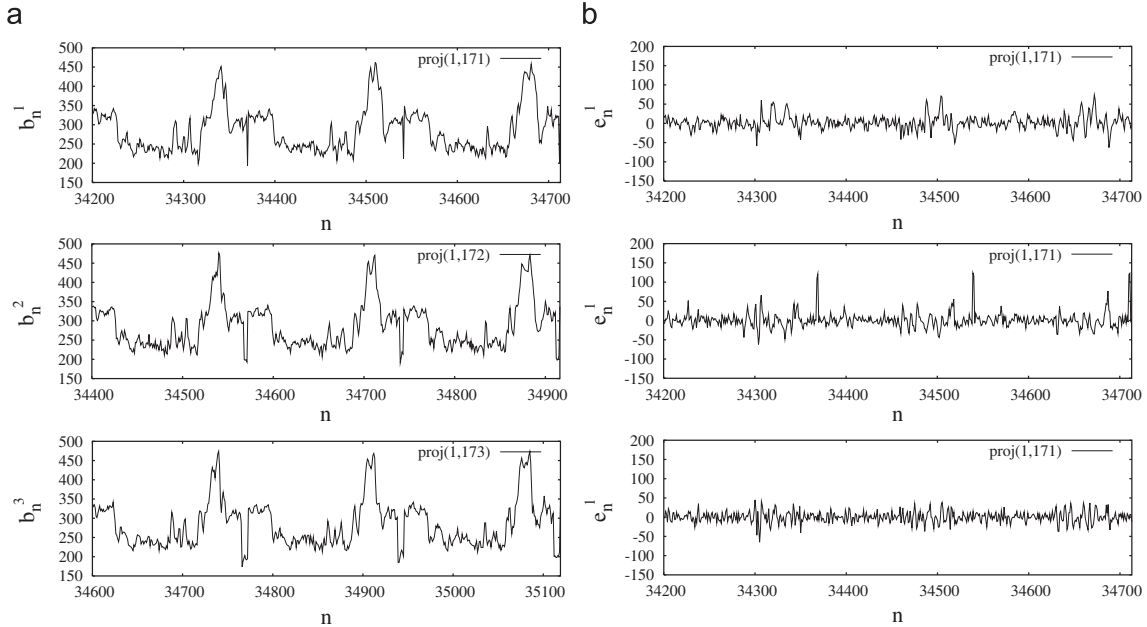


Fig. 3. (a) A zoom of three projections of ‘Lena’ with projection vectors $(1, q_i)$ where q_i is (i) 171, (ii) 172, and (iii) 173 showing the periodicity within a projection and the correlation between the projections. (b) (i) Intra-projection coding applied to the section of $\text{proj}_{1,171}$ given in (a(i)). (ii) Inter-projection coding using $\text{proj}_{1,172}$ (a(ii)) as a reference projection. (iii) Both intra- and inter-projection coding of $\text{proj}_{1,171}$.

9.2. Lossless image compression via predictive coding of discrete Radon projections

141

A. Kingston, F. Autrusseau / Signal Processing: Image Communication 23 (2008) 313–324

317

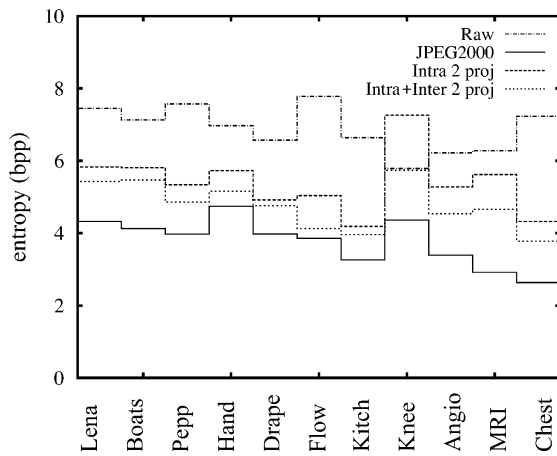


Fig. 4. (a) The compression results for intra- and inter-projection compression for two projections as outlined in [1] compared with JPEG2000 for the 11 test images depicted in (b) numbered left to right, top to bottom. The first row contains 512×512 natural images: 'Lena', 'Boats', 'Peppers'. The second row contains C2RMF images: 'Hand' 1200×1854 , 'Drape' 2376×3542 , 'Flowers' 1405×1125 , 'Kitchen' 3822×3333 . The last row contains 256×256 medical images: 'Knee', 'Angio', 'MRI', 'Chest'.

$\sum |p_i| \geq P$). In a compression context, this restriction is more dominant than selecting projection directions according to texture orientation. A method to take advantage of this is a subject for future research.

The results of applying (5) and (7) followed by entropy coding to 11 test images has been presented in Fig. 4. Compression results have been given as the final entropy in bits per pixel (bpp) and have been compared with the original entropy of the image and the compression results from applying JPEG2000. JPEG2000 has been selected for comparison even though it does not generally give optimal compression. It is robust to image type (i.e., natural, artificial, smooth, textured), and is commonly used due to its multi-resolutional capabilities. Likewise, the Mojette also has other capabilities in data storage and encryption so achieving compression at least comparable to JPEG2000 is desirable.

Fig. 4 shows that some degree of compression is generally achieved (apart from the high contrast 'Knee' image), however, the results are not comparable to those from JPEG2000. The next section seeks to identify more powerful prediction schemes which are appropriate for both intra- and inter-projection coding to improve compression results.



Fig. 4. (Continued)

4. Projection compression using multi-band image techniques

4.1. Intra-projection compression of projection images

Phillipé and Guédon [16] showed that the 2D image auto-correlation is retained in the Mojette projections. If the projection data are arranged in columns of length q_i or rows of length p_i (whichever is greater is preferable), this auto-correlation becomes apparent as the projection appears as a 'folded' image. This has been depicted in Fig. 5 for three projections of 'Lena'. The remapping of projection bins is performed by projecting pixel value, $f(k, l)$, to $\text{proj}_{p_i, q_i}(n_k, n_l)$ according to

$$\text{if } |p_i| \geq q_i \begin{cases} n_k = \left\lfloor \frac{k}{|p_i|} \right\rfloor, \\ n_l = l - \frac{(k - n_k)}{p_i} q_i \end{cases}$$

$$\text{otherwise } \begin{cases} n_l = \left\lfloor \frac{l}{|q_i|} \right\rfloor, \\ n_k = k - \frac{(l - n_l)}{q_i} p_i \end{cases} \quad (8)$$

(where $\lfloor x \rfloor$ gives the greatest integer less than or equal to x), such that the corresponding bin in the 1D projection, $\text{proj}_{p_i, q_i}(b)$, is found as

$$b = n_l p_i - n_k q_i. \quad (9)$$

This implies that 2D image compression schemes can be applied when performing intra-projection coding. The prediction complexity can be increased from a simple DPCM of order 1 to DPCM order to 3 and further to well known ADPCM techniques with context coding such as LOCO (of order 3) [17] and CALIC (of order 7) [18] and Glicbawls [12] which uses the entire set of causal data. As an example, the compression results in bpp for each of

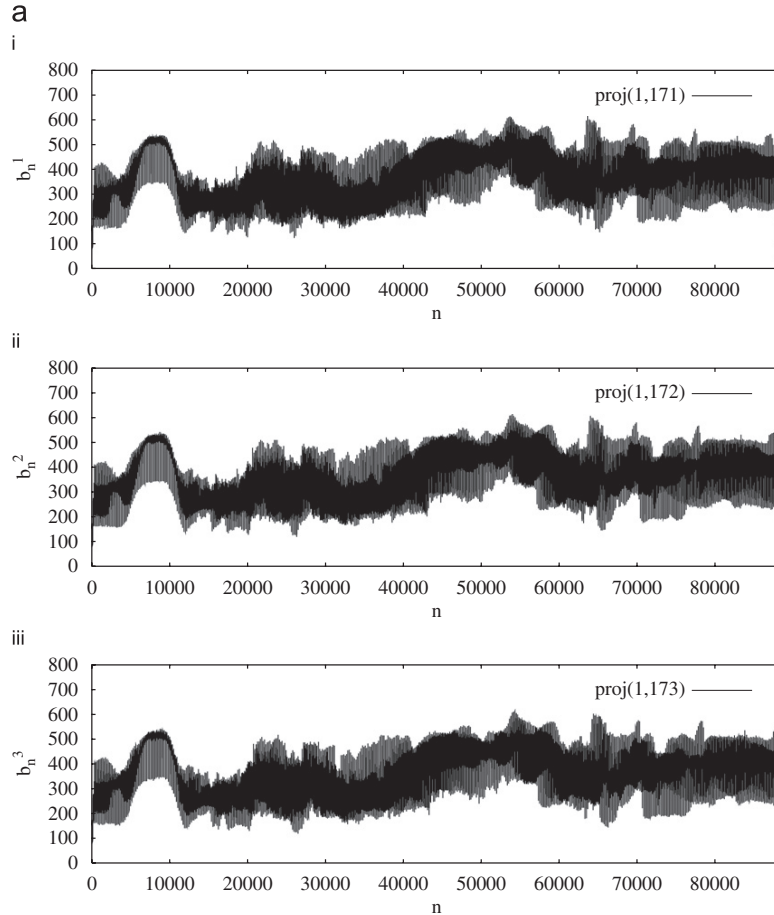


Fig. 5. (a) Three 1D projections of 'Lena' with projection vectors $(1, q_i)$ where q_i is (i) 171, (ii) 172, and (iii) 173. (b) The same projection data displayed as images with columns heights of q_i and a width of 514.



Fig. 5. (Continued)

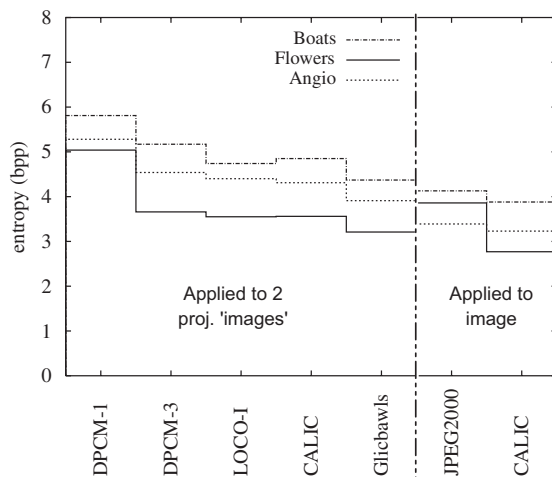


Fig. 6. The entropy results for intra-projection compression of two projection 'images' using prediction schemes: DPCM-1, DPCM-3, LOCO, CALIC, and Glicbawls. This is compared with applying JPEG2000 and CALIC directly to the image.

these respective methods to encode two projections of three different images using the direction vectors $\{(P/2, 1), (P/2 + 1, 1)\}$ is given in Fig. 6 and compared with

the result of applying JPEG2000 to the image. (Similar results are achieved using the direction vectors $\{(1, Q/2), (1, Q/2 + 1)\}$). This shows that compression results better than JPEG2000 can be achieved in some cases with the use of more sophisticated image coding techniques without the need for inter-projection coding. This is an important result for distributed storage. However, the more complex techniques require more memory and computation time. A good trade-off (by design) is CALIC.

The implications of these results are that redundancy can be inserted for distributed storage or transmission very efficiently. For the example above, an extra projection with a direction vector $(-P/2 - 1, 1)$ can be also included. The resulting compression will still be less than 6.0 bpp (as shown in Fig. 7 for 'Flowers') but with the added advantage that any two of the three projections is sufficient to recover the image according to (4). If each projection is stored on a unique server or transmitted over a unique channel, this is a very secure distribution scheme. Of course the number of projections required for reconstruction and the degree of redundancy can be tuned as required.

Fig. 7 gives the compression achieved with the number of projections required to recover the 'Flowers' image, N . At each N there are three values, the first is the compression attained without redundancy in the

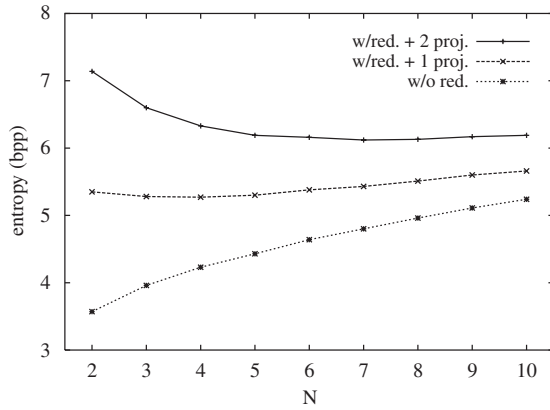


Fig. 7. A plot of compression against the number of projections, N , required to reconstruct the 'Flowers' image; without redundancy, with one extra projection (so any N of the $N + 1$ projections is sufficient for reconstruction), and with two extra projections (so any N of the $N + 2$ projections is sufficient for reconstruction).

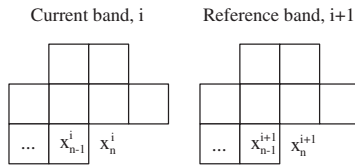


Fig. 8. The causal 'neighbourhood' in both the current spectral band and the reference spectral band used in inter-band CALIC prediction.

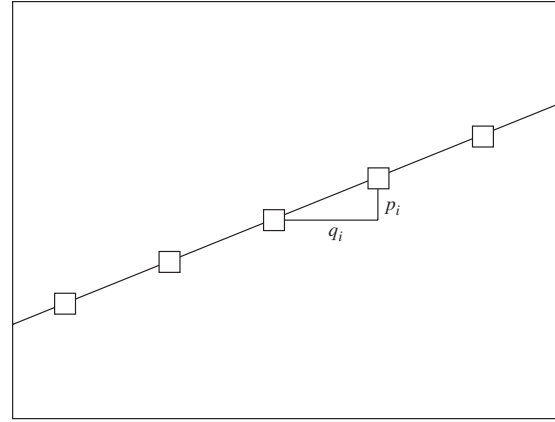
projections, i.e., N projections. The next two give the compression attained with some redundancy, where 1 (resp. 2) extra projection(s) are included such that any N of the $N + 1$ (resp. $N + 2$) projections are sufficient to recover the image, i.e., an additional redundancy of $1/N$ (resp. $2/N$).

Since the projections can be represented as images, inter-projection coding could be considered to be similar to inter-band image coding, e.g., coding between the RGB components of an image. The next section investigates inter-projection compression issues and techniques associated with this idea.

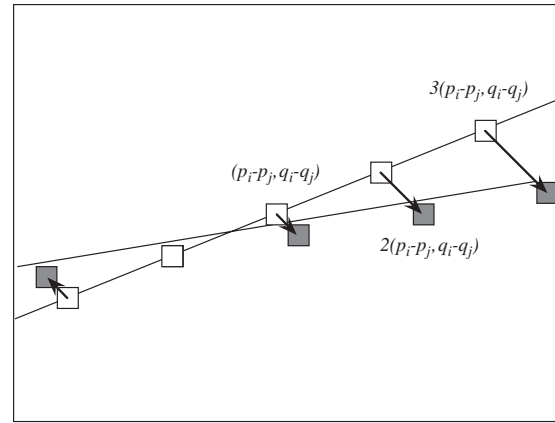
4.2. Inter-projection compression

A simple but effective inter-band prediction method is presented in [19] to extend CALIC to multi-spectral images. The essential idea is to compute the cross-correlation between the current band, x^i , and the next band, x^{i+1} , over the causal 'neighbourhood' data used in CALIC prediction. These regions are depicted in Fig. 8; note that x_n^i and x_n^{i+1} have the same spatial location but are in different bands. If these two regions are highly correlated then the intra-band CALIC prediction of the n th pixel value of the current band, $\hat{x}_n^i$, can be improved upon by using information from the next band, x^{i+1} . This is true since the actual value of the n th pixel in the reference band, x_n^{i+1} , is known.

a



b



c

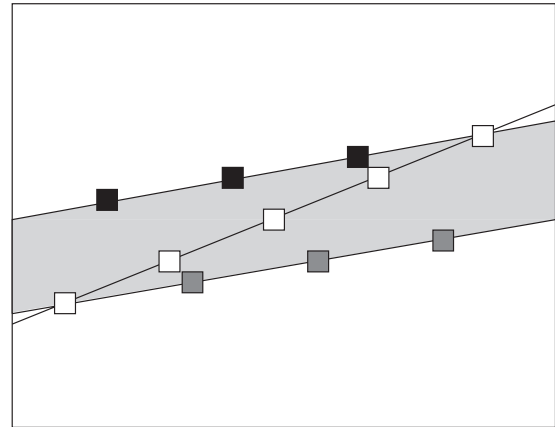


Fig. 9. (a) Relative position of pixel values summed to give a projection bin value. (b) The difference in sampled pixels between two raysums of different projections including a common pixel. (c) The shaded region contains raysums from proj_{p_i, q_i} that intersect the given raysum from proj_{p_i, q_i} whose contributing pixels are shown as white squares.

A similar idea can be applied between Mojette projections of an image. In inter-band CALIC the different bands contain images that are spatially consistent but with different intensities according to the band of the spectrum. With inter-projection compression, however, the same pixel intensities, i.e., $f(k, l)$, exist in all projections but with a different phase or period of ‘mixing’ with other intensities. Therefore, the best phase to use when predicting each bin value must be determined. In other words, which projection, proj_{p_i, q_i} , should be used as the reference projection and which bin in this projection best corresponds to the current bin required to be predicted in the current projection, $\text{proj}_{p_{i+1}, q_{i+1}}$?

Let us investigate the position of pixels summed in a ‘raysum’ to give a bin value, $\text{proj}_{p_i, q_i}(b)$. Adjacent pixels in this set are separated by (p_i, q_i) as shown in Fig. 9a. Consider the pixels sampled by two raysums from different projections as depicted in Fig. 9b. This figure shows that the distance between the previous and subsequent pixels sampled increases by $(p_i - p_j, q_i - q_j)$ each step. Therefore when coding proj_{p_i, q_i} the obvious choice of projection to utilise as a reference for inter-projection compression is $\text{proj}_{p_{i+1}, q_{i+1}}$ such that the length of this difference vector is minimised. This is denoted as $\text{proj}_{p_{i+1}, q_{i+1}}$.

The question of which raysum from this projection should be used for prediction of $b_n^i = \text{proj}_{p_i, q_i}(n)$ is not so obvious since it is more content dependant. A selection criterion is required to determine the best raysum from $\text{proj}_{p_{i+1}, q_{i+1}}$ in the shaded region of Fig. 9c that will give the best prediction for the raysum, b_n^i , shown. There are two questions: 1. How many candidate raysums are there in $\text{proj}_{p_{i+1}, q_{i+1}}$? and 2. What are their bin indices?

To address the first question, assume there are a maximum of r_i image pixels sampled in the raysum to give b_n^i . The value of r_i is found as $\min(\lceil P/p_i \rceil, \lceil Q/q_i \rceil)$ where $\lceil x \rceil$ gives the smallest integer greater than or equal to x . The longest possible vector between corresponding sampled pixels of the two lines is therefore $(r_i - 1)(p_i - p_{i+1}, q_i - q_{i+1})$. This also gives the maximum possible vector between pixels on the two bounding lines in Fig. 9c. Thus, from (2) there are $M = (r_i - 1)(q_i - q_{i+1})p_{i+1} - (p_i - p_{i+1})q_{i+1} + 1$ candidate raysums in the shaded region of Fig. 9c. For this example assume $M = 5$ as shown in Fig. 10a.

To determine the indices of these candidate bins are straightforward. The raysum that samples pixels closest to the l axis (dark grey pixels in Fig. 10a) intersects with the raysum giving b_n^i (white pixels in Fig. 10a) in first region of the image, $m = 0$, as indicated in Fig. 11a. Therefore, the (n_k, n_l) position of these bins is identical in the two projection ‘images’ (as depicted in Fig. 10b), i.e.,

$$\tilde{b}_{n-m}^{i+1} = \text{proj}_{p_{i+1}, q_{i+1}}(n_l p_{i+1} - n_k q_{i+1}). \quad (10)$$

This is for a set of (p_i, q_i) direction vectors where $|p_i| \geq q_i$. The same equation applies to the case where $|p_i| < q_i$, however, here it is the raysum the samples pixels closest to the k axis (black pixels in Fig. 10a).

In summary, the bin values from $\text{proj}_{p_{i+1}, q_{i+1}}$ that are candidates to be used in the prediction of b_n^i include $\tilde{b}_{n-m}^{i+1}$

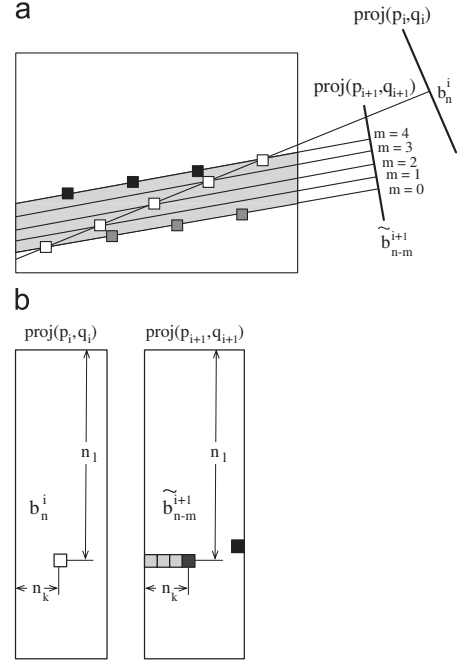


Fig. 10. (a) Depiction of the five candidate raysums from $\text{proj}_{p_{i+1}, q_{i+1}}$ in the shaded region labelled by phase, m . Any of these could be used when predicting the given bin value in proj_{p_i, q_i} . (b) The position of these bins in the respective projection ‘images’.

from (10) and the preceding $M - 1$ bin values, i.e., $\tilde{b}_{n-m}^{i+1}$ for $m \in [0, M - 1]$. These bins have been depicted in Fig. 10b for the example. Given the optimal candidate bin ‘phase’, m , the intra-projection prediction can be improved by

$$\hat{b}_n^i = \text{pred}(b_n^i) + \tilde{b}_{n-m}^{i+1} - \text{pred}(\tilde{b}_{n-m}^{i+1}), \quad (11)$$

where $\text{pred}()$ is any of the prediction models introduced in Section 4.1.

Selecting phase, m , in the reference projection for the prediction, i.e., using $\tilde{b}_{n-m}^{i+1}$, gives the best prediction for the m th section of the image, as labelled in Fig. 11a for the example given in Fig. 10a and b. The prediction is successively worse for the regions $m \pm 1, m \pm 2$, and so on. A sensible choice for m is therefore the raysum in the centre of the shaded region, $M/2$, and this does in general give the best result. However, if the central region of the image is smoother than towards the image boundaries, the correct prediction of texture edge positions is more critical than an overall minimum distance between corresponding sampled pixels of the two raysums.

This has been demonstrated for several 512×512 images using the $\text{proj}_{104,1}$ to predict projection $\text{proj}_{103,1}$. In this case (as in the previous diagrams), there are $M = 5$ candidate raysums to investigate which correspond to the five sections of the image as shown in Fig. 11a. In these regions the sampled pixels of the raysums from b_n^i and $\tilde{b}_{n-m}^{i+1}$ are identical (and hence the data of these sections is removed entirely according to (11)). Fig. 11b gives the inter-projection compression rates achieved using (11)

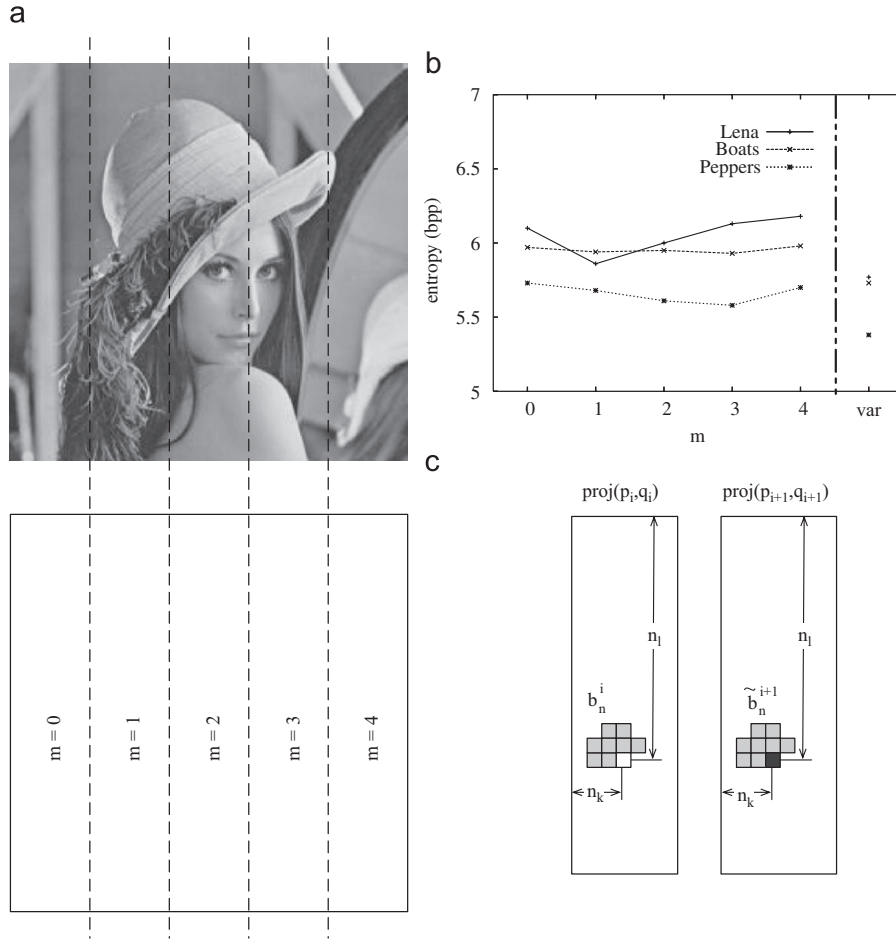


Fig. 11. (a) The five regions of the image which in selecting phase, m , in prediction will be removed entirely from the current projection after coding. (b) The compression results for inter-projection compression of $\text{proj}_{103,1}$ from $\text{proj}_{104,1}$ using constant m and variable m , i.e., 'var'. (c) The causal neighbours from the current projection and the reference projection used firstly to determine cross-correlation and also for prediction.

with constant phase, m , for all predictions. Here $\text{pred}()$ is the gradient adjusted prediction (GAP) introduced in [18]. It is expected that these plots attain a minimum at $m = 2$, however, this is not true in practise for any of the plots.

By selecting phase, m , the contribution of the m th section of the data is removed. Thus, better compression results are achieved when the phase corresponds to the most textured region of the image. For example, the second region of 'Lena', labelled $m = 1$ in Fig. 11a, contains the highly textured hair. Removing this textured region from the projection by setting $m = 1$ for inter-projection coding gives the lowest entropy as seen in Fig. 11b. These results demonstrate that the accurate location of texture edges is very important in the predictors performance.

A content dependant method is desired to select the optimum raysum from the M candidate raysums of $\text{proj}_{p_{i+1}, q_{i+1}}$. A fast method that attains near optimal performance in inter-projection coding is to determine the vertical region of the image (or horizontal for a set of $(1, q_i)$ direction vectors) with maximum 'activity' and select the phase of $\tilde{b}_{n-m}^{i+1}$, accordingly.

A slower but more effective selection criteria can be achieved by investigating the 2D cross-correlation between the causal neighbourhoods of b_n^i and $\tilde{b}_{n-m}^{i+1}$ in the two projections for all phases, $m \in [0, M - 1]$. These causal neighbourhoods have been depicted in Fig. 11c. The phase, m , of the candidate bin selected, $\tilde{b}_{n-m}^{i+1}$, is that with the greatest cross-correlation with b_n^i . If this cross-correlation is above some threshold it should be useful to improve the intra-projection prediction, otherwise only intra-projection coding is considered. This gives the best overall compression results as shown for the 'var', (i.e., variable), column in Fig. 11b with compression consistently lower than any constant phase, m .

Fig. 12b shows, for the example projections, that inter-projection coded data (12b(i) and (ii)) has been effectively decorrelated from the basis projection (12b(iii)). Fig. 12a shows the compression rates using GAP for inter-projection prediction and CALIC for intra-projection coding are comparable to JPEG2000. Fig. 13a and b summarises the results achieved in this paper for intra- and inter-projection coding, respectively. They plot the average

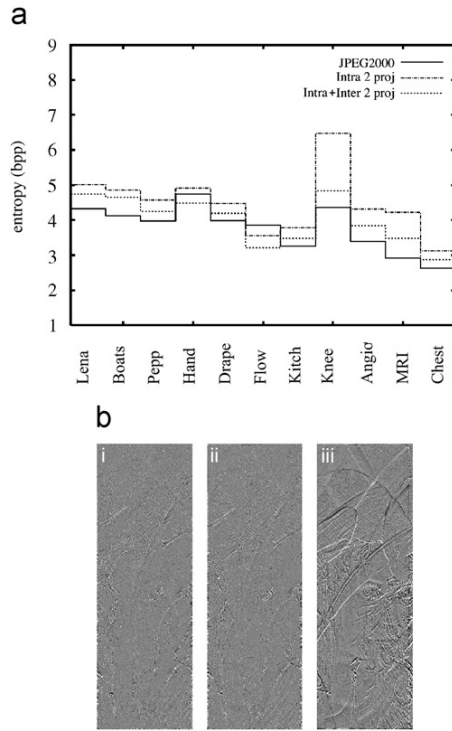


Fig. 12. (a) The compression results for inter-(GAP) and intra-(CALIC) projection compression of two projections using variable phase, m . Compared with the inter-projection compression results from the previous section using scheme B. (b) A visualisation of the decorrelation achieved using inter-projection coding. Three prediction error projections (i) $\text{proj}_{170,1}$, (ii) $\text{proj}_{171,1}$, and (iii) $\text{proj}_{172,1}$ of 'Lena'. Projections (i) and (ii) have been inter-coded and effectively decorrelated from projection (iii). Note that the grey levels are centered about 0 with a window of 64.

compression ratios of each type of image, i.e., natural, art, and medical, using the compression techniques from the preliminary study (Fig. 4a), and the multi-spectral band image coding scheme (Fig. 12a). The average entropy of each scheme is compared with the entropy of the original image and the compression achieved using JPEG2000. Results show that this scheme is particularly suited to the intended application to the scanned art image database of the C2RMF. For the art images, the compression achieved using solely intra-projection coding is similar to JPEG2000 and including inter-projection coding is more effective than JPEG2000.

5. Conclusions and future research

The technique to losslessly compress images via linear prediction of the Mojette projection presented in [1] has been improved upon here. Average compression achieved using intra-projection coding of two projections was improved from 5.38 to 4.48 bpp using a fast lossless image coding technique. Average inter-projection compression entropy was also improved from 4.77 to 4.00 bpp using a lossless inter-spectral band image compression technique. Fig. 13 shows that these improved results are comparable to those achieved using JPEG2000 applied directly to the image and that these techniques are particularly suited to the intended purpose of this work on compressing scanned art images for the TSAR project.

The image coding techniques have been adapted to fully exploit the nature of the Mojette projection data. This periodic nature is present since the Mojette projections preserve the 2D auto-correlation of the image and implies that image compression and inter-spectral band image compression can be applied. The prediction method selected for intra-projection coding should be selected

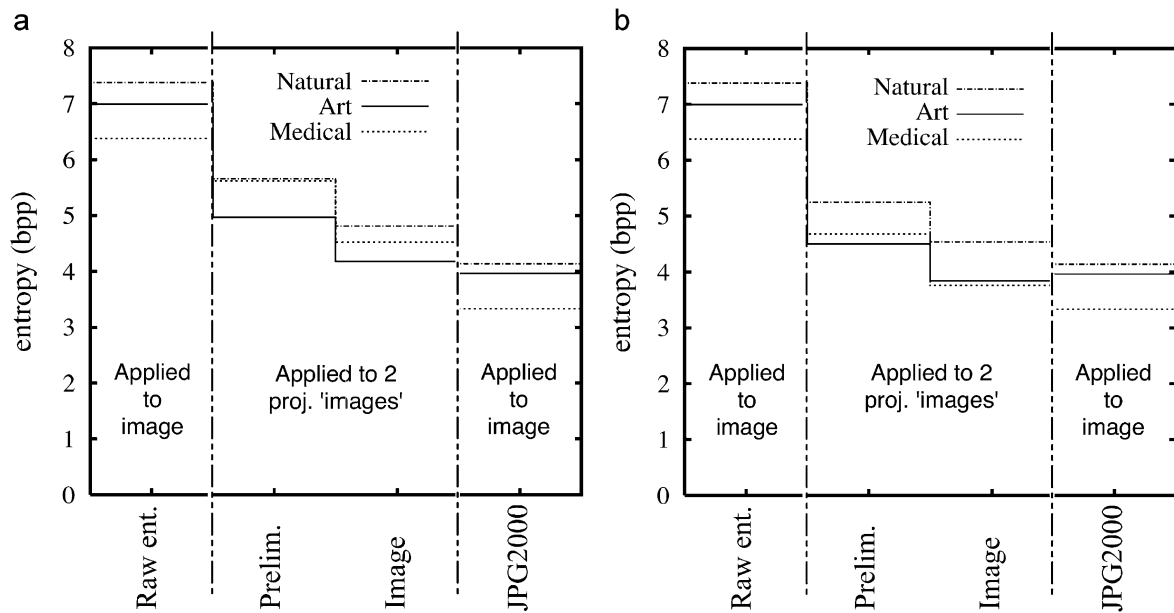


Fig. 13. A plot of average compression ratio over the 11 test images for both (a) intra- and (b) inter-projection coding using: the DPCM-1 method from the preliminary study, and the CALIC image scheme. Results are compared with the average entropy of the raw image and the compression achieved by applying JPEG2000 to the image.

depending on the requirements for compression and implementation time. A method to select the optimal raysum from all candidate raysums of a reference projection to use for inter-projection prediction has also been presented.

Compression rates comparable to JPEG2000 are achieved using image coding techniques, however, the image coding techniques take advantage of the 2D correlation and hence, by design, have less complexity, require less memory, and thus have a lower implementation time. Another possible compression scheme that may be applicable is video inter-frame coding with motion estimation and is a direction for future research. Another area for future investigation is that of coding colour images with predictions using inter-projection and inter-band correlation simultaneously.

These results imply that the Mojette projections which have applications in distributed storage and encryption of images in databases can also be effectively losslessly compressed and many of the requirements of an image database outlined in Section 1 can be achieved using solely the Mojette transform.

The methods explored here concentrated on predictive coding. Techniques using scalable transforms such as DWT, DCT and FFT may prove useful. If a block based approach could be made feasible, then it may also be beneficial to investigate projections directed along and orthogonal to texture orientation.

Acknowledgements

The majority of this work was conducted while A.K. held a postdoctoral position at l'Université de Nantes supported by a grant from the Région Pays de la Loire, France. This work is partially supported by projects Miles, Région Pays de la Loire, and ANR-TSAR.

References

- [1] F. Autrusseau, B. Parrein, M. Servièrès, Lossless compression based on a discrete and exact radon transform: a preliminary study, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. II, May 2006, Toulouse, France, pp. 425–428.
- [2] H. Bin, G.F. Elmasry, C.N. Manikopoulos, Joint lossless-source and channel coding using ARQ/go-back-(N, M) for image transmission, *IEEE Trans. Image Process.* 12 (12) (2003) 1610–1617.
- [3] G. Davis, J. Danskin, Joint source and channel coding for image transmission over lossy packet networks, in: SPIE Conference on Wavelet Applications of Digital Image Processing XIX, August 1996, pp. 376–387.
- [4] S.R. Deans, *The Radon Transform and Some of its Applications*, revised ed., Krieger, Malabar, FL, 1993.
- [5] J.P. Guédon, D. Barba, N. Burger, Psychovisual image coding via an exact discrete Radon transform, in: T.W. Lance (Ed.), *Proceedings of the Visual Communications AND Image Processing (VCIP)*, May 1995, Taipei, Taiwan, pp. 562–572.
- [6] J.P. Guédon, N. Normand, The Mojette transform: the first ten years, in: G.É. Andres, G. Damiand, P. Lienhardt (Eds.), *Proceedings of the 12th International Conference on Discrete Geometry for Computer Imagery, Lecture Notes in Computer Science*, vol. 3429, Springer, Poitiers, France, 2005, pp. 79–91.
- [7] J.P. Guédon, B. Parrein, N. Normand, Internet distributed image information system, *Integrated Comput. Aided Eng.* 8 (2001) 205–214.
- [8] M. Katz, *Questions of Uniqueness and Resolution in Reconstruction from Projections*, *Lecture Notes in Biomathematics*, Springer, Berlin, 1977.
- [9] B.T. Kelley, V.K. Madiseti, The fast discrete Radon transform. I: theory, *IEEE Trans. Image Process.* 2 (3) (1993) 382400.
- [10] A. Kingston, S. Colosimo, P. Campisi, F. Autrusseau, Lossless image compression and selective encryption using a discrete radon transform, in: *IEEE International Conference on Image Processing*, September 16–19 2007, San Antonio, TX, USA.
- [11] F. Matúš, J. Flusser, Image representation via a finite Radon transform, *IEEE Trans. Pattern Anal. Mach. Intell.* 15 (10) (1993) 996–1006.
- [12] B. Meyer, P. Tischer, Glicbawls—gray level image compression by adaptive weighted least squares, in: *IEEE Computer Society, Proceedings of the Data Compression Conference*, Snowbird, Utah, USA, March 2001, p. 503.
- [13] Y. Ming, N. Bourbakis, An overview of lossless digital image compression techniques, in: *48th Midwest Symposium on Circuits and Systems*, vol. 2, 7–10, August 2005, pp. 1099–1102.
- [14] N. Normand, J.P. Guédon, O. Philippé, D. Barba, Controlled redundancy for image coding and high-speed transmission, in: R. Ansari, M.J. Smith (Eds.), *Proceedings of the SPIE Visual Communications and Image Processing*, vol. 2727, February 1996, pp. 1070–1081.
- [15] N. Normand, A. Kingston, P. Évenou, A geometry driven reconstruction algorithm for the Mojette transform, in: A. Kuba, L.G. Nyúl, K. Palágyi (Eds.), *Proceedings of the 13th International Conference on Discrete Geometry for Computer Imagery, Lecture Notes in Computer Science*, vol. 4245, Springer, Szeged, Hungary, 2006, pp. 122–133.
- [16] O. Philippé, J.P. Guédon, Correlation of the Mojette representation for non-exact image reconstruction, in: *Proceedings of Picture Coding Symposium*, vol. 1, Berlin, Germany, September 1997, pp. 237–241.
- [17] M.J. Weinberger, G. Seroussi, G. Sapiro, LOCO-I: a low complexity, context-based, lossless image compression algorithm, in: *Proceedings of the Data Compression Conference*, IEEE Computer Society, Los Alamitos, CA, USA, March–April 1996, pp. 140–149.
- [18] X. Wu, N. Memon, Context-based, adaptive, lossless image coding, *IEEE Trans. Comm.* 45 (4) (1997) 437–444.
- [19] X. Wu, N. Memon, Context-based lossless interband compression-extending CALIC, *IEEE Trans. Image Process.* 9 (6) (2000) 994–1001.
- [20] F. Zhai, Y. Eisenberg, A.K. Katsaggelos, Joint source-channel coding for video communications, in: A.L. Bovik (Ed.), *Handbook of Image and Video Processing*, second ed., Elsevier Academic Press, USA, 2005.

9.3 Perceptual DFT watermarking with improved detection and robustness to geometrical distortions

Perceptual DFT watermarking with improved detection and robustness to geometrical distortions

Matthieu Urvoy, Dalila Goudia, Florent Autrusseau

Abstract—More than ever, the growing amount of exchanged digital contents calls for efficient and practical techniques to protect intellectual property rights. During the past two decades, watermarking techniques have been proposed to embed and detect information within these contents, with four key requirements at hand: robustness, security, capacity and invisibility. So far, researchers mostly focused on the first three, but seldom addressed the invisibility from a perceptual perspective and instead mostly relied on objective quality metrics.

In this paper, a novel DFT watermarking scheme featuring perceptually-optimal visibility versus robustness is proposed. The watermark, a noise-like square patch of coefficients, is embedded by substitution within the Fourier domain; the amplitude component adjusts the watermark strength, and the phase component holds the information. A perceptual model of the Human Visual System (HVS) based on the Contrast Sensitivity Function (CSF) and a local contrast pooling is used to determine the optimal strength at which the mark reaches the visibility threshold.

A novel blind detection method is proposed to assess the presence of the watermark. The proposed approach exhibits high robustness to various kind of attacks, including geometrical distortions. Experimental results show that the robustness of the proposed method is globally slightly better than state-of-the-art. A comparative study was conducted at the visibility threshold (from subjective data) and showed that the obtained performances are more stable across various kinds of contents.

Index Terms—Watermarking, visibility, robustness, contrast sensitivity, Fourier, subjective experiment, Grubbs' test

I. INTRODUCTION

FACING the ever-growing quantity of digital documents transmitted over the internet, it is more than ever necessary for efficient and practical data hiding techniques to be designed in order to protect intellectual property rights. Watermarking is one such technique and has been extensively studied for the past two decades; applied to still images, it comes down to embedding an invisible information, called watermark, that can be retrieved and matched even when the watermarked image was attacked to some degree.

Four key requirements have been driving researchers in designing watermarking algorithms: the invisibility, the robustness, the capacity and the security. Any watermarking algorithm should ideally provide the best tradeoff between these four aspects. Lately, security has been widely studied

[1]; as for the invisibility, the robustness and the capacity, they influence each other and often must be addressed together. For instance, when the robustness is increased, the perceptual quality of the watermarked image inevitably decreases.

Yet, few works address the problem of invisibility from a perceptual perspective. Watson's visual models [2], [3], in particular, have been used to compute Just Noticeable Differences (JND) masks, thus providing a perceptual model of the visibility of the watermark [4], [5] which was embedded either in the Discrete Cosine Transform (DCT) or the Discrete Wavelet Transform (DWT) domains. More recently, another Human Visual System (HVS) model was used for DWT embedding [6]. Some of these methods are non-blind [4] or require side information to be transmitted in order to reconstruct the JND mask at the receiver [5], [6]. Some other perceptual methods are fully blind but they use heuristics instead of HVS models to derive JND masks [7], [8], [9].

Alternatively, some approaches based on statistical Objective Quality Metrics (OQMs) tune the embedding strength based on the computed quality score [10], [11] and do not require any side information at the receiver either. However, two main problems emerge: these metrics often provide wrong estimates of the perceived quality; more importantly, OQMs provide a continuous quality scale and were not intended to scenarios, such as watermarking, targeting the visibility threshold. In practice, the only reliable way to assess the invisibility is to conduct a subjective experiment in which observers are asked whether they can notice the watermark or not. In addition, most comparisons between watermarking techniques found in the literature are also based on OQMs – most often the Peak Signal to Noise Ratio (PSNR) – which, as will be shown in this paper, can be misleading.

Besides invisibility, the robustness is also a key aspect in watermarking. None of the aforementioned techniques are robust to common attacks such as geometrical distortions: they cannot efficiently withstand their desynchronization effects. Rotation-Scale-Translation (RST) invariant techniques have been proposed to address this issue through the use of appropriate transform domains. In [12], the watermark is embedded in the Fourier-Mellin domain, to the cost of an important algorithmic complexity. Later, in [13], log-polar mapping of the Fourier domain is used to make the watermark robust against RST attacks. Another log-polar Fourier embedding technique is proposed in [14]. Besides these log-polar mapping techniques – that commonly exhibit an important computational load –, some directly embed the watermark within the modulus of the Discrete Fourier Transform (DFT). In [15], the watermark is made rotation invariant thanks to its circular

Copyright (c) 2013 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org

M. Urvoy, D. Goudia and F. Autrusseau are with LUNAM Université, Université de Nantes, IRCCyN UMR CNRS 6597, Institut de Recherche en Communications et Cybernétique de Nantes, Polytech Nantes, rue Christian Pauc BP 50609 44306 Nantes Cedex 3.

Manuscript received ; revised .

shape. A similar technique is proposed in [16] where circular dots are embedded in the Fourier transformed coefficients.

The success of a watermarking algorithm not only depends on its embedding strategy, but also on its detection strategy. In particular, properly estimating the optimal detection threshold is crucial and thus has been extensively studied. Usually, the problem is stated in terms of two hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$, which respectively denote scenarios with and without embedded watermark in the considered host content. In [17], [18], statistical models for the signal and the watermark are used to derive a theoretical threshold. Typically, these are Gaussian or binomial models; the threshold is chosen so as to limit the probability of false detections. However, the statistics of both the watermark and the host image are likely to change with various parameters (e.g. watermark size, embedded message, etc), thus impacting the parameters of the fitted models, which in turn requires for the detection threshold to be re-evaluated. Moreover, if the statistics of both the host image and the watermark do not quite strictly fit the derived models, the obtained threshold is likely not to be optimal.

In this work, a robust, blind, substitutive and perceptual watermarking technique is proposed. Various characteristics of the HVS are used to determine and adjust the visibility level of the embedded watermark, thus resulting in an optimal invisibility versus robustness tradeoff. The proposed technique is designed to be robust against various kinds of attacks (including geometrical distortions). In addition, a subjective experiment is conducted in order to assess the invisibility of the watermark. Finally, the proposed detection strategy is efficient and adapts to the statistics of the watermark.

II. OVERVIEW OF THE CONTRIBUTION

Similarly to [15], [16], it is proposed to embed the watermark in the Fourier domain; the magnitude is used to control the energy of the watermark while the phase is used to hold its information. For best robustness, the watermarked coefficients are grouped into two symmetrical square patches, which can be expressed as a sum of sinewaves. Once displayed, this results in a combination of sine gratings at various visual frequencies and orientations: a perceptual model is used to adjust their amplitude at the visibility threshold. The watermarked image is then obtained by inverse Fourier transform. Sec. III describes the perceptual model, then Sec. IV focuses on the embedding process. Sec. V details the detection process.

The proposed method is designed to be robust against multiple kinds of attacks. Sec. VI investigates the robustness of the proposed method in comparison to [15] and [16]. Sec. VII reports the results of a subjective experiment assessing the visibility of the watermark. Sec. VIII then re-evaluate the robustness when the strength has been adjusted according to the observers feedback. Finally, Sec. IX validates the proposed approach in the context of low quality Print & Scan (P&S).

III. REACHING THE VISIBILITY THRESHOLD

Choosing the appropriate watermark strength is a delicate but crucial step while designing a watermarking technique. Over-estimations of the strength are likely to result in visible

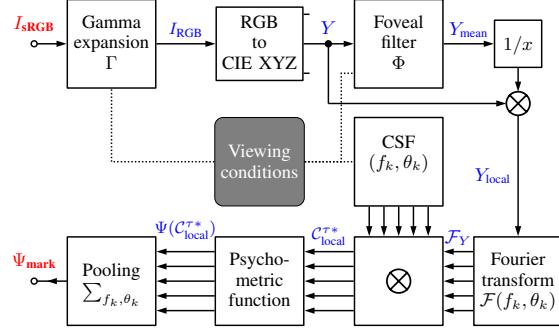


Fig. 1. The proposed HVS model estimates the probability Ψ_{mark} that a watermark – embedded in an image I_{sRGB} at visual frequencies f_k and orientations θ_k – is visible. Please refer to Section III for employed notations.

watermarks; on the contrary, under-estimations are likely to reduce the robustness performances. In this paper, it is proposed to use some properties of the HVS to automatically determine the perceptually optimal watermarking strength, at which the embedded watermark appears at the visibility threshold.

Computational models of the HVS providing estimates for the visibility of differences between an original and a distorted image have been proposed, such as in [19], [20], [21], [22] and [23]. They generally implement some of the following elements: (1) non-linear sensitivity to the amplitude of the luminance, (2) conversion to contrast, (3) contrast sensitivity to visual frequencies, (4) oblique effect, (5) subband decomposition into visual channels, (6) masking effects, (7) psychometric probability of detection and (8) error pooling. A survey of these models can be found in [24].

Practical applications, however, may not implement all of these HVS properties; visual channels, for instance, may be discarded due to their high complexity and their rather small contribution to the visibility estimates [23]. In a preliminary subjective experiment involving 4 expert observers, it was reported that the proposed watermark is first noticed in uniform image areas where masking effects do not occur (see Sec. VII-B for experimental details). Therefore, the proposed HVS model is simplified and discards the perceptual channel decomposition. Moreover, excluding the masking effects from the model underestimates the visibility threshold and thus the watermark strength: the invisibility requirement is met anyway. The proposed model is illustrated in Fig. 1.

A. Modeling viewing conditions

Typically, HVS models require both the viewed image and the viewing conditions to be input. Let $I_{\text{sRGB}}(x, y)$ denote an image to be watermarked, $0 \leq x < R_x$, $0 \leq y < R_y$, where R_x and R_y are respectively its horizontal and vertical resolutions. Let S_x and S_y denote I_{sRGB} 's displayed size in meters. The display illumination is noted L ; it is set to 280 cd.m^{-2} in the proposed model, a typical value for modern LCD monitors (slightly above the advised illumination level in [25]). The viewing distance d is generally expressed as a

multiplicative factor of the active image area's height, such that $d = \bar{d} \cdot S_y$, where $\bar{d}$ is the normalized viewing distance.

B. From pixel values to perceived contrast

Contrast sensitivity models generally take physical luminance in cd.m^{-2} as an input; digital images on the other hand are gamma encoded for display purposes. The proposed model assumes that a typical monitor ($\gamma = 2.2$) is used for display; the standard RGB (sRGB) color space [26] will thus be used. Gamma expansion (Γ) is first applied to I_{sRGB} in order to transform the original sRGB pixel values into linear RGB values: $I_{\text{RGB}} = \Gamma(I_{\text{sRGB}})$. In a second step, I_{RGB} is converted into the CIE XYZ color space, with D65 white point. Let I_{XYZ} denote the obtained image; its component Y is proportional to the physical luminance and will thus be used for contrast computation and watermark embedding.

Michelson's formula for contrast [27], when applied to a sine grating of peak amplitude A_{peak} , defines the contrast as

$$\mathcal{C}(x, y) = \frac{A_{\text{peak}}}{Y_{\text{mean}}(x, y)} \quad , \quad (1)$$

where Y_{mean} is the mean illumination of the area supporting the sine grating. In typical images, the illumination varies locally; Y_{mean} thus needs to be computed locally as well. As proposed in [28], a raised cosine filter Φ of diameter one degree of visual angle is applied to Y to obtain $Y_{\text{mean}}(x, y)$; this provides an estimate of the average foveal illumination over the entire image domain. The Fourier domain, which is later used to embed the watermark, is not spatially localized. Prior to applying such non-localized processing steps, it is proposed to normalize Y with respect to the local luminance Y_{mean} . The locally normalized luminance map $Y_{\text{local}}(x, y) = Y(x, y)/Y_{\text{mean}}(x, y)$ may then be input to Fourier computations. Within this normalized luminance space, the local contrast of a sine grating writes as

$$\mathcal{C}_{\text{local}} = A_{\text{peak}} \quad . \quad (2)$$

C. Contrast sensitivity

The sensitivity to contrast is influenced by numerous factors [29]. Intensive research has made available a large number of computational models known as Contrast Sensitivity Functions (CSFs), describing our sensitivity to contrast levels as a function of the visual frequency: the CSF returns the inverse of the threshold contrast above which a sine grating becomes visible. In [29], an accurate CSF is built by interlinking models for the optical Modulation Transfer Function, the photon noise, the lateral inhibition, and the neural noise. Some aspects of this model, though, are not theoretically supported; authors in [30] addressed these issues and incorporated better models for cone and midget ganglion cell densities. However, the resulting models are computationally intensive and do not fit practical applications such as ours.

CSFs with medium to low complexities have been made available as well. In [31], the CSF proposed by Mannos and Sakrison features a single parameter, the visual frequency. However, such a model lacks important factors such as light adaptation and stimulus size. In, the CSF proposed by Daly

features key factors, but was designed to be used within his Visible Differences Predictor (VDP) quality model and should be used with caution in other contexts. In [32], finally, Barten provides a simplified formula for his initial CSF, which also incorporates the oblique effect and the influence of the surround luminance. In this paper, the proposed Fourier watermark embedding technique (as will be further discussed in section IV), modifies frequency coefficients whose orientations are oblique. For this reason, it is necessary to implement the oblique effect as our sensitivity varies with the orientation of the visual pattern. Barten's simplified CSF formula [32], at binocular viewing, will thus be used in the proposed model:

$$\begin{aligned} \text{CSF}(f, \theta) = & 5200 \cdot e^{-0.0016 \cdot f^2 \left(1 + \frac{100}{L}\right)^{0.08}} \\ & \cdot \left(1 + \frac{144}{\Theta^2(I)} + 0.64 \cdot f^2 \cdot \left(1 + 3 \sin^2(2\theta)\right)\right)^{-0.5} \\ & \cdot \left(\frac{63}{L^{0.83}} + \frac{1}{1 - e^{-0.02f^2}}\right)^{-0.5} \end{aligned} \quad (3)$$

where f is the visual frequency in cycles per degree (cpd), L is the adaptation luminance in cd.m^{-2} and is assumed to be equivalent to the display illumination (see Sec. III-A), $\Theta^2(I)$ is the square angular area of the displayed image I in square visual degrees, and θ the orientation angle.

From Eqs. (2) and (3), one may now obtain the threshold amplitude $A_{\text{peak}}^\tau(f, \theta)$ of a sine grating

$$A_{\text{peak}}^\tau(f, \theta) = \mathcal{C}_{\text{local}}^\tau(f, \theta) = \frac{1}{\text{CSF}(f, \theta)} \quad (4)$$

where $\mathcal{C}_{\text{local}}^\tau$ is the local contrast threshold.

D. Psychometric function

The psychometric function Ψ is typically used to relate the parameter of a physical stimulus to the subjective responses. When applied to contrast sensitivity, Ψ may describe the relationship between the contrast level and the probability that such contrast can be perceived [20]. In the proposed approach, Daly's Weibull parametrization [20] is used:

$$\Psi(\mathcal{C}_{\text{local}}^{\tau*}) = 1 - e^{-(\mathcal{C}_{\text{local}}^{\tau*})^\beta} \quad , \quad (5)$$

where $\mathcal{C}_{\text{local}}^{\tau*} = \mathcal{C}_{\text{local}}/\mathcal{C}_{\text{local}}^\tau$ is the ratio between the locally normalized contrast and its threshold value given in Eq. (4). β is the slope of the psychometric curve; its value is usually obtained by fitting from experimental data. Typical values for β range from 1.3 (contrast discrimination) up to 4 (contrast detection) [33]. Moreover, the psychometric function is typically applied locally, contrary to our model which applies it in the Fourier domain, hence globally. For both these reasons, it is proposed to use $\beta = 2$, a rather low value which provides a large transition area between invisible and visible domains.

E. Watermark frequency pooling

In this paper, the watermark is embedded into multiple Fourier coefficients, which results into the superimposing of multiple sine gratings G_k , $0 \leq k < N$, with visual frequencies

f_k and orientations θ_k . The CSF solely provides an estimate for the visibility of a single sine grating; a summation model is thus required to estimate the combined visibility level of all embedded gratings. Numerous models have been proposed [23], [20]. Recent findings suggest that Minkowski summation [23] may not be adequate [34]; the proposed model will thus use probability summation as in [20].

Let $C_{\text{local}}^{\tau*}(f_k, \theta_k)$ denote the contrast threshold ratio of G_k . As in [20], the probability Ψ_{mark} that the watermark is perceived by an observer computes as

$$\Psi_{\text{mark}} = 1 - \prod_{k=0}^{N-1} \left(1 - \Psi(C_{\text{local}}^{\tau*}(f_k, \theta_k)) \right), \quad (6)$$

where Ψ_{mark} is set according to Eq. (5): $1 - e^{-(1)^\beta} = 0.632$.

Let us now work our way back from the entire watermark to individual gratings. In the proposed approach, it is assumed that all embedded gratings share the same probability of detection $\Psi(C_{\text{local}}^{\tau*}(f_k, \theta_k)) = \psi, \forall k \in [0, N[$. This probability can be derived from Ψ_{mark} by solving Eq. (6):

$$\psi = 1 - (1 - \Psi_{\text{mark}})^{1/N}. \quad (7)$$

The corresponding normalized contrast, which is equal to the optimal grating peak amplitude, may thus be obtained by inversion of the psychometric function (Eq. (5))

$$A_{\text{peak}}(f_k, \theta_k) = C_{\text{local}}^{\tau}(f_k, \theta_k) \cdot \left[-\ln \left((1 - \Psi_{\text{mark}})^{1/N} \right) \right]^{1/\beta} \quad (8)$$

IV. MAGNITUDE-PHASE SUBSTITUTIVE EMBEDDING

The proposed embedding technique relies on both the amplitude and the phase Fourier components to modulate a binary watermark. The amplitude allows to adjust the watermark strength whereas the phase holds the binary information.

Let an N-bits binary sequence $\mathcal{M}_{1..N}$ be the message to embed. A binary watermark $\mathcal{W}(i, j), 0 \leq i < M, 0 \leq j < M$, is generated from $\mathcal{M}_{1..N}$ using a PRNG and arranged into a $M \times M$ matrix.

Let $\mathcal{F}_Y(u, v)$ denote the Fourier transform of Y_{local} over frequency domain $\Omega = \{(u, v) : -R_x/2 \leq u < R_x/2, -R_y/2 \leq v < R_y/2\}$, where u and v are the horizontal and vertical frequencies. Let $\Omega_W^+ = \{(u, v) : u_w \leq u < u_w + M, v_w \leq v < v_w + M\}$ and $\Omega_W^- = \{(u, v) : (-u, -v) \in \Omega_W^+\}$ be two subsets of Ω within which the watermark will be embedded; u_w and v_w are called watermark modulation frequencies. The watermarked spectrum is obtained by substitution as follows:

$$\begin{aligned} \widetilde{\mathcal{F}}_Y(u, v) = & \\ \begin{cases} \alpha/2 \cdot A_{\text{peak}}^*(u, v) \cdot e^{\pi \cdot \mathcal{W}(u-u_w, v-v_w)}, & (u, v) \in \Omega_W^+ \\ \alpha/2 \cdot A_{\text{peak}}^*(u, v) \cdot e^{\pi \cdot \mathcal{W}(-u-u_w, -v-v_w)}, & (u, v) \in \Omega_W^- \\ \mathcal{F}_Y(u, v), & \text{elsewhere} \end{cases} \end{aligned} \quad (9)$$

where α is a weighting factor that controls the watermark energy relatively to the predicted visibility level. $A_{\text{peak}}^*(u, v)$

equals to $A_{\text{peak}}(f(u, v), \theta(u, v))$ and denotes the embedded grating's optimal amplitude obtained in Eq. (8) for discrete frequency (u, v) ; $f(u, v)$ and $\theta(u, v)$ respectively denote its visual frequency and orientation. In this paper, the values for u_w and v_w are both set to 80% of the highest horizontal and vertical Fourier frequencies, thus ensuring low sensitivity to watermarked frequencies, which induces high threshold amplitudes, hence high watermark energy. Lower values for u_w and v_w were also experimented (20% of the maximum frequency), with similar robustness performances.

The proposed embedding scheme given in Eq. (9) is a substitutive technique. Here, the choice for amplitude substitution is straightforward as it allows precise control over the energy of the watermark. Moreover, as can be seen in Eq. (9), phase substitution is performed in order to encode the binary watermark information: zeros are coded as null phases, ones as π phases. Eventually, the watermarked image is obtained by inverse Fourier transform, local luminance de-normalization, and transform into the original color space.

V. TWO-STEPS BLIND WATERMARK DETECTION

Let $\widehat{I}_{\text{sRGB}}$ be a supposedly watermarked image; it is assumed to have the same resolution and dimensions as the original image I_{sRGB} , otherwise it is re-scaled. Let $Y_{\text{local}}(\widehat{I})$ denote its locally normalized luminance map (see Section III-B).

The proposed detection algorithm is blind, thus neither requires the original image I_{sRGB} nor the original Fourier coefficients nor the CSF weighting coefficients. It is performed in two steps. At first, Template Matching (TM) is used to compute a 2D correlation matrix between the Fourier transform of $Y_{\text{local}}(\widehat{I})$ – denoted $\widehat{\mathcal{F}}_Y$ – and the watermark to be detected $\mathcal{W}(i, j)$. Then, outlier detection is performed to assess the presence of matching locations within the correlation matrix, thus providing a binary decision (zero-bit watermarking).

A. A 2D-correlation algorithm based on template matching

Often, detection schemes are solely based on a correlation coefficient that is compared to a predefined threshold value (e.g. [4], [15]). Detectors based on cross-correlation, cross-covariance and their optimized variants [35] generally perform better, at the cost of an increased computational load. Most often, however, one-dimensional correlation is performed [16], while the searched pattern is likely to be two-dimensional. Moreover, geometric distortions affect the location and shape of the watermark in the Fourier domain.

Not only does 2D cross-correlation account for the two-dimensional structure of the embedded watermark, but the displacement of its correlation peaks allows to estimate the geometrical distortions undergone by the input image $\widehat{I}_{\text{sRGB}}$. Therefore, it is proposed to perform two-dimensional TM similarly to [14]. Furthermore, no additional re-synchronisation watermark is required (contrary to [36], [37], [38]); the TM is able to retrieve the watermark as long as it is located, even partially, within the searched area.

Typically, TM involves: (1) a supporting signal $\mathcal{S}$, (2) a searched template $\mathcal{T}$, (3) a search area Ω_S and (4) a matching

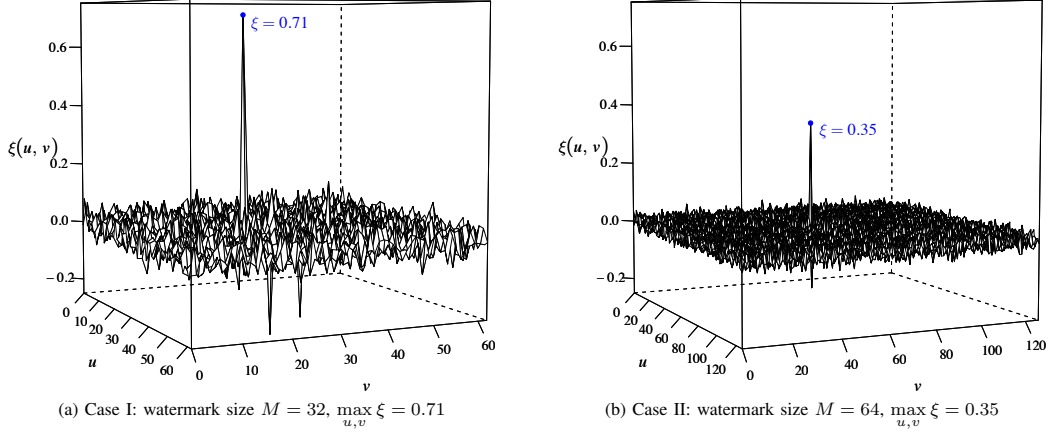


Fig. 2. Correlation matrices $\xi(u, v)$: how the dynamic range of the detection peaks and the surrounding noise varies with the watermark size M .

criterion ρ . In our case, $\mathcal{S}$ is derived from $\widehat{\mathcal{F}}_Y$:

$$\mathcal{S}(u, v) = \begin{cases} \left| \widehat{\mathcal{F}}_Y(u, v) \right|, & -\pi/2 < \arg(\widehat{\mathcal{F}}_Y(u, v)) \leq \pi/2 \\ -\left| \widehat{\mathcal{F}}_Y(u, v) \right|, & \pi/2 < \arg(\widehat{\mathcal{F}}_Y(u, v)) \leq 3\pi/2 \end{cases} \quad (10)$$

which encompasses both the Fourier energy and the phase information into a single signed signal. This stands in contrast to Kang *et al.* [14], who perform phase-only correlation to cope with the large dynamic range of the energy of the watermarked coefficients. In our scenario, this range depends on the CSF; therefore it is rather low, which allows us to take the energy level into account as well. The template $\mathcal{T}$ is derived from the binary watermark $\mathcal{W}$ as follows:

$$\begin{aligned} \forall (i, j) \in [0, M]^2, \\ \mathcal{T}(i, j) = A_{\text{peak}}^* (u_w + i, v_w + j) \cdot (2 \cdot \mathcal{W}(i, j) - 1) \end{aligned} \quad (11)$$

which maps $\mathcal{W}$'s binary values from $\{0, 1\}$ to $\{-1, 1\}$ and accounts for the optimal amplitude A_{peak}^* of the embedded gratings (see Eq. (9)). The search area Ω_S is centered around the original watermark modulation frequencies, such that $\Omega_S = \{(u, v) : u_w - M < u < u_w + M, v_w - M < v < v_w + M\}$. Finally, Pearson's correlation coefficient serves as matching criterion ρ , so that the final 2D correlation matrix is given by $\xi(u, v) = \rho(\mathcal{T}, \mathcal{S}(\Omega_{u,v}^{M \times M}))$, where $(u, v) \in \Omega_S$ and $\Omega_{u,v}^{M \times M} = \{(u', v') : u \leq u' < u + M, v \leq v' < v + M\}$.

B. Peak detection

Typically, watermark detection is granted when the correlation score exceeds some given threshold value. It is generally obtained by measuring [15] (experimentally) or estimating [18] (theoretically) the distribution of the values of the correlation matrix under true or false detection scenarios. On the one hand, theoretical approaches are based on statistical analysis, and thus make assumptions that may not hold in practice. On

the other hand, experimental thresholds are only valid within the scope of the test signals. In any case, the obtained threshold is constant, and might not be optimal.

In contrast, the dynamic range of a correlation matrix varies with numerous parameters. To illustrate this, the proposed embedding method was used to watermark the Lena image. The watermarked image was then rotated (1.5°) and blurred (Gaussian noise, $\sigma = 2.0$) to simulate an attack. Fig. 2 plots the obtained correlation matrices, for $M = 32$ (Fig. 2a) and $M = 64$ (Fig. 2b). Although the detection peaks are obvious in both cases, their amplitude, as well as the amplitude of the surrounding noise, differ significantly. In other words, the detection decision should be driven by the relative difference in amplitude between the correlation peak(s) and the surrounding noise. In Fig. 2, for the same image, the same embedding method, and against the same attack, a fixed detection threshold (e.g. 0.4) would properly detect the correlation peak from Fig. 2a and miss the one from Fig. 2b.

Instead, detection peaks may be seen as outliers. Numerous methods for outlier detection have been proposed in the context of statistical analysis [39]. Grubbs' test [40] is one such method, both robust and low computational, and solely requires that input correlation matrices are (approximately) normally distributed. This assumption is quite common [18]; Kolmogorov-Smirnov tests further confirmed Gaussianity of the correlation matrices in practical experiments. The proposed detection method features an iterative implementation of Grubbs' test that removes one outlier value at a time, up to a predefined maximum number of outliers. Alternatively, the performances of the Extreme Studentized Deviate (ESD) test were also investigated and proved to be identical in practical experiments. Such an approach prevents us from fixing the threshold, and adapts the detection to the observed correlation matrix. Moreover, it can be applied to any correlation-based watermark embedding method. Last but not least, the test's significance level α_G can be used to control the tradeoff between detection capacity and false alarm rate. High (resp. low) values bring higher (resp. lower) True Positive (TP) and

False Positive (FP) rates. In terms of complexity, the proposed method proves to be very fast: with 1024×1024 images and 64×64 watermarks, both detection and embedding steps took approximately 0.2s to perform on a mid-2012 Mac Book Pro (2.3 GHz Intel Core i7 processor, 8Gb of RAM), of which 0.26ms was needed to apply Grubbs' test.

VI. EXPERIMENTAL RESULTS

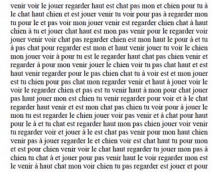
Most experimental results were obtained from dataset $\mathcal{D}_a$ containing nine images with resolution 512×768 , including six natural color images, two text images and one cartoon color image. They are listed in Table I. When a larger number of images were required (e.g. threshold selection), a second dataset $\mathcal{D}_b$ featuring 1000 natural color images was used. The results are compared to those of [15] and [16]. Datasets, watermarked images and additional data are available online¹. Fig. 3 shows examples of watermarked images at default embedding strength α_0 ($\alpha = 1$, see Eq. (9)).

TABLE I
DATASET $\mathcal{D}_a$ USED IN THE EXPERIMENTS

#	Name	Type	Description
1	ED	Cartoon	Elephant dream (652 nd frame, cropped)
2-7	k03 – k23	Natural	Images 3, 7, 15, 20, 21 and 23 from [41]
8	dsc	Text	Scanned payslip, cropped
9	wilk	Text	Wilkins text [42]



(a) Cartoon image ED



(b) Text image wilk

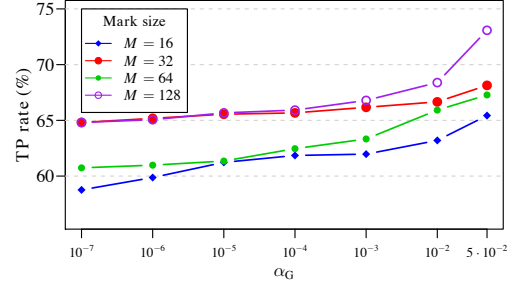
Fig. 3. Examples of atypical images (cartoon and text): watermarked images.

A. Grubbs significance level

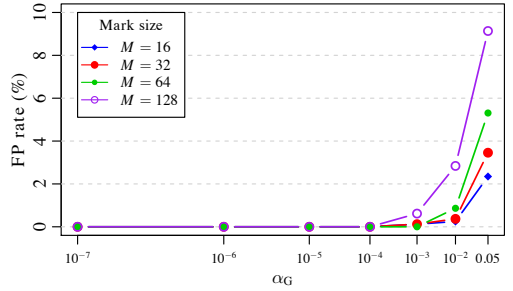
In order to determine the optimal significance level for Grubbs' test, $\mathcal{D}_a$'s images were watermarked with the proposed algorithm. 90 attacks, from the Stirmark benchmark [43], were then applied to both original (hypothesis $\mathcal{H}_0$) and watermarked (hypothesis $\mathcal{H}_1$) images. Detection was finally run on each resulting image with varying values for α_G . This process was repeated for several watermark sizes $M \in \{16, 32, 64, 128\}$. The results obtained for both $\mathcal{H}_0$ and $\mathcal{H}_1$ scenarios are shown in Fig. 4.

As can be seen from Fig. 4a, the TP rate slightly varies with α_G , and mostly with the size of the embedded watermark. As for the FP rate, it also increases with the size of the watermark, but exponentially increases with α_G . Therefore, the value of α_G may be adjusted according to the desired false alarm rate. In the next experiments, we set M to 64 – in order to provide high detection performances – and α_G to $1e^{-4}$ – in order to ensure a low FP rate (null in experimented images) –.

¹<http://www.irccyn.ec-nantes.fr/~autrusse/DFTWmking/>



(a) $\mathcal{H}_1$: TP detection rate versus α_G



(b) $\mathcal{H}_0$: FP detection rate versus α_G

Fig. 4. True Positive and False Positive detection rates against Stirmark attacks on $\mathcal{D}_a$: influence of the watermark size (M) and α_G

B. False alarms and true detections

To ensure optimal detection performances, the gap between the detection scores in watermarked images (hypothesis $\mathcal{H}_1$) and those in un-marked images (hypothesis $\mathcal{H}_0$) need to be as large as possible. Detection scores were collected on dataset $\mathcal{D}_b$ (1000 images). Fig. 5 plots the histograms (solid lines) of the detection scores in both scenarios $\mathcal{H}_0$ and $\mathcal{H}_1$, along with the fitted Gaussian models (colored areas). As can be seen, $\mathcal{H}_0$ and $\mathcal{H}_1$ distributions are clearly disjoint in the proposed technique, similarly to [15]. Conversely, Fig. 5 shows that the distributions of $\mathcal{H}_0$ and $\mathcal{H}_1$ are much closer in [16], thus resulting in respectively higher and lower FP and TP rates.

C. Robustness to attacks

The robustness of the proposed watermarking algorithm to various attacks was measured with the Stirmark benchmark [43]. All 90 Stirmark attacks were applied on dataset $\mathcal{D}_a$, thus resulting in 810 attacked images. The robustness is reported in terms of maximum of correlation against experimented attacks; yet, the proposed detection scheme depends on the output of Grubbs' test (absence/presence of outliers). Therefore, the robustness of the proposed algorithm is reported in terms of percentage of images with a positive detection (detection rate). The obtained results are compared with [15] and [16], where detection rates correspond to the percentage of images for which correlation is higher than the predefined thresholds.

The obtained results are plotted in Fig. 6: details for each attack are available online. For each attack, the dark bar(s) correspond to the best performing algorithm(s) and the light

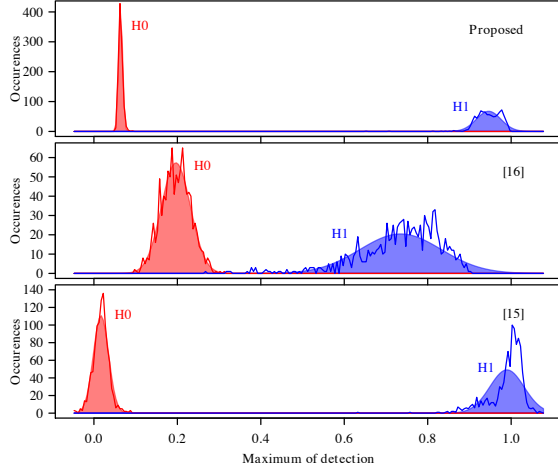


Fig. 5. $\mathcal{H}_0$ (original) versus $\mathcal{H}_1$ (marked): comparison with [15] and [16] on dataset $\mathcal{D}_b$. Solid lines represent the experimental histograms, and colored areas represent the fitted Gaussian models.

bar(s) to the least performing algorithm(s). Overall, it appears that the proposed technique performs best with an average detection rate of 62.5%, while [15] and [16] respectively reach 52.1% and 53.1%. It should be noted that [16] performs well against most kinds of attacks, although seldom performing best out of the three algorithms. The proposed technique is especially robust to geometric distortions such as rotations (despite the fact that [15] and [16] embed a circular watermark contrary to the proposed approach) and shearing. Still, only [16] withstands severe rotations and cropping. In addition, the proposed technique is nearly as robust as [15] to scaling and filtering operations, and nearly as robust as [16] to cropping. Moreover, it is significantly more robust to random bending, which randomly adds local geometrical distortions, as can be induced by severe P&S effects. On the downside, it appears that the proposed technique is not robust to severe JPEG compression, although this can be explained by the fact that the chosen watermark modulation frequencies are high and thus very likely to be strongly affected by compression. Better robustness against coding artifacts could be achieved by simply shifting the embedded watermark towards lower frequencies. Further information can be found in Table II which lists average detection rates amongst groups of attacks.

VII. SUBJECTIVE EXPERIMENT

Previously, we assessed the robustness of the proposed watermarking scheme at default strength $\alpha_0 = 1$ and compared it to [15], [16] at suggested strengths. Yet, one may wonder whether these strengths are equivalent in terms of visibility, hence whether previous results are truly comparable. For this reason, a subjective experiment was conducted to determine the optimal strengths at which the watermarks appear at visibility threshold. Section VIII will re-assess the robustness of the compared algorithms at the obtained strengths.

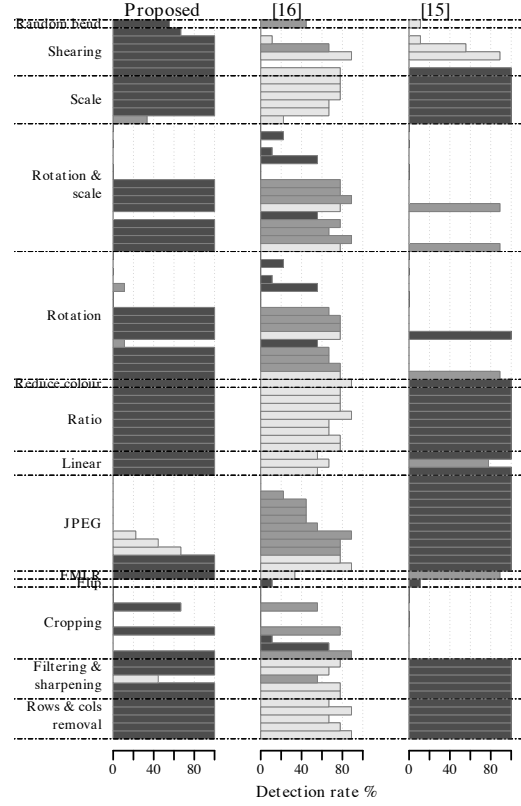


Fig. 6. Robustness to Stirmark attacks at default embedding strength (α_0): detection rates on dataset $\mathcal{D}_a$

TABLE II
STIRMARK BENCHMARK AT DEFAULT EMBEDDING STRENGTH: AVERAGE DETECTION RATES^a (%) AMONGST GROUPS OF ATTACKS ON DATASET $\mathcal{D}_a$.

Alg.	Overall	Rows & cols removal	Filtering & sharpening	Cropping	JPEG	Linear	Ratio	Rotation	Rotation & scale	Scale	Shearing	Random bend
Prop.	62.5	100	88.9	29.6	27.8	100	100	51.4	50.0	88.9	94.5	55.6
[16]	53.1	77.8	71.1	33.3	51.9	59.3	76.4	45.8	48.6	64.8	40.7	44.4
[15]	52.1	100	100	0.0	100	92.6	100	11.8	11.1	100	42.6	11.1

^a In each column, the bold value designates the best performing algorithm.

A. Apparatus and methodology

Observers were seated in a standardized room [44] and watched experimental stimuli on a 40" TV Logic LVM401 display, with 1920×1080 resolution. Screen brightness was set to 200 cd.m^{-2} . Calibration was performed with an Eye One Pro luminance meter: gamma correction was set to 2.2 and white point to 6600 K. Room illumination was set to 30 cd.m^{-2} behind the screen, hence 15% of the perceived screen brightness. Finally, the viewing distance was set to six times

the height of the presented images (as in [6]).

The Two Alternative Forced Choice (2AFC) protocol was used as it is best suited to estimate visibility thresholds. 37 naive observers were recruited to participate to the experiment; screening tests ensured that they had a perfect visual acuity (Snellen chart) and no color deficiencies (Ishihara plates). The 2AFC methodology [45] (chap. 8, p. 258) was used to assess the visibility of the embedded watermarks. Nine anchor images were first presented for the observers to familiarize themselves with the sources and the type of degradations. All 243 images were then shown randomly to each observer. Observers were showed each pair of images (the watermarked image and its original version) during 10 seconds, at which point images were hidden. Observers were forced to vote for either one or the other image they thought was containing the watermark, before they could resume the experiment.

B. Stimuli set

Source images were taken from dataset $\mathcal{D}_a$. The nine images were then watermarked with three different algorithms (proposed, [15] and [16]). This watermarking was repeated for various embedding strengths $\alpha_i = w_i \cdot \alpha_0$, $1 \leq i \leq 9$, where w_i are weighting coefficients, and α_0 is the default strength. $\alpha_0 = \alpha = 1$ according to Eq. (9) in the proposed algorithm; $\alpha_0 = 0.3$ in [15]; finally, α_0 is image-adaptive in [16].

In 2AFC experiments, it is especially important to make sure that presented stimuli span the entire visibility range, therefore from completely invisible to severely distorted: the obtained results can thus be best fitted to the expected psychometric curve. Table III lists the values for w_i , whose suitability was assessed by 4 expert observers who performed the proposed experiment prior to naive observers. In total, 243 images were included in the subjective experiment.

TABLE III
DATASET $\mathcal{D}_a$: WATERMARK STRENGTH WEIGHTING COEFFICIENTS w_i

Algorithm	Img.	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9
Proposed	1-7	0.13	0.25	0.50	0.75	1.00	1.50	2.00	4.00	6.00
	8-9	0.25	0.50	0.75	1.00	1.25	1.38	1.50	2.00	4.00
[15]	1-7	0.06	0.13	0.19	0.25	0.38	0.50	1.00	2.00	3.00
	8-9	0.03	0.06	0.09	0.13	0.15	0.17	0.25	0.50	1.00
[16]	1-7	0.25	0.37	0.50	0.75	1.00	2.00	5.00	10.00	20.00
	8-9	0.50	0.75	1.00	2.00	3.00	5.00	7.00	10.00	20.00

C. Analysis

Let $\mathcal{V}(i, j)$ denote the vote of the j^{th} observer at the i^{th} image: it is equal to 1 if the watermarked image was correctly identified, and to 0 otherwise. For each of the 243 watermarked images, the percentage of observers who correctly identified the watermarked image $\mathcal{R}(i) = 1/N_{\text{obs}} \cdot \sum_{j=1}^{N_{\text{obs}}} \mathcal{V}(i, j)$ is called the 2AFC detection rate, where $N_{\text{obs}} = 37$ is the number of observers. A clustering-based analysis revealed 6 observers whose votes were inconsistent, these were removed from subsequent analysis, thus reducing N_{obs} to 31.

Fig. 7 plots the 2AFC detection rates $\mathcal{R}(i)$ obtained for image ED from dataset $\mathcal{D}_a$ (circles) against the strength weighting coefficients w_i . Typically, the visibility threshold is assumed to be obtained for a 2AFC score of 75% [45]. Experimental data are fitted to a sigmoidal Weibull curve to obtain the corresponding psychometric curve, from which the threshold value and therefore optimal watermarking strength α^* can be obtained. Fig. 7 also shows the corresponding Weibull fit (solid line), along with the 75% threshold which is reached for strength $\alpha^* = w^* \cdot \alpha_0$, with $w^* = 0.62$.

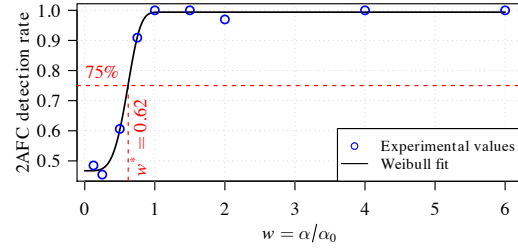


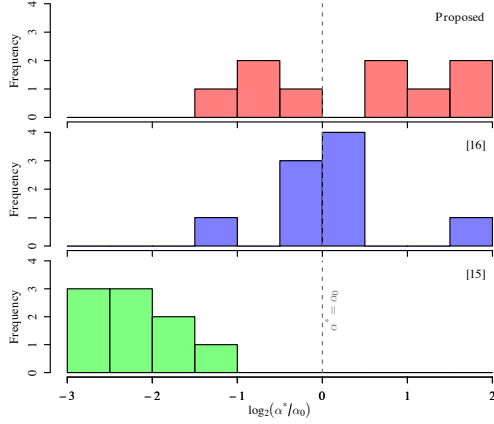
Fig. 7. Proposed method: 2AFC analysis on image ED - dataset $\mathcal{D}_a$

The same fitting process is repeated for each image and each algorithm, therefore resulting in 27 perceptually optimized strengths $\alpha^* = w^* \cdot \alpha_0$. Table IV lists the obtained weights w^* . Fig. 8 plots w^* 's distributions for each of the three algorithms; for plotting purposes, the histogram was computed on the binary logarithm of w^* : $\log_2(w^*) = 0$ therefore corresponds to $\alpha = \alpha_0$, $\log_2(w^*) = 1$ to $\alpha = 2 \cdot \alpha_0$ and $\log_2(w^*) = -1$ to $\alpha = 0.5 \cdot \alpha_0$. These strengths α^* , in the scope of experimented images, optimize the robustness versus visibility tradeoff. The further away the default strength α_0 is from its optimum α^* , the more (in)visible is the watermark. When the default strength α_0 is greater than its optimum α^* (i.e. $\log_2(w^*) < 0$), the watermark is likely to become visible. Conversely, when α_0 is less than α^* (i.e. $\log_2(w^*) > 0$), the watermark is invisible but the strength is not maximized, hence likely resulting in a loss of robustness.

TABLE IV
DATASET $\mathcal{D}_a$: PERCEPTUALLY OPTIMAL STRENGTH WEIGHTING COEFFICIENTS w^*

Algorithm	ED	k03	k07	k15	k20	k21	k23	dsc	wilk
Proposed	0.62	2.10	2.98	1.46	0.46	1.96	3.00	0.84	0.55
[15]	0.32	0.48	0.30	0.20	0.18	0.14	0.22	0.14	0.16
[16]	0.88	1.08	0.72	1.01	1.04	0.49	0.96	3.64	1.30

Fig. 8 shows that algorithm [15] overestimates the watermarking strength by factors ranging from 2 to 8: the embedded watermark is always visible at default strength. The other two algorithms perform much better in this regard: the distribution of their w^* are nearly centered around the default position; misestimation factors remain within the $[0.46; 3.64]$ range (Table V). In addition, both of these algorithms are more likely to under-estimate the strength than to over-estimate it, which is preferable as it keeps the watermark invisible.

Fig. 8. Distribution of perceptually optimized weights $w^* = \alpha^*/\alpha_0$ on $\mathcal{D}_a$.TABLE V
PERCEPTUALLY OPTIMAL STRENGTH WEIGHTING FACTORS: A SUMMARY.

Algorithm	Optimal weight w^*			Estimation Adequacy (%)	
	min.	max.	average	over-est.	under-est.
Proposed	0.46	3.00	1.55	44.4	55.6
[15]	0.14	0.48	0.24	100.0	0.0
[16]	0.49	3.64	1.24	44.4	55.6

Above all, these results tend to show that the proposed HVS model provides good estimates for the visibility threshold. While [16] does not feature any kind of psychophysical model, the embedded watermark also nears the visibility threshold.

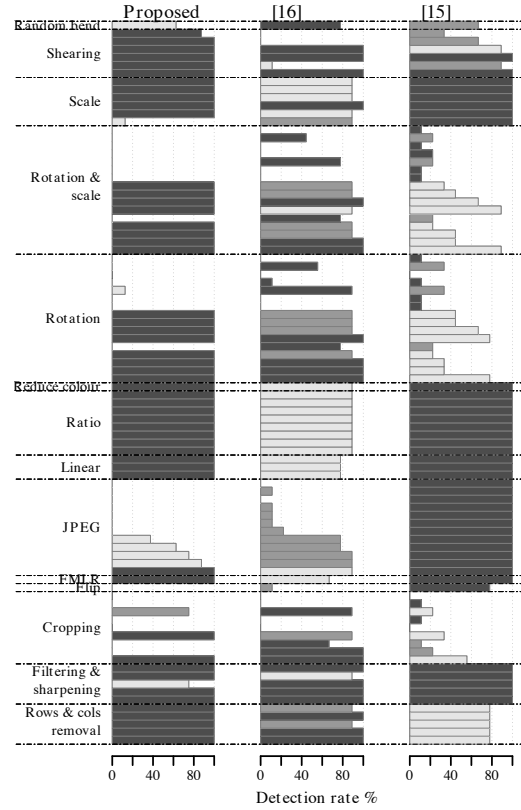
VIII. PERFORMANCES AT THE VISIBILITY THRESHOLD

Here, the algorithms are re-evaluated with the obtained perceptually optimized strengths α^* . Images from dataset $\mathcal{D}_a$ are thus watermarked using α^* instead of α_0 .

A. Robustness to attacks

The robustness benchmark performed in section VI-C was thus rerun at α^* . The obtained results are plotted in Fig. 9; again, the darkest bar(s) correspond to the best performing algorithm(s) and the light bar(s) to the least performing algorithm(s). Table VI summarizes average detection rates amongst groups of attacks.

In average, the three algorithms now feature similar robustness percentages: [15] and [16] tie at 65.3%, and the proposed method reaches 63.2%. In the proposed method, the robustness to individual attacks is very similar to the scores obtained at default strength (see Fig. 6). Still, the robustness to filtering, shearing and random bending moderately improves with optimal strengths. In [16], the robustness is significantly improved for all kinds of attacks; it even slightly outperforms our method against JPEG, rotation, and the rotation & scaling. As a consequence, the introduction of a perceptual model into [16] would be very likely to significantly enhance its robustness instead of targeting a given PSNR.

Fig. 9. Robustness to Stirmark attacks at visibility threshold (α^*): detection rates on dataset $\mathcal{D}_a$ TABLE VI
STIRMARK BENCHMARK AT VISIBILITY THRESHOLD: AVERAGE DETECTION RATES^a (%) AMONGST GROUPS OF ATTACKS COMPARED TO BASELINE RATES^b

Alg.	Overall	Rows & cols removal	Filtering & sharpening	Cropping	JPEG	Linear	Ratio	Rotation	Rotation & scale	Scale	Shearing	Random bend
Prop.	63.2	100	95.0	30.6	30.2	100	100	50.8	50.0	85.4	97.9	62.5
Diff.	+0.7	-	+6.1	+0.9	+2.4	-	-	-0.6	-	-3.5	+3.5	+6.9
[16]	65.3	95.6	97.8	49.4	40.7	77.8	88.9	61.8	59.0	90.7	51.9	77.8
Diff.	+12.2	+17.8	+26.7	+16.1	-11.1	+18.5	+12.5	+16.0	+10.4	+25.9	+11.1	+33.3
[15]	65.3	77.8	100	18.5	100	100	100	33.3	35.4	100	79.6	66.7
Diff.	+13.0	-22.2	-	+18.5	-	+7.4	-	+21.5	+24.3	-	+37.0	+55.6

^a In each column, the bold value designates the best performing algorithm.

^b Rows "Diff." indicate the difference in detection rate w.r.t. baseline rates obtained at default embedding strengths – see Table II.

Surprisingly in [15], there is no overall loss in robustness despite the strong reduction in the embedding strength ($\alpha^* = 0.24 \cdot \alpha_0$ in average), but on the contrary a drastic improvement. This can be explained by the fact that the

normalized correlation value in [15] is inversely proportional to the embedding strength. The detection threshold from [15] being fixed, the FP rate is likely to rise.

Besides high detection capabilities, providing stable performances over a rather large range of embedding strengths is also essential. In the proposed technique, variations in the embedding strength do not significantly affect the robustness thanks to the efficiency of the perceptual model that accurately set the watermark strength with respect to its visibility threshold. Conversely in [15] and [16], small variations in the embedding strength strongly affect their robustness, in other words, unstable performances.

B. Objective quality

Among the three tested algorithms, three different strategies are observed. In the proposed method, a perceptual model automatically adjusts the watermark strength at the visibility threshold. In [15], the embedding strength is fixed and is thus independent of the image content. In [16], the embedding strength is progressively increased, so as to reach a target PSNR of 40 dB. Although it is widely accepted that the PSNR does not reflect image quality, it is still a recurrent quality metric when assessing the visibility of a watermark.

Fig. 10 plots the 2AFC detection rate as a function of the PSNR for image k21. The numerical values next to each point within the figure gives the detection performances for various w_i against 90 Stimark attacks (in percents). At a constant PSNR of ~ 37.9 dB (see the vertical dimension line **A** in Fig. 10), the proposed watermark is invisible (2AFC percentage of 27.3%), while the watermark from [15] is highly visible (2AFC percentage of 93.9%). Focusing now on data points close to the visibility threshold (see the horizontal dimension line **B** in Fig. 10), it appears that the obtained PSNRs significantly differ between [16] (43.2 db) and the proposed method (29.3 dB), while the watermark in [16] is slightly above the visibility threshold (81.8%) and the proposed watermark is slightly under the threshold (69.7%). In addition, the robustness of the proposed method is higher (62%) than the one of [16] (56%). Therefore, assessing the quality of experimented algorithms, from their PSNRs only, would conflict with the subjective results (ground truth), which in turn would lead to erroneous conclusions.

Concerning the robustness performances, one can see that experimented algorithms behave differently with varying w_i . As it was already seen in section VIII-A, slight variations in embedding strength have a significant impact on the robustness in [16] and [15], but have little impact on the proposed approach. In the proposed method, performances remain stable across experimented w_i ; in [16], they are dramatically reduced (from 78% down to 5%) at low w_i ; surprisingly in [15], they increase (from 52% up to 88%) with decreasing w_i .

This effect was already observed in Sec. VIII-A (see Fig. 9). We further investigated this unexpected behavior and studied the Receiver Operating Characteristics (ROCs) of [15]. The ROC curves obtained for various strengths $w \cdot \alpha_0$, with w ranging from 0.24 (the average w^* across dataset $\mathcal{D}_a$ in [15]) to 0.14 (the minimum w^*). While the ROC curve remains near-optimal for $w = 0.24$, it quickly declines at lower strengths

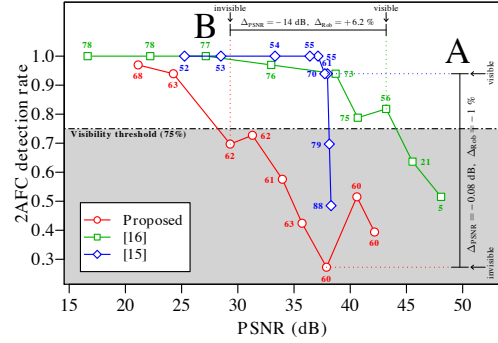


Fig. 10. 2AFC visibility percentage versus PSNR for $\mathcal{D}_a$'s image k21

(Fig. 11). At $w = 0.14$, the TP rate starts decreasing at high FP rates (10^{-1}) and drops to 0.1 for an FP rate of 10^{-16} . Neither in the proposed technique nor in [16] was this observed.

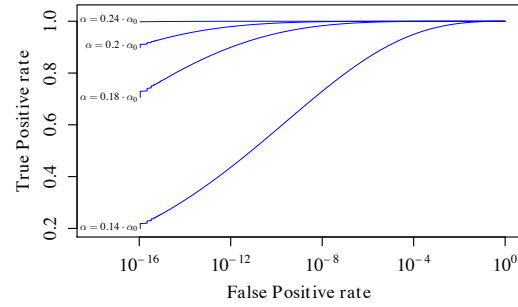


Fig. 11. Receiver Operating Characteristic curves of [15] for various strengths.

C. Content-dependency: some aspects

Fig. 12 plots, for each algorithm, the average robustness to Stimark attacks in each image belonging to dataset $\mathcal{D}_a$, at default strength (dashed line) and optimal strength (solid line). In addition, the bars in Fig. 12 plot the absolute difference in robustness between default and optimal strengths, *i.e.* the gain or loss in robustness when switching from α_0 to α^* . It can be seen that the performances are consistent across all images in the proposed approach. Conversely in [15] and [16], performances are much more inconsistent. This can be explained by the fact that, thanks to the accurate estimation of the visibility threshold provided by the proposed perceptual model, the watermark strength is properly adapted to the visual contents whilst this is not the case in [15] and [16].

IX. ROBUSTNESS TO PRINT & SCAN

In order to validate the proposed scheme in a more realistic scenario, the robustness to Print & Scan (P&S) of the proposed method was evaluated on seven 512×512 standard images (Baboon, Barbara, Boats, Fruits, Lena, Monarch and Peppers denoted as $\mathcal{D}_s^i$, $1 \leq i \leq 7$, in Table VII) and nine images of the $\mathcal{D}_a$ database (denoted $\mathcal{D}_a^j$, $1 \leq j \leq 9$).

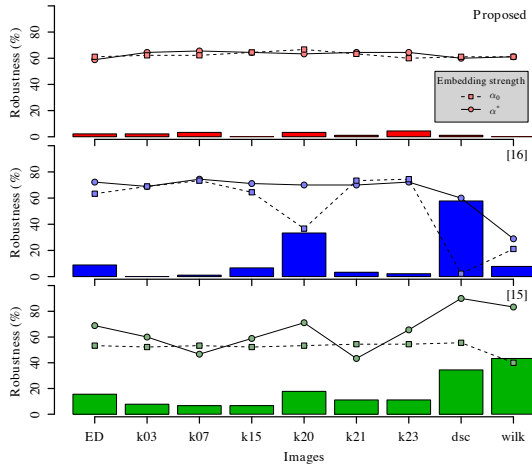


Fig. 12. Average robustness per image for dataset $\mathcal{D}_a$ at default strength α_0 (dashed lines) and optimal strength α^* (solid line). The bars plot the corresponding difference in robustness.

As demonstrated in Sec. VII-C, default (α_0) and optimal (α^*) strengths are close; here, images were watermarked at α_0 . To ensure best reproducibility, each watermarked image was first printed five times at 600 ppi (with a 80% downscaling) on a Dell 2335dn laser printer. Printed images were then scanned at 75, 150, 200 and 300 ppi with a Lexmark CX410de scanner, thus leading to 320 scanned images (images were placed approximately straight on the scanner glass).

For each watermarked image and scanning resolution, Table VII lists the number of images (out of five) for which the watermark was detected. As can be seen, the proposed approach is not robust to very low scanning quality (75 ppi), but is robust to slightly higher resolutions 150 to 300 ppi. The detection fails only in text images *wilk* (see Fig. 3b) and *dsc*. In comparison, [16] is robust to a combination of print (600 ppi) and scan (150 ppi). The technique in [37] is not robust to scanning resolutions below 600 ppi, while the technique in [14] can withstand resolutions as low as 100 ppi.

An informal 2AFC subjective test was run on the 16 prints with 8 observers. The results show that the watermark remains under the visibility threshold (53.1% of detection).

X. CONCLUSION

This paper proposes a new watermarking method. The watermark, a square patch of coefficients, is embedded within the Fourier domain by substitution of both the magnitude (energy) and the phase (information). The watermark strength is perceptually optimized. The detection features both template matching and outlier detection, the latter being applied to the obtained correlation matrix. The decision is positive if at least one outlier is detected, and negative otherwise.

The proposed method was extensively compared to two competing algorithms from the literature. A subjective experiment was conducted in order to determine the perceptually optimized watermarking strengths (*i.e.* at the visibility threshold). The performances of both the proposed and the compared

algorithms were evaluated twice: at default then perceptually optimal strengths. Experimental results showed that:

- 1) the proposed perceptual model accurately sets the watermark to its visibility threshold and is stable across all kinds of experimented contents;
- 2) the template matching accurately locates the watermark for slight geometrical distortions;
- 3) Grubbs' test for outlier performs very well both in terms of True Positives and False Positives;
- 4) the proposed method is robust against Print & Scan and shows state-of-the-art performances.

ACKNOWLEDGMENT

This work was partially supported by project OSEO A1202004 R, from which the author MU was funded.

REFERENCES

- [1] F. Cayre, C. Fontaine, and T. Furon, "Watermarking security: theory and practice," *IEEE Trans. on Sig. Proc.*, vol. 53, no. 10, pp. 3976–3987, Oct. 2005.
- [2] A. B. Watson, "DCT quantization matrices visually optimized for individual images," in *Proc. SPIE 1913, Human Vision, Visual Processing, and Digital Display IV*, vol. 1913, Sep. 1993, pp. 202–216.
- [3] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. D. Villasenor, "Visual thresholds for wavelet quantization error," in *Proc. of SPIE, Human Vision and Electronic Imaging*, vol. 2657, Apr. 1996, pp. 382–392.
- [4] C. I. Podilchuk and W. Zeng, "Image-adaptive watermarking using visual models," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 4, pp. 525–539, May 1998.
- [5] H.-H. Tsai, C.-C. Liu, and K.-C. Wang, "Blind Wavelet-based Image Watermarking Based on HVS and Neural Networks," in *Proc. of the 9th Joint Conference on Information Sciences (JCIS)*, 2006, pp. 5–8.
- [6] C.-H. Chou and K.-C. Liu, "A perceptually tuned watermarking scheme for color images," *IEEE Trans. on Image Process.*, vol. 19, no. 11, pp. 2966–2982, Nov. 2010.
- [7] F. Bartolini, M. Barni, V. Cappellini, and A. Piva, "Mask building for perceptually hiding frequency embedded watermarks," in *Proc. Intl. Conf. on Image Processing 1998. ICIP98*, vol. 1, 1998, pp. 450–454.
- [8] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun, "A stochastic approach to content adaptive digital image watermarking," in *Proc. of 3rd Intl. Workshop on Info. Hiding, IH99*, 1999, pp. 211–236.
- [9] A. Mairgiotis and N. Galatsanos, "Bayesian watermark detection and new perceptual mask based on a spatially weighted Total Variation image prior," in *2010 IEEE Intl. Workshop on Information Forensics and Security*, Dec. 2010, pp. 1–6.
- [10] F. Zhang, W. Liu, W. Lin, and K. N. Ngan, "Spread spectrum image watermarking based on perceptual quality metric," *IEEE Trans. on Image Process.*, vol. 20, no. 11, pp. 3207–18, Nov. 2011.
- [11] P. B. Nguyen, M. Luong, and A. Beghdadi, "Statistical Analysis of Image Quality Metrics for Watermark Transparency Assessment," in *Advances in Multimedia Information Processing - PCM 2010*, 2010, vol. 6297, pp. 685–696.
- [12] J. O'Ruanidh and T. Pun, "Rotation, scale and translation invariant digital image watermarking," in *Proceedings of International Conference on Image Processing*, vol. 1, IEEE, 1997, pp. 536–539.
- [13] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, M. L. Miller, and Y. M. Lui, "Rotation, scale, and translation resilient watermarking for images," *IEEE Trans. on Image Process.*, vol. 10, no. 5, pp. 767–82, Jan. 2001.
- [14] X. Kang, J. Huang, and W. Zeng, "Efficient General Print-Scanning Resilient Data Hiding Based on Uniform Log-Polar Mapping," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 1, pp. 1–12, Mar. 2010.
- [15] V. Solachidis and I. Pitas, "Circularly symmetric watermark embedding in 2-D DFT domain," *IEEE Trans. Image Process.*, vol. 10(11), pp. 1741–53, Jan. 2001.
- [16] A. Poljicak, M. Mandic, and D. Agic, "Discrete Fourier transformbased watermarking method with an optimal implementation radius," *Journal of Electronic Imaging*, vol. 20, no. 3, pp. 033 008–1–8, Jul. 2011.
- [17] A. Piva, M. Barni, F. Bartolini, and V. Cappellini, "Threshold Selection for Correlation-Based Watermark Detection," in *Proc. COST254 Workshop on Intelligent Communications*, 1998, pp. 5–6.

9.3. Perceptual DFT watermarking with improved detection and robustness to geometrical distortions

161

IEEE Transactions on Information Forensics and Security, VOL. , NO. ,

12

TABLE VII
ROBUSTNESS TO PRINT & SCAN: NUMBER OF DETECTED IMAGES^a IN FIVE STANDARD 512 × 512 IMAGES AT SEVERAL SCANNING RESOLUTIONS

Image	$\mathcal{D}_s^1$	$\mathcal{D}_s^2$	$\mathcal{D}_s^3$	$\mathcal{D}_s^4$	$\mathcal{D}_s^5$	$\mathcal{D}_s^6$	$\mathcal{D}_s^7$	$\mathcal{D}_a^1$	$\mathcal{D}_a^2$	$\mathcal{D}_a^3$	$\mathcal{D}_a^4$	$\mathcal{D}_a^5$	$\mathcal{D}_a^6$	$\mathcal{D}_a^7$	$\mathcal{D}_a^8$	$\mathcal{D}_a^9$	Average
75 ppi	0	1	3	4	2	0	1	0	0	0	0	0	2	0	0	0	16.25 %
150 ppi	5	5	5	5	4	5	5	4	5	5	5	5	5	5	0	0	85.00 %
200 ppi	5	5	5	5	4	5	5	3	5	5	5	5	5	5	0	0	83.75 %
300 ppi	5	5	5	5	4	5	5	4	5	5	5	5	5	5	0	0	85.00 %

^a Out of five printed & scanned copies.

- [18] M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 783–91, Jan. 2001.
- [19] P. Barten, "Evaluation of subjective image quality with the square-root integral method," *JOSA A*, vol. 7, no. 10, pp. 2024–2031, 1990.
- [20] S. J. Daly, "Visible differences predictor: an algorithm for the assessment of image fidelity," in *Digital Images and Human Vision*, A. B. Watson, Ed. The MIT Press, Aug. 1993, pp. 179–206.
- [21] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2," in *ACM SIGGRAPH 2011*. ACM Press, 2011, p. 1.
- [22] A. B. Watson and J. A. Solomon, "Model of visual contrast gain control and pattern masking," *J. Opt. Soc. Am. A, Optics, image science, and vision*, vol. 14, no. 9, pp. 2379–91, Sep. 1997.
- [23] A. B. Watson and A. J. Ahumada, "A standard model for foveal detection of spatial contrast ModelFest experiment," *Journal of vision*, vol. 5, pp. 717–740, 2005.
- [24] L. Zhang-ge, "Numerical Observers for the Objective Quality Assessment of Medical Images," Ph.D. dissertation, LUNAM University, Université de Nantes, IRCCyN lab., 2012.
- [25] ITU-R BT.2022, "General viewing conditions for subjective assessment of quality of SDTV and HDTV television pictures on flat panel displays BT Series Broadcasting service," Tech. Rep., 2012.
- [26] IEC-61966-2-1:1999, "Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space - sRGB," International Electrotechnical Commission (IEC), Tech. Rep., 1999.
- [27] A. A. Michelson and H. B. Lemon, *Studies in Optics*. University of Chicago Science Series, University of Chicago Press, 1927.
- [28] R. A. Frazor and W. S. Geisler, "Local luminance and contrast in natural images," *Vision Res.*, vol. 46(10), pp. 1585–98, May 2006.
- [29] P. Barten, *Contrast sensitivity of the human eye and its effect on image quality*. SPIE Press, 1999.
- [30] A. Roka, P. Galambos, and P. Baranyi, "Contrast sensitivity model of the human eye," in *2009 4th Intl. Symposium on Computational Intelligence and Intelligent Informatics*, vol. 1. IEEE, 2009, pp. 93–99.
- [31] J. Mannos and D. Sakrison, "The effects of a visual fidelity criterion of the encoding of images," *IEEE Trans. Inf. Theory*, vol. 20, no. 4, pp. 525–536, Jul. 1974.
- [32] P. Barten, "Formula for the contrast sensitivity of the human eye," *Electronic Imaging 2004*, vol. 5294, pp. 231–238, 2003.
- [33] S. a. Wallis, D. H. Baker, T. S. Meese, and M. a. Georgeson, "The slope of the psychometric function and non-stationarity of thresholds in spatiotemporal contrast vision," *Vision Res.*, vol. 76, pp. 1–10, 2013.
- [34] T. S. Meese and R. J. Summers, "Theory and data for area summation of contrast with and without uncertainty: evidence for a noisy energy model," *Journal of vision*, vol. 12, no. 11, pp. 1–28, Jan. 2012.
- [35] J. P. Lewis, "Fast normalized cross-correlation," in *Vision interface*, 1995, pp. 120–123.
- [36] K. Solanki, U. Madhow, B. Manjunath, and S. Chandrasekaran, "Estimating and undoing rotation for print-scan resilient data hiding," in *IEEE Intl. Conf. on Image Process.*, 2004, vol. 1, 2004, pp. 39–42.
- [37] K. Solanki, U. Madhow, B. S. Manjunath, S. Chandrasekaran, and I. El-Khalil, "'Print and Scan' Resilient Data Hiding in Images," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 4, pp. 464–478, Dec. 2006.
- [38] S. Pereira, J. O'Ruanaidh, F. Deguillaume, G. Csuska, and T. Pun, "Template based recovery of Fourier-based watermarks using log-polar and log-log maps," in *Proc. IEEE International Conference on Multimedia Computing and Systems*, vol. 1, 1999, pp. 870–874.
- [39] V. Hodge and J. Austin, "A Survey of Outlier Detection Methodologies," *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85–126, Oct. 2004.
- [40] F. E. Grubbs, "Procedures for detecting outlying observations in samples," *Technometrics*, vol. 11, no. 1, pp. 1–21, 1969.

- [41] R. Franzen. (2004) Kodak Lossless True Color Image Suite, <http://r0k.us/graphics/kodak/>.
- [42] A. J. Wilkins, R. J. Jeane, P. D. Pumfrey, and M. Laskier, "Rate of Reading Test: its reliability, and its validity in the assessment of the effects of coloured overlays," *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, vol. 16, no. 6, pp. 491–7, Dec. 1996.
- [43] F. A. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Attacks on copyright marking systems," in *Proc. of Information Hiding*, vol. 1525, April 1998, pp. 218–238.
- [44] ITU-R-BT.500-11, "Methodology for the subjective assessment of the quality of television pictures," Intl. Telecom. Union, Tech. Rep., 2004.
- [45] I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking and Steganography*, 2nd ed. San Francisco, CA, USA: The Morgan Kaufmann Series in Mult. Inf. and Sys., 2007.



Matthieu Urvoy received the Eng. degree in Electronics and Computer Engineering from the INSA in Rennes (France) and the M.Sc. degree in Electronics and Electrical Engineering from Strathclyde University (Glasgow, Scotland) in 2007. He received the Ph.D. degree on Signal and Image Processing at the University of Rennes in 2011. Since then, he is a post-doctoral fellow at Polytech'Nantes (IRCCyN lab). His research interests include video compression, human vision, 3D QoE and watermarking.



Dalila Goudia was born in Algeria in 1970. She received M.Sc. degree in Software Engineering and MS degree in Electronics, Vision and Pattern Recognition from the University of Science and Technology of Oran, Algeria, in 1994 and 2006 respectively. She also received the Ph.D. degree in Computer Science in 2011 from the University of Montpellier 2, France. Her current research interests include image processing, compression, and digital watermarking.



Florent Autrusseau received the M.Sc. degree in Image Processing from the University of Nantes (France) in 1999 and the Ph.D. degree on digital watermarking and perceptual modeling in 2002 at the IRCCyN lab in Nantes. In 2003–2004, he was a post-doctoral fellow at the Visual Science Labs in the University of Chicago. Since 2004, he is a research engineer at Polytech'Nantes, IRCCyN lab. His research interests include digital watermarking, HVS models, Color perception and Image analysis.

9.4 Characterization of 3D Bifurcations in Micro-scan and MRA-TOF Images of Cerebral Vasculature for Prediction of Intra-Cranial Aneurysms

Characterization of 3D Bifurcations in Micro-scan and MRA-TOF Images of Cerebral Vasculature for Prediction of Intra-Cranial Aneurysms

A.Nouri^a, F. Autrusseau^{b,*}, R. Bourcier^c, A. Gaignard^c, V. L'allinec^d, C. Menguy^c, J. Véziers^e, H. Desal^c, G. Loirand^c, R. Redon^c

^aENSC, Ecole Nationale Supérieure de Chimie, LASTID laboratory, Department of Physics, Faculty of Science, Ibn Tofail University, BP 133, 14000 Kénitra, Morocco

^bInserm, UMR 1229, RMeS, Regenerative Medicine and Skeleton, & Laboratoire de Thermique et Energie de Nantes, LTEN, U6607, University of Nantes, F-44042, France

^cDepartment of Diagnostic and Interventional Neuroradiology, Hospital Guillaume et René Laennec; INSERM, UMR1087, l'institut du thorax, CHU de Nantes, France

^dDepartment of Diagnostic and Interventional Neuroradiology, Hospital Guillaume et René Laennec; INSERM, UMR1087, l'institut du thorax, CHU de Nantes & Angers University Hospital, Radiology Department, Angers, France

^eInserm, UMR 1229, RMeS, Regenerative Medicine and Skeleton, University of Nantes, ONIRIS, F-44042, France

Abstract

An aneurysm is a vascular disorder where ballooning may form in a weakened section of the wall in the blood vessel. The swelling of the aneurysm may lead to its rupture. Intra-cranial aneurysms are the ones presenting the higher risks. If ruptured, the aneurysm may induce a subarachnoid haemorrhage which could lead to premature death or permanent disability. In this study, we are interested in locating and characterizing the bifurcations of the cerebral vascular tree. We use a 3D skeletonization combined with a graph-based approach to detect the bifurcations. In this work, we thus propose a full geometric characterisation of the bifurcations and related arteries. Aside from any genetic predisposition and environmental risk factors, the geometry of the brain vasculature may influence the chance of aneurysm formation. Among the main achievements, in this paper, we propose accurate, predictive 3D measurements of the bifurcations and we furthermore estimate the risk of occurrence of an aneurysm on a given bifurcation.

Keywords: 3D Graph, Geometry, 3D skeleton, Aneurysm, Bifurcation characterization, Micro-CT, MRA-TOF

1. Introduction

1.1. Context

The cardiovascular system (also called circulatory system) is composed of all blood vessels that carry the blood through the entire human body. On certain organs, the vascular system becomes denser, the arteries, capillaries or veins split into several branches, forming a vascular tree. Several environmental factors or a genetic anomaly (Bourcier et al., 2018) can lead to a weakened vascular system. In this work, we particularly focus on the formation of Intra Cranial Aneurysms (ICA). A weakened wall of the blood vessel may lead to the formation of an aneurysm. In the brain, aneurysms may take several forms, but ninety percent of the ICA are saccular aneurysms. Here, we will particularly pay attention to the ones occurring at the bifurcations and more specifically between the two “children” branches forming the bifurcation.

Often an aneurysm may remain benign and never evolve into a dangerous state. The main complication induced by an aneurysm is when it does rupture, and thus provoke a subarachnoid haemorrhage that may lead to the death or a permanent

disability (Ajiboye et al., 2015). ICAs must be closely monitored, as the risk of rupture is prevalent: the risk of rupture is higher along a sub set of arteries called the “Circle of Willis” (Bogunovic et al., 2013). Eighty-five percent of the saccular ICAs occur along the Circle of Willis (Keedy, 2006).

This work is part of a wide national research project (Bourcier et al., 2017) (the ICAN project) focusing on the understanding of the multiple factors that may encourage the formation of saccular intra-cranial aneurysms along the Circle of Willis. The ICAN project aims at determining the reasons why an aneurysm would appear for a given patient at a particular bifurcation depending of many different factors (such as patient habits, family history, genetic predisposition, bifurcation geometry). The ICAN Project has several active components; a study of the genetics of aneurysm formation, automatic detection of aneurysms and here the automated measurement of arterial properties. Medical image processing tools are thus required. During the first part of the project, a study was conducted on mice. Micro Computed Tomography scanner (Micro-CT) acquisition of injected mouse brains was then done *post-mortem*. In the second part of this project, the imaging tools developed for aneurysm detection were applied to existing MRI (TOF) acquisitions from human subjects.

In this study, we make use of classical MRI acquisitions (acquired on human patients) and we also test our model on mice

*Corresponding author

Email address: Florent.Autrusseau@univ-nantes.fr (F. Autrusseau)

vasculature. We thus test the accuracy of our various measurements using two distinct modalities with significantly different resolutions. The resolution of the MRI-TOF volumes used in this work were ranging from $(290 \times 520 \times 168)$ to $(696 \times 768 \times 168)$ voxels. The resolution of images acquired with the Micro-CT were $(1008 \times 1141 \times 1008)$. We have at our disposal the Micro-CT acquisition of 22 mice brains. Once the mice were euthanized and injected with contrast media, the Micro-CT renders visible fine details of the vascularization in the brain. We also have collected 39 MRI acquisitions on humans, among which 28 present an unruptured aneurysm (25 saccular aneurysms located onto a bifurcation and 3 fusiform aneurysms). Our main ambition in the framework of this study is to estimate the risk of *de novo* aneurysm formation on a given bifurcation, given its geometrical configuration.

1.2. State of the art

Although blood vessel detection has been widely studied in the literature, fewer works have focused precisely on the cerebral bifurcation characterization. As pointed out in (Zhao and Hamarneh, 2014), the detection of bifurcations in a vascular tree is very important. In their work, those authors proposed a parametric geometric deformable model to match the actual 3D medical images of bifurcations. Among the various studies on vascular tree segmentation, some work has focused on specific organs of the human body, for instance, liver vascularization was targeted in (Friman et al., 2010), whilst (Orkisz et al., 2014) focused on the pulmonary vascular tree (lungs), and (Calvo et al., 2011) specifically addressed segmentation of the blood vessels in retinal images. Interested readers may refer to two interesting reviews (Kirbas and Quek, 2004; Lesage et al., 2009) for details on methods for vessel segmentation.

A common method for detecting blood vessel direction, and hence to possibly detect bifurcations within CT or MRA images, is to use the eigenvalues of the Hessian matrix (Frangi et al., 1998; Krissian et al., 2000). In (Krissian et al., 2000), the authors use multi-scale analysis to extract blood vessels of various diameters. Using the eigenvectors and eigenvalues of the Hessian matrix, the authors then estimate vessel center-lines and were further able to detect the junctions formed by the modelled vessels. Some authors opt for the use of variational region growing methods, such as in (Orkisz et al., 2014), where the authors had to confront the delineation of lung tissue within CT images, in order to properly extract the pulmonary vascular tree. Such methods may suffer from weaker performance for vessels below a given size.

For cerebral vasculature, (Robben et al., 2016) proposed a machine learning method to automatically segment and label the cerebral vasculature along the Circle of Willis. When cerebral vasculature is considered, it is actually rather difficult to use learning-based methods, as there is a significant structural variability in the vasculature of individuals (Bogunovic et al., 2013; Robben et al., 2016). In (Bogunovic et al., 2013) a skeletonization method is used to extract the vessel center-lines, but there the seed points were determined manually.

(Bor et al., 2008) showed that the geometry of the bifurcation correlates quite well with various diseases, as it influences the

fluid mechanics of the blood flow. Multiple factors affect the fluid dynamics of blood flow. The angles of vessels at bifurcations, the distance separating two consecutive bifurcations, the tortuosity of the artery branch and the diameters of the arteries at the bifurcation are all the parameters that have an impact on the blood speed, and thus modulate the risk of formation of an aneurysm.

For most of the previously cited work on blood vessel segmentation, the acquisition was performed using either Computed Tomography Angiography (CTA) scans (Lidayová et al., 2016; Cetin et al., 2013), or Computerized Tomography (CT) Scans (Orkisz et al., 2014; Zhao and Hamarneh, 2014). In a part of our study, we use Micro-CT acquisitions, where we have the added benefit of a reasonably high image definition. This permits us to explore the accuracy of our methods at various scales.

Unlike the majority of the previously cited studies, where the goal was to study the shape or the position of the aneurysm, in this work, we intend to propose a full characterization of the vascular tree (within the vicinity of the bifurcations), so that we can further predict the chances that a saccular ICA may occur on a specific bifurcation.

We have identified some properties of the vascular tree that can infer a risk of aneurysm formation. We propose methods to automatically collect the following information: 1) bifurcation angles, 2) bifurcation thickness (cross-section area, minimum and maximum diameters), 3) the length of the arteries between two consecutive bifurcations and 4) a measure of the arteries' tortuosity between two consecutive bifurcations.

This current study intends to tackle the problem of aneurysm prediction from a different perspective than similar works (Bacigaluppi et al., 2014), as here, we compare mirror bifurcations within the same patients' vasculature, i.e. we can control the other factors (genetic predisposition and environmental risk factors), and study only the geometric properties. One can observe an example of mirror bifurcations on an MRA-TOF slice in Figure 2.

This paper is organized as follows: in section 2, we propose a bifurcation detection method based on a skeletonization combined with a graph-based approach. We also present various methods to fully characterize these detected bifurcations. Section 3 is dedicated to the presentation of our experimental results. We compare the efficiency of our bifurcation detection with two prior methods in the literature. When possible, we compare our estimations with some ground truth measurements. In section 3, the correlations between the risk of aneurysm occurrence and bifurcation geometry are examined through some experimental results. We discuss the main advantages, and some limitations of our work in section 4. Finally, section 5 concludes this work.

2. Material and Methods

In this work, we propose a method for bifurcation characterisation and aneurysm occurrence prediction. It follows a two steps process. At first, we need to accurately detect the vascular tree, and reliably locate the bifurcations. Once the bifurcations

can be isolated, each one goes through a full characterisation process. This section will thus be composed of these two steps, we will first present the proposed method for bifurcation detection, and will then present the geometric characterisation step.

2.1. Automatic detection of 3D bifurcations

The first stage of the image-processing pipeline proposed in this paper is the detection of the 3D bifurcations in cerebral vasculatures. We present in this section a new approach to automatically localize the bifurcation centers in the 3D image space.

Given the 3D cerebral vasculature U of a mouse or a human brain (acquired respectively using a Micro-CT and an MRA-TOF scan), we begin by computing its 3D skeleton S using the method described in (Lee et al., 1994). To each voxel s_i of S are associated its 3D coordinates $\vec{s}_i = (x_i, y_i, z_i)$. Exploiting the full 3D skeleton instead of 2D image slices reduces the total processing time. More importantly, the better connectivity between 3D voxels simplifies the construction of the corresponding graph structure. In order to skeletonize the volume U , an octree structure is used to analyze the 26-connectivity ($3 \times 3 \times 3$ neighborhood of each voxel). The algorithm proceeds by iteratively sweeping over the image, and removing voxels at each iteration until the image stops changing. After each iteration, a list of removable candidate voxels is formed. Each voxel of this list is checked sequentially to ensure its removal preserves the connectivity of the image. Based on this 3D skeleton, a non-oriented graph $G = (V, E)$ is constructed, where V represents the set of nodes and $E \in V \times V$ the set of edges. Each edge $e(v_i, v_j) \in E$ connects the nodes v_i and v_j . As for S , each node $v_i \in V$ of the graph is associated its 3D coordinates $\vec{v}_i = (x, y, z)$. The significant benefits of using a graph structure can be summarized in three main points:

- The graph allows scanning of the skeleton (and hence the volume) in 3D thus avoiding the 2D slice-by-slice analysis. This preserves the 3D information that may be lost when projecting from 3D to 2D.
- The graph provides an accurate localization of the bifurcation center in the 3D space.
- The graph can be restricted to locate bifurcations in any region of interest. Here we restrict our attention to the Circle of Willis, since this is where most intra-cranial aneurysms occur (Robben et al., 2016). For this we use the cross-section area characteristic described in section 2.2.1.

Figure 1 presents the 3D skeleton along with its associated graph for a Micro-CT cerebral vasculature of a mouse brain. Unlike Micro-CT acquisitions where the mice required a barium sulfate injection in the vascular tree prior to any cerebral image acquisition, the MRA-TOF images associated to human brains have to be segmented in order to extract the cerebral vasculature from the subcutaneous tissue. For this, we use a “Skull Stripping” method, the *Brain Extraction Tool* (Smith, 2002) coupled with an entropy-based approach (Kapur et al., 1985) selected for its low complexity and reliability. The pre-processing step needed to segment MRA-TOF images performs

well, it exhibits comparable results to recent state-of-the-art approaches such as (Merveille et al., 2014). After the Skull Stripping and segmentation, we obtain a clean volume with its skeleton as presented in Figure 2. To achieve this, short edges of the graph (and their associated parts of the 3D skeleton) are suppressed, as are graph nodes with one unique neighbor. Figure 2 also depicts one 2D slice from an original MRA-TOF volume (prior to the skull stripping).

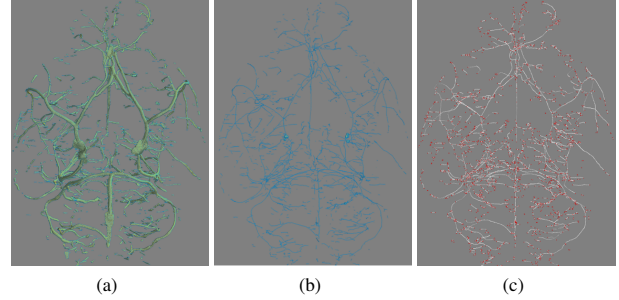


Figure 1: Micro-CT cerebral vasculature of a mouse brain. a) Micro-CT volume, b) 3D skeleton of (a) and c) Graph constructed on (b). The red dots represent the graph nodes and the white segments are the graph edges.

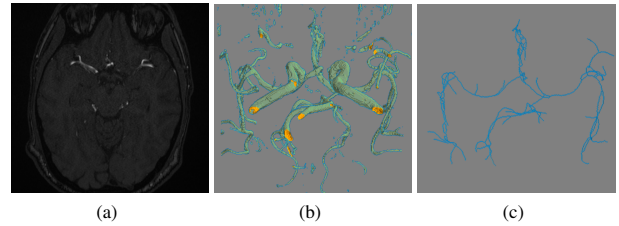


Figure 2: Clean segmentation of MRA-TOF 3D stacks. a) One slice of the MRA-TOF, b) 3D segmented stack and c) 3D skeleton of (b).

A 3D bifurcation has a Y-junction shape (for either two merging or splitting arteries) and its graph contains four nodes (see Figure 3), one node at the center of the 3D skeleton bifurcation and the remaining three located at the end points of the three branches. Hence, if any of the graph nodes has three neighbors, this node is considered as the center of the bifurcation.

It sometimes happens that the vasculature splits into more than two “daughters” arteries, we did encounter such scenarios, and our method can reliably detect such configurations. Nevertheless, although in our experiments we did find some “trifurcations” among all the tested TOF stacks, no trifurcations were found to be harbouring an aneurysm on the Circle of Willis.

2.2. Geometric characterization of 3D arterial bifurcations

When blood flow is subjected to a sudden speed decrease at a bifurcation, the impact on the artery inner layer generates repeated mechanical stresses that may weaken the artery

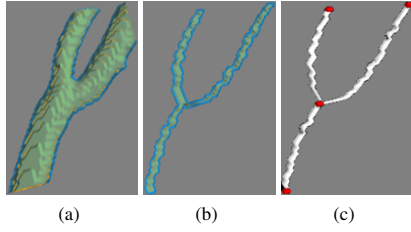


Figure 3: Images (a), (b) and (c) respectively represent a bifurcation from a Micro-CT acquisition, its 3D skeleton and the associated graph with red nodes and white edges.

wall. The more the artery structure deviates from linear, parallel walls, the greater the stress impact. Decreased tortuosity of the artery reduces the friction of the blood flow on the artery walls, and thus increases the blood speed, which is expected to correlate highly with these larger stress impacts when the blood flow reaches the bifurcation. Wider bifurcation angles and a strong reduction in arterial diameter between the mother and daughter branches would exacerbate this effect. In this section, we propose to measure various geometrical properties of the intracranial arterial tree which may play a role in the development of saccular aneurysms. The analysis of the intrinsic attributes of vasculature geometry as listed above may help identify patients with an increased risk of developing aneurysms (Bor et al., 2008). To develop these concepts toward achieving an eventual Computer Aided Diagnosis, here we model various relevant characteristics such as cross-section area, cross-section diameter, bifurcation angles, branch tortuosity and the geodesic distance between consecutive bifurcation centers. We compute these structural biomarkers in a fully automatic manner.

2.2.1. Measurement of arterial thickness

Arteries are composed of various cellular layers, the innermost layer, called the tunica intima is in direct contact with the blood flow. Among other anatomical factors, the size of the artery may influence the speed of the blood flow, it is thus crucial to estimate the thickness of the artery. Two approaches were considered, we have computed both the arterial branch cross-sectional area, as well as the diameter of the arterial branch.

The area of a branch in a 3D bifurcation represents an important parameter to quantify the thickness of a branch. This thickness value helps to identify the mother branch of a bifurcation where blood flows before dispersing into the two daughter branches. It is mandatory to correctly identify the mother branch before computing the angle between the down-stream two daughter branches (see section 2.2.2). Neuro-radiologist researchers suspect that an important cross-sectional area of the mother branch may lead to the formation of aneurysms because the high blood pressure from strong flow pushes on the arterial wall that forms the junction of the daughter branches. Given images of a set of 3D cerebral vasculature, we detect all possible bifurcation centers using the graphical approach presented

earlier. Each bifurcation B of a center c is extracted from the volume within a $60 \times 60 \times 60$ block. Then, the 3D skeleton S_B and its associated graph G_B are computed. The bifurcations' branches are delimited by the graph nodes $(b_1, b_2, b_3) \in G_B$. For a target branch cb_1 , its cross-sectional area is computed by considering a voxel n belonging to the branch cb_1 located at a distance of 10 voxels from the bifurcation center c . The vector $\vec{cn}$ is then considered as the directional vector of the branch cb_1 . The volume B is subsequently rotated in order to align $\vec{cn}$ co-linear with the unit vector $\vec{Z}$. To obtain the area of the perpendicular section of the branch cb_1 , we extract a 2D slice of the 3D volume rotated parallel to the (XY) plane that passes through the voxel n which necessarily lies perpendicular to cb_1 . Performing a contour detection inside the resulting 2D slice eliminates sections to which the voxel n does not belong. Finally, the cross-sectional area is defined as the number of detected pixels on the 2D slice. Figure 4 and Figure 5 respectively represent an illustration of the branch cross-section extraction and some results on various 3D MRA bifurcations. We identify high-risk target bifurcations as those where the mother branch cross-sectional area is large compared to that for both of the two daughter branches that constitute a bifurcation (see section 3.2.1).

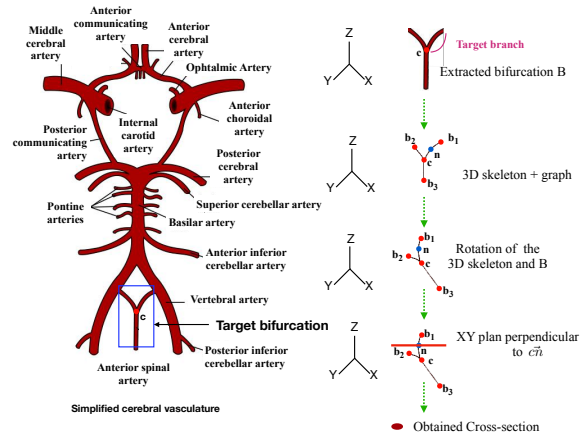


Figure 4: Illustration of the cross-section computation of an arbitrary branch bifurcation.

We have also proposed an approach for the computation of maximal and minimal arterial diameters as recommended in (Bourcier et al., 2017). Once the artery has been aligned along the z-axis, and a 2D slice of the TOF volume has been extracted (the extracted slice is thus perpendicular to the artery orientation), our model next computes a set of oriented projections of this 2D slice image onto 1D projection bins. A discrete Radon transform is applied (Guédon, 2013). The projections span a 180° range of view angles with a 1° step. For each view, the width of the projected vessel is computed. This permits an angle-sensitive determination of both the minimum and maximum thickness of the blood vessel. Figure 6 illustrates this transformation, where the projection $\mathcal{R}f(\theta)$ at angle $\theta = 153^\circ$

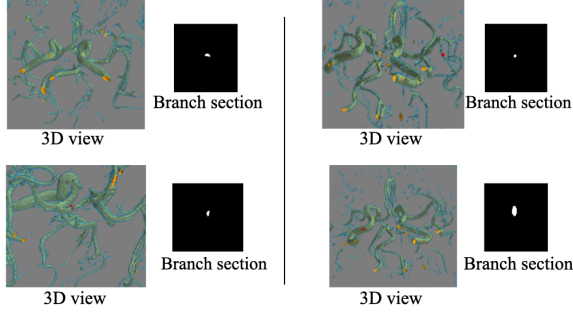


Figure 5: Different branch sections results. Red dots in the 3D views represent the section points.

leads to the measure of the smallest artery diameter ($\mathcal{D}_{153} = 9$ bins), whereas angle $\theta = 44^\circ$ produces the largest diameter estimate (21 projection bins).

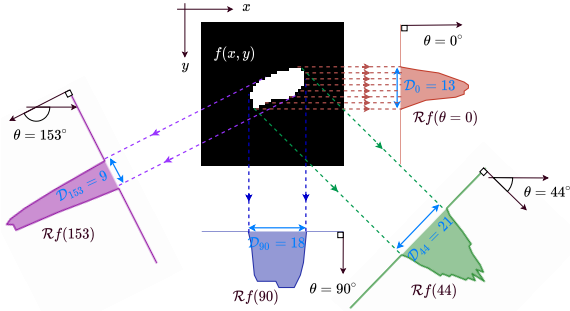


Figure 6: Projections at various angles of a 2D image cross-section into 1D bins for the computation of the minimum and maximum diameters. Here, the minimum and maximum diameters of the artery are respectively 9 and 21 bins (at angles $\theta = 153^\circ$ and $\theta = 44^\circ$).

2.2.2. Computation of arterial branch angles

The angle between the central axes of two daughter branches of a bifurcation and in particular the two angles between each daughter branch and the mother branch are important geometrical parameters which may help to predict the risk of aneurysm formation. Indeed, neuro-radiologist researchers have noticed that saccular aneurysms often arise at bifurcations where significant angles occur between daughter branches (Bor et al., 2008). Accurate automatic computation of these angles then becomes relevant for aneurysm prediction. We aim to automatically compute the opposite angles $\hat{A}_1$ and $\hat{A}_2$ as illustrated in Figure 7(b). Using the 3D skeleton and its associated graph, the angle $\hat{A}_1$ can be defined as:

$$\hat{A}_1 = \text{atan2}(\|\vec{CN} \cdot \vec{CL}\|_2, -\vec{CN} \cdot \vec{CL}) \quad (1)$$

where C is a graph node representing the bifurcation center and L, M, N are the remaining graph nodes delimiting the branches of the bifurcation. $\vec{CN}$ and $\vec{CL}$ are defined respectively

as $\vec{CN} = \vec{N} - \vec{C}$ and $\vec{CL} = \vec{L} - \vec{C}$. The function $\text{atan2}(y, x)$ computes the arc tangent of all four quadrants, providing a result in the interval $[-\pi, \pi]$. The symbols $\bullet$ and $\cdot$ respectively represent the cross and dot products. The notation $\|\cdot\|_2$ refers to the second norm. The same method is applied for the estimation of both $\hat{A}_1$ and $\hat{A}_2$.

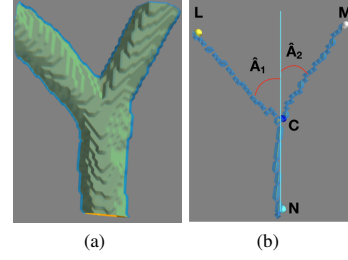


Figure 7: Target angles to compute: a) 3D bifurcation and b) 3D skeleton with estimated daughter angles.

2.2.3. Geodesic distance between arterial branches

A large geodesic distance between two adjacent bifurcations, coupled with a significant mother branch cross-section area and large daughter angles enhances the risk of saccular aneurysm forming. A large geodesic gap between bifurcations may result in an increased speed of the blood flow into the mother branch of a bifurcation, hence weakening and distorting the junction vessel wall and leading to the formation of a bulge (Figure 8). In our graphical representation of arterial structure as a 3D skeleton, the centers of bifurcations and their end points are associated with graph nodes. Lists of voxels (each with their corresponding 3D coordinates) link each node to its neighbors and are stored within the graph structure. Hence, the geodesic distance between any two detected bifurcation centers is defined as the cardinality of the voxel list between them. This estimation of the geodesic distance between two consecutive bifurcations is then quite straightforward. There is no need here to validate our model with any ground truth measurements, if the bifurcations centers have been accurately located, the geodesic distance must be correctly estimated.

2.2.4. Arterial tortuosity

The arteries' tortuosity also influences the fluid mechanics of the blood flow and hence correlates significantly with the formation of intracranial aneurysms (Labeyrie et al., 2017). Tortuosity can be defined by the degree of curvature of an arterial segment. In this section, we propose a new method to compute the tortuosity of any artery in the cerebral vasculature based on the curvature of its voxels. To the best of our knowledge, there are currently no existing methods to measure the global curvature of a cerebral blood vessel. We are thus unable to compare our tortuosity measurements with observations in the literature. However, in order to validate our curvature estimations, we have collected ground truth measurements from human ob-

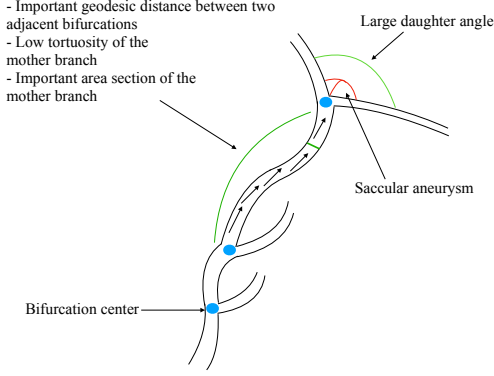


Figure 8: Influence of geometrical characteristics in the occurrence of aneurysms in a cerebral vasculature.

servers. Twenty observers have been asked to rate the degree of tortuosity of 27 3D arterial images¹.

Given an artery A and its 3D skeleton S_A , 3D normal vectors of the skeleton binary voxels are first calculated. Tangent vectors for each voxel in the artery are computed and used to derive the normal vectors. These are oriented perpendicular to the 3D tangent vectors. Figure 9 illustrates the 3D normal vectors obtained for an example artery.

To compute the curvature of a target voxel v_i , we assess variations between its normal vector $\vec{n}_{v_i}$ and the normal vectors of its neighbors. For this, we consider four voxels at each side (left and right) of the target voxel. Both the mean of these normal vectors and their vector positions on each side are computed to derive two normal vectors and two vector positions at the left and right of the target voxel. Figure 10 illustrates this process. The curvature $C(v_i)$ of a target voxel v_i is defined as:

$$C(v_i) = \frac{C_{left} + C_{right}}{2} \quad (2)$$

where C_{left}^k is given by:

$$C_{left} = \frac{(\vec{n}_{v_i} - \text{mean}(\vec{n}_{v_{left}})) \cdot (\vec{v}_i - \text{mean}(\vec{v}_{left}))}{\|\vec{v}_i - \text{mean}(\vec{v}_{left})\|_2^2} \quad (3)$$

where $\text{mean}(\vec{n}_{left}) = \frac{\sum_{i=0}^{|N_{left}|} \vec{n}_i}{|N_{left}|}$ and $\text{mean}(\vec{v}_{left}) = \frac{\sum_{i=0}^{|N_{left}|} \vec{v}_i}{|N_{left}|}$. $|N_{left}|$ refers to the cardinality of the neighborhood. C_{right}^k is defined similarly.

Once the curvature of each voxel artery has been computed, the aim now is to compute a scalar value representing the global tortuosity ($\mathcal{T}$) over the full extent of the artery. To do this, we average the resulting voxel curvatures using a weighted Minkowski sum:

$$\mathcal{T} = \left(\frac{\sum_{i=0}^{|v|} C(v_i)^p}{|v|} \right)^{1/p} \quad (4)$$

¹ http://webseubjexp.polytech.univ-nantes.fr/Tortuosity/index_IsolatedArteries.php

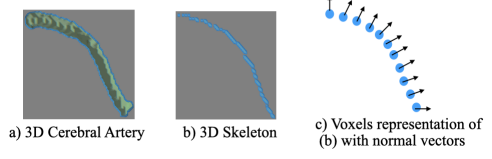


Figure 9: Illustration of normal vectors of a 3D artery.

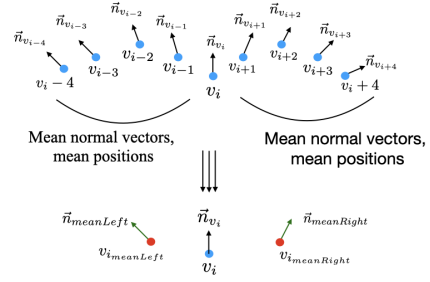


Figure 10: Curvature computation of a target voxel v_i .

where the weight factors p and r have been empirically set to 49.9 and 23.6 respectively, as these values provided the most robust estimation of the overall tortuosity.

3. Experimental Results

The previous section gave some details on the two steps process. We went through the vascular tree detection, and the bifurcations location. We have seen an in-depth description of the bifurcation characterization. We were thus able to collect the bifurcation geometrical properties. Here, we will present the results in terms of both bifurcation detection, and characterization. When possible, we will collect the ground truth data and provide an in-depth comparison with our models.

3.1. 3D bifurcations detection

Intra-cranial aneurysms arise most frequently in the arteries located along the Circle of Willis (Robben et al., 2016). We have tested our approach on different 3D cerebral vasculatures to ensure that all 3D bifurcations located on the Circle of Willis are detected correctly. Figure 11 presents the bifurcations centers that have been detected (small red spheres) on a mouse Micro-CT acquisition and shows a closer view of the bifurcation arteries associated to the Circle of Willis. Our proposed approach correctly detected all bifurcations of interest (9 bifurcations located on the Circle of Willis). This method was also able to successfully detect all bifurcations within any different region of the full target volume.

Figure 12 presents the results for our bifurcation detection method in two different examples of human cerebral vasculature (MRA-TOF images). After the Skull Stripping process (mandatory for separating the cerebral vasculature from the

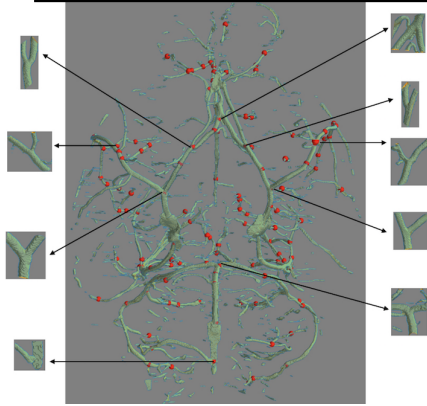


Figure 11: Centers of 3D bifurcations detected in a Micro-CT volume. 3D bifurcations associated to the Circle of Willis are enlarged.

subcutaneous tissue as explained above), the proposed graph-based approach is applied as for the Micro-CT volumes. Our approach succeeds in detecting all bifurcations (on or beyond the Circle of Willis) for this method of data acquisition. Figure 12, shows that there is significant variability amongst different vasculatures, as explained in (Bogunovic et al., 2013; Robben et al., 2016).

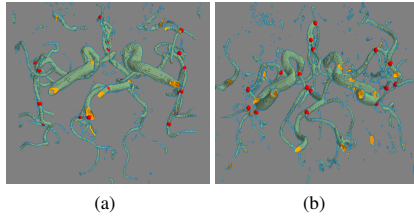


Figure 12: Images (a) and (b) present the bifurcations detected on two different MRA-TOF human cerebral vasculatures.

Now, let us provide a comparison of the proposed method with two reference approaches of the literature (Macedo et al., 2013; Zhao and Hamarneh, 2014). For this, we consider a ground truth data-set, “*Vascusynth*” (Hamarneh and Jassi, 2010; Jassi and Hamarneh, 2011), which consists of 10 groups of data. In this database, each group contains 12 randomly generated volumes with the number of bifurcations incrementing from 1 to 56 in steps of 5. 3D coordinates (x, y, z) of the bifurcation centers are provided. These locations represent the ground truth coordinates for each bifurcation center. We applied our proposed model to this test data with the aim of reproducing these known coordinates with the highest accuracy. To compare the three approaches, we considered a subset of the *Vascusynth* database comprised of 10 volumes with 16 bifurcations in each volume. Cumulative histograms of the measured distances between the location of the predicted bifurcation centers (DB) and the ground truth centers (GT) are used. Figure 13

presents the resulting cumulative histograms. We observe that our proposed approach detects all possible bifurcations, with a maximum 19 pixel error from the ground truth centers (Figure 13(a)) whereas the two competing approaches (Macedo et al., 2013; Zhao and Hamarneh, 2014) required more than 50 pixel distance errors to achieve complete bifurcation detection. Furthermore, the distribution of distance errors from our approach (Figure 13(a)) diminishes faster than those for the comparison methods (Figure 13(b)). This result confirms the high precision of our approach to detect and localize 3D bifurcations. An accurate localization of the bifurcations is important, as any measurement inaccuracy might induce erroneous geodesic distance computation during the bifurcation characterisation (described later).

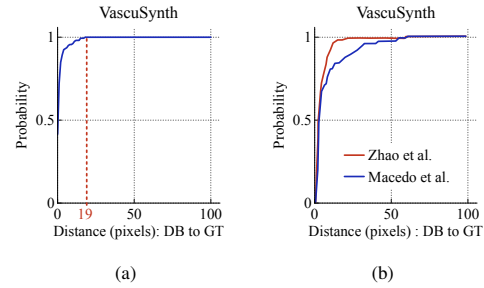


Figure 13: Cumulative histograms of distances between the predicted bifurcations’ localizations and the ground truth localizations of 10 volumes with 16 bifurcations each belonging to the *Vascusynth* database (Hamarneh and Jassi, 2010; Jassi and Hamarneh, 2011): (a) The proposed approach and (b) approaches of (Macedo et al., 2013; Zhao and Hamarneh, 2014) (from (Zhao and Hamarneh, 2014)).

3.2. Characterization of the bifurcations and arteries

In this section, we present our experimental results regarding the bifurcation characterization. The performances of our method are given with regard to three distinct geometrical parameters: the arterial thickness, the bifurcation angles, and the tortuosity.

3.2.1. Arterial thickness

In order to ensure that our model properly estimates arterial diameters, we have collected subjective thickness measurements. Ten people were asked to manually measure the smaller and larger diameters of ten cerebral arteries. The measurements were performed using the *ImageJ* software.

Figure 14 shows both the average user-defined measurements (gray boxes), and the model outputs (black boxes). Solid lines stand for the minimum artery diameters, whereas dotted lines represent their maximum diameters. Our model accurately predicted both diameters (as evident by alignment of the solid, *resp.* dotted boxes). The Pearson correlation between the model prediction and the ground truth (user-defined measurements) was 0.93 for the minimum diameters and 0.92 for the maximum diameters.

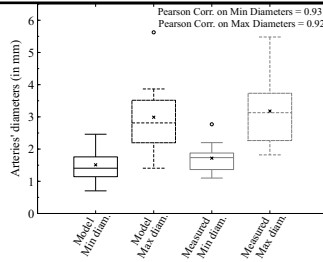


Figure 14: Distribution of the min and max diameters according to the model (in black) and measures from humans (in gray).

3.2.2. Bifurcation angles

To attest the relevance of our graphical approach for the automated measurement for pairs of daughter branch angles, we have compared our computed angle values with those obtained with the *ImageJ* software. Two subjects among the authors (AN and FA) manually measured these angles, we used the average of these two independent observer measures. Figure 15 presents the results for the two angles $\hat{A}_1$ and $\hat{A}_2$. We note that the angle values estimated by our model are highly correlated with the manual measurements. Moreover, using our model, the angles are automatically computed, without requiring any user intervention. The insets below the bar plots (in Figure 15) show the differences between the model and the human measures. The differences, being mostly positive, mean that our model slightly overestimates the bifurcation angles. However, the errors are acceptably low over all twenty measurements.

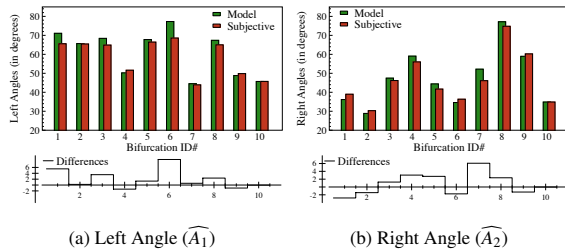


Figure 15: Comparison between the angle values measured using the *ImageJ* software and those obtained by our approach.

3.2.3. Tortuosity

In this section, we compare the results of the proposed tortuosity measure against the subjective scores provided by human observers. To verify this, we have constructed a ground truth which consists of 27 human cerebral arteries. The tortuosity indexes of these branches were evaluated by 20 human observers. The observers were able to interact with the 3D arteries (3D rotation, zoom in/out) and were asked to assign a representative tortuosity score between 0 and 100 (where 0 refers to a low tortuosity degree and 100 represents extreme tortuosity). Figure 16 presents four example arteries belonging to the constructed

ground truth with their associated Mean Opinion Score (MOS) of tortuosity.

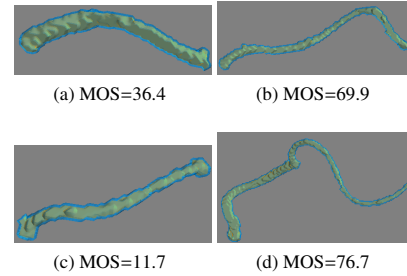


Figure 16: Example of 3D arteries belonging to the constructed ground truth with their associated Mean Opinion Score (MOS) of tortuosity.

The objective tortuosity results provided by our measure and the subjective results collected from human observers showed strong agreement, with a Pearson correlation coefficient of 90.03% and a normalized RMSE (Root Mean Square Error) between the objective and subjective results of 0.13. These indicators attest the precision of the proposed measure of tortuosity.

3.3. Linking arterial bifurcation geometry and the risk of aneurysms

The risk of aneurysm formation depends on several factors. A genetic predisposition may account for a significant portion of the probability of aneurysm formation. Environmental interactions, such as smoking habits or hypertension may also increase the risk of developing an aneurysm. We wish to avoid the complications and interference from these external factors on our study of the influence of the bifurcation geometry on aneurysm formation. To do this, we analyzed the effects of arterial geometry when those external factors remain constant by analyzing intra-patient examples. We thus compared the geometry of matching bifurcations located on the same arteries on the left and right side of the patient's brain (mirror bifurcations) to evaluate the risk of aneurysms. We gathered the MRA-TOF volumes from 25 patients. Among the 25 patients, some presented an aneurysm on the left Middle Cerebral Artery (MCA) whereas the others exhibited an aneurysm on the right MCA. The proposed model of bifurcation characterization was tested on this data base of 50 bifurcations (including the 25 aneurysm-free mirror bifurcations).

Figure 17 shows the geometrical parameters derived using our automated graphical method. We compare those bifurcations that did exhibit an aneurysm (black boxes), and those without aneurysms (red boxes). As expected, the bifurcations that actually exhibited an aneurysm presented larger angles, lower tortuosity, and longer distance of the mother artery. However, as can be seen on Figure 17(c), it appears that the cross-section of the mother artery may not be a strong feature to help estimate the risk of occurrence of an aneurysm. For each bifurcation, we also subsequently computed the difference between the cross-sections of the two daughter arteries, as Neuro-

9.4. Characterization of 3D Bifurcations in Micro-scan and MRA-TOF Images of Cerebral Vasculature for Prediction of Intra-Cranial Aneurysms

171

radiologists have suspected that this may have some impact of the risk of aneurysm formation (Can et al., 2015). Unfortunately, our results couldn't confirm this hypothesis, as can be observed on Figure 17(d). A significant difference between daughters' arterial cross-section did not seem to be related with the occurrence of an aneurysm. Nevertheless, as pointed out in (Can et al., 2015): "In univariate analysis, Basilar Tip Aneurysms were associated with smaller basilar artery diameter (...) and larger difference in the diameters of P1 segments of the PCAs (...), but these relationships were not statistically significant.". In that study the authors focused specifically on the basi-

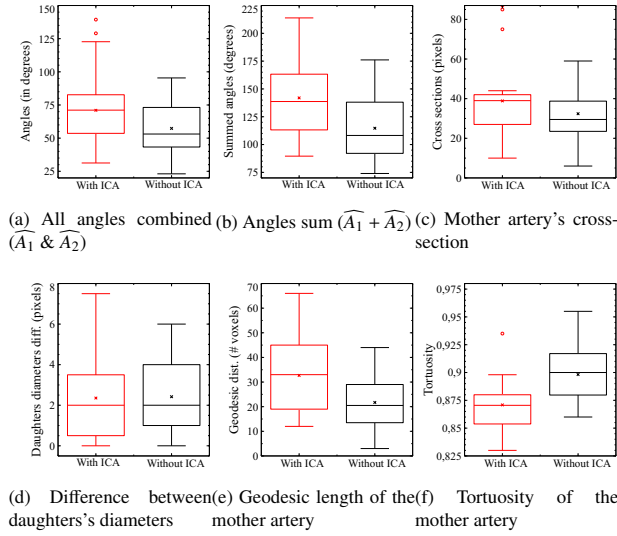


Figure 17: Distribution of various geometrical properties (angles, cross-sections, geodesic distance, and tortuosity).

lar artery, whereas here our results (Figure 17) pertained to the MCA. Overall, except for the diameter feature which, at best, is weakly linked to the risk of aneurysm formation, a strong relationship was found for all three other features. Importantly, the results from our proposed model closely mimic the chances of occurrence of an aneurysm based on three geometric considerations: the angle formed by the daughter branches, the tortuosity and length of the mother branch.

4. Discussion

With regard to the weak link we have found between the risk of aneurysm formation and the diameter of the arteries, we have investigated this issue further. A closer examination was specifically conducted on the diameter feature, as this particular marker appeared to be unrelated to the aneurysm formation. We show in Figure 18 how two bifurcations having unbalanced daughter diameters may either develop or not an aneurysm. We suspect that a weighted combination of both diameters and angles or maybe diameters and tortuosity might more efficiently predict the risk of aneurysm formation.

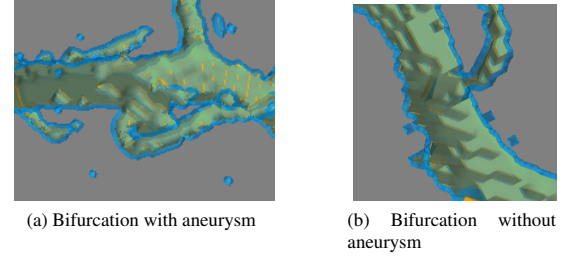


Figure 18: Two distinct (mirror) bifurcations of the same patient. Despite a similar configuration (equivalent unbalanced diameters of the daughter arteries), an aneurysm developed only onto one bifurcation.

Overall, our model proved to efficiently characterize the bifurcations, and we were thus able to link the various geometrical properties of the bifurcations to the risk of aneurysm formation. We were able to show that three geometric properties of the bifurcations may have an impact on the aneurysm formation. However, the whole process we have presented is strongly dependent on the vascular tree segmentation. A better and more suited segmentation of the vascular tree might ensure a more generalizable characterisation method. Effectively, our characterization process is dependent on its very first step, the image binarization. Similarly, another limitation of our work is the lack of an explicit link between a detected bifurcation and its anatomical correspondence, i.e. once a bifurcation is detected, we cannot ensure it belongs to the circle of Willis. In our experiments, we hypothesized that the larger arteries belong to the circle of Willis, nevertheless, we have no means to ensure from an anatomical point of view that this assumption holds. Combining our bifurcation detection and characterization to a brain vascular atlas might be very useful in clinical application scenarios.

5. Conclusions and perspectives

This research explored the structure of cerebral vasculature (in particular at vessel bifurcations in the Circle of Willis) and its relationship to the onset of saccular aneurysms in the brain. We proposed several methods to fully characterize these bifurcations by testing novel algorithms that automatically measure the angles formed by the two daughter branches, the normal arterial cross-sections and the distance between two consecutive bifurcations. We provided a new estimate for the mother arterial tortuosity. We collated subjective manual measurements to confirm the accuracy for some of these proposed methods, using a database of 3D scanner images of both human and mice brains. The measured vessel angles, diameters and tortuosity estimations were well correlated to the measurements made by human observers. Full characterization of these bifurcations enabled us to associate these features with the possible occurrence of a saccular aneurysm. Our results show, given control of other contributing factors (for genetic differences and environmental factors), that bifurcation geometry can be stated to

play an important role in the risk of developing an aneurysm. Amongst the geometric features we quantified, the bifurcation angles, the tortuosity and the distance to the previous bifurcation appeared to be the three most important determinants.

Among the improvements we can bring to the 3D bifurcation characterization, future investigations might be devoted to the improvement of the segmentation and bifurcation detection step. One might also try to combine this characterization process to an anatomical Atlas, which would allow us to benefit from a precise localization of the aneurysms and ensure that we only process the ones located onto the circle of Willis.

Acknowledgments

The authors wish to thank Pr. Imants Svalbe who assisted in the proof-reading of the manuscript. This work received funding from grant ANR-15-CE17-0008.

References

- Ajiboye, N., Chalouhi, N., Starke, R.M., Zanaty, M., Bell, R., 2015. Unruptured Cerebral Aneurysms: Evaluation and Management. *The Scientific World Journal* 2015, 1–10.
- Bacigaluppi, S., Piccinelli, M., Antiga, L., Veneziani, A., Passerini, T., Rampini, P., Zavanone, M., Severi, P., Tredici, G., Zona, G., Krings, T., Boccardi, E., Penco, S., Fontanella, M., 2014. Factors affecting formation and rupture of intracranial saccular aneurysms. *Neurosurgery Review* 37, 1–14. doi:10.1007/s10143-013-0501-y.
- Bogunovic, H., Pozo, J.M., Cardenas, R., Roman, L.S., Frangi, A.F., 2013. Anatomical labeling of the circle of willis using maximum a posteriori probability estimation. *IEEE Transactions on Medical Imaging* 32, 1587–1599.
- Bor, A.S.E., Velthuis, B.K., Majoie, C.B., Rinkel, G.J., 2008. Configuration of intracranial arteries and development of aneurysms. *Neurology* 70, 700–705.
- Bourcier, R., Chatel, S., Bourcereau, E., Jouan, S., Marec, H., Dumas-Duport, B., Sevin-Allouet, M., Guillon, B., Roualdes, V., Riem, T., Isidor, B., Lebranchu, P., Connault, J., Tourneau, T., Gaignard, A., Loirand, G., Redon, R., Desal, H., 2017. Understanding the Pathophysiology of Intracranial Aneurysm: The ICAN Project. *Neurosurgery* 80, 621–626.
- Bourcier, R., Scouarnec, S.L., Bonnaud, S., et al., 2018. Rare coding variants in angptl6 are associated with familial forms of intracranial aneurysm. *American Journal of Human Genetics* 102, 133–141.
- Calvo, D., Ortega, M., Penedo, M.G., Rouco, J., 2011. Automatic detection and characterisation of retinal vessel tree bifurcations and crossovers in eye fundus images. *Computer Methods and Programs in Biomedicine* 103, 28–38.
- Can, A., Mouminah, A., Ho, A.L., Du, R., 2015. Effect of Vascular Anatomy on the Formation of Basilar Tip Aneurysms. *Neurosurgery* 76, 62–66.
- Cetin, S., Demir, A., Yezzi, A., Degertekin, M., Unal, G., 2013. Vessel Tractography Using an Intensity Based Tensor Model With Branch Detection. *IEEE Trans Med Imaging* 32, 348–363.
- Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M., 1998. Multiscale vessel enhancement filtering. *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98. LNCS, vol 1496* 1496, 130–137.
- Friman, O., Hindennach, M., Kühnel, C., Peitgen, H.O., 2010. Multiple hypothesis template tracking of small 3D vessel structures. *Medical Image Analysis* 14, 160–171.
- Guédon, J., 2013. *The Mojette Transform: Theory and Applications*. ISTE, Wiley.
- Hamarneh, G., Jassi, P., 2010. Vascusynth: Simulating vascular trees for generating volumetric image data with ground truth segmentation and tree analysis. *Computerized Medical Imaging and Graphics* 34, 605–616.
- Jassi, P., Hamarneh, G., 2011. Vascusynth: Vascular tree synthesis software. *Insight Journal January-June*, 1–12.
- Kapur, J., Sahoo, P., Wong, A., 1985. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing* 29, 273 – 285.
- Keedy, A., 2006. An overview of intracranial aneurysms. *McGill Journal of Medicine* 9, 141–146.
- Kirbas, C., Quek, F., 2004. A review of vessel extraction techniques and algorithms. *Computing Surveys* 36, 81–121.
- Krissian, K., Malandain, G., Ayache, N., Vaillant, R., Troussset, Y., 2000. Model-Based Detection of Tubular Structures in 3D Images. *Computer Vision and Image Understanding* 80, 130–171.
- Labeyrie, P., Braud, F., Gakuba, C., Gaberel, T., Orset, C., Goulay, R., Emery, E., Courthéoux, P., Touzé, E., 2017. Cervical artery tortuosity is associated with intracranial aneurysm. *International Journal of Stroke* 12, 549–552.
- Lee, T.C., Kashyap, R., Chu, C.N., 1994. Building Skeleton Models via 3-D Medial Surface/Axis Thinning Algorithms, in: *CVGIP: Graphical Models and Image Processing*, pp. 462–478.
- Lesage, D., Angelini, E.D., Bloch, I., Funka-Lea, G., 2009. A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes. *Medical Image Analysis* 13, 819–845.
- Lidayová, K., Frimmel, H., Wang, C., Bengtsson, E., Smedby, O., 2016. Fast vascular skeleton extraction algorithm. *Pattern Recognit. Lett.* 76, 67–75.
- Macedo, M.M.G., Galarreta-Valverde, M.A., Mekkaoui, C., Jackowski, M.P., 2013. A centerline-based estimator of vessel bifurcations in angiography images, in: *SPIE, Medical Imaging : Computer-Aided Diagnosis*, pp. 8670–8670.
- Merveille, O., Talbot, H., Najman, L., Passat, N., 2014. Tubular structure filtering by ranking orientation responses of path operators, in: *Computer Vision – ECCV 2014, Springer International Publishing, Cham*. pp. 203–218.
- Orkisz, M., Hernández Hoyos, M., Pérez Romanello, V., Pérez Romanello, C., Prieto, J.C., Revol-Muller, C., 2014. Segmentation of the pulmonary vascular trees in 3D CT images using variational region-growing. *Irbm* 35, 11–19.
- Robben, D., Türetken, E., Sunaert, S., Thijs, V., Wilms, G., Fua, P., Maes, F., Suetens, P., 2016. Simultaneous segmentation and anatomical labeling of the cerebral vasculature. *Medical Image Analysis* 32, 201–215.
- Smith, S.M., 2002. Fast robust automated brain extraction. *Human Brain Mapping* 17, 143–155.
- Zhao, M., Hamarneh, G., 2014. Bifurcation Detection in 3D Vascular Images Using Novel Features and Random Forest, in: *IEEE International Symposium on Biomedical Imaging (IEEE ISBI)*, pp. 421–424.

Bibliographie

- [Autrusseau 1999] Florent Autrusseau. Modélisation psychovisuelle pour l'aquamarquage des images. Master's thesis, Université de Nantes, Rue Ch. Pauc, Nantes, France, Nov. 1999. (Cité en page 22.)
- [Autrusseau 2002a] F. Autrusseau et Jp. Guédon. *Image watermarking for copyright protection and data hiding via the Mojette transform*. In Security and Watermarking of Multimedia Contents IV, SPIE Electronic Imaging,, volume 4675, pages 378–386, San Jose, CA, January 2002. (Cité en page 84.)
- [Autrusseau 2002b] F. Autrusseau et Jp. Guédon. *Image Watermarking in the Fourier Domain Using the Mojette Transform*. In 14th IEEE International Conference on Digital Signal Processing (DSP2002), volume 2, pages 725–728, Santorini Greece, July 2002. (Cité en page 84.)
- [Autrusseau 2002c] Florent Autrusseau. *Modélisation Psychovisuelle pour le tatouage des images*. PhD thesis, Université de Nantes, Rue Ch. Pauc, Nantes, France, 9 2002. (Cité en page 22.)
- [Autrusseau 2003a] F. Autrusseau et Jp. Guédon. *Chiffrement Mojette d'images médicales*. ISI journal, Ingénierie des Systèmes d'Information, Health information systems, vol. 8, no. 1, pages 113–134, March 2003. (Cité en page 29.)
- [Autrusseau 2003b] F. Autrusseau et Jp. Guédon. *A joint multiple description-encryption image algorithm*. In IEEE International Conference on Image Processing, ICIP, 2003, volume 3, pages 269–272, Barcelona, SPAIN., 2003. (Cité en page 29.)
- [Autrusseau 2003c] F. Autrusseau et Jp. Guédon. *Perceptual image watermarking using a secure Mojette transmission scheme*. In 4th Workshop COST 276, Transmitting processing and watermarking multimedia contents, Bordeaux, France, March 31st - April 1st 2003. (Cité en page 84.)
- [Autrusseau 2003d] F. Autrusseau, Jp. Guédon et Y. Bizais. *Watermarking and cryptographic schemes for medical imaging*. In SPIE Medical Imaging, Image processing, volume 5032(105), pages 958–965, San Diego California, USA, 15-20 February 2003. (Cité en page 29.)
- [Autrusseau 2006a] F. Autrusseau, B. Parrein et M. Servieres. *Lossless Compression Based on a Discrete and Exact Radon Transform : A Preliminary Study*. In IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'06, volume 2, pages 425–428, Toulouse, France, 2006. (Cité en pages 29, 76, 77, 80, 107 et 117.)
- [Autrusseau 2006b] F. Autrusseau et S. K. Shevell. *Temporal nulling of induction from spatial patterns modulated in time*. Cambridge University press, Visual Neuroscience, vol. 23, no. 3-4, pages 479–482, May 2006. (Cité en pages 23, 28 et 59.)

- [Autrusseau 2007a] F. Autrusseau et P. Le Callet. *A robust image watermarking technique based on quantization noise visibility thresholds*. Elsevier Signal Processing, vol. 87, no. 6, pages 1363–1383, June 2007. (Cité en pages 26, 34, 37 et 116.)
- [Autrusseau 2007b] F. Autrusseau, P. Le Callet et A. Ninassi. *A study of content based watermarking using an advanced HVS model*. In IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pages 485–488, Nov 26-28 2007. (Cité en page 8.)
- [Autrusseau 2008] F. Autrusseau. *Fourier sub-band embedding*. Online, 2008. (Cité en pages 27 et 45.)
- [Autrusseau 2009a] F. Autrusseau et M. Babel. *Subjective Quality assessment of LAR coded art images*. Online, 2009. (Cité en pages 27, 45, 47 et 111.)
- [Autrusseau 2009b] F. Autrusseau et P. Bas. *Subjective Quality assessment of the Broken Arrows watermarking technique*. Online, 2009. (Cité en pages 27, 45 et 47.)
- [Autrusseau 2009c] F. Autrusseau et Peter Meerwald. *DT-CWT versus DWT watermark embedding*. Online, 2009. (Cité en pages 27, 45 et 47.)
- [Autrusseau 2010a] F. Autrusseau, S. David et Vinod Pankajakshan. *A Subjective Study of Visibility Thresholds for wavelet domain watermarking*. In IEEE International Conference on Image Processing, ICIP'2010, pages 201–204, Hong Kong, Sept 26-29 2010. (Cité en pages 8 et 26.)
- [Autrusseau 2010b] F. Autrusseau, T. Stütz et V. Pankajakshan. *Subjective quality assessment of selective encryption techniques*. Online, 2010. (Cité en pages ix, 7, 27, 45, 48 et 49.)
- [Autrusseau 2011a] F. Autrusseau, S. David, V. Pankajakshan et P. Campisi. *A perceptually driven hybrid additive-multiplicative watermarking technique in the wavelet domain*. In Electronic Imaging : Media Watermarking, Security and Forensics XIII, volume 7880, San Francisco, USA, 23-27 January 2011. (Cité en pages 7, 8 et 26.)
- [Autrusseau 2011b] F. Autrusseau, L. Thibos et S. K. Shevell. *Chromatic and Wavefront Aberrations : L-, M- and S-cone Stimulation with Typical and Extreme Retinal Image Quality*. Elsevier Vision Research, vol. 51, pages 21–22, 2011. (Cité en pages 23, 28, 59, 63, 113, 116 et 121.)
- [Autrusseau 2013a] F. Autrusseau. *Method for inserting in the Fourier domain, a digital watermark in an image and corresponding method for detecting a digital watermark in an image to be analyzed, (Patent)*, 2013. (Cité en pages 34, 36, 108, 112 et 113.)
- [Autrusseau 2013b] F. Autrusseau et D. Goudia. *2AFC evaluation of wavelet watermarking in the smooth areas*. online, 2013. (Cité en pages 7, 27 et 45.)
- [Autrusseau 2013c] F. Autrusseau et D. Goudia. *Non linear hybrid watermarking for High Dynamic Range Images*. In IEEE International Conference on Image

- Processing, ICIP'2013, pages 4527–4531, Melbourne, Australia, Sept 15-18 2013. (Cité en page 7.)
- [Autrusseau 2013d] F. Autrusseau et M. Urvoy. *2AFC evaluation of Fourier watermarked images/documents*. online, 2013. (Cité en pages 7, 27, 45 et 47.)
- [Autrusseau 2016] F. Autrusseau, Y. Amouriq et Jp. Guédon. *Linking bone microarchitecture to projections texture analysis*. In SPIE Medical Imaging, San Diego, CA, USA, 27 Feb. - 3 Mar 2016. (Cité en page 111.)
- [Autrusseau 2020] F. Autrusseau, A. Nouri et R. Bourcier. *Method for Segmenting Digital Images, Devices and Systems for the Same, (Patent)*, Feb. 2020. (Cité en pages 7, 103, 113 et 117.)
- [Barten 1999] PGJ Barten. Contrast sensitivity of the human eye and its effect on image quality. SPIE Press, 1999. (Cité en page 37.)
- [Bay 2006] H. Bay, T. Tuytelaars et L. Van Gool. *SURF : Speeded Up Robust Features*. In 9th European Conference on Computer Vision, Graz, Austria, 2006. (Cité en pages 55 et 109.)
- [Bléry 2014] P. Bléry, F. Autrusseau, E. Crauste, E. Freuchet, P. Weiss, Jp. Guédon et Y. Amouriq. *Bone Vascularization : A Way To Study Bone Microarchitecture ?* In SPIE Medical Imaging, San Diego, CA, USA, Feb. 15-20 2014. (Cité en pages 23 et 111.)
- [Bouchakour 2008] M. Bouchakour, G. Jeannic et F. Autrusseau. *JND Mask Adaptation for Wavelet Domain Watermarking*. In IEEE Intl. Conf. on Multimedia & Expo, ICME 2008, pages 201–204, Hannover, Germany, June 23-26 2008. (Cité en pages 8, 26 et 36.)
- [Bourcier 2017a] R. Bourcier, S. Chatel, E. Bourcereau, S. Jouan, HL. Marec, B. Daumas-Duport, M. Sevin-Allouet, B. Guillon, V. Roualdes, T. Riem, B. Isidor, P. Lebranchu, J. Connault, TL. Tourneau, A. Gaignard, G. Loirand, R. Redon et H. Desal. *Understanding the Pathophysiology of Intracranial Aneurysm : The ICAN Project*. Neurosurgery, vol. 80, no. 4, pages 621–626, 2017. (Cité en page 30.)
- [Bourcier 2017b] R. Bourcier, C. Lenoble, B. Guyomarch-Delasalle, B. Daumas-Duport, C. M Papagiannaki, R. Redon et H. Desal. *Is there an inherited anatomical conformation favoring aneurysmal formation of the anterior communicating artery ?* Neurosurgery, vol. 126, pages 1598–1605, 2017. (Cité en page 30.)
- [Carnec 2008] M. Carnec, P. Le Callet et D. Barba. *Objective quality assessment of color images based on a generic perceptual reduced reference*. Signal Processing : Image Communication, vol. 23(4), pages 239–256, 2008. (Cité en page 48.)
- [Carosi 2010] M. Carosi, V. Pankajakshan et F. Autrusseau. *Toward a simplified perceptual quality metric for watermarking applications*. In Proceedings of the SPIE conference on Electronic Imaging, volume 7542, 2010. (Cité en pages 7, 8, 27, 44, 45, 46, 107, 113 et 116.)

- [Cermak 2009] G. Cermak, L. Thorpe et M. Pinson. *Test Plan for Evaluation of Video Quality Models for Use with High Definition TV Content*, 2009. (Cité en page 51.)
- [Chandra 2014] S.S. Chandra, N. Normand, A. Kingston, JP. Guédon et I. Svalbe. *Robust Digital Image Reconstruction via the Discrete Fourier Slice Theorem*. IEEE Signal Processing Letters, vol. 21(6), June 2014. (Cité en page 75.)
- [Chen 2008] W. Chen, F. Autrusseau et P. Le Callet. *A Error-Propagation-Free Perceptual Watermarking Algorithm for H.264/AVC encoded video*. In 1st Workshop on Visual Signal Processing and Analysis (VISPA), Paris, France., Nov. 12-13 2008. (Cité en pages 8 et 41.)
- [Chen 2014] W. Chen, Z. Shahid, T. Stütz, F. Autrusseau et P. Le Callet. *Robust drift-free bit-rate preserving H.264 watermarking*. Springer, Multimedia Systems, vol. 20 (2), no. ISSN 0942-4962, pages 179–193, Apr 09 2014. (Cité en pages ix, 7, 8, 41, 43, 107 et 116.)
- [CIE 2006] CIE. *Fundamental Chromaticity Diagram with Physiological Axes – Part I*. Cie technical report, CIE, 2006. (Cité en page 68.)
- [Crauste 2015] E. Crauste, F. Autrusseau, Jp. Guédon, P. Pilet, Y. Amouriq, P. Weiss et B. Giumelli. *Bone vascularization and bone microarchitecture characterizations according to the microCT resolution*. In SPIE Medical Imaging, Orlando, FL, USA, Feb. 21-26 2015. (Cité en pages 23 et 111.)
- [Daly 1993] Scott J. Daly. *Visible differences predictor : an algorithm for the assessment of image fidelity*. In Andrew B. Watson, editeur, Digital Images and Human Vision, pages 179–206. The MIT Press, Aug. 1993. (Cité en pages 26 et 37.)
- [Deans 1983] Stanley Roderick Deans. The Radon transform and some of its applications. Numéro 9780471898047 de Dover Books on Mathematics Series. Wiley, The University of Michigan, second edition édition, 1983. (Cité en page 71.)
- [Dong 2012] J. Dong, L. Su, Y. Zhang, F. Autrusseau et Z. Yang. *Illumination direction estimation of three-dimensional surface texture based on active basis and Mojette transform*. Journal of Electronic Imaging, vol. 21, page 013023, Apr 09 2012. (Cité en pages 23, 72, 112 et 117.)
- [Evenou 2006] P. Evenou, F. Autrusseau et T. Hamon. *Secure Distributed Storage based on the Mojette transform*. In Hermes-Lavoisier editors, editeur, NOTERE'06, pages 161–170, Toulouse, France, 2006. (Cité en page 8.)
- [Evenou 2019] P. Evenou, D. Feron, S. David et J-P. Monchanin. *System and Method for Processing Data, (Patent)*, Nov. 2019. (Cité en page 28.)
- [Falconer 2003] Kenneth Falconer, editeur. Fractal geometry : Mathematical foundations and applications. John Wiley & Sons, 2003. (Cité en page 96.)
- [Freuchet 2015] E. Freuchet, B. Recur, Jp. Guédon, A. Kingston, F. Autrusseau et Y. Amouriq. *Building a bone micro-CT images atlas for micro-architecture*

- recognition*. In SPIE Medical Imaging, Orlando, FL, USA, Feb. 21-26 2015. (Cité en page 7.)
- [Furon 2008] T. Furon et P. Bas. *Broken arrows*. In EURASIP Journal on Information Security, pages 1–13, 2008. (Cité en page 47.)
- [Grubbs 1969] Frank E. Grubbs. *Procedures for detecting outlying observations in samples*. Technometrics, vol. 11, no. 1, pages 1–21, 1969. (Cité en page 39.)
- [Guédon 1995] JP. Guédon, D. Barba et N. Burger. *Psychovisual image coding via an exact discrete Radon transform*. In Lance T. Wu, editeur, Proc. Visual Communications & Image Processing (VCIP), pages 562–572, Taipei, Taiwan, May 1995. (Cité en pages 28 et 76.)
- [Guédon 2012] Jp. Guédon, F. Autrusseau, Y. Amouriq, P. Bléry, Jm. Bouler, P. Weiss, Fx. Barbarin, T. Dallet et V. Dallerit. *Exploring relationships between fractal dimension and trabecular bone characteristics*. In SPIE Medical Imaging, volume 8317(7), San Diego, USA, 4-9 February 2012. (Cité en pages 23 et 111.)
- [Guédon 2013] JP. Guédon, editeur. *The mojette transform : Theory and applications*. ISTE. Wiley, 2013. (Cité en pages 71, 72 et 75.)
- [Guillou 2020] J. Guillou, F. Autrusseau et R. Bourcier. *Brain vasculature segmentation based on human perception criteria*. In SPIE Medical Imaging, Houston, TX, USA, 2020. (Cité en pages 7, 103, 112 et 117.)
- [Hofbauer 2018] H. Hofbauer, F. Autrusseau et A. Uhl. *To See or Not To See : Determining the Recognition Threshold of Encrypted Images*. In 7th European Workshop on Visual Information Processing (EuVIP’18), Tampere, Finland, 2018. (Cité en pages 27, 51 et 52.)
- [Hofbauer 2019] H. Hofbauer, F. Autrusseau et A. Uhl. *University of Salzburg Encryption Evaluation Database (USEE DB)*. online, 2019. (Cité en page 45.)
- [Hofbauer 2020a] H. Hofbauer, F. Autrusseau et A. Uhl. *To Recognize or not to Recognize — A Database of Encrypted Images with Subjective Recognition Ground Truth*. Accepted for publication in Elsevier Information Sciences, Nov. 2020. (Cité en pages 27, 52 et 116.)
- [Hofbauer 2020b] H. Hofbauer, F. Autrusseau et A. Uhl. *University of Salzburg Encryption and Quality Evaluation Database (USEEQ DB)*. online, 2020. (Cité en page 55.)
- [IEC1999 1999] IEC1999. *Multimedia systems and equipment - Colour measurement and management - Part 2-1 : Colour management - Default RGB colour space*, 1999. (Cité en page 51.)
- [ITU-R BT.2022 2012] ITU-R BT.2022. *General viewing conditions for subjective assessment of quality of SDTV and HDTV television pictures on flat panel displays BT Series Broadcasting service*, 2012. (Cité en page 51.)
- [ITU-R-BT.500-11 2004] ITU-R-BT.500-11. *Methodology for the subjective assessment of the quality of television pictures*, 2004. (Cité en page 51.)

- [Jia 2008] P. Jia, J. Dong, L. Qi et F. Atrusseau. *Directionality Measurement and Illumination Estimation of 3D Surface Textures by Using Mojette Transform*. In 19th International Conference on Pattern Recognition, ICPR., 2008. (Cité en page 112.)
- [Katz 1978] M.B. Katz. Questions of uniqueness and resolution in reconstruction from projections. Lecture Notes in Biomathematics. Springer Berlin Heidelberg, 1978. (Cité en page 72.)
- [Kingston 2007] A. Kingston, S. Colosimo, P. Campisi et F. Atrusseau. *Lossless Image Compression and Selective Encryption Using a Discrete Radon Transform*. In IEEE International Conference on Image Processing, ICIP'2007, pages 465–468, 2007. (Cité en pages 7, 8, 29, 86, 107, 108 et 113.)
- [Kingston 2008a] A. Kingston et F. Atrusseau. *Lossless Image Compression via Predictive Coding of Discrete Radon Projections*. Elsevier Signal Processing : Image Communication, vol. 23, no. 4, pages 313–324, May 2008. (Cité en pages 29, 108, 117 et 121.)
- [Kingston 2008b] A. Kingston, B. Parrein et F. Atrusseau. *Redundant Image Representation via Multi-Scale Digital Radon Projection*. In IEEE International Conference on Image Processing, ICIP'2008, 2008. (Cité en pages 7, 29 et 108.)
- [Kingston 2009a] A. Kingston et F. Atrusseau. The mojette transform : Theory and applications, chapitre Lossless compression, Chap 9, pages 215–238. ISBN : 978-1-84821-080-6. ISTE-Wiley, 2009. (Cité en pages 7, 29, 108 et 117.)
- [Kingston 2009b] A. Kingston, F. Atrusseau, E. Grall, T. Hamon et B. Parrein. The mojette transform : Theory and applications, chapitre Mojette based security, Chap 10, pages 239–268. ISBN : 978-1-84821-080-6. ISTE-Wiley, 2009. (Cité en pages 8, 29 et 108.)
- [Kingston 2009c] A. Kingston, B. Parrein et F. Atrusseau. The mojette transform : Theory and applications, chapitre Multiresolution Mojette transform, Chap 6, pages 115–148. ISBN : 978-1-84821-080-6. ISTE-Wiley, 2009. (Cité en pages 29, 76, 81, 108 et 117.)
- [Koh 2011] K-J. Koh et K-A. Kim. *Utility of the computed tomography indices on cone beam computed tomography images in the diagnosis of osteoporosis in women*. Imaging Sci Dent., vol. 41(3), page 1016, 2011. (Cité en pages 30 et 97.)
- [L'Allinec 2020] V. L'Allinec, S. Chatel, M. Karakachoff, E. Bourcereau, Z. Lamoureux, A. Gaignard, F. Atrusseau, S. Jouan, A.C. Vion, G. Loirand, H. Desal, O. Naggara, R. Redon, M. Edjlali et R. Bourcier. *Cerebral Aneurysm, prediction of evolution : The UCAN Project*. Neurosurgery, Oxford University Press, Feb. 2020. (Cité en page 30.)
- [Le Callet 2005] P. Le Callet et F. Atrusseau. *Subjective quality assessment IRC-CyN/IVC database*. Online, 2005. (Cité en page 27.)

- [Le Callet 2008] P. Le Callet, F. Autrusseau et P. Campisi. Visibility control and quality assessment of watermarking and data hiding algorithms, chapitre Chap. IX : Visibility control and Quality assessment of watermarking and data hiding algorithms, pages 164–193. ISBN : 978-1-59904-869-7. Idea Group Publishing, April 2008. (Cité en pages 26, 45, 113 et 116.)
- [Lewis 1995] J. P. Lewis. *Fast normalized cross-correlation*. In Vision interface, pages 120–123, 1995. (Cité en pages 37, 55 et 109.)
- [Lin 2008] WH Lin, SJ Horng et TW Kao. *An Efficient Watermarking Method Based on Significant Difference of Wavelet Coefficient Quantization*. IEEE Transactions on Multimedia, vol. 10, no. 5, pages 746–757, Aug. 2008. (Cité en pages 40 et 41.)
- [Lowe 1999] David G. Lowe. *Object recognition from local scale-invariant features*. In Proceedings of the International Conference on Computer Vision, volume 2, pages 1150–1157, 1999. (Cité en pages 55 et 109.)
- [MacLeod 1979] D.I.A. MacLeod et R.M. Boynton. *Chromaticity diagram showing cone excitation by stimuli of equal luminance*. Journal of the Optical Society of America, vol. 69, pages 1183–1185, 1979. (Cité en page 60.)
- [MacLeod 1983] D.I.A. MacLeod et M.A. Webster. Factors influencing the color matches of normal observers, pages 81–92. Academic London, 1983. (Cité en page 68.)
- [Mahé 2016] L. Mahé, F. Autrusseau, H. Desal, Jp. Guédon, H. Der Sarkissian, Y. Le Teurnier et S. Davila. *Quality metrics can help the expert during neurological clinical trials*. In SPIE Medical Imaging, San Diego, CA, USA, 27 Feb. - 3 Mar 2016. (Cité en pages 6 et 111.)
- [Maillet 2008a] A. Maillet, F. Rivard, F. Autrusseau, R. Lehn, M. Gelgon et P. Bruneau. *"Safim@ge"*. In NEM Summit, Networked & Electronic Media, Saint Malo, France, Oct 13-15 2008. (Cité en pages 41 et 107.)
- [Maillet 2008b] A. Maillet, F. Rivard, R. Lehn, M. Gelgon et F. Autrusseau. *"Safim@ge : projets de recherche"*, Jan 29-30 2008. (Cité en page 41.)
- [Maillet 2009] A. Maillet, F. Rivard, F. Autrusseau, R. Lehn, M. Gelgon et P. Bruneau. *"Safim@ge"*, Jan 27-28 2009. (Cité en pages 41 et 112.)
- [Maillet 2011] A. Maillet, F. Rivard, F. Autrusseau, R. Lehn, M. Gelgon et P. Bruneau. *Safim@ge*. WISG'2011 - Workshop Interdisciplinaire sur la Sécurité Globale, Jan 25-26 2011. (Cité en pages 41, 107 et 112.)
- [Mao 2006] Y. Mao et M. Wu. *A joint signal processing and cryptographic approach to multimedia encryption*. IEEE Transactions on Image Processing, vol. 15, no. 7, pages 2061–2075, 2006. (Cité en page 49.)
- [Marimont 1994] D.H. Marimont et B.A. Wandell. *Matching color images : The effects of axial chromatic aberration*. Journal of the Optical Society of America A, vol. 11, pages 3113–3122, 1994. (Cité en pages x, 63 et 65.)

- [Marini 2007a] E. Marini, F. Autrusseau et P. Le Callet. *Evaluation of standard watermarking techniques*. Online, 2007. (Cité en pages 8, 27 et 45.)
- [Marini 2007b] E. Marini, F. Autrusseau, P. Le Callet et P. Campisi. *Evaluation of standard watermarking techniques*. In SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX, San Jose , USA, January 2007. (Cité en pages 8, 107 et 113.)
- [Meyer 2001] B. Meyer et P. Tischer. *Glicbawls - Grey Level Image Compression by Adaptive Weighted Least Squares*. In IEEE Computer Society, Proc. Data Compression Conference, page 503, Snowbird, Utah, USA, 2001. (Cité en page 79.)
- [Michel 2018] P. Michel, V. Prezelin, P. Bléry, Y. Amouriq, Jp. Guédon et F. Autrusseau. *Automatic Generation of the Dental Scheme based on 2D radiographs*. In SPIE Medical Imaging, San Diego, CA, USA, 2018. (Cité en pages 30, 94 et 111.)
- [Monnier 2004] P. Monnier et S. K. Shevell. *Chromatic induction from S-cone patterns*. Vision Research, vol. 44(9), pages 849–856, April 2004. (Cité en page 60.)
- [Ninassi 2006] A. Ninassi, P. Le Callet et F. Autrusseau. *Pseudo No Reference image quality metric using perceptual data hiding*. In SPIE Human Vision and Electronic Imaging, volume 6057, pages 146–157, San Jose, CA, USA, January 2006. (Cité en pages 8, 35 et 107.)
- [Noorkami 2007] M. Noorkami et R. M. Mersereau. *A Framework for Robust Watermarking of H.264-Encoded Video With Controllable Detection Performance*. IEEE Transactions on Information Forensics and Security, vol. 2, no. 12, pages 14–23, 2007. (Cité en pages ix, 42 et 43.)
- [Normand 1996a] N. Normand, JP. Guédon, O. Philippé et D. Barba. *Controlled redundancy for image coding and high-speed transmission*. In R. Ansari et MJ. Smith, éditeurs, Proc. SPIE Visual Communications and Image Processing 1996, volume 2727, pages 1070–81. SPIE, Feb 1996. (Cité en page 28.)
- [Normand 1996b] N. Normand, JP. Guédon, O. Philippé et D. Barba. *Controlled redundancy for image coding and high-speed transmission*. In R. Ansari et MJ. Smith, éditeurs, Proc. SPIE Visual Communications and Image Processing, volume 2727, pages 1070–81, Feb. 1996. (Cité en page 73.)
- [Normand 2006a] N. Normand, A. Kingston et P. évenou. *A Geometry Driven Reconstruction Algorithm for the Mojette transform*. In A. Kuba, LG. Nyúl et K. Palágyi, éditeurs, Proc. 13th International Conference on Discrete Geometry for Computer Imagery, volume LNCS4245, pages 122–33, Szeged, Hungary, Oct 2006. Springer-Verlag. (Cité en page 28.)
- [Normand 2006b] N. Normand, A. Kingston et P. Évenou. *A Geometry Driven Reconstruction Algorithm for the Mojette Transform*. In Attila Kuba, László G. Nyúl et Kálmán Palágyi, éditeurs, Discrete Geometry for Computer Imagery,

- pages 122–133, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg. (Cité en page 74.)
- [Nouri 2018] A. Nouri, F. Autrusseau et R. Bourcier. *Method for Locating and Characterizing Bifurcations of a Cerebral Vascular Tree, Associated Methods and Devices, (Patent)*, Aug. 2018. (Cité en pages 7, 99 et 113.)
- [Nouri 2019] A. Nouri, F. Autrusseau, R. Bourcier, A. Gaignard, V. L’Allinec, C. Menguy, J. Veziers, H. Desal, G. Loirand et R. Redon. *3D bifurcations characterization for intra-cranial aneurysms prediction*. In SPIE Medical Imaging, San Diego, CA, USA, 2019. (Cité en pages 7, 99 et 112.)
- [Nouri 2020] A. Nouri, F. Autrusseau, R. Bourcier, A. Gaignard, V. L’Allinec, C. Menguy, J. Veziers, H. Desal, G. Loirand et R. Redon. *Characterization of 3D Bifurcations in Micro-scan and MRA-TOF Images of Cerebral Vasculature for Prediction of Intra-Cranial Aneurysms*. Computerized Medical Imaging and Graphics, Aug. 2020. (Cité en pages 7, 31, 99, 112, 118 et 121.)
- [Pankajakshan 2010] V. Pankajakshan et F. Autrusseau. *A Multi-purpose Objective Quality Metric for Image Watermarking*. In IEEE International Conference on Image Processing, ICIP’2010, pages 2589–2592, Hong Kong, Sept 26-29 2010. (Cité en pages 7, 27, 44, 45, 48, 107 et 116.)
- [Parrein 2003] B. Parrein, JP. Guédon et N. Normand. *Multimedia forward error correcting codes for wireless LAN*. Annals of Telecommunications, vol. 3, no. 4, pages 448–463, 2003. (Cité en page 29.)
- [Petitcolas 1998] F.A.P. Petitcolas, R.J. Anderson, et M.G. Kuhn. *Attacks on copyright marking systems*. In Second Workshop on information hiding, pages 218–238, 1998. (Cité en page 35.)
- [Petitcolas 2000] Fabien A. P. Petitcolas. *Watermarking schemes evaluation*. IEEE Signal Processing, vol. 17 (5), pages 58–64, September 2000. (Cité en page 35.)
- [Pitzalis 2008] D. Pitzalis, G. Aitken, F. Autrusseau, M. Babel, F. Cayre et W. Puech. *TSAR : Secure Transfer of High Resolution Art Images*. In Electronic Imaging and the Visual Arts, EVA 2008, Firenze, Italy, 2008. (Cité en pages 36 et 111.)
- [Poljicak 2011] A. Poljicak, M. Mandic et D. Agic. *Discrete Fourier transform-based watermarking method with an optimal implementation radius*. Journal of Electronic Imaging, vol. 20, no. 3, pages 033008–1–8, Juillet 2011. (Cité en pages 38 et 39.)
- [Saadane 1999a] A. Saadane et F. Autrusseau. *Aquamarquage des documents audiovisuels pour leur transmission, diffusion, circulation, distribution en toute sécurité sur les réseaux hétérogènes de telecommunication. Analyse et specifications*. Rnrt-aquamrars, IRESTE Nantes, 1999. (Cité en page 22.)
- [Saadane 1999b] A. Saadane et F. Autrusseau. *Aquamarquage des documents audiovisuels pour leur transmission, diffusion, circulation, distribution en*

- toute sécurité sur les réseaux hétérogènes de telecommunication. *Modélisation psycho-visuelle pour le tatouage des images*. Rnrt-aquamrars, IRESTE Nantes, 1999. (Cité en page 22.)
- [Sarkar 2009a] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, J. Stauder et P. Morvan. *Study of Observer Variability on Modern Display Colorimetry : An Analysis of CIE 2006 Model*. In 11th Congress of the International Colour Association (AIC), Sydney, Australia., 27 Sept - 2 Oct 2009 2009. (Cité en pages 6 et 59.)
- [Sarkar 2009b] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, J. Stauder et P. Morvan. *Study of Observer Variability on Modern Display Colorimetry : Comparison of CIE 2006 Model and 10 deg Standard Observer*. In 11th Congress of the International Colour Association (AIC), Sydney, Australia., 27 Sept - 2 Oct 2009 2009. (Cité en pages 6 et 59.)
- [Sarkar 2010a] A. Sarkar, L. Blonde, P. Le Callet, F. Autrusseau, P. Morvan et J. Stauder. *A color matching experiment using two displays : design considerations and pilot test results*. In Proceedings of the Fifth European Conference on Color in Graphics, Imaging and Vision, CGIV, Joensuu, Finland, 2010. (Cité en pages 6 et 59.)
- [Sarkar 2010b] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, P. Morvan et J. Stauder. *Toward Reducing Observer Metamerism Issue in Industrial Applications : Colorimetric Observer Categories and Observer Classification*. In in the eighteenth Color Imaging Conference, Color Science and Engineering Systems, Technologies, and Applications, CIC18, San Antonio, Texas, November 8–12 2010. (Cité en pages 6 et 59.)
- [Sarkar 2010c] A. Sarkar, L. Blonde, P. Le Callet, F. Autrusseau, J. Stauder et P. Morvan. *How Modern Displays Push Conventional Colorimetry to its Limit*. In Colour in art, science, design, conservation, research, printmaking, digital technologies, textiles (CREATE), Gjøvik, Norway, 2010. (Cité en pages 6 et 59.)
- [Sarkar 2010d] A. Sarkar, L. Blonde, P. Le Callet, F. Autrusseau, J. Stauder et P. Morvan. *Modern displays : Why we see different colors, and what it means ?* In 2nd European Workshop on Visual Information Processing, EU-VIP'2010, Paris, July, 5th-7th, France, 2010., 2010. (Cité en pages 6 et 59.)
- [Sarkar 2011a] A. Sarkar, F. Autrusseau, F. Vienot, P. Le Callet et L. Blonde. *From CIE 2006 Model to Improved Age-Dependent and Average Colorimetric Observers*. Journal of the Optical Society of America A, JOSA-A, vol. 28 (10), pages 2033–2048, 2011. (Cité en pages 6, 59 et 116.)
- [Sarkar 2011b] Abhijit Sarkar. *Identification and Assignment of Colorimetric Observer Categories and Their Applications in Color and Vision Sciences*. PhD thesis, Université de Nantes, August 2011. (Cité en page 27.)
- [Sheikh 2005] H.R. Sheikh et A.C. Bovik. *Image Information and Visual Quality*. In IEEE Trans. on Image Processing, May 2005. (Cité en page 48.)

- [Smith 1983] V.C. Smith, J. Pokorny et Q. Zaidi. How do sets of color-matching functions differ, pages 93–105. Academic London, 1983. (Cité en page 68.)
- [Solachidis 2001] V Solachidis et I Pitas. *Circularly symmetric watermark embedding in 2-D DFT domain*. IEEE Trans. Image Process., vol. 10(11), pages 1741–53, Janvier 2001. (Cité en pages 38, 39, 40 et 41.)
- [Stockman 1993] A. Stockman, D.I.A. MacLeod et N.E. Johnson. *Spectral sensitivities of the human cones*. J. Opt. Soc. Am. A, vol. 10, no. 2, pages 2491–2521, 1993. (Cité en page 68.)
- [Strauss 2009] C. Strauss, F. Pasteau, F. Autrusseau, M. Babel, L. Bedat et O. Deforges. *Subjective and Objective Quality Evaluation of LAR coded art images*. In IEEE Intl. Conf. on Multimedia & Expo, ICME 2009, pages 674–677, New York City, NY, June 28 - July 3. 2009. (Cité en page 111.)
- [Stütz 2010] T. Stütz, V. Pankajakshan, F. Autrusseau, A. Uhl et H. Hofbauer. *Subjective and Objective Quality Assessment of Transparently Encrypted JPEG2000 Images*. In 12th ACM Workshop on Multimedia and Security, ACM-MMSEC'2010, pages 247–252, Rome, Italy, September 9-10 2010. (Cité en pages 7, 27, 45, 48, 52 et 112.)
- [Stütz 2011] T. Stütz et F. Autrusseau. *H.264/CAVLC Structure preserving watermarking*. online, 2011. (Cité en pages 7, 27, 43, 45 et 112.)
- [Stütz 2014] T. Stütz, F. Autrusseau et A. Uhl. *Non-Blind Structure-Preserving Substitution Watermarking of H.264/CAVLC Inter-Frames*. IEEE Transactions on Multimedia, vol. 16 (5), pages 1337–1349, Aug. 2014. (Cité en pages 7, 43, 112 et 116.)
- [Thibos 1999] L. N. Thibos et A. Bradley. Modeling the refractive and neuro-sensor systems of the eye, chapitre New York : McGraw-Hill, pages 101–159. Idea Group Publishing, 1999. (Cité en page 63.)
- [Thibos 2002] L. N. Thibos, A. Bradley et X. Hong. *A statistical model of the aberration structure of normal, well-corrected eyes*. Ophthalmic and Physiological Optics, vol. 22, pages 427–433, 2002. (Cité en page 63.)
- [Urvoy 2013] M. Urvoy et F. Autrusseau. *Print and scan robust image watermarking algorithm*. Journal of South China University of Technology, JSCUT, vol. 7, pages 120–125, 2013. (Cité en pages 7 et 37.)
- [Urvoy 2014a] M. Urvoy et F. Autrusseau. *Application of Grubbs' test for outliers do the detection of watermarks*. In 2nd ACM Workshop on Information Hiding and Multimedia Security, pages 49–60, Salzburg, Austria, June 11-13 2014. (Cité en pages 7, 34, 39 et 108.)
- [Urvoy 2014b] M. Urvoy, D. Goudia et F. Autrusseau. *Perceptual DFT watermarking with improved detection and robustness to geometrical distortions*. in IEEE Transactions on Information Forensics and Security, IEEE-TIFS, vol. 9(7), pages 1108–1119, July 2014. ISSN : 1556-6013. (Cité en pages 7, 26, 34, 37, 39, 53, 108, 116 et 121.)

- [Viénot 1980] Françoise Viénot. *Relations between inter- and intra-individual variability of color-matching functions. experimental results*. J. Opt. Soc. Am. A, vol. 70, no. 12, pages 1476–1483, 1980. (Cité en page 68.)
- [Voloshynovskiy 2000] S. Voloshynovskiy, S. Pereira, A. Herrigel, N. Baumgartner et T. Pun. *Generalized watermarking attack based on watermark estimation and perceptual remodulation*. In SPIE Security and Watermarking of Multimedia Contents, volume 3971, pages 358–370, 2000. (Cité en page 48.)
- [VQEG_Contributors 2008] VQEG_Contributors. *Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, phase II (FR-TV2)*, 2008. (Cité en page 51.)
- [VQEG_Contributors 2010] VQEG_Contributors. *Hybrid Perceptual/Bitstream Group TEST PLAN - Draft Version 1.9*, 2010. (Cité en page 51.)
- [VQEG_Contributors 2011] VQEG_Contributors. *Hybrid Perceptual/Bitstream Group TEST PLAN - Draft Version 2.9*, 2011. (Cité en page 51.)
- [Wang 2004] Z. Wang, A.C. Bovik, H.R. Sheikh et E.P. Simoncelli. *Image Quality Assessment : From Error Visibility to Structural Similarity*. IEEE Transactions on Image Processing, vol. 13(4), pages 600–612, 2004. (Cité en page 26.)
- [Wang 2006] Z. Wang et X. Shang. *Spatial pooling strategies for perceptual image quality assessment*. In IEEE International Conference on Image Processing, Atlanta, GA, Oct. 8-11 2006. (Cité en page 47.)
- [Watson 1993] A. Watson. *DCT Quantization Matrices Visually Optimized for Individual Images*. In Proc. SPIE Society of Photo-Optical Instrumentation Engineers, volume 1913, pages 202–216, San Jose, CA, USA, September 8 1993. (Cité en page 41.)
- [Weinberger 1996] M.J. Weinberger, G. Seroussi et G. Sapiro. *LOCO-I : a low complexity, context-based, lossless image compression algorithm*. In Proc. Data Compression Conference, IEEE Computer Society, pages 140–149, Los Alamitos, CA, USA, 1996. (Cité en page 79.)
- [Wu 1997] X. Wu et N. Memon. *Context-based, adaptive, lossless image coding*. IEEE Transactions on Communications, vol. 45(4), pages 437–444, 1997. (Cité en page 79.)
- [Wu 2000] X. Wu et N. Memon. *Context-based lossless interband compression-extending CALIC*. IEEE Transactions on Image Processing, vol. 9(6), pages 994–1001, 2000. (Cité en page 80.)
- [Zhang 2010] C. Zhang, J. Dong, J. Li et F. Atrousseau. *A New Information Hiding Method for Image Watermarking Based on Mojette Transform*. In Second International Symposium on Networking and Network Security, ISNNS, pages 124–128, Jinggangshan, China, 2-4 April 2010. (Cité en page 112.)
- [Zhang 2014] F. Zhang, W. Jiang, F. Atrousseau et W. Lin. *Exploring V1 by modeling the perceptual quality of Images*. Journal of Vision, vol. 14(1), no. 1, page 14, Jan. 2014. (Cité en page 112.)

-
- [Zou 2009a] D. Zou et J. Bloom. *H.264/AVC substitution watermarking : a CAVLC example*. In Proceedings of the SPIE, Media Forensics and Security, volume 7254, Jan Jose, CA, USA, Janvier 2009. SPIE. (Cité en page 44.)
- [Zou 2009b] D. Zou, N. Prigent et J. Bloom. *Compressed video stream watermarking for peer-to-peer based content distribution network*. In Proceedings of the IEEE International Conference on Multimedia & Expo, ICME '09, pages 1390–1393, New York, NY, USA, Juin 2009. IEEE. (Cité en page 44.)