

N° d'ordre : 634

THÈSE

présentée à

l'Université de Bretagne Occidentale
U.F.R. de Sciences & Techniques

pour obtenir le titre de

Docteur
Mention Électronique

par

Emmanuel Quémener

Synthèse et Analyse de Signatures Temporelles par Filtrage Spatial

—

Application à l'Identification d'Objets et la Reconnaissance d'Attitudes

soutenue le 20 janvier 1999 devant la commission d'examen composée de :

Rapporteurs :	P. AMBS C. OZKUL	Professeur à l'Université de Haute-Alsace Professeur à l'Université de Rouen
Examineurs :	L. CALVEZ G. BUREL	Professeur à l'Université de Bretagne Occidentale Professeur à l'Université de Bretagne Occidentale
Directeur :	J.-L. DE BOUGRENET DE LA TOCNAYE	Professeur à l'ENST Bretagne
Invités :	G. KERYER G. COPPIN Y. PELLIARD	Chercheur à l'ENST Bretagne Ingénieur à Thomson-CSF/RCM Ingénieur DSP/STTC

*Travail effectué au sein du Département d'Optique
de l'École Nationale Supérieure de Télécommunications de Bretagne
UMR CNRS 6616*

Table des matières

Remerciements	11
Introduction	13
1 Le Filtrage : Outil de Reconnaissance de Formes	15
1.1 Introduction	15
1.2 Introduction au filtrage spatial	15
1.2.1 Filtrage dans l'espace de Fourier	15
1.2.2 La transformée de Fourier optique	16
1.2.3 Corrélation & filtrage adapté	17
1.3 Corrélation & reconnaissance de formes	17
1.4 La nécessité d'un critère de corrélation	18
1.5 Le choix d'un filtre de phase	19
1.6 La nécessité du contrôle de la sélectivité	23
1.7 D'une base de références au filtre composite	26
1.7.1 La synthèse d'un filtre composite	26
1.7.2 Optimisation du filtre composite	26
1.7.3 Robustesse d'un filtre composite	29
1.7.4 Nombre de références intégrées au filtre & nature du spectre	29
1.8 Contrôle de positionnement des pics de corrélation	36
1.9 Limitation de la saturation par filtrage multi-voies	38
1.10 La synthèse d'un filtre composite-multivoies	43
1.11 Le Prétraitement comme opération de filtrage	43
1.11.1 Intérêt d'un prétraitement	44
1.11.2 Un prétraitement intégrant la résolution	44
1.11.3 Intérêt d'un prétraitement linéaire	45
1.12 Conclusion	50
2 Synthèse et analyse de signatures	53
2.1 Introduction	53
2.2 Problématique	53
2.3 Le caractère temporel d'un signal	53
2.4 Vers la formation d'une signature	54
2.5 Synthèse de signature	55
2.5.1 Introduction	55
2.5.2 Signature d'une antenne synthétique	56
2.5.3 Signature de séquences d'images	59

2.5.4	Synthèse de séquences d'images par une base de références	60
2.5.5	Synthèse «auto-suffisante» d'une séquence d'images	63
2.6	Analyse de Signatures	69
2.6.1	Signaux d'antenne synthétique	69
2.6.2	Signatures de séquences d'images	74
2.7	Architecture générique de traitement	74
2.7.1	L'architecture générique, de haut en bas	75
2.8	Conclusion	78
3	Applications	79
3.1	Signaux Visible & Infrarouge couplés	79
3.1.1	Présentation des séquences utilisées	80
3.1.2	Approche «PCE»	80
3.1.3	Approche «Parcours Fractal»	86
3.1.4	Comparaison des deux approches	92
3.2	Signaux Radar «Temps-Fréquence»	93
3.2.1	Importance de l'horizon temporel	93
3.2.2	Le choix du «Parcours Fractal»	94
3.2.3	Expérimentation	94
3.3	Conclusion	99
4	Implantations & Prospectives	101
4.1	Introduction	101
4.2	Un traitement simple pour des implantations duales	101
4.3	Implantation Optique	102
4.3.1	Généralités	102
4.3.2	Implantation des différentes opérations	104
4.3.3	Perte de dynamique & effets de binarisation	107
4.3.4	Quelques éléments d'implantation	110
4.3.5	Éléments d'une validation expérimentale	112
4.4	Implantation Numérique	114
4.4.1	Un coût de calcul non-négligeable	115
4.4.2	Une implantation sur DSP	115
4.5	Comparaison entre implantations	115
4.6	Prospectives	117
4.7	Conclusion	117
	Conclusion	119
A	Les ondelettes continues	121
A.1	Définition	121
A.2	Illustrations	122
B	Un Parcours Fractal : la Courbe de Hilbert	125
B.1	La courbe de Hilbert	125
B.2	Un algorithme de synthèse récursif	126
B.3	Application	126

C Réalisation d'une Base de Données Multi-Senseurs	127
C.1 Protocole de prise de vue	127
C.2 Matériel de prise de vue & de numérisation	128
D Performances d'un DSP Moderne	133
D.1 Un exemple de DSP : la famille <i>TMS320C6x</i>	133
Bibliographie	137

Table des figures

1.1	Corrélation entre les images de deux visages identiques.	20
1.2	Corrélation entre les images de deux visages différents.	21
1.3	Rapport entre les <i>PCE</i> issus de corrélations entre visages identiques ou différents.	21
1.4	Résultats de corrélations pour deux visages.	22
1.5	Comparaison d'une empreinte digitale à une base de 99 empreintes.	24
1.6	Comparaison d'une empreinte en rotation avec elle-même.	25
1.7	Réponses du filtre pour une empreinte en rotation.	28
1.8	réponses de filtres composites comportant 5, 11, 21, 31 et 61 références à une empreinte en rotation.	29
1.9	Réponses de filtres composites comportant 5, 11, 21, 31 et 61 références à une base d'empreintes prises au hasard.	30
1.10	Décroissance du <i>PCE</i> en fonction du nombre de références intégrées dans le filtre composite.	31
1.11	Exemples de spectres associés à deux types d'image : une empreinte digitale et un visage.	32
1.12	Réponse du filtre pour une attitude de visage.	33
1.13	Réponses de filtres composites comportant 5, 11 et 21 références à une empreinte en rotation.	34
1.14	Décroissance du <i>PCE</i> en fonction du nombre de références intégrées dans le filtre composite.	35
1.15	Exemple de fonction de phase intégrée au filtre.	37
1.16	Application d'un filtre composite avec porteuse.	39
1.17	Duplication d'un spectre.	40
1.18	Synthèse d'un filtre multivoies.	41
1.19	Corrélation par un filtre multivoies.	42
1.20	Exemple de filtrage par une base d'ondelettes de niveaux a compris entre $1/4$ et 4	46
1.21	Corrélation avec prétraitement linéaire à base d'ondelettes.	49
2.1	Codage de l'information dans une signature.	55
2.2	Disposition des paramètres pour la synthèse d'une signature d'antenne.	57
2.3	Signature d'antenne de différentes cibles.	58
2.4	Premier codage d'une signature à partir d'une séquence.	60
2.5	Formation d'une ligne de la signature.	61
2.6	Corrélation d'une référence avec un ensemble de séquences.	62
2.7	Second codage d'une signature à partir d'une séquence.	64

2.8	Exemple de transformation 2D→1D par un parcours de Hilbert.	67
2.9	Influence d'un décalage sur la transformation 2D→1D.	68
2.10	Restitution d'une cible par analyse de sa signature d'antenne.	72
2.11	Pointage d'une cible par analyse de sa signature d'antenne.	73
2.12	Schéma général de synthèse puis d'analyse par filtrage spatial.	76
3.1	Exemples d'attitudes de visages.	81
3.2	Synthèse & Analyse de signatures à partir d'une Base de Références. . . .	82
3.3	Exemples de signatures par la méthode «PCE».	84
3.4	Synthèse & Analyse de signatures avec extraction par parcours fractal. . .	87
3.5	Étapes dans la formation d'une ligne de signature.	88
3.6	Formation d'une signature à partir des plans d'intercorrélation.	89
3.7	Exemples de signatures d'attitudes de visages obtenues par parcours fractal.	90
3.8	Synthèse & Analyse de signatures avec extraction par parcours fractal. . .	95
3.9	Exemples de signaux «Temps-Fréquence» de bâtiments de surface.	95
3.10	Exemples de signatures de bateaux obtenues par l'approche du filtrage non-linéaire.	97
4.1	Le montage «4f», base du filtrage optique.	103
4.2	Montage optique de l'approche «PCE».	107
4.3	Montage optique de l'approche «Parcours fractal».	108
4.4	Réduction de dynamique sur les éléments.	109
4.5	Couplage de deux EASLM pour un double filtrage.	111
4.6	Montage expérimental de réalisation d'un corrélateur compact.	113
4.7	Implantation d'un filtre composite avec contrôle de positionnement.	113
A.1	Ondelette du «Chapeau Mexicain».	123
B.1	Parcours de Hilbert pour trois images.	125
C.1	Utilitaire Tcl/Tk.	130
C.2	Exemple d'attitude.	131
D.1	Performances d'un DSP pour le calcul d'une FFT.	134

Liste des tableaux

1.1	<i>PCE</i> avec prétraitement linéaire.	50
3.1	Table présentant les <i>PCE</i> issus des corrélations entre signatures.	85
3.2	Table présentant les <i>PCE</i> issus des corrélations entre signatures.	91
3.3	Table présentant les <i>PCE</i> issus des corrélations entre signatures.	98
4.1	Influence des réductions de dynamique sur les <i>PCE</i>	108
4.2	«Pipelining» de DSP.	116
C.1	Séquences numérisées.	128
C.2	Base d'apprentissage.	129
C.3	Base de test.	129

Remerciements

Mes remerciements se tournent tout d'abord vers Jean-Louis de Bougrenet de la Tocnaye, chef du département d'optique de l'ENST BRETAGNE, lequel m'a accueilli au sein de son laboratoire puis m'a guidé, durant plus de trois années, dans mes activités de recherche.

Je tiens ensuite à exprimer ma profonde gratitude auprès du professeur P. Ambs et du professeur C. Ozkul, pour leur disponibilité, en cette fin d'année 1998, à juger le travail présenté dans ce document.

Je remercie le professeur L. Calvez et le professeur G. Burel d'avoir accepté de participer à ce jury.

Enfin, je remercie Y. Pelliard, pour l'intérêt qu'il porte, par sa présence lors de ma soutenance, à mon travail.

Sans les industriels ou les financeurs, ces travaux n'auraient pu être menés : je tiens donc à remercier Idmatics pour sa base d'empreintes digitales, SFIM-ODS pour m'avoir permis de réaliser les prises de vues de visages dans leurs locaux, avec leurs instruments. Je remercie également Thomson-CSF/RCM pour les signaux radar de bâtiments de surface. Un grand merci tout particulier pour G. Coppin, lequel a beaucoup œuvré, par son dynamisme, à l'«initialisation» de ce projet commun entre Thomson et l'ENST BRETAGNE, entre autres. Je me devais également d'exprimer ma gratitude auprès du «département développement» de la «Communauté Urbaine de Brest», lequel a participé au financement de mon allocation de recherche pendant trois ans.

Durant ces trois années, mes activités de recherche m'ont amené à travailler au département de traitement de signal de l'École Navale puis au département d'optique de l'ENST BRETAGNE. Je remercie tout ceux qui, par leur disponibilité, leur sympathie ou leur intérêt, m'ont aidé à poursuivre voire achever ces travaux. Dans cet esprit, je suis infiniment reconnaissant à Gilles Keryer pour son encadrement, ses précieux conseils et sa volonté de partager ses connaissances.

Comme doctorant, je remercie tous ceux qui ont donné un peu de leur temps de recherche pour améliorer notre qualité de vie au sein de ce campus de l'ENST BRETAGNE, notamment par le «*Bureau des Thésards*».

Je remercie ensuite tous ceux sans qui cette expérience de recherche, démarrée en

DEA à Meudon en 1993, achevée à Brest en 1999, n'aurait pu se réaliser au travers de disciplines aussi variées : B. Fort, Y. Mellier, P. Léna et tous ceux qui ont jalonné ces cinq années d'études.

Enfin, parmi les «gourous» ou anonymes ayant participé à mon travail sans le savoir, je remercie B. Kernighan, D. Ritchie, B. Stroustrap, D. Knuth, L. Lamport, R. Stallman, J. Ousterhout, L. Torvalds et tous les autres développeurs de logiciels libres. Sans leurs réalisations, mes investigations dans ce gouffre temporel qu'est l'informatique auraient été (ô combien) plus périlleuses.

Pour conclure, je remercierai ma famille pour ces études scientifiques qu'elle a su encourager, mes amis ainsi que mon épouse, Noëlle, pour son soutien inconditionnel durant ces (presque) trois années.

Introduction

Le département d'optique de l'ENST Bretagne poursuit, depuis plusieurs années, au travers de projets, européen (ESPRIT III-HICOPOS¹) ou DRET [29], des investigations dans le domaine du traitement optique de l'information [17]. Les développements récents dans le domaine des composants électro-optiques (FLC-SLM pour «Ferroelectric Liquid Crystals Spatial Light Modulators») ont permis la réalisation de corrélateurs rapides d'une fréquence supérieure au kilo-kertz.

Cependant, d'autres applications – l'analyse d'une scène en évolution durant un laps de temps donné par exemple – ont demandé d'étendre l'utilisation du filtrage spatial pour d'autres traitements plus complexes. La corrélation est, de nos jours, une opération de calcul trop sommaire : elle ne peut faire face aux tâches exigées pour le traitement de signaux de natures spectrale (différents capteurs), spatiale (images) et temporelle (observation sur un horizon temporel).

Le traitement de telles masses de données hétérogènes ne pouvant se réduire à un simple filtrage spatial, nous nous intéresserons à la formation, par ce même outil de filtrage, d'une image intermédiaire, la signature temporelle, caractéristique autant de l'objet que de son comportement.

Suivant la synthèse de cette signature temporelle, nous aborderons ensuite son analyse, toujours par filtrage spatial, nous permettant de conclure sur la nature du comportement, voire l'identité de la forme observée.

Pour cela, nous commencerons par détailler le «comment» avant le «pourquoi» de la signature. En effet, le premier chapitre sera destiné à développer le filtrage suivant plusieurs aspects, notamment sur la possibilité d'insérer un prétraitement au sein du filtre pour introduire une nouvelle dimension dans le filtrage : la notion de détail.

Le «comment» du filtrage spatial établi, nous nous pencherons sur le «pourquoi» de son utilisation dans la synthèse puis l'analyse d'une signature temporelle. Une généralisation de la synthèse d'une signature d'antenne synthétique nous conduira vers la formation d'une signature temporelle à partir de séquences d'images.

Les procédés de synthèse et d'analyse développés dans un cadre très général, nous reviendrons aux applications en nous penchant tout d'abord sur des attitudes de visage, observées conjointement dans les domaines visible et infrarouge, pour lesquelles nous

¹Rapport final HICOPOS n°7050

dispositions d'une base de données significatives. Ensuite, nous éprouverons notre approche face à des échantillons de signaux radar en représentation «temps-fréquence».

Enfin, dans un dernier chapitre, les différentes implantations possibles du filtrage spatial seront développées, suivant les axes optique et numérique. La comparaison de ces deux approches sera ainsi l'occasion de présenter les avantages respectifs de ces deux processeurs souvent opposés. Le cadre opérationnel fixant les contraintes quant aux données à traiter justifiera ainsi la simplicité apparente de notre approche.

Chapitre 1

Le Filtrage : Outil de Reconnaissance de Formes

1.1 Introduction

Le filtrage demeure le principal outil que nous utiliserons dans la synthèse puis l'analyse de signatures. Toutefois, il apparaît intéressant de se pencher sur son aspect fonctionnel, notamment en reconnaissance de formes.

Pour cela, nous proposons de détailler la nature du filtrage, partant de ses origines et cheminant vers les récentes applications en reconnaissance de formes.

L'occasion nous sera ainsi donnée de développer les différents filtres susceptibles de nous aider, soit dans une opération de reconnaissance, soit dans une extraction de caractéristiques, liée par exemple à des paramètres tels que la «résolution».

1.2 Introduction au filtrage spatial

1.2.1 Filtrage dans l'espace de Fourier

Si nous nous intéressons au filtrage dans son sens étymologique, nous découvrons que deux sens lui sont généralement attribués : soit «passage au travers d'un filtre» (dispositif ne laissant passer que certains éléments), soit «contrôle». En traitement de signal, ces deux sens se mêlent, donnant au filtrage la fonction de contrôle par la sélection de certains de ces éléments.

L'analyse de Fourier donne à tout signal sommable la possibilité de s'exprimer à partir de son spectre. Le passage entre l'espace originel, que nous qualifierons de direct, et l'espace des fréquences (espace de Fourier), est assurée par la transformée de Fourier, notée \mathcal{F} (expression de la TF pour une fonction $f(t)$ donnée par 1.1), se généralisant pour une fonction $f(x, y)$ à deux dimensions (expression 1.2).

$$F(\nu) = \mathcal{F}_{f(t)}(\nu) = \int_{-\infty}^{+\infty} f(t)e^{-2i\pi\nu t} dt \quad (1.1)$$

$$F(\nu_x, \nu_y) = \mathcal{F}_{f(x,y)}(\nu_x, \nu_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{-2i\pi(\nu_x x + \nu_y y)} dx dy \quad (1.2)$$

Aussi, le filtrage s'exprime dans l'espace de Fourier par un produit de spectres.

Par conséquent, dans l'espace direct, la représentation d'un filtrage prend la forme d'une convolution (expression 1.3), se généralisant à deux dimensions (expression 1.4).

$$(f * g)(t) = \int_{-\infty}^{+\infty} f(t-u) g(u) du \quad (1.3)$$

$$s_f(x, y) = (s * f)(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s(x-u, y-v) f(u, v) du dv \quad (1.4)$$

Pour résumer, le résultat du filtrage entre une fonction s , de spectre S et un filtre f , de spectre F , se trouve être un produit de convolution (noté $*$) dans l'espace direct et un produit dans l'espace de Fourier (équation 1.5).

$$\begin{cases} \text{Espace Direct} & : & s_f & = & (s * f) \\ \text{Espace Spectral} & : & \mathcal{F}_{s_f} & = & (S F) \end{cases} \quad (1.5)$$

Toutefois, les signaux que nous devons analyser ne se trouvent généralement pas dans l'espace de Fourier. Il est donc indispensable, pour réaliser le filtrage, de passer dans l'espace des fréquences à l'aide de la TF , de réaliser le produit des deux spectres ainsi obtenus avant de revenir dans l'espace direct à l'aide de la TF inverse. L'intérêt d'un passage dans l'espace de Fourier n'existe donc *a priori* que pour des fonctions simples à transformer, ou si nous sommes en mesure de réaliser de telles opérations rapidement. Dans le chapitre destiné à présenter nos implantations, nous rappellerons que, numériquement, le filtrage utilisant des FFT, ou «*Fast Fourier Transform*», demeure plus intéressant en terme de temps de calcul qu'une convolution directe.

1.2.2 La transformée de Fourier optique

La physique propose, dans certaines conditions, une transformation de Fourier naturelle. En effet, la diffraction, dans le cadre de l'approximation de Fraunhofer (diffraction à l'infini) [10], réalise la TF à l'infini. Une lentille utilisée en «transformateur de phase» ramène cette figure de diffraction à l'infini dans son propre plan focal. Il est donc possible d'observer le spectre, ou plus précisément le carré du module du spectre d'un objet éclairé par une onde plane. L'expérience d'Abbe et Porter [13] se base sur tel montage, dit «4f» : une grille, placée dans le plan focal objet de la première lentille, est éclairée par une onde plane. Dans le plan focal objet se trouve le spectre. Un masque posé dans ce plan permet de ne sélectionner que quelques fréquences particulières. Une seconde TF permet de revenir à la grille filtrée. L'expérience originelle avait utilisé un masque ne laissant passer que les fréquences suivant l'axe horizontal : la grille se «transformait» alors en barreaux.

De cette expérience dérivèrent d'autres applications, comme le «déflouage» ou le «détramage». Cependant, le filtrage dans l'espace des fréquences se réduisait à un simple

filtrage «tout ou rien» de certaines zones du spectre. La transformée de Fourier d'un signal étant en général complexe, il s'avérait nécessaire de disposer, d'une meilleure dynamique, à la fois en amplitude, mais aussi en phase. Les travaux de Gabor sur l'holographie permirent bientôt de réaliser un enregistrement sur plaque photographique d'un signal, autant en amplitude qu'en phase par des procédés interférométriques. Vander-Lugt [37] proposa alors d'utiliser de telles techniques en reconnaissance de formes [5], où le filtre présenté serait adapté ou non à la scène présentée, et dont le filtrage permettrait de réaliser s'il y avait ou non adéquation entre scène présentée et référence. C'est sur ce principe que fonctionne la majorité des corrélateurs optiques actuels, dont les performances sont limitées par la capacité de codage de leur filtre [11]. Tel est donc le cadre que nous allons désormais étudier.

1.2.3 Corrélation & filtrage adapté

La définition «mathématique» associée à la corrélation, notée \otimes , est présentée comme une convolution particulière. La corrélation $c_{fg}(t)$ entre deux fonctions $f(t)$ et $g(t)$ s'exprime à partir de la convolution (équation 1.6). \tilde{g}^* représente le complexe conjugué du transposé de g .

$$c_{fg}(t) = (f \otimes g)(t) = (f * \tilde{g}^*)(t) = \int_{-\infty}^{+\infty} f(t+u) g^*(u) du \quad \text{avec} \quad \tilde{g} = g(-t) \quad (1.6)$$

Cette expression prend une forme sensiblement simplifiée dans l'espace des fréquences, en prenant $C_{fg} = \mathcal{F}_{c_{fg}}$ (équation 1.7).

$$C_{fg} = \mathcal{F}_{c_{fg}} = \mathcal{F}_{(f \otimes g)} = (F G^*) \quad (1.7)$$

La corrélation entre f et g suivant cette définition est appelée «filtrage adapté». Ce filtrage, bien que très utilisé en traitement de signal [4], n'est pas nécessairement le plus adapté au problème nous préoccupant, à savoir la reconnaissance de formes.

1.3 Corrélation & reconnaissance de formes

Le filtrage adapté, ou corrélation, optimise le rapport signal à bruit gaussien (SNR). Toutefois, à cette connotation «signal» de la corrélation peut être associée une corrélation «généralisée» [9], laquelle constitue plus un opérateur de comparaison. Réalisée par une opération de filtrage (dont la nature reste à définir), cette corrélation sera notée \otimes , de manière à ne pas la confondre avec le filtrage adapté traditionnel.

Dans ce vaste domaine qu'est la reconnaissance déterministe de formes [3], le traitement généralement présenté se compose de deux phases distinctes :

- l'extraction de caractéristiques ;
- le classement suivant ces caractéristiques.

La reconnaissance de formes par filtrage spatial, lorsqu'elle s'applique donc à des images, ne déroge pas à ce schéma général : une première étape recherche à extraire un ensemble de caractéristiques, la seconde prendra une décision, sous forme d'un classement, à partir de ces mêmes caractéristiques.

Jusqu'à présent, les missions confiées au filtrage spatial, dans le domaine de la reconnaissance de formes étaient :

- de reconnaître la nature d'un objet dans une scène se composant d'un seul objet (reconnaissance d'un visage) ;
- de reconnaître la nature d'un objet dans une scène composée d'un grand nombre d'objets (reconnaissance d'un caractère dans une page) ;
- de reconnaître un objet se déplaçant dans une scène et déterminer en permanence sa position (poursuite), ou pointage d'un objet situé dans une scène.

De ces trois applications du filtrage spatial découlent évidemment bon nombre d'applications. Toutefois, dans celles liées à la sécurité, l'objectif fixé est de savoir si une personne a l'autorisation de pénétrer dans une zone d'accès restreint ou non. Pour cela, une caractéristique exclusive va être sélectionnée pour chaque individu qui servira à la constitution d'une base (une empreinte, un visage, un fond de rétine, etc...). La comparaison vérifiera qu'un individu se présentant correspond bien, par sa caractéristique, à une personne «connue» de la base.

Ainsi, l'opération de filtrage spatial permettant la reconnaissance de formes se scinde-t-il en trois composantes :

- la scène ;
- le filtre intégrant la référence ;
- le plan de corrélation.

Nous commencerons tout d'abord par évoquer le critère de corrélation qui va déterminer la nature du filtre.

Ensuite, la création du filtre intégrant la référence sera abordée ; occasion nous sera donnée de choisir la forme générale du filtre avant d'en proposer, après analyse de ses performances, quelques améliorations : du «compositage» de références vers la synthèse d'un filtre multi-voies, passant par les étapes intermédiaires que sont le contrôle de la position du pic de corrélation et une meilleure utilisation de la capacité de codage.

Enfin, revenant à la scène, nous développerons un des prétraitements qu'il nous est possible d'appliquer sur l'image originelle (et donc les références). Ce prétraitement prendra la forme d'une extraction de caractéristiques, telles que la *résolution*.

1.4 La nécessité d'un critère de corrélation

Dans le cadre d'une opération de corrélation, la conception du filtre reste intimement liée au critère de corrélation, le premier étant censé optimiser le second. Ainsi, le choix du critère que nous allons poser conditionnera en un certain sens le type de filtre qu'il nous faudra par la suite considérer dans notre filtrage.

Nous nous proposons donc de choisir un critère énergétique, lié à la localisation de l'énergie. Ainsi, plus l'énergie du résultat de la corrélation sera localisée et sa fraction à l'énergie totale importante, plus le critère sera important.

Un tel critère est connu sous le nom de *PCE* ou «*Peak to Correlation Energy*» [19] [23]. Il se calcule en divisant le maximum d'intensité à l'énergie totale du plan de corrélation (équation 1.8).

$$PCE_{fg} = \frac{|\max(f \otimes g)|^2}{|\iint (f \otimes g)|^2} \quad (1.8)$$

Certaines variantes du *PCE* existent, la plus commune étant de ne pas considérer uniquement le maximum mais une zone autour de ce maximum.

1.5 Le choix d'un filtre de phase

Le critère pour évaluer l'adéquation entre scène et référence étant le *PCE*, il nous reste désormais à choisir le filtre optimal permettant d'obtenir le meilleur *PCE* possible dans le cas d'une reconnaissance ou d'un rejet.

Aussi, pour une référence r de spectre R , nous noterons le filtre optimal f , de spectre F . L'opération de corrélation avec une scène s de spectre S sera donc la suivante :

Opération	Espace Direct	Espace Spectral
En terme de corrélation	$c_{sr} = (s \otimes r)$	$C_{sr} = S g(R)$
En terme de filtrage	$c_{sr} = (s * f)$	$C_{sr} = S F$

Nous devons donc rechercher quelle fonction g doit être appliquée au spectre de la référence R pour former le filtre F .

Dans cette optique, revenons à la définition du *PCE*. Ce dernier étant le rapport entre l'intensité du maximum sur l'image sur l'énergie totale de l'image, sa valeur est nécessairement réelle, positive et inférieure à l'unité. Le filtre optimal serait donc le filtre qui permet de placer la totalité de l'énergie du résultat de la corrélation dans le pic. Ce filtre porte un nom, c'est le filtre inverse.

En effet, assimilons ce pic à la distribution de Dirac δ . La *TF* inverse de ce pic est la fonction unité $\mathbf{1}$: le produit entre la *TF* du filtre F et le spectre de la scène S doit donc être identique à l'unité, ce qui implique que le filtre F associé au spectre de la référence R n'est autre que son inverse (équation 1.9).

$$F = \frac{1}{R} = \frac{R^*}{|R|^2} \quad (1.9)$$

Ainsi, les phases des filtres inverse ou adapté sont identiques. Seul le module est différent. Ces filtres peuvent donc s'écrire sous une seule et unique forme en introduisant une puissance 2γ au module du spectre du dénominateur (équation 1.10) : le filtre adapté correspondant à une valeur de γ nulle et le filtre inverse à une valeur unité.

$$F_\gamma = \frac{R^*}{|R|^{2\gamma}} \quad (1.10)$$

Ainsi, sondons cette valeur γ entre 0 et 1 afin de trouver le coefficient optimal à appliquer à notre filtre. Pour cela, nous allons sélectionner deux images distinctes, l'une servira de référence puis de scène : nous pourrons ainsi tester, par l'adjonction de bruit additif gaussien, la robustesse du filtre ; l'autre image servira de scène : elle permettra de juger de la sélectivité.

Les résultats d'une telle simulation sont présentés sur les figures 1.1 et 1.2. De façon à comparer les *PCE* suivant les différentes réalisations de bruit, examinons le rapport entre les *PCE* issus de corrélations entre visages identiques ou différents (figure 1.3). Nous constatons tout d'abord que le rapport entre les *PCE* des corrélations entre deux visages sont proches, dès lors que l'on considère un filtrage adapté ($\gamma = 0$). Ce rapport augmente progressivement et atteint, hormis le cas où aucun bruit n'est ajouté, une valeur maximale pour γ compris entre 0,5 et 0,7. Nous constatons ainsi que les filtres proches du filtre inverse ($\gamma = 1$) ne sont pas adaptés à la reconnaissance d'objets soumis à un bruit additif. Un bon compromis consiste donc à considérer un filtrage plus robuste que le filtre inverse au bruit mais également plus sélectif que le filtre adapté [15], au regard du critère, le *PCE* que nous nous sommes fixés. Le filtre caractérisé par une valeur de γ égale à 0,5 s'avère être le meilleur compromis entre robustesse et discrimination : c'est désormais ce filtre, communément appelé filtre de phase pure (POF pour «*Phase Only Filter*»), très largement étudié [15], que nous utiliserons désormais dans nos filtrages spatiaux. De plus, ne perdons pas de vue que notre choix doit nécessairement se plier aux contraintes liées aux composants que nous comptons utiliser pour réaliser notre filtrage.

Il a cependant été démontré la sous-optimalité du POF, par la synthèse d'autres filtres plus généraux [31], dotés de compromis ajustables (en robustesse au bruit et en sélectivité).

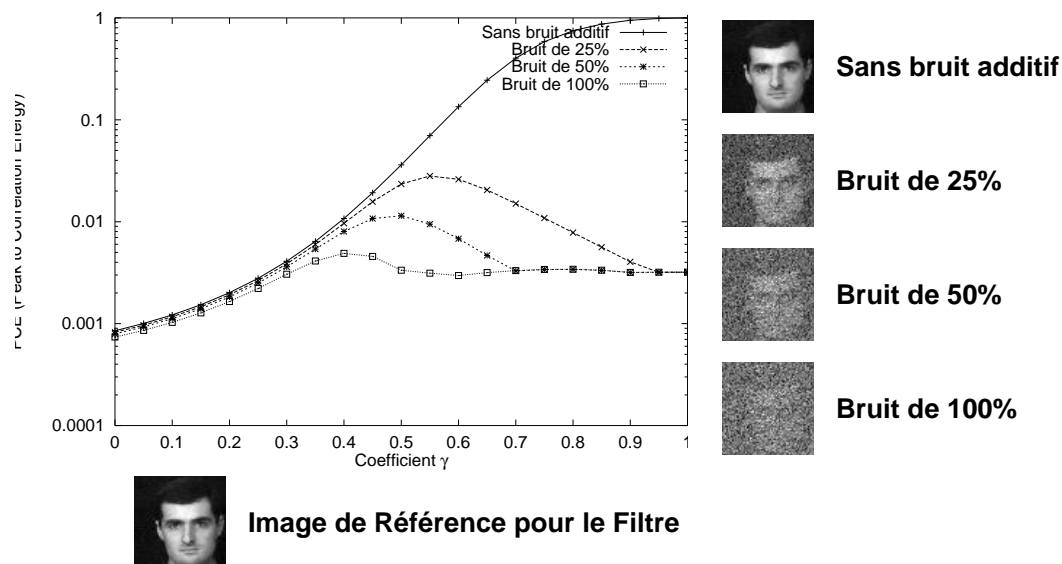


FIG. 1.1 – **Corrélation entre les images de deux visages identiques** : le visage au bas de la figure présente l'image à partir de laquelle se construit le filtre. A droite apparaissent les valeurs du *PCE* pour 4 réalisations différentes de bruit : sans bruit et des écarts-types de bruit allant de 25% à 100% de l'intensité maximale de l'image originelle.

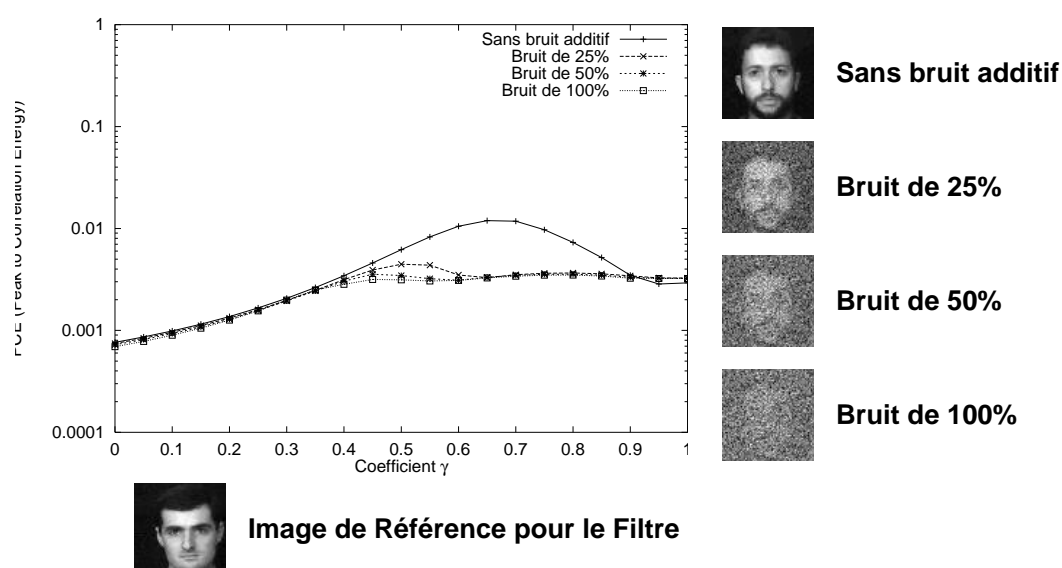


FIG. 1.2 – **Corrélation entre les images de deux visages différents** : le visage au bas de la figure présente l'image à partir de laquelle se construit le filtre. A droite apparaissent les valeurs du PCE pour 4 réalisations différentes de bruit : sans bruit et des écarts-types de bruit allant de 25% à 100% de l'intensité maximale de l'image originelle.

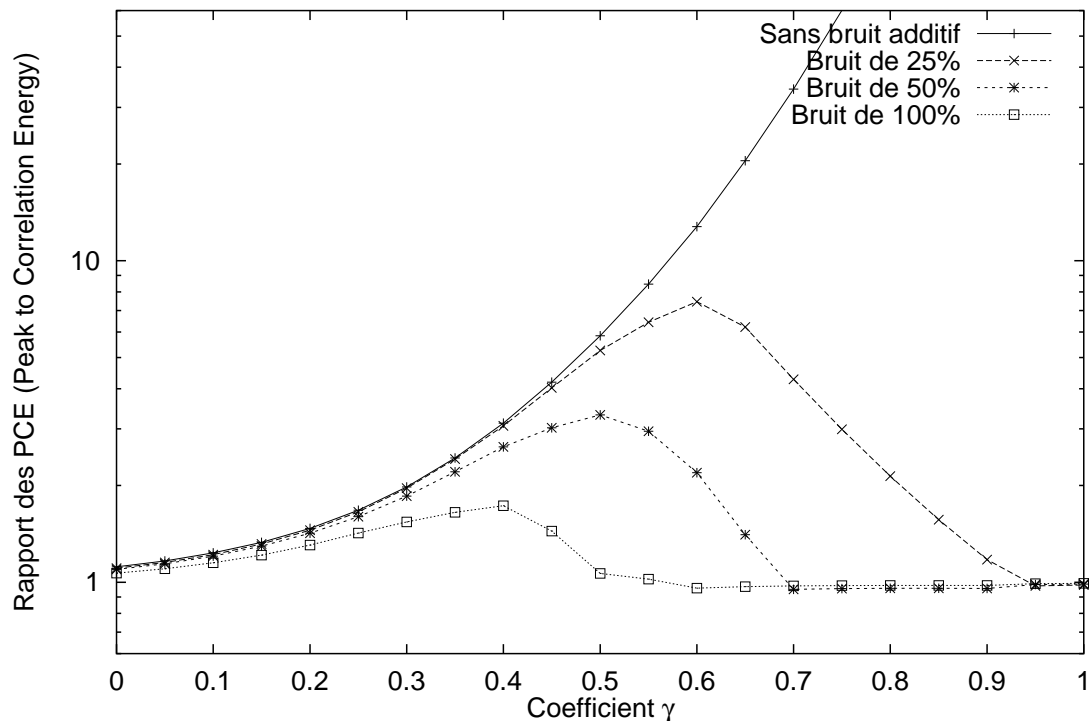


FIG. 1.3 – **Rapport entre les PCE issus de corrélations entre visages identiques ou différents** : les résultats présentés sont issus du rapport entre les résultats présentés figure 1.1 et figure 1.2.

De façon à établir de manière plus précise l'adéquation entre notre critère, le *PCE*, comme critère d'une «bonne» corrélation, il est intéressant, de présenter les résultats des corrélations dans leur ensemble. Ainsi, sur la figure 1.4 apparaissent les plans de corrélations pour deux visages différents. Nous constatons que, plus le coefficient γ se rapproche de l'unité, plus le pic de corrélation s'affine; à l'intérieur du pic se situe de plus en plus d'énergie. Le *PCE* étant le rapport entre l'énergie du pic par l'énergie du plan, il demeure donc un bon critère. Toutefois, à l'examen de la figure 1.4, nous pouvons également constater qu'une zone centrale reste assez énergétique dans le cas où les images sont différentes, alors que le *PCE* par rapport à une «auto-corrélation» se situe deux ordres de grandeur au dessous; la raison en est plus à la représentation qu'à la physique. En effet, tous les résultats de corrélation sont présentés après une normalisation à l'unité par le maximum de l'image (la zone d'intensité maximale se voit attribuer la valeur 1, quelle que soit l'image). Cette méthode de normalisation permet de constater que le filtre inverse, même s'il présente des pics, n'est pas adapté à la recherche ou la comparaison d'un élément dans une scène.

Coefficient γ

0,7	0,75	0,8	0,85	0,9	0,95	1
0,35	0,4	0,45	0,5	0,55	0,6	0,65
0	0,05	0,1	0,15	0,20	0,25	0,3

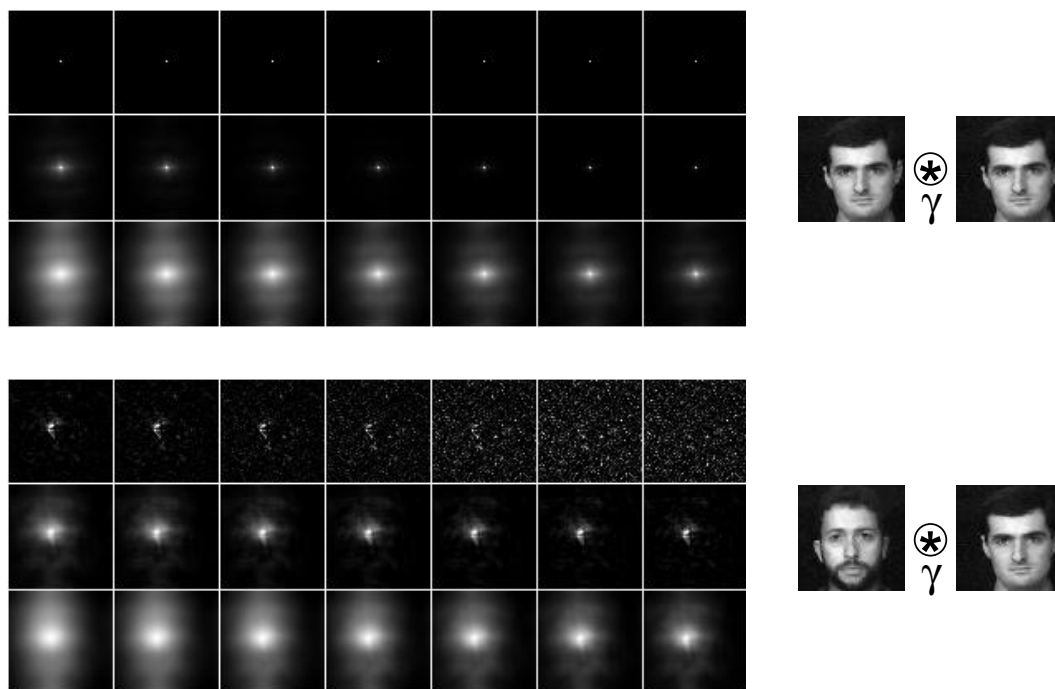


FIG. 1.4 – Résultats de corrélations pour deux visages : un visage est comparé à lui-même dans un premier temps, puis à un autre visage. Les mosaïques présentent les plans de corrélations pour une valeur de γ comprise entre 0 et 1.

Les résultats présentés sur les figures 1.1, 1.2, 1.3 illustrent les choix de critère et de filtre. Ils confirment également ce qui est généralement obtenu dans la littérature [29]. Il apparaît maintenant intéressant d'aborder d'autres aspects, telle la robustesse d'un signal à des transformations géométriques.

1.6 La nécessité du contrôle de la sélectivité

D'un point de vue pratique, le filtre de phase demeure un acceptable compromis entre robustesse et sélectivité, lorsque les bruits appliqués sont des bruits aussi standards que sont les bruits gaussiens. Lorsque nous considérons la «robustesse» d'un filtre, il s'agissait d'une robustesse dans la corrélation lorsque ce même signal était perturbé par un bruit.

Toutefois, la robustesse peut revêtir un sens différent si les signaux que nous devons reconnaître sont soumis à des transformations géométriques comme la rotation et l'homothétie. Un filtre «robuste», dans ce cadre-là, ne sera pas un filtre robuste dans un sens où il résiste à des bruits additifs, mais robuste parce qu'il est capable de reconnaître une référence, même si cette dernière subit une transformation géométrique. Dans cette optique, seule la translation conserve intactes les qualités d'un filtre. Dans notre contexte, il n'en est pas de même des autres transformations.

Une identification à partir d'empreintes digitales sera le cadre de notre étude sur la sélectivité et la robustesse du filtre de phase. Pour cela, nous disposons d'une base d'empreintes digitales [28]. Nous cherchons à reconnaître une empreinte parmi cette base. Nous allons donc procéder à la corrélation entre cette empreinte à reconnaître et toutes les empreintes reconnues. Prenons 100 empreintes au hasard formant notre base et corrélons-la avec un des éléments de cette base, nous obtenons la courbe présentée sur la figure 1.5 : nous constatons la sélectivité du filtre de phase. En effet, deux ordres de grandeur existent entre une «auto-corrélation» et une «inter-corrélation». L'identification, dans ce cadre, pourra être réalisée par un simple seuillage, par exemple autour d'un PCE de 0,01.

Examinons maintenant la robustesse d'un filtre de phase. Pour cela, nous allons prendre une empreinte à laquelle nous allons appliquer une rotation comprise entre -30° et $+30^\circ$, par pas de 1° . Cette base en rotation est alors comparée par corrélation à l'empreinte originelle. La figure 1.6 confirme la faible robustesse d'un filtre de phase à la rotation (la coupure à 3 dB n'est que de 2° tout au plus).

Pour pallier ce manque de robustesse face à certaines transformations, certaines méthodes existent. Nous allons nous intéresser à la plus générale réalisée par apprentissage, ce par opposition à d'autres méthodes, comme celle, par exemple des harmoniques circulaires.

La première approche que nous allons développer se base sur un assemblage de références au sein d'un seul et unique filtre. En effet, dans l'exemple des empreintes digitales, il paraît impensable d'intégrer à la base permettant la comparaison toutes les déformations possibles susceptibles d'affecter chaque empreinte de la base. La base de

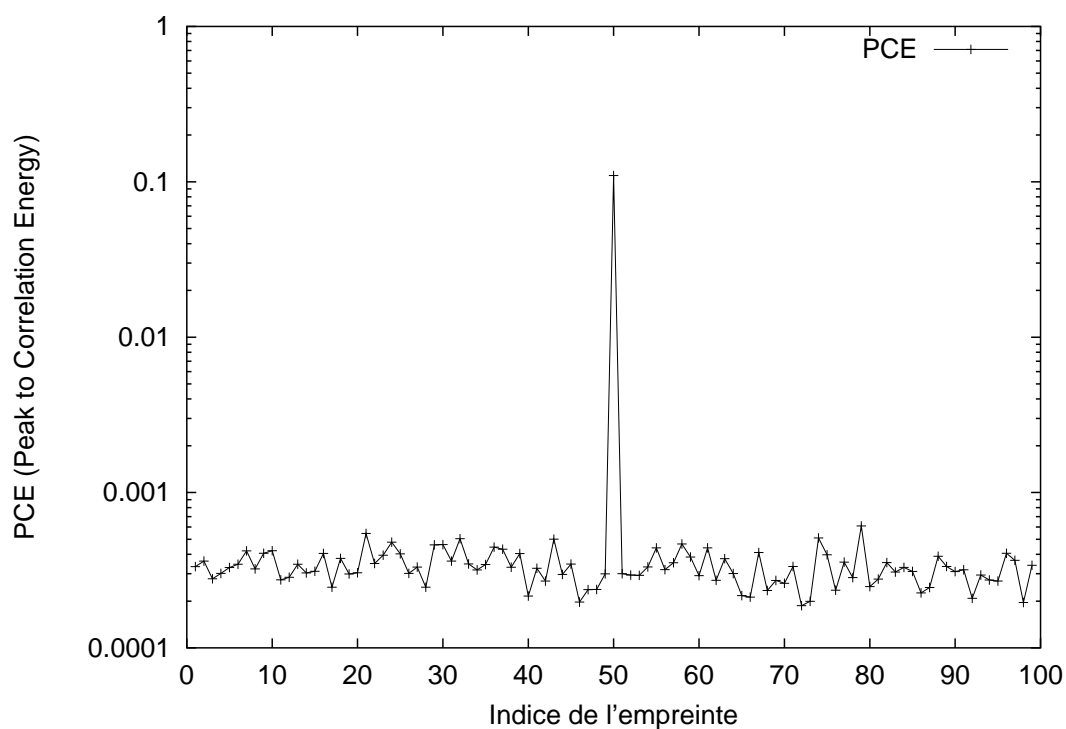


FIG. 1.5 – **Comparaison d'une empreinte digitale à une base de 99 empreintes** : la corrélation est réalisée par un filtre de phase. Le pic permet de revenir à l'indice de l'empreinte reconnue.

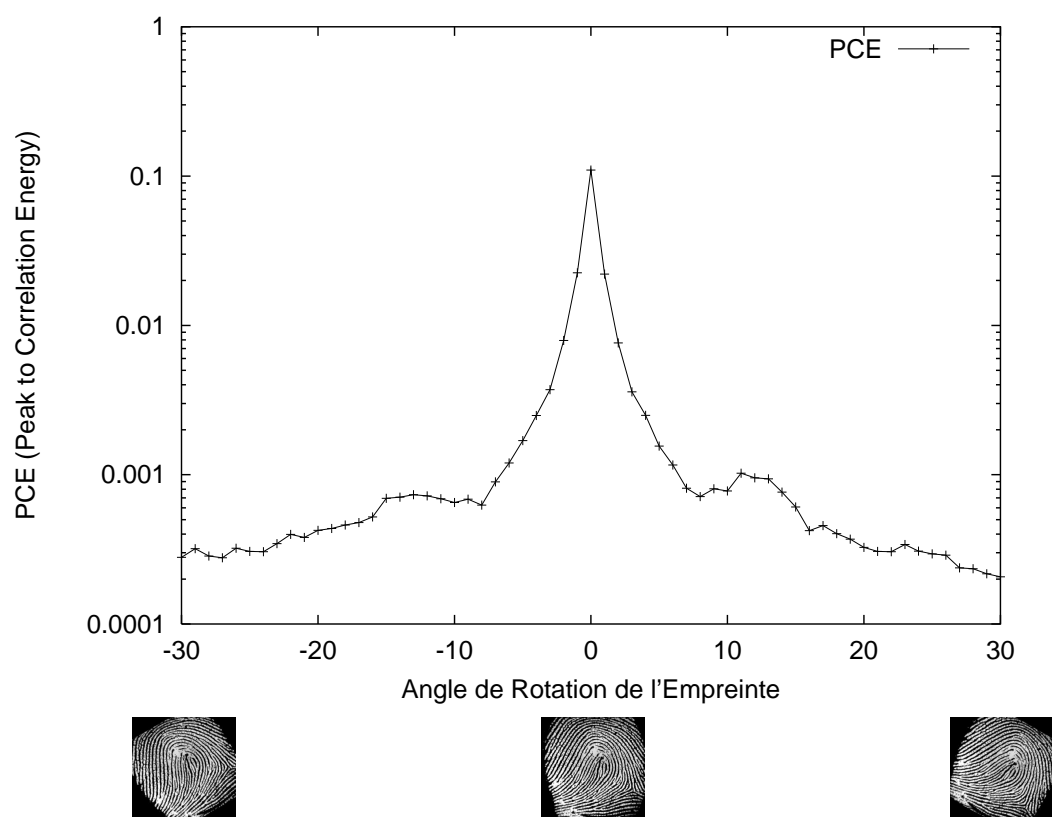


FIG. 1.6 – **Comparaison d'une empreinte en rotation avec elle-même** : le pic indique l'angle pour lequel l'adéquation entre les deux images comparées sont les plus proches. La robustesse du filtre issu de l'empreinte servant de référence s'estime à partir de la largeur du pic.

références y gagnerait certes en exhaustivité, mais le problème se reporterait d'un point de vue plus pragmatique sur l'intégration du filtrage, qu'il soit numérique ou optique. Aussi est-il intéressant de réaliser un filtre intégrant plusieurs références.

L'exhaustivité de la base de références, garantissant la reconnaissance, se trouve donc opposée à la cadence imposée par l'application. Pour contourner cette difficulté, nous nous proposons de réduire la base de références à présenter par la synthèse d'un filtre composite, dont les utilisations seront les suivantes :

- l'insertion de plusieurs références d'une même classe dans un même filtre : compositage de références ;
- l'insertion de plusieurs références de classes différentes : compositage de références et contrôle de positionnement des pics ;
- l'insertion d'un prétraitement contrôlant la résolution des détails intervenant dans la corrélation : intégration d'un prétraitement linéaire dans le filtre.

Ce dernier aspect est certainement le plus original des trois précités. C'est d'ailleurs sur ce dernier point que nous allons nous baser pour synthétiser nos signatures à des niveaux de détails différents.

1.7 D'une base de références au filtre composite

1.7.1 La synthèse d'un filtre composite

Le filtre composite est un filtre dont la caractéristique consiste à reconnaître un ensemble de références données [21]. Le type de filtre dont nous avons justifié l'utilisation est un filtre de phase pure, donc de transmittance unité. La fonction de codage s'intègre donc dans la phase du filtre.

La réalisation de ce filtre composite s'obtient en combinant linéairement les spectres R_i des références r_i suivant des poids différents. Nous obtenons ainsi une combinaison linéaire de toutes les références :

$$R = \sum_{i=1}^N \sigma_i R_i \quad (1.11)$$

Enfin, le codage du filtre F à partir de R se réalise en estimant sa phase conjuguée :

$$F = \frac{R^*}{|R|} \quad (1.12)$$

1.7.2 Optimisation du filtre composite

A ce stade, nous devons toutefois élaborer une technique nous permettant d'estimer les poids respectifs σ_i . Pour cela, nous devons établir un critère, lequel jugera si les poids choisis sont adéquats. La nature composite du filtre sous-entend que chacune des références qu'il intègre répond de manière identique lorsque l'une d'entre elles est présentée ; il paraît donc important que notre critère de corrélation, le *PCE*, soit identique quelle que soit la référence intégrée au filtre. Nous allons pour cela rechercher quelle est la combinaison permettant la meilleure uniformité sur les *PCE* issus des

corrélations entre le filtre et les références le composant.

Le critère d'uniformité de notre filtre composite s'établit par le rapport entre la moyenne des PCE par leur écart-type. Plus ce critère sera élevé, plus les PCE seront élevés et proches les uns des autres :

$$c = \frac{\overline{PCE_i}}{\sigma(PCE_i)} \quad (1.13)$$

où PCE_i représente le PCE associé à la corrélation entre la référence de spectre R_i et le filtre F .

Quelle stratégie adopter pour estimer les poids optimaux ?

Nous supposons tout d'abord qu'il existe une étroite relation entre le poids σ_i et le PCE_i . Notre première étape consiste donc à observer les $PCE_{i(0)}$ lorsque tous les poids sont mis à 1.

Ensuite, nous supposons, de façon intuitive, qu'il existe une puissance optimale β telle que, pour tous les éléments i , $PCE_{i(0)}^{-\beta}$ représente le poids optimal σ_i .

Cette approche impose ensuite une manœuvre itérative destinée à estimer cette puissance β . D'un point de vue pratique, elle est assez variable selon le type de signal utilisé, selon que son spectre est concentré ou étalé.

Il demeure donc indispensable de procéder à un certain nombre de simulations pour parvenir à l'optimum d'uniformité. Dans un registre très pragmatique, un pas de 0,1 entre 0,5 et 1.5 est amplement suffisant.

Cette condition d'uniformité sur les PCEs impose une stratégie sur la recherche des poids optimaux.

La fonction de codage du filtre se réalise en ne conservant que la phase conjuguée de R . Ainsi, le filtre résultant d'un ensemble de références r_i de spectre R_i se résume à l'expression :

$$F = \frac{\sum_{i=1}^N \sigma_i R_i^*}{\left| \sum_{i=1}^N \sigma_i R_i \right|} \quad (1.14)$$

La figure 1.7 présente l'évolution des PCE en fonction de la puissance β appliquée dans le calcul des poids optimaux. L'exemple présenté est un filtre composite comportant 21 références en rotation par rapport à une empreinte. Nous constatons, dans ce cas, que le coefficient β optimal était égal à 1. Le filtre ainsi obtenu présentera approximativement le même PCE pour l'empreinte identifiée, ce pour un angle compris entre -10° et $+10^\circ$.

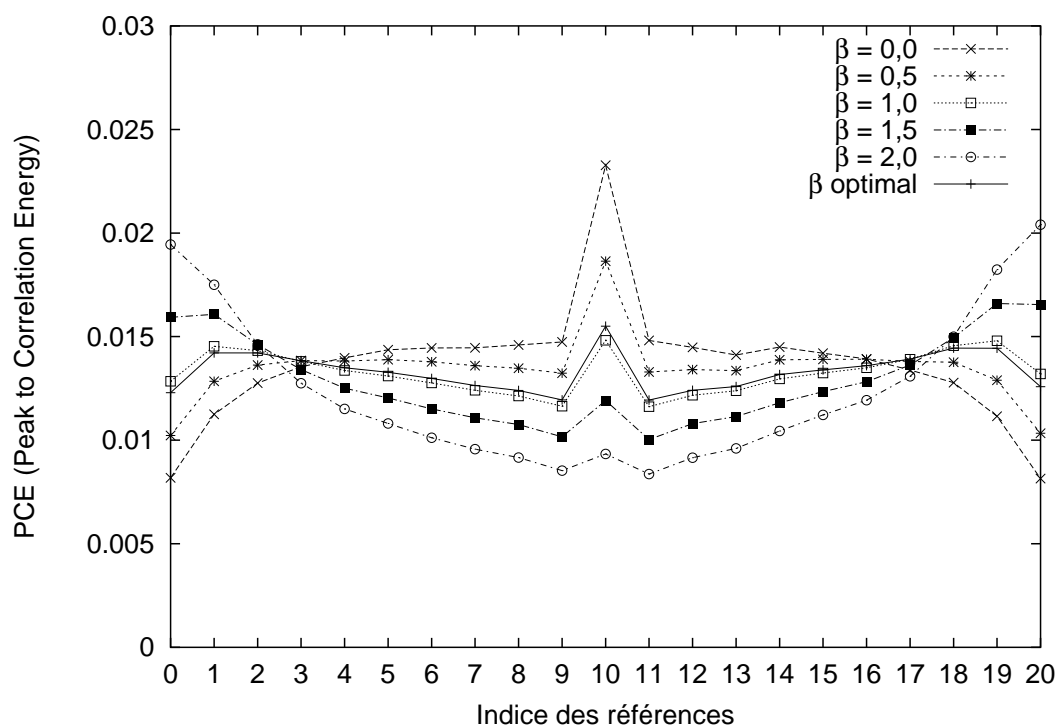


FIG. 1.7 – **Réponses du filtre pour une empreinte en rotation** : les courbes représentent les PCE obtenus par application du filtre à chacune des références qui le compose. La première courbe ($\beta = 0$) correspond à la réponse du filtre pour des poids unitaires. Les suivantes pour différentes valeurs de puissance β appliquées sur l'inverse des PCE obtenus pour $\beta = 0$. L'exposant optimal, dans ce cas, se situe autour de 1.

1.7.3 Robustesse d'un filtre composite

Intéressons-nous maintenant à la synthèse de filtres composites disposant de 5° , 10° , 20° , 30° et 60° d'invariance. La figure 1.8 indique la robustesse des filtres ainsi obtenus : sur leur domaine de «compétence», les filtres composites répondent de façon uniforme aux références dont ils sont composés.

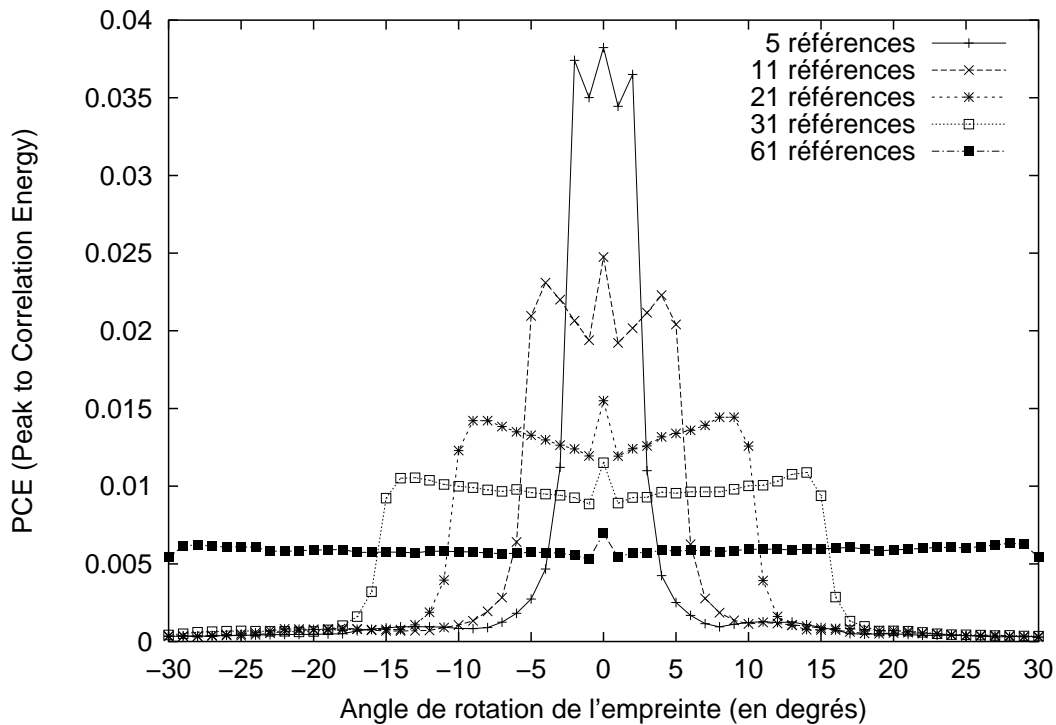


FIG. 1.8 – réponses de filtres composites comportant 5, 11, 21, 31 et 61 références à une empreinte en rotation. : Le filtre est «activé» uniquement si l'empreinte comparée possède un angle de rotation correspondant au support du filtre.

Une fois la robustesse de ces filtres établie, il nous reste à contrôler que la sélectivité des 5 filtres présentés auparavant est toujours aussi forte, en leur appliquant une base d'empreintes prises au hasard. La figure 1.9 présente de tels résultats. Nous pouvons constater que la sélectivité, même dans le cas d'un filtre comportant 61 références, reste encore très significative : un facteur de l'ordre de 20 les sépare.

1.7.4 Nombre de références intégrées au filtre & nature du spectre

Nous avons donc remarqué que le *PCE* se réduisait de façon significative [20] lorsque le nombre de références intégrées dans le filtre composite augmente. Il convient d'observer quelle forme prend cette décroissance dans notre application aux empreintes digitales. En effet, comme le montre la figure 1.10, le comportement de la décroissance est donné par

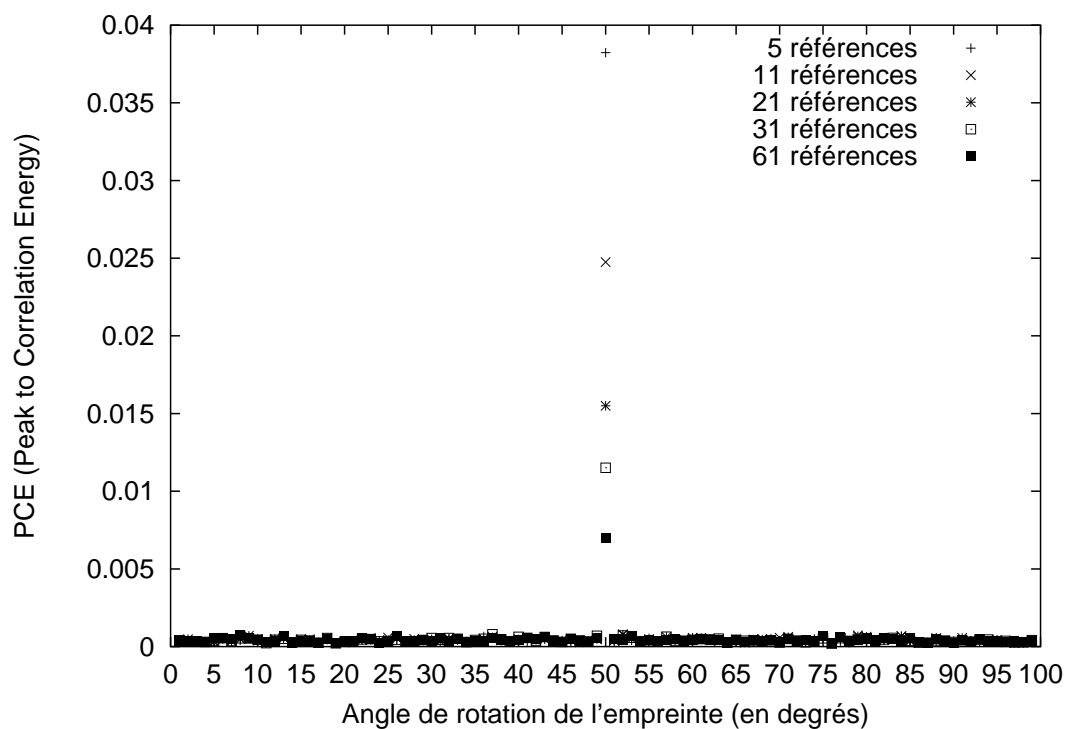


FIG. 1.9 – Réponses de filtres composites comportant 5, 11, 21, 31 et 61 références à une base d'empreintes prises au hasard : seule la référence indiquée 50 présente un PCE supérieur au $1/100^\circ$: il est donc très vraisemblable que cette dernière corresponde à celle utilisée pour la synthèse du filtre

l'équation empirique 1.15, avec pour paramètres $a = 0,11$, $b = 0,63$ et c négligeable.

$$f(x) = \frac{a}{x^b} + c \quad (1.15)$$

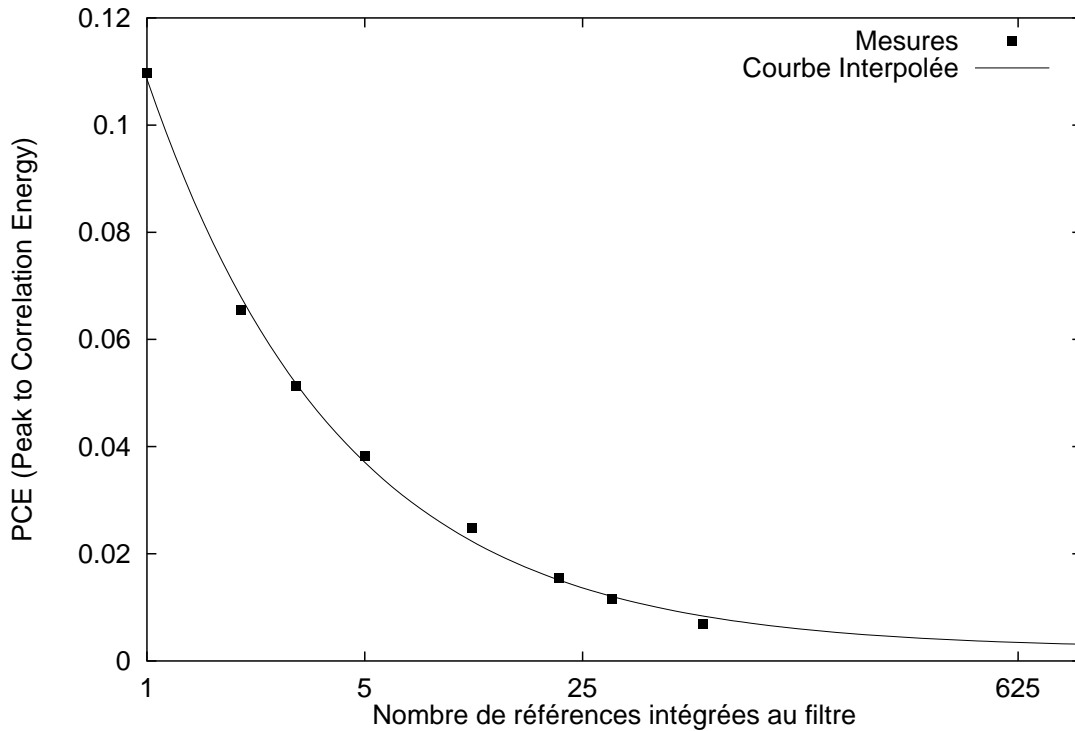


FIG. 1.10 – **Décroissance du PCE en fonction du nombre de références intégrées dans le filtre composite** : la loi de décroissance est fixée par la fonction $0,11 x^{-0.69} - 0.002$.

Toutefois, les résultats intéressants que nous venons de développer ne sont pas «généralisables» à tout type d'images. En effet, les empreintes digitales, que nous avons prises comme exemples, disposent d'un spectre très étalé (figure 1.11-a). La structure même de l'empreinte «ressemble» à un réseau de pas plus ou moins irrégulier, suivant deux directions. Il était donc nécessaire que son spectre comprenne préférentiellement les fréquences associées aux pas des réseaux des empreintes. *A contrario*, les visages, quelle que soit leur orientation, disposent de spectres très concentrés (figure 1.11-b). Cette propriété découle donc de la nature des textures présentes dans l'image : sur les visages, les hautes fréquences ne sont générées que par les contours liés aux visages.

Nous pouvons donc, à partir d'une attitude de visage, par exemple un visage en rotation, créer une base de filtres composites comprenant 5, 11 et 21 références. La figure 1.12 illustre pour un filtre de 21 références l'évolution des PCE en fonction d'une puissance β croissante.

Nous devons nous assurer que les filtres composites ainsi obtenus ont une bonne robustesse sur les références qu'ils représentent (figure 1.13). Nous constatons que le

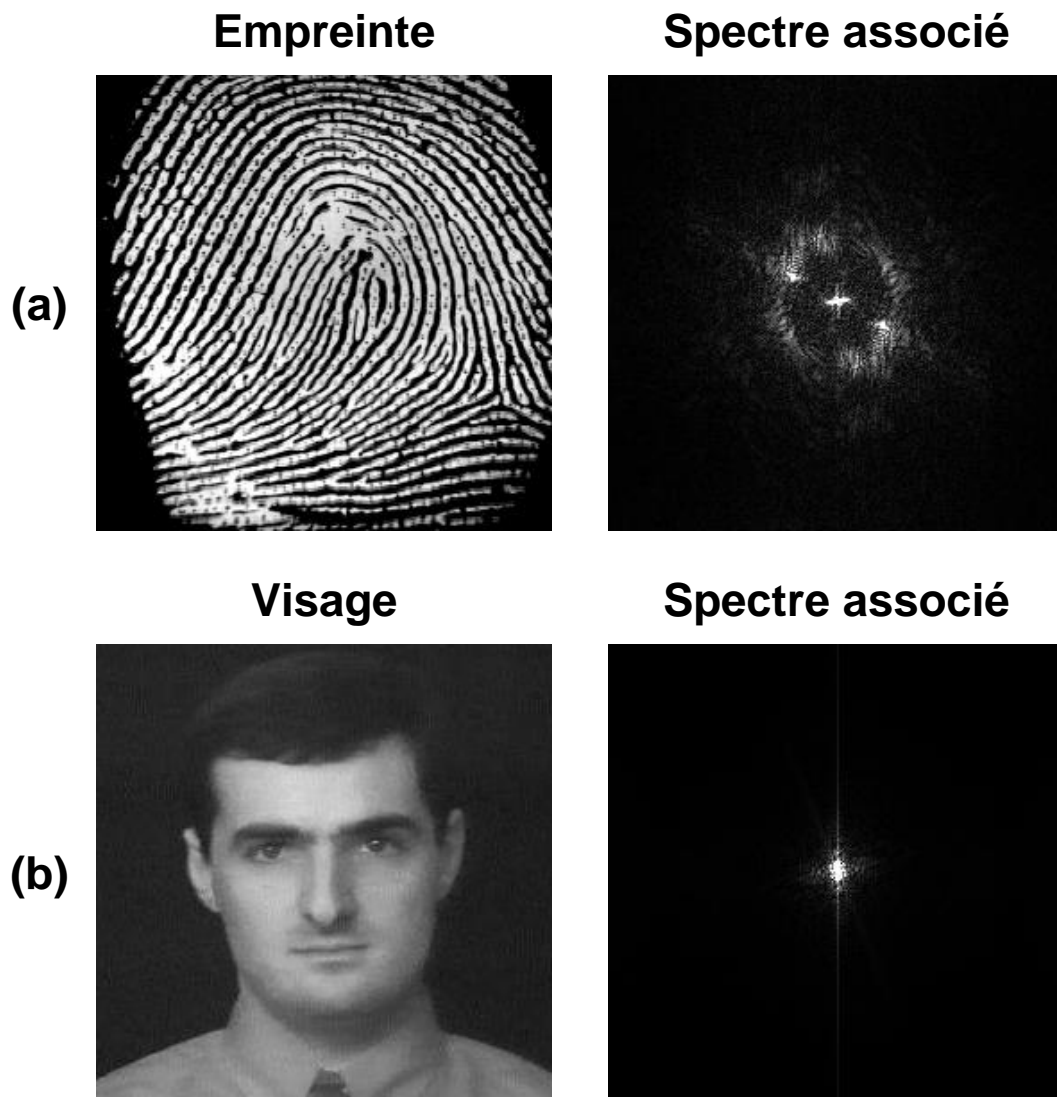


FIG. 1.11 – Exemples de spectres associés à deux types d'image : une empreinte digitale et un visage : pour une empreinte digitale (a), le spectre est «éclaté», comparé à un visage (b), dont le spectre est plus concentré. Pour rehausser les zones de haute fréquence les plus actives, chaque pixel blanc correspond à une valeur égale ou supérieure au $1/1000^{\circ}$ du maximum d'intensité sur l'image.

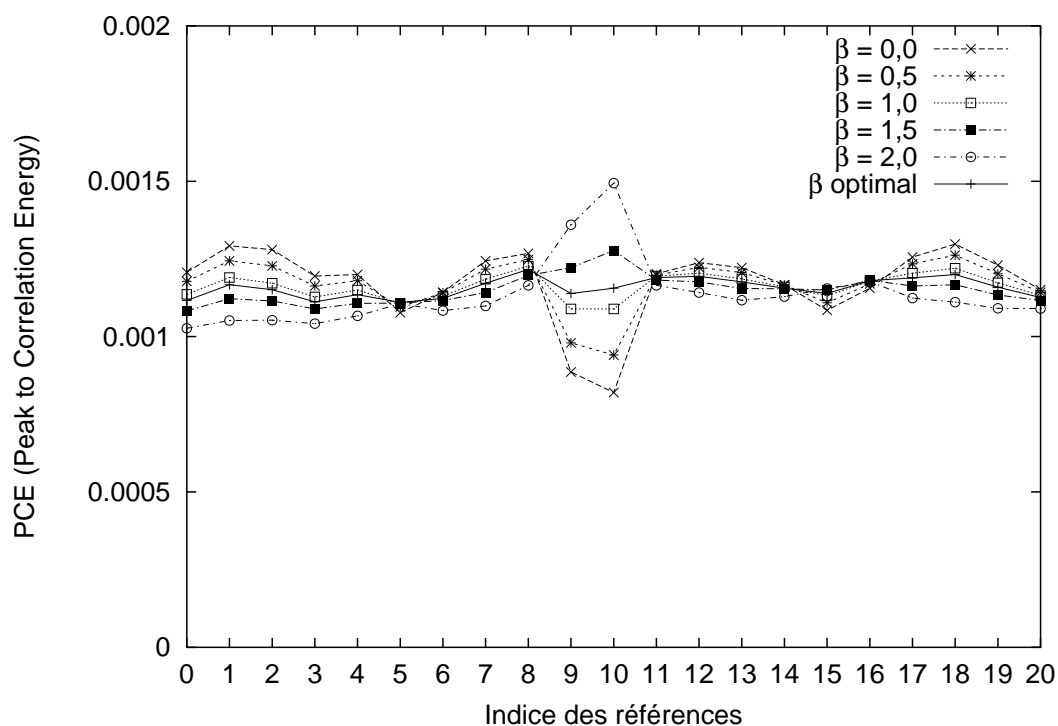


FIG. 1.12 – **Réponse du filtre pour une attitude de visage** : les courbes représentent les PCE obtenus par application du filtre à chacune des références qui le compose. La première courbe ($\beta = 0$) correspond à la réponse du filtre pour des poids unitaires. Les suivantes pour différentes valeurs de puissance β appliquées sur l'inverse des PCE obtenus pour $\beta = 0$. L'exposant optimal, dans ce cas, se situe autour à 1.2.

PCE moyen obtenu dans le cas d'une «activation» est bien inférieur à celui obtenu pour une empreinte digitale (typiquement, pour un filtre composite de 5 références, le *PCE* passe de 0,04 à 0,002, soit une chute d'un facteur 20).

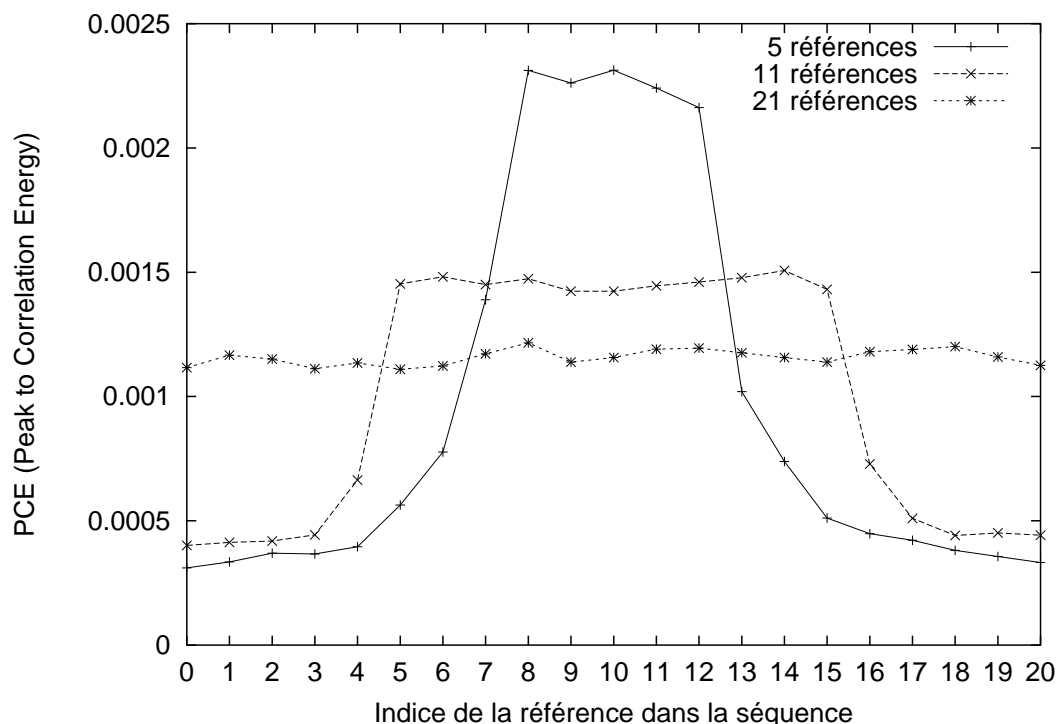


FIG. 1.13 – Réponses de filtres composites comportant 5, 11 et 21 références à une empreinte en rotation : le filtre est «activé» uniquement si la référence correspond à une de celles qu'il intègre.

Comme nous l'avons réalisé dans l'application liée aux empreintes, il est intéressant de constater la décroissance du *PCE* en fonction du nombre de références qui sont intégrées au filtre. La figure 1.14 présente de tels résultats. Si nous cherchons à approcher ces points par une fonction de la forme présentée en 1.15, nous trouvons comme coefficients $a \approx 0,005$, $b \approx 0,49$ et c négligeable. Nous constatons de plus que le rapport 20 entre les intensités du *PCE* pour des empreintes ou visages se retrouvent quel que soit le nombre de références intégrées.

En conclusion, l'opération de synthèse du filtre est rapide, dès lors que la puissance β , permettant un contrôle de l'uniformité, est approximativement connue. Toutefois, deux problèmes se posent désormais à son utilisation :

- comment lever l'ambiguïté sur une référence ?

Si un filtre composite «répond» par la présence d'un pic, il demeure impossible de savoir laquelle des références individuelles a été «activée» (ce point peut cependant être traité par une approche hiérarchique) ;

- comment éviter les problèmes de saturation ?

La saturation intervient lorsque les spectres trop centrés de chacune des références

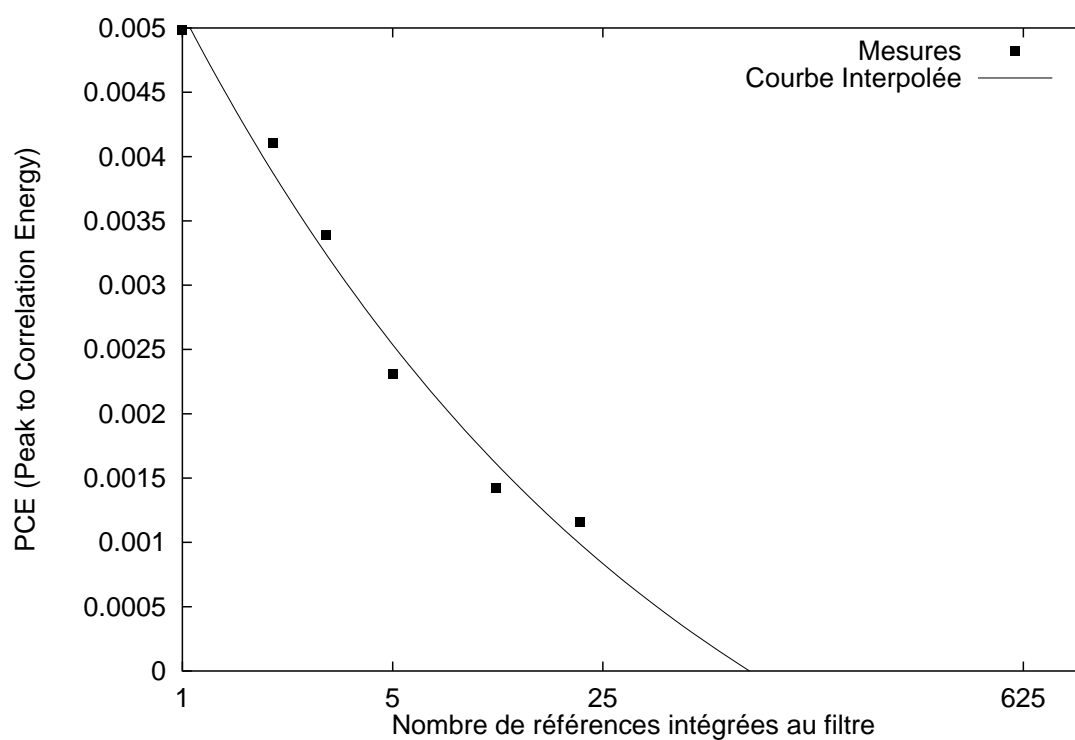


FIG. 1.14 – Décroissance du PCE en fonction du nombre de références intégrées dans le filtre composite : la loi de décroissance est fixée par la fonction $0,008x^{-0.26} - 0.002$.

se superposent et dégradent le filtre résultant ;

Pour répondre à ces deux problèmes, nous proposons deux solutions :

- un contrôle du positionnement des pics de corrélation ;
- une meilleure utilisation de la capacité de codage du filtre ;

Une fois associées, ces deux améliorations visent à transformer le filtre composite en filtre composite-multivoies.

1.8 Contrôle de positionnement des pics de corrélation

Le filtre composite nous permet donc d'intégrer un certain nombre de références dans le filtre, donc de reconnaître, à l'aide d'un seul et unique filtre, une base donnée. Toutefois, il faut préciser que, si l'une des références est «activée», il est pour l'instant impossible de savoir directement laquelle. Cette condition n'est pas d'une importance extrême lorsque les références sont associées à une même classe (par exemple, invariance par rapport à une distorsion). Il n'en est évidemment pas de même pour des références associées à des classes différentes.

En effet, si l'espace de codage est suffisant pour coder des classes différentes, il est nécessaire de savoir, lorsqu'un pic apparaît, à quelle référence il est associé. Cette ambiguïté peut être levée, par exemple, en contrôlant la position du pic de corrélation pour chacune des classes intégrées au filtre. Pour réaliser pareille opération, il suffit d'introduire auprès des spectres références composant chaque classe une phase complexe ou «porteuse» : une fois le passage dans l'espace direct effectué, cette phase introduite déplace le pic de la quantité voulue (figure 1.15).

Ainsi, si nous voulons qu'une image soit déplacée d'une certaine quantité (d_x, d_y) , son spectre doit être multiplié par une fonction de phase, une «porteuse», dépendant du couple (d_x, d_y) , de la dimension totale de l'image N_x, N_y . Ainsi, pour une image $s(i, j)$ de spectre $S(i, j)$, nous obtenons $S'(i, j)$ telle que :

$$S'(i, j) = S(i, j) e^{-i2\pi\left(\frac{i d_x}{N_x} + \frac{j d_y}{N_y}\right)} \quad (1.16)$$

De même, si nous désirons réaliser un filtre composite comprenant i références correspondant à des classes distinctes (différentes empreintes par exemple), nous allons légèrement modifier l'expression du filtre composite en la forme suivante :

$$F = \frac{\sum_{i=1}^N \sigma_i R_i^*}{\left| \sum_{i=1}^N \sigma_i R_i' \right|} \quad \text{avec} \quad R_i' = R(i, j) e^{-i2\pi\left(\frac{i d_x}{N_x} + \frac{j d_y}{N_y}\right)} \quad (1.17)$$

La figure 1.16 illustre les plans de corrélation obtenus par application d'un filtre composite avec porteuse sur les 5 références qui le composent. L'interprétation toutefois du plan de corrélation s'en trouve compliquée : en effet, notre critère, le *PCE* réduisait un plan de corrélation à un scalaire. Ici, non seulement le *PCE* garde son importance, mais sa position devient un élément primordial dans la connaissance de la référence activée. Enfin, la réponse du filtre à chacune des références peut, par exemple, être optimisée de façon à satisfaire un critère de réponse au filtre, tel que l'uniformité sur

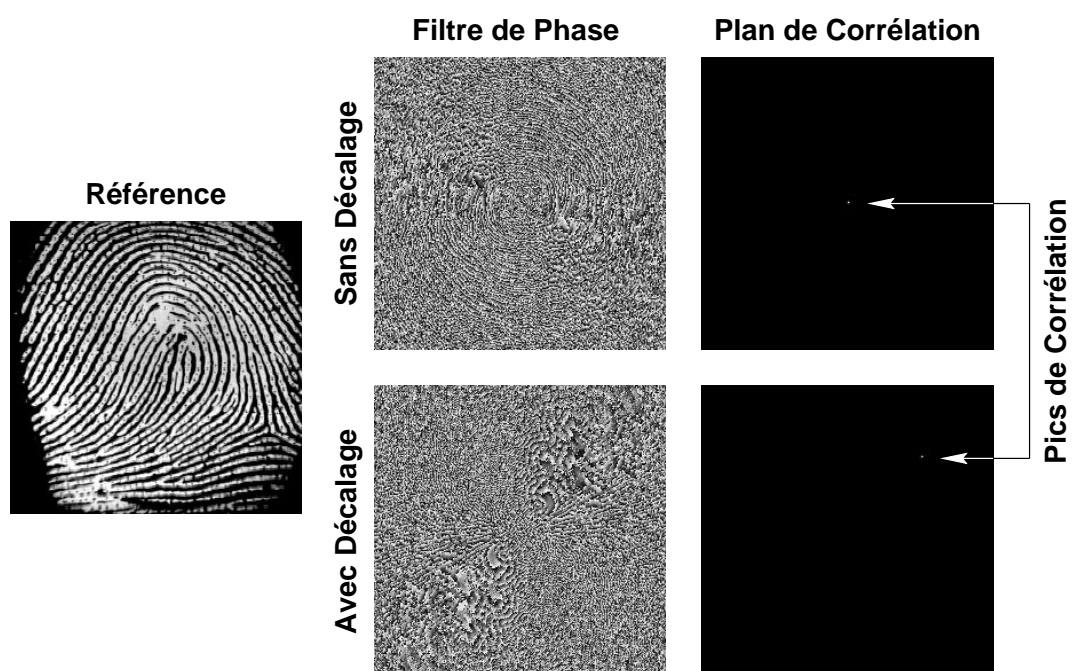


FIG. 1.15 – Exemple de fonction de phase intégrée au filtre : le filtre initial est un filtre de phase. En haut, le filtre sans intégration de la fonction de décalage présente un pic de corrélation centré. En bas, le filtre intègre une fonction de phase correspondant à un décalage de 64 pixels suivant les deux axes ; le pic de corrélation se trouve déplacé en conséquence suivant la première diagonale de 64 pixels.

le *PCE*. En tout état de cause, la décision ne peut se réduire à un unique scalaire, et l'analyse pondérée des différentes réponses s'avère ici plus que jamais indispensable.

C'est pourquoi, dans l'application telle que nous l'avons illustrée, les classes sont exclusives : deux références au sein d'un même filtre composite ne peuvent être «activées». Dans un cadre plus général, il est vraisemblable que plusieurs références soient activées simultanément, imposant dans le plan de corrélation plusieurs pics ; la première étape est une localisation des différents pics sur l'image, la seconde est une estimation de l'énergie pour chacun de ces pics, enfin la troisième est une estimation de l'énergie du plan de corrélation dans son ensemble. Ainsi donc pouvons-nous rechercher autant de *PCE* qu'il existe de classes intégrées dans le filtre composite. Si par contre nous utilisons le filtre pour connaître la position d'un objet déterminé dans une scène, la tâche est plus ardue : le déplacement d'un objet reconnu par le filtre dans la scène occasionne un déplacement du pic de corrélation. Si le pic se déplace d'une valeur trop importante, il peut «sortir» de la zone estimée comme celle étant associée à sa référence : une ambiguïté naît.

En conclusion, le contrôle du positionnement des pics de corrélation est une approche intéressante pour savoir à quelle classe nous devons associer un pic de corrélation. Ce codage supplémentaire exige l'introduction d'une phase complexe différente pour chacune des classes intégrées au filtre ; il agit donc sur la capacité de codage globale du filtre.

1.9 Limitation de la saturation par filtrage multi-voies

Le second problème concerne la saturation. Les exemples que nous avons développés, à base d'empreintes ou de visages, ont illustré de quelle façon la nature du spectre influençait la corrélation : les filtres composites issus de visages présentaient, non seulement une décroissance plus rapide, mais aussi un *PCE* en cas d'activation bien plus faible que leurs homologues calculés à partir d'empreintes digitales.

La raison invoquée étant la répartition spectrale de l'information, nous proposons donc une meilleure utilisation de la capacité de codage offerte par le filtre. En effet, si les superpositions saturaient rapidement certaines zones communes du filtre, pourquoi ne pas séparer les références dans l'espace de filtrage ?

L'alternative proposée correspond au concept de filtre multi-voies. En effet, dans l'architecture classique, le filtre composite dispose d'un espace de codage commun à toutes les références. Dans le cas d'un filtre multi-voies, un espace de codage est réservé à chaque référence, ou, plus généralement, à chaque classe de références : le filtre multi-voies se compose donc de filtres composites séparés. Pour réaliser le filtrage correspondant à une corrélation, il nous est nécessaire de présenter devant chaque voie un exemplaire du spectre à filtrer. Cette opération, la duplication de spectres, se réalise par le produit d'une fonction de phase pour l'ensemble des voies. Ainsi, ayant une scène $s(x, y)$ dont nous voulons reproduire le spectre 4 fois (suivant une mosaïque 2×2), c'est-à-dire obtenir :

$$\begin{aligned} S'(X, Y) &= S(X+a, Y+b) + S(X-a, Y+b) + S(X-a, Y-b) + S(X+a, Y-b) \\ &= S(X, Y) * (\delta_{-a, -b} + \delta_{a, -b} + \delta_{-a, b} + \delta_{a, b}) \end{aligned}$$

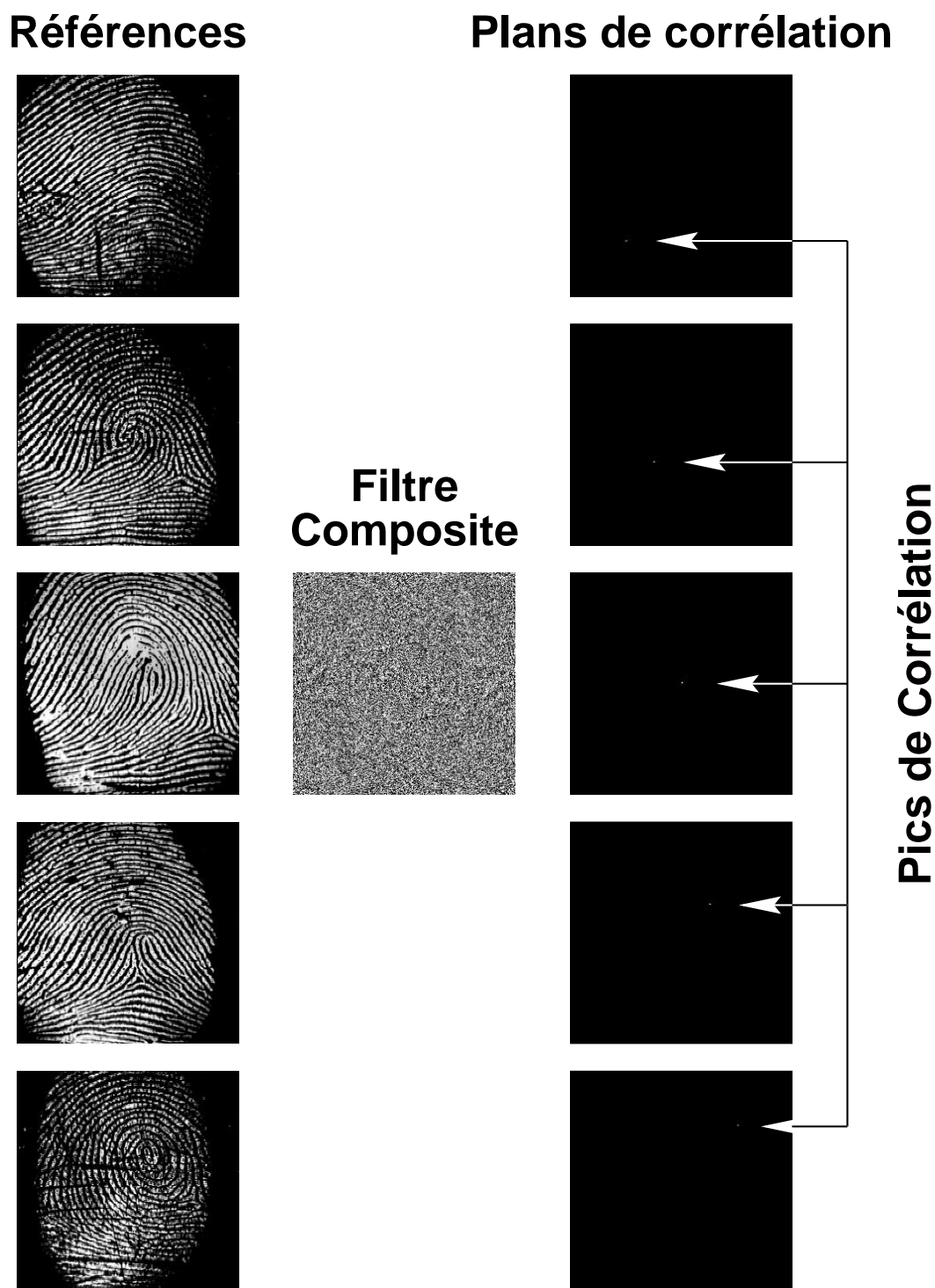


FIG. 1.16 – Application d'un filtre composite avec porteuse : le filtre composite présenté au centre a été construit à partir de cinq empreintes de références différentes. Lors de la synthèse du filtre, une porteuse différente a été appliquée à chaque spectre de référence. Il s'en suit un déplacement du pic de corrélation lorsque l'empreinte présentée correspond à une référence.

dans l'espace direct, nous aurons :

$$\begin{aligned} s'(x, y) &= (e^{2i\pi(ax+by)} + e^{2i\pi(ax-by)} + e^{-2i\pi(ax-by)} + e^{-2i\pi(ax+by)}) s(x, y) \\ &= 2(\cos(ax + by) + \cos(ax - by)) s(x, y) \\ &= d_{a,b}(x, y) s(x, y) \end{aligned}$$

Il est également possible de simplifier cette fonction de duplication $d_{a,b}(x, y)$ en posant devant chacun de ses cosinus (cos) une fonction signe (sgn).

La figure 1.17 illustre la fonction de duplication 2×2 ainsi que son action sur une empreinte digitale. Nous pourrions noter les 4 spectres formant une mosaïque 2×2 , identiques à celui présenté sur la figure 1.11.

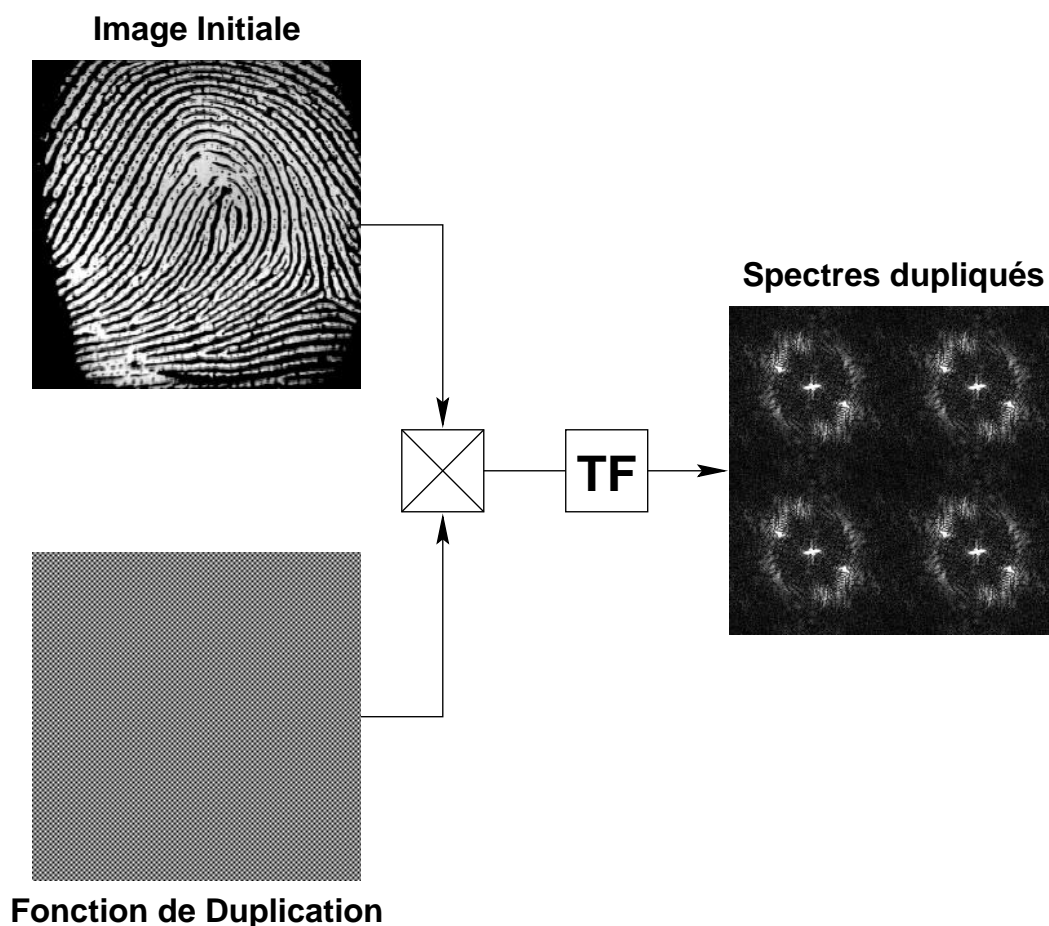


FIG. 1.17 – **Duplication d'un spectre** : elle se réalise par le produit entre l'image dont le spectre doit être dupliqué et la fonction à valeurs réelles $2(\cos(ax + by) + \cos(ax - by))$. a et b représentent le déplacement subi par les spectres dupliqués.

Une fois cette duplication de spectres effectuée, il ne nous reste plus qu'à multiplier chacun d'eux par le filtre désiré. Ayant divisé l'espace des fréquences en quatre quadrants, le codage de chacun des filtres ne se réalisera que sur $1/4$ de la surface totale. Nous allons donc calculer le filtre correspondant à la référence par les méthodes que nous avons déjà

développées, puis ajouter au filtre une fonction de déflexion de façon à déplacer le pic de corrélation. Nous sélectionnons ensuite uniquement la partie centrale. Le filtre (figure 1.18) issu de l'association de 4 références se construit donc en effectuant successivement pour chacune d'elles :

- le calcul de leur TF $\boxed{\text{TF}}$;
- le calcul du complexe conjugué du spectre $\boxed{\text{C}}$;
- le produit par une porteuse $\boxed{\text{P}}$;
- le produit par un masque passe-bas sélectionnant le quart de la surface $\boxed{\text{S}}$;
- le déplacement du spectre vers un des quadrants $\boxed{\text{D}}$.

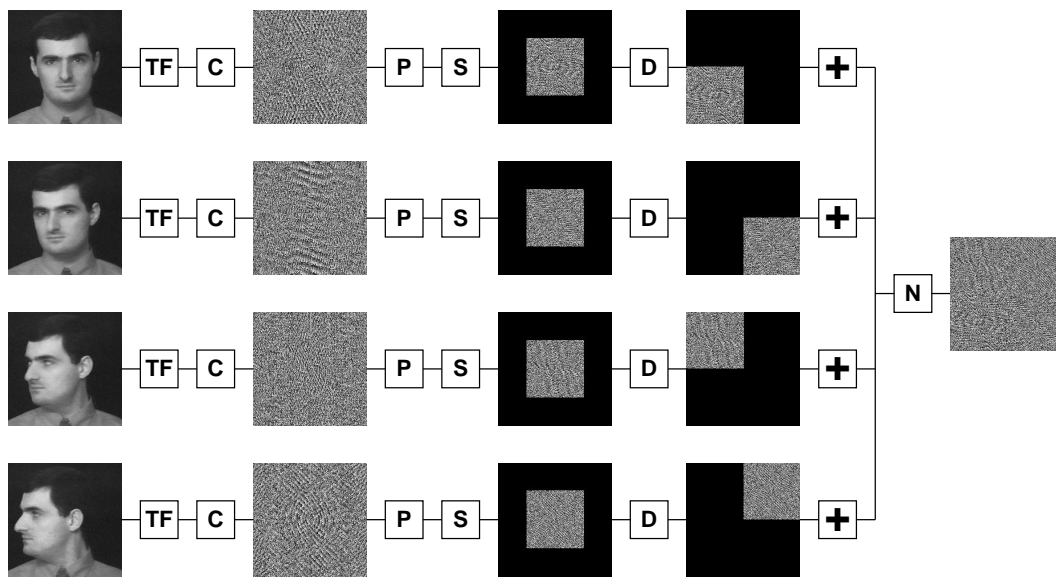


FIG. 1.18 – **Synthèse d'un filtre multivoies** : le filtre de phase finale (à droite) se compose de 4 filtres différents construit indépendamment, suivant la même méthode : après le passage dans l'espace des fréquences (TF), le spectre est conjugué (C), puis multiplié par une porteuse (P). Un masque passe-bas ne retient que le quart de la surface du filtre (S). Le reste est ensuite déplacé vers l'un des quadrants (D). Ces quatre filtres sont ajoutés, puis forment, après normalisation à l'unité (N), un filtre de phase multivoies. Remarque : seule la phase est présentée pour chacun des spectres.

Ensuite, les spectres issus de ces quatre traitements successifs sont additionnés $\boxed{+}$ puis les intensités normalisées à l'unité. Nous obtenons ainsi un filtre de phase multivoies, dont les quatre voies sont séparées. Pour appliquer un tel filtre à une image, il convient tout d'abord de dupliquer le spectre de l'image à l'aide de la fonction $d(x, y)$, puis d'en faire le produit avec le filtre de phase calculé précédemment. Une TF inverse permet de revenir dans l'espace de Fourier (figure 1.19).

Nous avons mieux utilisé la capacité de codage du filtre, en distinguant 4 espaces de codage dans le filtre. Ces quatre espaces de codage devaient intégrer, en plus du codage du filtre de phase correspondant aux références, la fonction de déplacement nécessaire à distinguer les différents plans de corrélation.

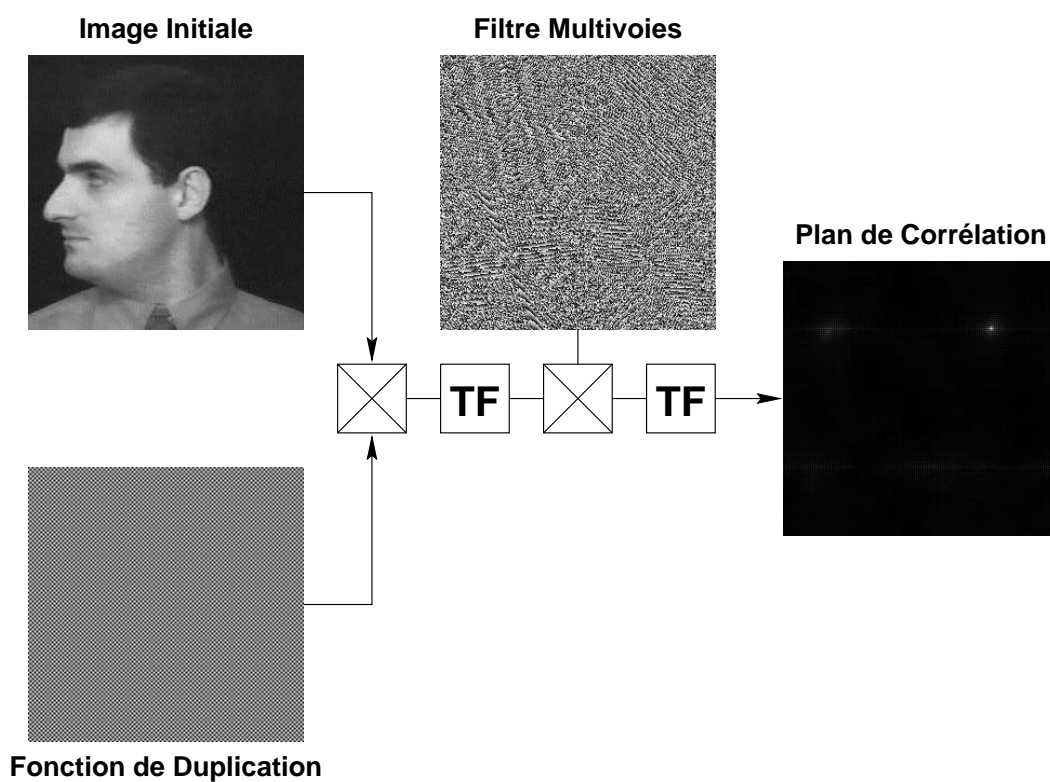


FIG. 1.19 – **Corrélation par un filtre multivoies** : l'image initiale est multipliée par une fonction de duplication. Le spectre dupliqué est ensuite multiplié au filtre avant de former dans l'espace direct un plan de corrélation.

Les gains en terme de corrélation d'une telle approche sont certains : ils permettent l'obtention d'un *PCE* entre 2 et 4 fois supérieur pour un nombre équivalent de références intégrées au filtre. Toutefois, ce résultat est à nuancer : le calcul du *PCE*, dans ce cas, est indépendant pour chaque quadrant. En effet, il est indispensable que nous disposions d'un *PCE* par canal. Ainsi, dans le cas initial, toute l'énergie n'étant pas comprise dans le pic était considérée comme bruit. Dans notre cas, l'énergie totale nécessaire au calcul du *PCE* se réduit au quart du plan de corrélation.

Ainsi, pour conclure, l'amélioration du codage du filtre divise le plan de filtrage en quatre composantes distinctes. Il divise également, par l'intégration au sein de chaque filtre élémentaire d'un contrôle sur la position des pics, le plan de corrélation en quatre quadrants. Le filtre dont nous venons d'établir la nature dispose donc de quatre voies distinctes, du filtre au plan de corrélation : il s'agit donc d'un filtre multi-voies. Derrière ce terme multi-voies, nous dissimulons la notion de séparation, autant des codages que des résultats de filtrage.

1.10 La synthèse d'un filtre composite-multivoies

Parmi les deux améliorations que nous avons proposées, nous avons un contrôle de la position des pics et une meilleure utilisation de la capacité de codage du filtre. Cependant, dans la synthèse du filtre multivoies présentée, chaque filtre élémentaire se réduisait au quart de sa surface originelle, de façon à former l'un des quatre canaux du filtre multivoies [1]. Une fusion des approches composite et multivoies consiste à envisager la superposition des filtres, et donc à créer des recouvrements. Les quatre voies de filtrage d'un filtre multivoies seraient donc associées comme le sont les références d'un filtre composite. Nous formerons ainsi un filtre composite-multivoies [6].

Toutefois, cette fusion des spectres n'aura d'intérêt que si la nature des spectres le permet : si les spectres des différents canaux se chevauchent, le gain en capacité de codage sera naturellement réduit. Il convient donc de s'intéresser à la répartition des spectres avant d'utiliser pareille approche.

Face aux performances qu'offre un filtre multivoies, nous avons généralement constaté, lors de l'utilisation d'un filtre composite-multivoies un gain significatif dans une application à des visages ; une augmentation de la valeur du *PCE* de 5% est généralement constatée. Ne disposant pas de bases très hétérogènes, se limitant à celles de visages ou d'empreintes, nous n'avons pu trouver de cas dans lesquels un filtre composite-multivoies développe toute sa puissance.

1.11 Le Prétraitement comme opération de filtrage

Cette quête du filtre le plus adapté à nos applications conduit à définir et à tester le domaine des filtres composites, multivoies et composite-multivoies. Elle nous a sensibilisés sur le contrôle du positionnement des pics de corrélation, sur la nécessité de ne plus réduire trop sommairement le plan de corrélation au *PCE* mais à un ensemble de

PCE, dont il nous est indispensable de connaître la position. Des trois composantes du filtrage, la scène, le filtre et la décision, nous n'avons développé que les deux derniers. Intéressons-nous maintenant à la scène. En effet, les prétraitements sont connus pour avoir une influence sur les traitements par filtrage car ils permettent de présélectionner l'information utile parce que pertinente pour établir la comparaison. De ce fait, la corrélation extrait des caractéristiques plutôt qu'elle ne prend une décision [2].

Comme nous pouvons le constater à la lecture des différents résultats issus de simulations, le filtrage de phase est très sélectif. Il serait donc intéressant de «contrôler» cette puissance de discrimination à l'aide d'un prétraitement sur l'image.

1.11.1 Intérêt d'un prétraitement

La première étape consiste à déterminer le rôle que nous comptons confier aux prétraitements des scènes et références. En effet, étant dans une logique de filtrage entre scène et référence, notre objectif demeure un compromis entre sélectivité et robustesse.

Compte-tenu de la nature des filtres utilisés jusqu'à présent (filtres de phase), seule la phase du spectre était affectée lors du passage au travers du filtre. Ainsi, plus le spectre d'un objet à reconnaître est distribué spatialement, plus le filtre sera efficace [32].

De plus, parmi les prétraitements possibles, certains peuvent être assimilés à des filtrages (convolution par un noyau). Dans ce cas, nous montrerons facilement qu'une opération de prétraitement peut être reportée dans l'espace de Fourier, tout comme nous avons défini le filtrage.

Parmi ces prétraitements linéaires, nous nous intéresserons aux prétraitements susceptibles d'extraire des caractéristiques locales des objets. Le filtrage par ondelettes sera une bonne illustration du report dans le plan du filtre de cette fonction de prétraitement. Le chapitre destiné aux applications nous démontrera que cette approche s'applique aussi bien sur des attitudes de visage que sur des signaux «temps-fréquence» de radar.

1.11.2 Un prétraitement intégrant la résolution

Ainsi donc, notre choix va-t-il se porter sur un prétraitement permettant d'extraire d'une image des caractéristiques d'une certaine taille. Nous accédons alors à divers détails, lesquels dépendent d'un nouveau paramètre, la résolution. Pour cela, nous avons opté pour un prétraitement sous forme de filtrage, entre notre scène ou référence et une ondelette continue [25].

La transformée en ondelette continue (Annexe A) est équivalente à un filtrage adapté entre la fonction à filtrer et une ondelette ψ de paramètre a . Ce paramètre fixe le filtrage passe-bande effectué par l'ondelette dans l'espace spectral (Annexe A.1) : plus a est grand, moins les hautes fréquences sont coupées. Nous sélectionnerons donc les basses fréquences avec des valeurs importantes de a , les hautes fréquences avec les valeurs faibles

de a .

Si nous utilisons comme ondelette le «chapeau mexicain» (équation A.10), le filtrage adapté se réduit à un filtrage standard : la fonction ψ étant réelle et paire, la TF de cette dernière sera également réelle et paire. Ainsi, le filtre p de spectre P évoqué plus haut se réduit à l'ondelette ψ_a de spectre Ψ_a .

Comme nous l'avons brièvement développé dans l'appendice A, la généralisation à deux dimensions de notre filtre est immédiate. Nous obtenons ainsi :

$$\psi_a(x, y) = \psi_a(x) \times \psi_a(y) \quad (1.18)$$

$$= \frac{1}{|a|} \frac{4}{3\sqrt{\pi}} \left(1 - \frac{x^2}{a^2}\right) \left(1 - \frac{y^2}{a^2}\right) e^{-\frac{x^2+y^2}{2a^2}} \quad (1.19)$$

De manière à juger de l'action d'une telle ondelette sur une image, appliquons-la à une image de visage suivant les cinq niveaux présentés sur la figure A.1.

Pour cela, nous allons estimer la corrélation entre une image de visage (figure 1.20) et une base d'ondelettes associées aux différents niveaux exprimés par la variable a^1 . Nous notons que la taille du détail élémentaire extrait lors d'une corrélation est comparable à celle du pic central de l'ondelette. Nous notons que certaines zones, pour $a = 1$ notamment, correspondent à certaines zones du visage, notamment les yeux, le nez, le front, les contours du menton.

Ainsi donc, le prétraitement utilisant un filtrage par une ondelette permet de revenir à l'extraction de caractéristiques liées à l'image, donc à l'objet observé. Nous pouvons remarquer que cette approche combine une opération traditionnelle de reconnaissance de formes et d'extraction de caractéristiques.

1.11.3 Intérêt d'un prétraitement linéaire

Nous venons de présenter un prétraitement sous la forme d'un filtrage entre une ondelette et une image originelle. Nous avons vu de quelle manière le coefficient de contraction a permettait d'extraire d'une image certains détails, dépendant implicitement de la taille, comme les yeux, le nez, la bouche, le menton, les oreilles.

Nous nous proposons désormais de ne plus considérer la corrélation entre scène et référence, mais celle entre scène et référence préalablement prétraitées. Le prétraitement que nous venons de mettre en œuvre était un filtrage adapté. Ainsi, deux filtrages interviennent pour prétraiter scène et référence, puis un filtrage pour établir la corrélation entre les scène et référence. Les filtrages étant des opérations linéaires, il peut s'avérer

¹Les images que nous manipulons sont composées de pixels, donc échantillonnées. A cette dimension d'image exprimée en nombre de pixels peut être associée une dimension réelle telle que le pas d'échantillonnage Δ soit identique à la fois dans les espaces direct et spectral. En effet, considérons un signal composé de N échantillons de pas Δ , le spectre associé aura une fréquence d'échantillonnage de $1/\Delta N$. Ainsi, comme $\Delta = 1/\Delta N$, nous avons $\Delta = 1/\sqrt{N}$. Une image de dimensions 256×256 pixels aura une taille «réelle» de $256/\sqrt{256}$, soit 16. Les valeurs de x et y qui seront associées à l'image seront comprises entre -8 et $8 - 1/256$.

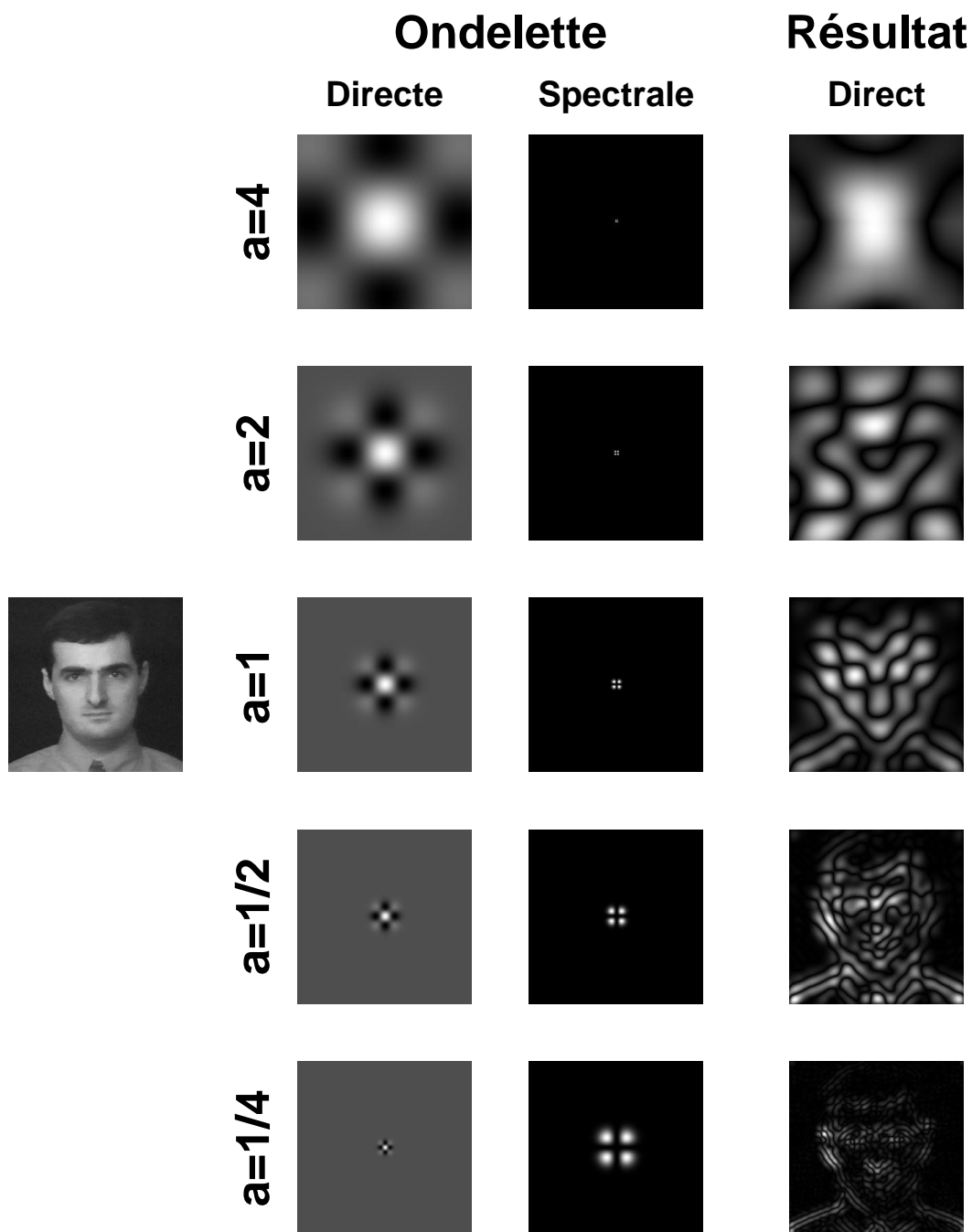


FIG. 1.20 – Exemple de filtrage par une base d'ondelettes de niveaux a compris entre $1/4$ et 4 : A gauche est représentée l'image originelle. A droite apparaissent les résultats de la corrélation entre l'image originelle et une ondelette représentée dans sa forme directe ou spectrale (les résultats présentent le module de l'image).

intéressant de les regrouper au sein d'un seul et unique filtre. Le gain, autant en terme de temps de calcul que de flexibilité, est évident : un seul filtrage suffit pour opérer la corrélation et le prétraitement sur les scènes et références, ce par une unique opération de multiplication dans l'espace des fréquences. De plus, les bases de données ne comprennent que des filtres issus d'un ensemble de références, auxquels sont associés les quelques filtres de prétraitement. Il n'est donc pas nécessaire de disposer de chacun des filtres intégrant chacun des prétraitements possibles.

Disposant désormais du filtre de prétraitement, dont nous décrirons plus en détail la structure, intéressons-nous aux formes que prennent les filtres intégrant le prétraitement linéaire et la référence. Pour cela, désignons par S le spectre de la scène, R le spectre de la référence et W le filtre de prétraitement. Quant au filtre résultant, il sera noté F .

Regardons pour les différents types de filtres étudiés précédemment le filtre équivalent dans lequel nous avons introduit les prétraitements. Nous désignerons par (s, S) la scène et son spectre associé, par (r, R) la référence et son spectre associé, enfin par (p, P) la fonction de prétraitement et son spectre associé. Nous avons en effet vu que la relation entre une référence r de spectre R et son filtre F avait pour expression :

$$F = \frac{R^*}{|R|^{2\gamma}} \quad (1.20)$$

Le filtrage permettant de prétraiter la scène S est le suivant :

$$S_P = S P \quad (1.21)$$

Le filtrage permettant de prétraiter la référence R est le suivant :

$$R_P = R P \quad (1.22)$$

Ainsi, la corrélation entre la scène et la référence prétraitées a pour expression :

$$(S P)F = (S P) \frac{(R P)^*}{|R P|^{2\gamma}} = S \frac{R^* \cdot |P|^2}{|R|^{2\gamma} |P|^{2\gamma}} \quad (1.23)$$

De cette expression, nous déduisons l'expression du filtre intégrant la comparaison et le prétraitement :

$$F = \frac{R^*}{|R|^{2\gamma}} |P|^{2(1-\gamma)} \quad (1.24)$$

Observons maintenant, pour différentes valeurs de γ courantes, comment ce filtre intégrant corrélation et prétraitement se comporte :

Valeur de γ	Type de filtrage	Filtre Correspondant
$\gamma = 0$	Filtrage adapté	$F = R^* \cdot P ^2$
$\gamma = 1$	Filtrage inverse	$F = R^* / R ^2$
$\gamma = 0,5$	Filtrage de phase	$F = R^* \cdot P / R $

Nous constatons tout d'abord que l'expression du filtrage inverse est identique si nous appliquons un prétraitement. Ce résultat est cohérent, compte tenu de la nature même du filtrage inverse.

La différence entre le filtrage adapté et celui de phase se situe dans la puissance du module du spectre du dénominateur et dans la puissance appliquée au filtre de prétraitement P .

Toutefois, notre objectif ne visait pas à retirer de cette image prétraitée les informations immédiates : nous nous étions fixés comme objectif de contrôler la sélectivité par un prétraitement. En effet, si nous nous plaçons à un niveau de détail, fixé par a , la corrélation comparera deux images à un niveau de résolution donné : si la résolution est faible, le détail sera grossier et la robustesse plus grande. La réciproque est immédiate : la sélectivité sera d'autant meilleure si le détail élémentaire est plus fin.

Pour juger du gain présenté par ce prétraitement, recherchons la corrélation entre une image de visage et des références de visages différents. A chaque référence est appliquée une même base d'ondelettes de niveaux a différents. La figure 1.21 illustre l'action d'un prétraitement linéaire sur la corrélation. Nous pouvons également adjoindre à ces plans de corrélation les estimations des PCE correspondants (table 1.1). Une étude à chaque niveau de résolution s'impose :

- Pour $a = 4$, les détails comparés sont très grossiers : le pic de corrélation est très évasé, comparable pour chacune des références. La sélectivité y est moindre. Ces résultats sont confirmés par les PCE correspondants.
- Pour $a = 2$, les pics de corrélations s'affinent. Nous remarquons une figure de corrélation symétrique pour la majorité des références, exceptée pour la 3^e : son maximum est cependant moins intense que celui des trois autres.
- Pour $a = 1$, seule la 2^e référence conserve une figure de corrélation symétrique et contrastée. Les 1^e et 4^e références comprennent un pic accompagné de plusieurs pics secondaires d'intensité comparable.
- Pour $a = 1/2$, le pic de corrélation s'affine encore pour la 2^e référence. Nous distinguons encore un faible pic pour la 4^e référence, observation confirmée par les valeurs de PCE .
- Pour $a = 1/4$, seule la 2^e référence présente un pic de corrélation.

Ainsi, grâce au prétraitement linéaire, nous améliorons la sélectivité de la corrélation, par son critère, le PCE , au travers du facteur de dilatation a . L'objectif d'une telle démarche de prétraitement est double :

- la recherche de caractéristiques particulières communes à plusieurs classes, à des résolutions différentes : à partir de l'ondelette, nous réalisons l'extraction et la localisation de détails liés à la dimension typique de l'ondelette. Sur la figure 1.20, nous remarquons en effet que la dimension typique des détails est comparable à la taille de l'ondelette.
- la hiérarchisation de la base de défilement : dans le cadre d'une corrélation classique, un grand nombre de références est comparé à la scène ; se pose alors le compromis entre l'exhaustivité de la classe et sa pertinence dans un cadre opérationnel. Si nous nous plaçons à de faibles niveaux de détail, certaines classes auront tendance à se confondre, par leurs PCE . Il est donc inutile de les conserver disjointes : une classe

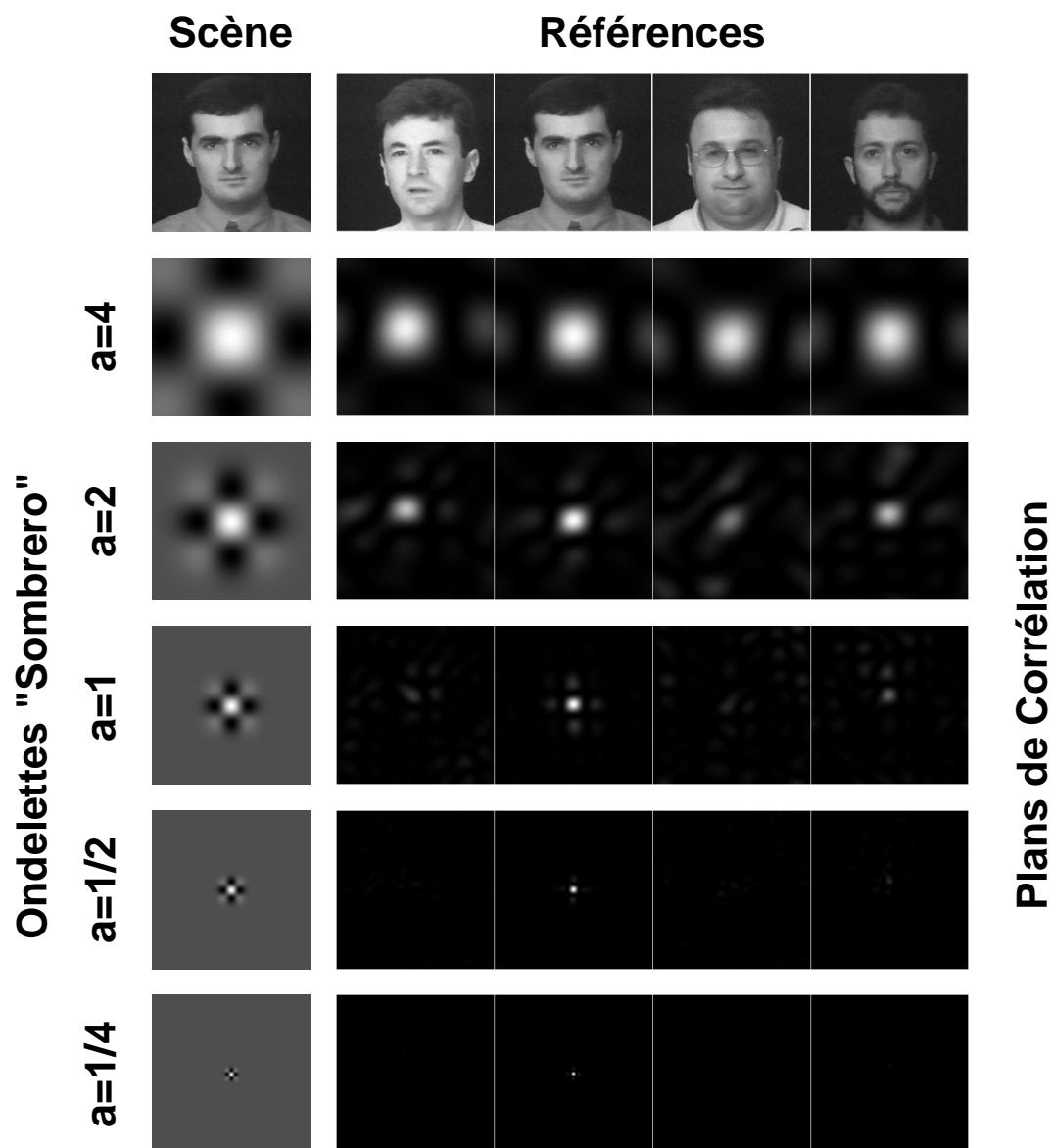


FIG. 1.21 – **Corrélation avec prétraitement linéaire à base d'ondelettes** : la colonne de gauche présente successivement l'image comparée à la base de références ainsi que les différentes ondelettes (partie réelle dans l'espace direct). En haut, les quatre références forment la base de comparaison. Enfin, une mosaïque de 4×5 plans de corrélation (représentés en énergie). Pour de faibles niveaux de détail (a grand), la discrimination n'est pas significative.

Niveau	Référence			
	1 ^e	2 ^e	3 ^e	4 ^e
4	0,14	0,15	0,13	0,14
2	0,3	0,4	0,18	0,29
1	0,28	1,3	0,19	0,48
1/2	0,3	6,2	0,36	1
1/4	0,64	24	0,28	1,2

TAB. 1.1 – *PCE* avec **prétraitement linéaire** : les *PCE* ($\times 10^{-3}$) représentés sont ceux issus des plans de corrélation entre une image scène et 4 images formant la base de référence. a représente le niveau de dilatation de l'ondelette. La 2^e référence contraste de plus en plus avec les trois autres références lorsque a diminue.

significative se substitue donc à chaque ensemble de classes proches à un niveau de résolution donné. A chaque saut de résolution correspond une nouvelle base de classes exclusives, donc une nouvelle base de défilement de références à présenter. Cette gestion impose de comparer la totalité des résultats à chaque résolution, de choisir la base de classes la plus vraisemblable et le cas échéant, de permettre un «saut arrière» de résolution [34].

Ces deux applications d'un contrôle de la corrélation semblent *a priori* contraires : dans un cas, un niveau de détail croissant permet de ne sélectionner à chaque étape que les références vraisemblables. Dans l'autre cas, le prétraitement linéaire consiste plus en une extraction de caractéristiques : chaque niveau de résolution permet d'accéder à une information supplémentaire. En fait, dans ces deux approches complémentaires, l'information que nous utilisons est locale dans un cas, globale dans l'autre cas.

1.12 Conclusion

Ainsi, le filtrage en reconnaissance de formes comporte trois éléments : la scène à reconnaître, le filtre et le critère issu du plan de corrélation. Nous nous sommes donc successivement intéressés :

- au critère du plan de corrélation : le *PCE* ;
- au filtre : un filtre de phase :
plusieurs évolutions successives dans la synthèse du filtre ont été ainsi développées : le filtre devenait composite (pour intégrer un ensemble de références), multivoies (pour éviter les effets de saturation). Un déplacement du pic dans le plan de corrélation permettait de lever l'ambiguïté sur le filtre activé. Ce codage supplémentaire exigeait une nouvelle analyse du plan de corrélation par une analyse des différents pics possibles associés.
- à la scène et aux références associées :
un prétraitement linéaire, à base d'ondelettes, permettait d'établir la corrélation entre scène et référence préalablement prétraitées. La propriété de linéarité d'un tel prétraitement permet de l'intégrer dans le filtre ;

Toutes les études sur ces trois plans l'ont été dans un seul et unique but : rendre le traitement le plus simple possible. Comme nous avons pu le voir, toutes ces améliorations ont

été ramenées dans le plan de filtrage, quel que soit l'objectif originel, et ceci afin de simplifier l'architecture (simple TF), la contrepartie se reportant sur la complexité du codage.

Nous disposons donc désormais d'un outil établissant une connexion entre des données brutes, des bases de références pour concourir à la création puis à l'utilisation de signatures particulières, lesquelles seront caractéristiques d'un comportement. C'est donc dans ce cadre applicatif que va s'insérer ce filtrage. Tel est donc l'objectif que nous allons poursuivre dans le prochain chapitre : la synthèse puis l'analyse de signatures.

Chapitre 2

Synthèse et analyse de signatures

2.1 Introduction

Ce chapitre a pour objectif de présenter les principes généraux de la synthèse et l'analyse de signatures, pour parvenir à identifier les traitements génériques, lesquels seront appliqués aux données réelles.

Pour cela, nous commencerons par nous pencher sur la nature des signaux que nous aurons à traiter, qu'ils soient unidimensionnels ou bidimensionnels, correspondant à des phénomènes déterministes ou stochastiques. Ensuite viendront une description des différentes approches de synthèse de signatures et l'implication du filtrage spatial dans un tel processus. Enfin, une fois cette signature établie, nous rechercherons comment le filtrage spatial, employé lors de la synthèse de la signature, pourra être utilisé dans la phase de reconnaissance.

2.2 Problématique

Il est tout d'abord nécessaire de replacer les signaux que nous allons traiter dans leur contexte originel, de nous intéresser à leur nature avant de développer plus en profondeur quelle signature serait la plus susceptible de les représenter au cours du temps.

2.3 Le caractère temporel d'un signal

Reprenons quelques-uns des thèmes brièvement évoqués dans l'introduction générale, à savoir «signaux à caractères temporels».

Tout système physique présente un comportement vis-à-vis du temps. Cette dépendance temporelle se scinde en deux catégories fort distinctes que sont les caractères déterministe ou stochastique. Dans ces deux cas, l'indépendance vis-à-vis du temps des caractéristiques d'un système permet de conclure à son comportement stationnaire. Toutefois, en télédétection, le comportement intrinsèque d'un système, que nous qualifierons désormais de «cible», se détermine par l'acquisition suivant la flèche temporelle d'un ensemble d'observables. Leur nature physique, leur variabilité sont autant d'informations complémentaires qu'il nous est indispensable d'analyser pour parvenir à identifier l'objet,

son attitude, ou simplement restituer une scène observée. Malheureusement, ces signaux que nous détectons se situent au dernier maillon d'une longue chaîne d'«opérateurs» de transformation que sont principalement la cible, le canal de transmission ou le détecteur. Ainsi donc, cette information «physique», sous forme de photons, de pression locale voire de particules, transformée en signal électrique, est-elle soumise à une variabilité temporelle dont les principaux acteurs sont :

- la cible ;
- le canal de transmission ;
- le détecteur ;
- un déplacement relatif entre cible et détecteur.

Chacune de ces composantes dispose de variabilités, aussi bien déterministes que stochastiques.

Ainsi, par exemple, l'évolution de signaux d'antenne synthétique active trouve-t-elle son origine dans le déplacement du détecteur par rapport à la cible, celle de l'antenne synthétique inverse dans le mouvement intrinsèque de la cible et la forte variabilité de deux échos successifs de sonar dans la perturbation du canal de transmission : l'eau.

2.4 Vers la formation d'une signature

Les signaux à variabilité temporelle sont donc la «matière» de notre opération de synthèse de signature : à partir de ces derniers nous devons construire notre signature, caractéristique non seulement de la variabilité du signal, mais aussi de sa structure à un instant donné. La difficulté de cette synthèse réside dans son caractère bivalent : elle doit être assez sélective pour séparer deux objets ou deux attitudes distinctes, tout en étant assez robuste pour ne pas considérer une fluctuation stochastique comme une évolution déterministe liée au système «cible-canal-détecteur». Ce compromis entre sélectivité et robustesse suggère l'utilisation d'un outil : le filtrage spatial.

Ce choix du filtrage spatial ou l'utilisation d'un certain type de capteurs impose cependant la structure de notre signature : image. Il est indispensable qu'elle intègre une dimension temporelle sous-jacente. Dans cette optique, nous nous proposons de réserver une des deux dimensions de l'image au codage de l'aspect temporel du signal, l'axe perpendiculaire étant destiné au codage de l'information en elle-même. Nous allons bientôt découvrir que ce choix, qui semble cependant limitatif pour des signaux bidimensionnels à caractère temporel évolutif – typiquement une attitude de visages – va nous contraindre à soigner notre transformation : si l'une des deux dimensions de la signature n'est autre que le temps, suivant l'autre direction, il nous sera indispensable de procéder à une réduction de 2D vers 1D (figure 2.1).

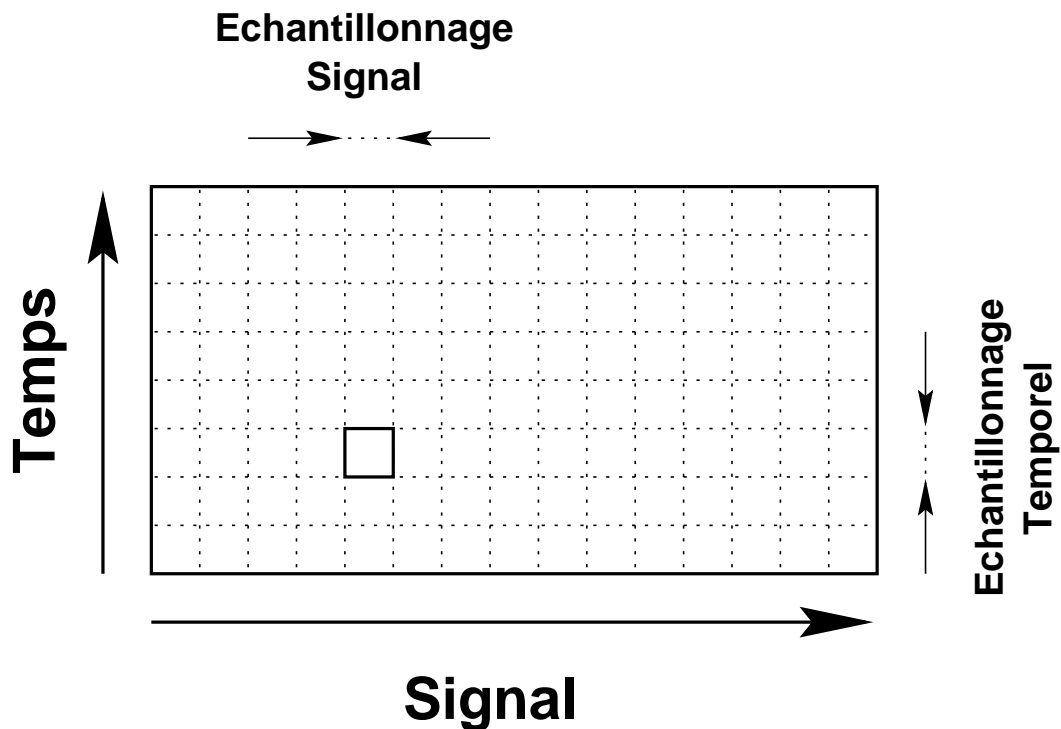


FIG. 2.1 – Codage de l'information dans une signature : Une signature temporelle comprend l'information physique issue du capteur et son évolution temporelle.

2.5 Synthèse de signature

2.5.1 Introduction

Nous avons eu l'occasion, dans la problématique, de revenir sur la nature de notre signature afin que nous puissions l'utiliser pour reconnaître une attitude. Nous avons, en effet, postulé que cette signature était bidimensionnelle : une image dont une des coordonnées (par exemple suivant l'axe vertical) était destinée au codage temporel de l'information. Suivant l'autre axe, nous disposons d'une liberté totale de codage de l'information.

Le schéma présenté figure 2.1 illustre le codage de la signature, sa structure. Parmi les mesures physiques à partir desquelles nous formerons notre signature, nous avons des signaux échantillonnés sous forme de vecteur (une mesure physique correspondant, par exemple, à une pression locale évoluant au cours du temps) ou de matrice (une séquence d'images issue d'une caméra).

Pour cela, nous partirons d'un signal unidimensionnel particulier, les signaux d'antenne synthétique. Nous appréhenderons la formation d'une signature conditionnée par la nature même de ce mode d'observation, dans le modèle de celle définie figure 2.1. Nous montrerons que cette signature d'antenne synthétisée naturellement, directement à partir des signaux reçus par l'antenne, peut se généraliser, moyennant quelques transformations, à des séquences d'images quelconques.

2.5.2 Signature d'une antenne synthétique

Comme nous allons le dévoiler, l'idée de synthèse d'une signature temporelle trouve son origine dans la formation naturelle d'une image dans le traitement d'un signal d'antenne à ouverture synthétique [30] [12].

En effet, nous ne reviendrons pas ici sur les principes d'une antenne synthétique et les améliorations qu'elle propose par rapport aux antennes classiques. Nous nous bornerons à préciser qu'une antenne synthétique se déplace relativement à la cible au cours du temps et que c'est cette évolution qui donne à l'antenne synthétique ses performances en terme de résolution.

De fait, le principal avantage de l'antenne synthétique active – par prudence de langage, nous précisons «active» pour ne pas les confondre avec les dispositifs d'interférométrie – réside dans sa principale voire paradoxale propriété : la résolution suivant la direction de déplacement ne dépend que de la taille de l'antenne.

Pour atteindre pareille caractéristique, l'antenne synthétique émet un signal périodique puis reçoit un ensemble d'échos provenant de la scène par rapport à laquelle elle se déplace. Chaque récurrence reçue comporte donc un instantané de la totalité des échos. Un empilement de ces récurrences permet pour chaque point de former une signature d'antenne, réponse qu'une antenne synthétique donne lorsqu'elle se déplace dans une scène. Pour un point unique, l'antenne se rapproche puis s'éloigne de la cible suivant une distance qui évolue de manière parabolique. Le délai séparant émission et réception étant proportionnel à la distance, il évolue suivant la même loi parabolique. Ainsi, la signature d'antenne d'une scène est-elle la superposition cohérente de la totalité des signatures d'antenne élémentaires des points de la scène[8].

Signature d'une cible ponctuelle

Le signal émis par l'antenne est un signal de porteuse f_e modulé linéairement en fréquence par b_e et de durée t_e . Les paramètres d'antenne que nous allons utiliser pour établir notre synthèse sont (les valeurs temporelles sont exprimées à partir des unités de distance, par l'utilisation de la célérité c sur milieu, égale approximativement à 1450 m/s dans l'eau de mer) :

- une émission de porteuse de fréquence f_e de 5 m^{-1} ;
- une modulation b_e de 5 m^{-1} ;
- une durée t_e de 4 m ;
- une dimension d'antenne L_a de 20 cm ;
- une vitesse de déplacement v_0 de 1 m/s ;
- une distance séparant antenne et centre de la cible D_0 de 10 m/s .

Nous notons que les paramètres d'émission de l'antenne sont rapportés à une distance.

Dans le cadre de notre simulation, nous utiliserons également les termes suivants :

- une profondeur d'observation en distance D_d de 10 m/s ;
- une taille en azimuth de l'image, quadruple de la largeur du lobe d'antenne à la distance d_0 .

Les paramètres concernant la scène sur laquelle nous allons travailler sont illustrés sur la figure 2.2. Nous allons présenter la signature d'antenne d'une cible ponctuelle située au centre de l'image. La réponse d'antenne rapportée à des distances les délais séparant émission et réception [38].

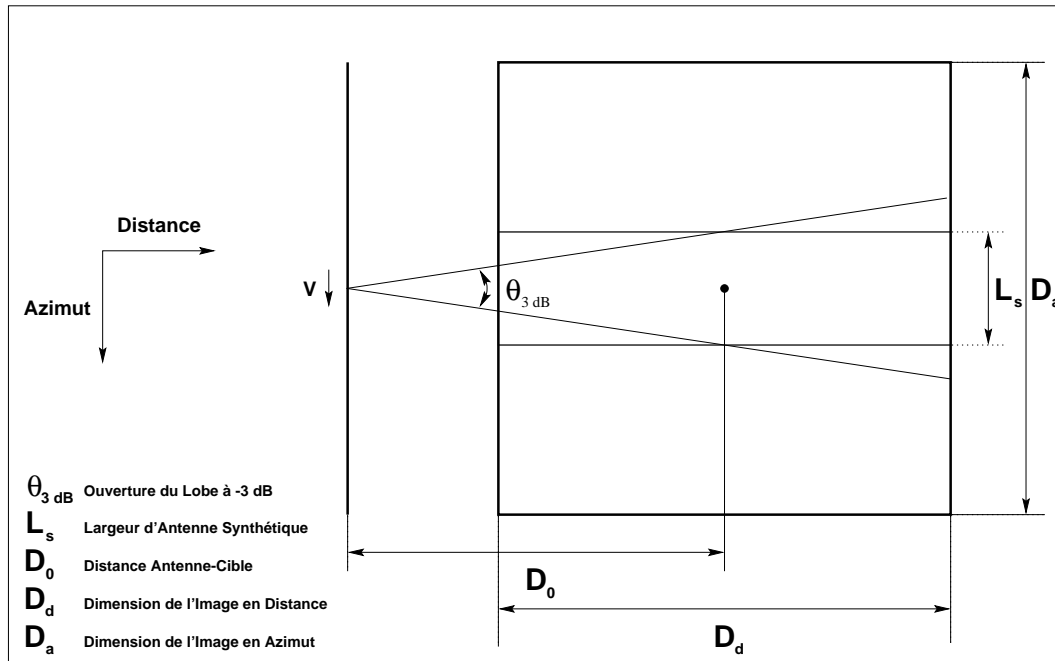


FIG. 2.2 – Disposition des paramètres pour la synthèse d'une signature d'antenne : cette figure illustre les différentes quantités citées dans le document.

De plus, pour une cible ponctuelle, selon le traitement que nous appliquons au signal reçu, la forme peut quelque peu évoluer. La première étape consiste à présenter directement dans la signature d'antenne le signal retour (figure 2.3-b), pour un azimut donné (caractérisant le mouvement de l'antenne). Il est également appliqué en général une descente en bande de base permettant de supprimer la porteuse du signal (figure 2.3-b).

Nous distinguons sur les figures 2.3-b et 2.3-c concernant une cible ponctuelle l'évolution parabolique de la distance la séparant de l'antenne au cours du déplacement de cette dernière.

Signature d'une cible complexe

La signature d'antenne pour une cible complexe se réalise en établissant la somme de la totalité des réponses d'antennes élémentaires de chacun des points constituant la cible. Chaque réponse d'antenne élémentaire est pondérée par :

- la réponse du point ;
- un coefficient d'atténuation égal au carré de la distance séparant le point de la cible à l'antenne ($\propto d^{-2}$).

Nous établissons donc la signature d'antenne pour deux scènes plus complexes, tout d'abord une lettre «F» composée de 17 points (figure 2.3-a) et une mine où apparaissent

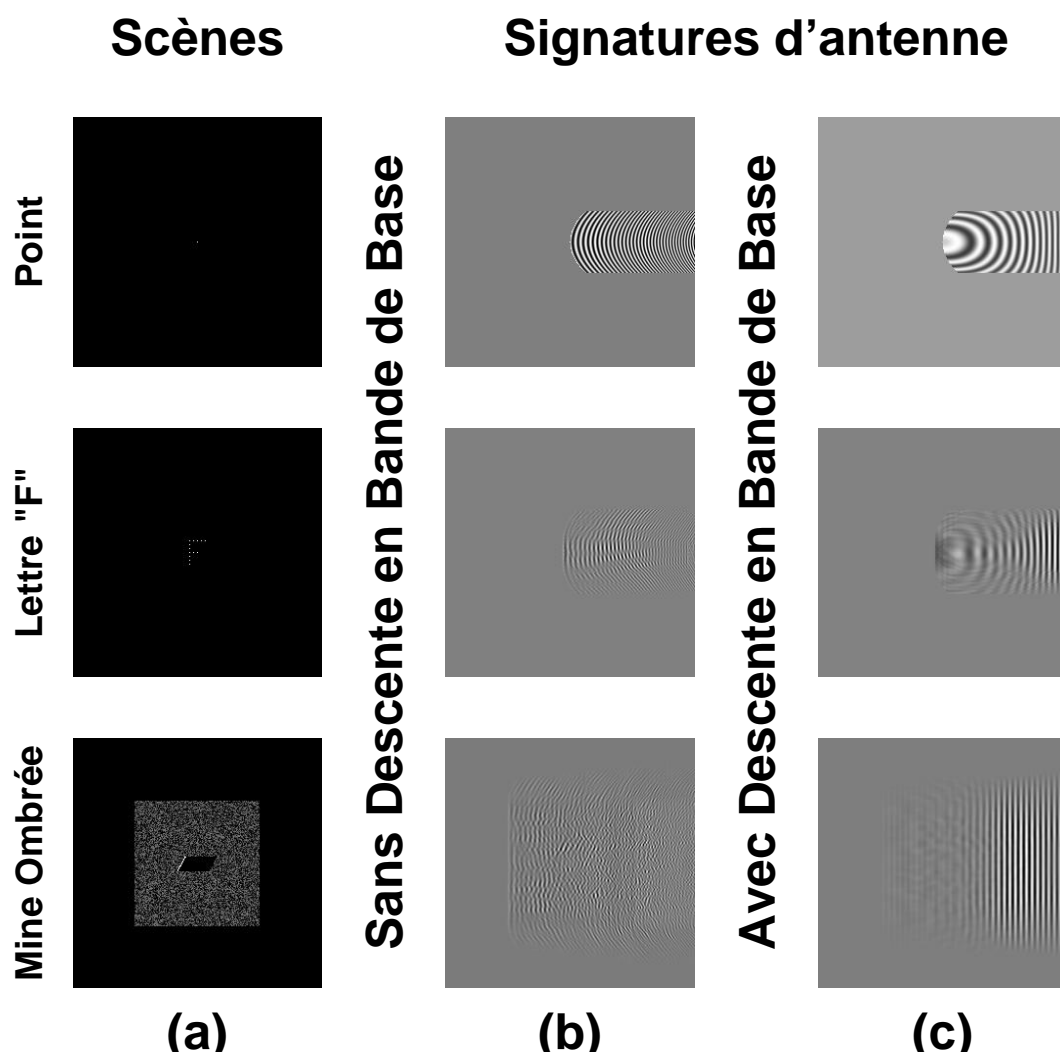


FIG. 2.3 – **Signature d'antenne de différentes cibles** : La signature est créée à partir d'un empilement de tous les échos retrodiffusés par l'ensemble des points constituant la cible. Trois cibles sont considérées : un point, une lettre «F» composée de points et une mine ombrée. La colonne (a) présente les images des cibles correspondantes, les figures (b) et (c) les signatures d'antenne selon qu'une descente en bande de base est appliquée (colonne c) ou non (colonne b).

la réflexion directe de la mine et son ombre 2.3-a sans et avec descente en bande de base (colonnes (b) et (c) de la figure 2.3).

Ces signatures d'antenne dont nous venons de présenter la synthèse sont les précurseurs des signatures que nous allons développer par la suite sur des séquences d'images.

2.5.3 Signature de séquences d'images

La synthèse d'une signature d'antenne synthétique a été l'occasion de nous familiariser avec la structure générique que nous comptons donner dès l'introduction à notre signature ; aussi allons-nous aborder la synthèse de signatures à partir d'une séquence d'images suivant deux axes complémentaires, selon que nous disposons ou non d'une base de références, ou que les données brutes sont ou ne sont pas «auto-suffisantes»¹. Nous disposerons ainsi de deux types de signatures : des «inter-signatures» générées avec la contribution d'une base de références, et des «auto-signatures», synthétisées par elles-mêmes.

Synthèse par filtrage spatial

Si nous reprenons la structure générique de notre signature (figure 2.1), nous remarquons que le codage du signal retour s'insère horizontalement, l'autre direction étant destinée à empiler les récurrences suivantes. Il s'agit donc, suivant cette direction horizontale, de réduire nos images à un unique vecteur.

Nous allons donc utiliser le filtrage pour extraire d'une image de la séquence l'information pertinente, autant spatialement que temporellement : les images disposent d'une répartition «énergétique» bidimensionnelle ; certaines des informations qu'elle contient sont redondantes dans la séquence. Un choix peut et même doit être fait sur leur conservation ou non, selon que nous cherchons à obtenir des signatures «stationnaires»² ou les plus structurées temporellement, donc présentant les motifs les plus caractéristiques.

Toutefois, le filtrage modifie le signal de la séquence sans en changer la structure : une image reste une image. Nous devons ainsi procéder à une transformation réduisant cette image à un vecteur, voire un point.

En effet, le corps général de notre synthèse à partir d'une séquence d'images est un filtrage, il se compose donc nécessairement :

- d'une TF permettant de passer dans l'espace des fréquences ;
- d'une base de filtres permettant de ne sélectionner qu'une information pertinente ;
- d'une base d'ondelettes permettant d'accéder à un niveau de détail donné ;
- d'une TF (ou TF inverse) permettant de revenir dans l'espace direct.

Par suite de ce filtrage, nous disposons d'une transformation permettant de réduire notre image filtrée.

¹A prendre dans le sens : *se suffisant à elle-même*

²stationnaire dans le sens où deux lignes horizontales de la signature sont très proches, quelle que soit leur position sur la signature. La signature ressemble alors plus à un code-barre.

A ce stade, deux méthodes que nous avons qualifiées de complémentaires vont être développées :

- l'une utilisera une base de références comme filtres ;
- l'autre n'utilisera pas de base de filtres, hormis ceux d'ondelettes et disposera une non-linéarité dans le filtrage : le spectre se réduit à son module, puis est élevé à une certaine puissance.

2.5.4 Synthèse de séquences d'images par une base de références

Si nous imposons une base supplémentaire de références, il est nécessaire que la signature prenne en considération, au sein même de sa structure, cette base de référence. Aussi, la structure de notre signature devient-elle celle présentée figure 2.4.

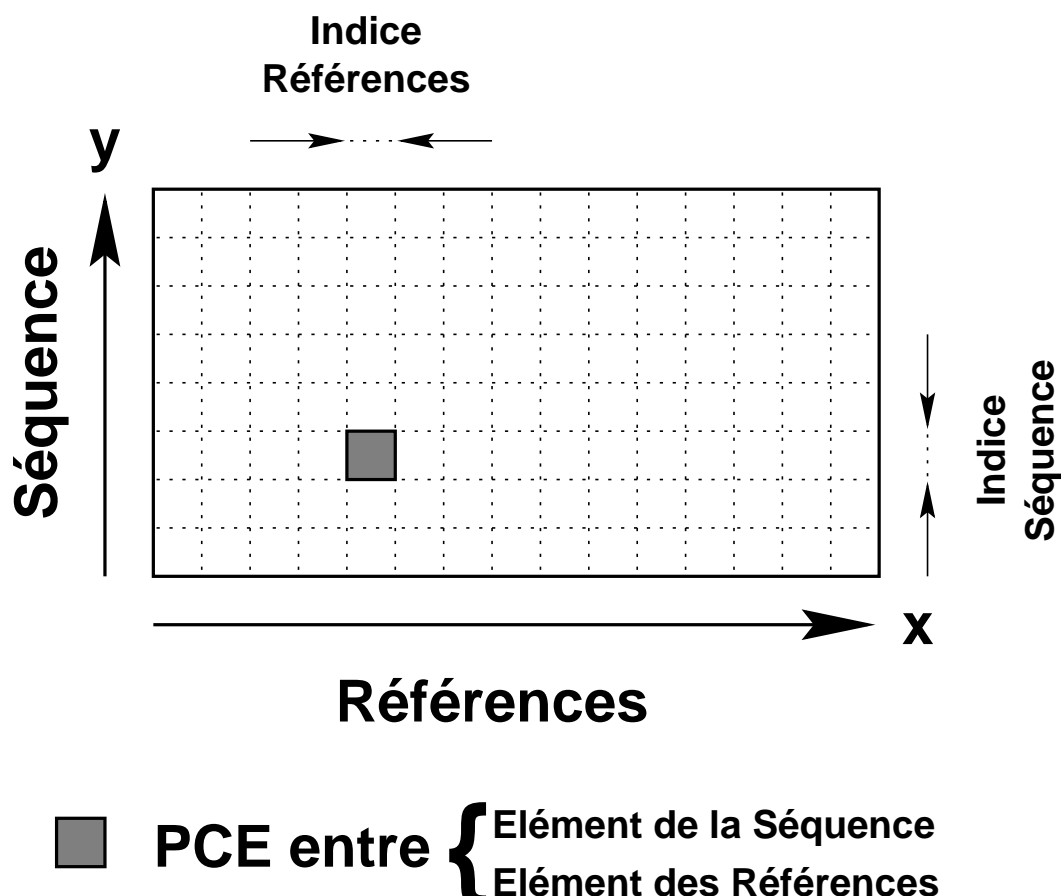


FIG. 2.4 – Premier codage d'une signature à partir d'une séquence : L'évolution de la séquence est présentée suivant l'axe y . Les références sont indicées suivant l'axe x . A chaque point (j, i) de la signature est associée la valeur du *PCE* entre l'élément i de la séquence et la référence j .

Ainsi le vecteur, lequel devait recevoir l'information pertinente issue d'un filtrage, est-il en fait composé d'éléments portant une information scalaire, le *PCE*, issue de la corrélation entre une référence de la base et l'image présentée.

En effet, dans le chapitre précédent, nous avons évoqué la méthode traditionnelle de reconnaissance : une scène, *a priori* inconnue, était comparée à un ensemble de références. La référence correspondant le plus à la scène (par l'estimation du *PCE* dans le plan de corrélation) permettait d'estimer la nature de la scène. Ainsi, une simple extraction du maximum des *PCE* indiquait l'indice de la référence susceptible de ressembler le plus à la scène. Une vision moins superficielle nous indique cependant à quel point cette méthode semble sommaire : en effet, il n'est en aucun cas (puisque le résultat se borne à extraire un seul point) pris en compte les résultats et l'intensité des autres *PCE* estimés. Pourtant, si ceux-ci se rapprochent de la valeur maximale de nos *PCE*, nous parlerons de «fausse alarme». La première approche que nous allons développer tient compte des informations que fournissent les autres corrélations, comme autant d'éléments supplémentaires.

Pour illustrer cette réflexion et juger de l'intérêt que nous avons à comparer une image de la séquence originelle à un ensemble de références de même nature, nous avons présenté une image d'attitude de visage, choisie au hasard, à l'ensemble des dix séquences composant notre base.

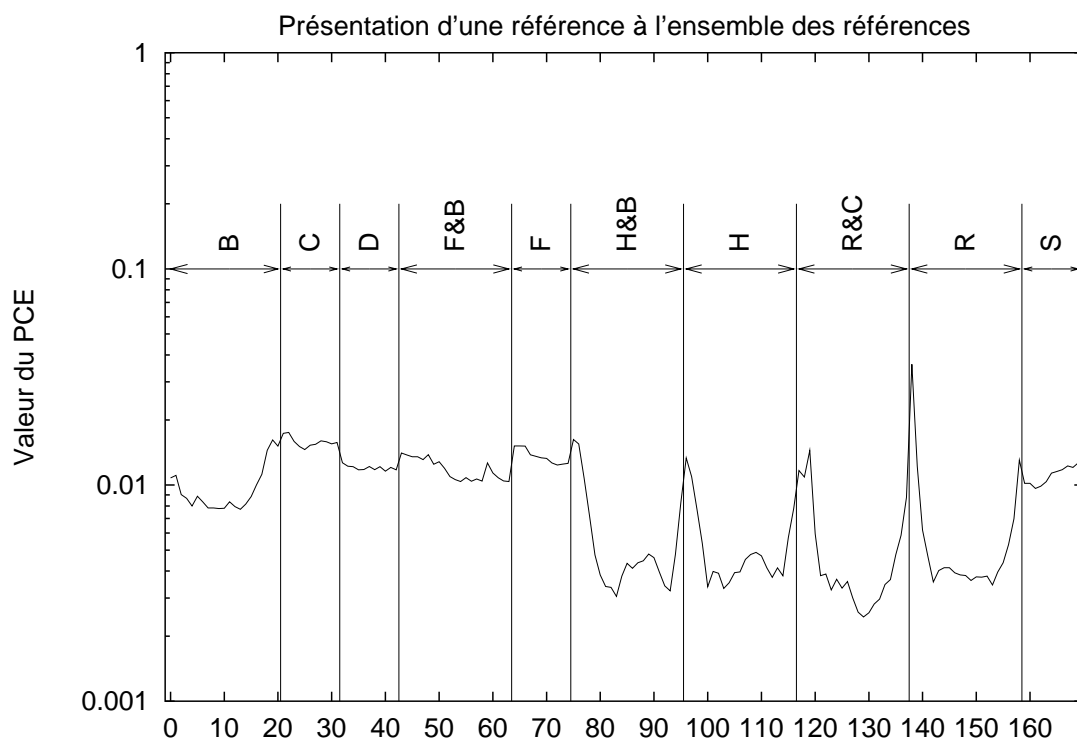


FIG. 2.5 – **Formation d'une ligne de la signature** : La première image d'attitude «Rotation» est comparée à l'ensemble des images des attitudes du visage «Emmanuel». Parmi les attitudes existent le bâillement (**B**), le clignement d'yeux (**C**), le sourire avec les dents visibles (**D**), le froncement avec bâillement (**F&B**), le froncement (**F**), le hochement avec bâillement (**H&B**), la rotation avec clignement d'yeux (**R&C**), la rotation (**R**) et le sourire (**S**).

L'examen de la figure 2.5 nous permet tout d'abord de tirer plusieurs conclusions :

- la référence correspondant à la scène est bien celle disposant du PCE maximal. Nous pourrions, par exemple, conclure que l'image présentée est la première image de la séquence «rotation» ;
- les références au début et à la fin de chaque séquence disposent généralement d'un PCE supérieur aux autres références. L'origine de ce comportement provient de la synthèse même des séquences : chaque attitude commence par un visage pris de face et se termine de façon identique. La scène étant la première image de la séquence de rotation \mathbf{R} , toutes les images proches présenteront un PCE comparable au PCE d'une intercorrélation (à cependant un facteur 3 près) ;

Cette courbe, présentée sur la figure 2.5, illustre ce que nous allons insérer dans notre signature : chaque PCE issu de chaque corrélation formera un pixel de la signature. Ainsi, sur une signature S_{ij} , l'indice j permet-il de revenir à l'indice de la référence présentée.

Dans le même esprit, tout comme nous l'avons proposé pour contrôler les performances de la corrélation, un prétraitement linéaire peut être intégré au filtre, de façon à mettre un contraste plus ou moins grand entre certaines références de la base. Ainsi, si nous reprenons la même base de références, pour des niveaux de contraction a entre 2 et $1/4$, nous formons 4 courbes (figure 2.6).

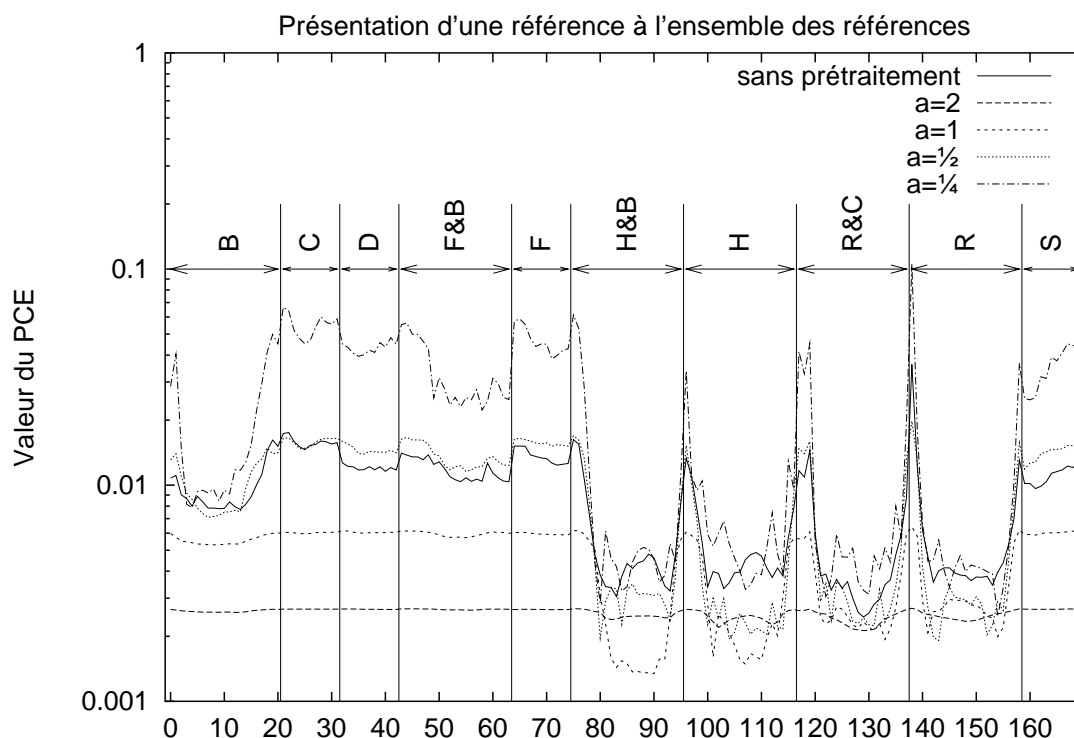


FIG. 2.6 – **Corrélation d'une référence avec un ensemble de séquences** : Un prétraitement linéaire intégré au filtre est appliqué pour quatre valeurs de contraction de a . Il s'en suit une évolution notable du PCE , notamment pour les a faibles. Les ondelettes correspondant au filtre de prétraitement linéaire sont présentées à droite d'une des références.

Plusieurs remarques s'imposent :

- la structure générale de la courbe est conservée, quelque soit le niveau a ;
- les contrastes, sont très sensibles pour des niveaux de dilatation a faibles, indiquant qu'à des fortes résolutions, seules les images de séquences proches de celles présentées donnent des *PCE* importants ;
- à de faibles résolutions (a élevée), le contraste est moins marqué entre les images.

La formation d'une signature élémentaire, pour un élément de la séquence, utilise le prétraitement développé dans le chapitre sur le filtrage. Nous formons ainsi une signature élémentaire pour chaque image de la séquence dont nous réalisons la signature (chaque indice i).

Cette première approche impose de disposer d'une base de références assez significative pour former une signature discriminante, que ce soit pour reconnaître l'attitude ou la forme. Nous pouvons également ajouter que l'introduction d'une base de références dans le traitement de la séquence originelle «enrichit», en terme d'informations, la séquence elle-même : il établit des comparaisons entre des éléments provenant d'attitudes ou de formes différentes.

Une application de la méthode que nous venons d'exposer sera développée dans le prochain chapitre. Les signaux considérés seront des vues de visages multi-capteurs visible et infrarouge.

2.5.5 Synthèse «auto-suffisante» d'une séquence d'images

La synthèse «auto-suffisante» que nous allons développer par un filtrage non-linéaire repose sur la transformation, dans l'espace des fréquences, du résultat du spectre réduite à son module, lequel est élevé à une puissance k . Cette variante n'est pas à proprement parler une approche exclusive de la précédente ; toutefois, l'introduction d'un filtrage visant à opérer une opération de comparaison n'a plus lieu d'être. En effet, nous avons toujours supposé utiliser des filtres de phase. La non-linéarité faisant disparaître la phase du spectre, il est inutile que nous la conservions.

De ce fait, si notre dessein était de former la signature la plus exclusivement caractéristique d'une séquence, nous supposons désormais que les signaux sont suffisants pour réaliser la synthèse d'une signature discriminante.

Par rapport à ce que nous avons présenté sur la figure 2.4, la signature s'en trouve légèrement modifiée, pour se rapprocher de celle que nous formions pour une signature d'antenne synthétique (figure 2.7).

Quel rôle allons-nous confier à ce filtrage non-linéaire ? Plaçons-nous dans le cas où nous disposons de deux sources de séquences entre lesquelles nous allons établir des liens par filtrage.

Pour cela, nous allons tout d'abord supposer que nous formons, dans l'espace direct, avant tout traitement, la jonction de deux images i_1 et i_2 (s'exprimant par l'équation 2.1). La distance séparant les centres des deux images est d

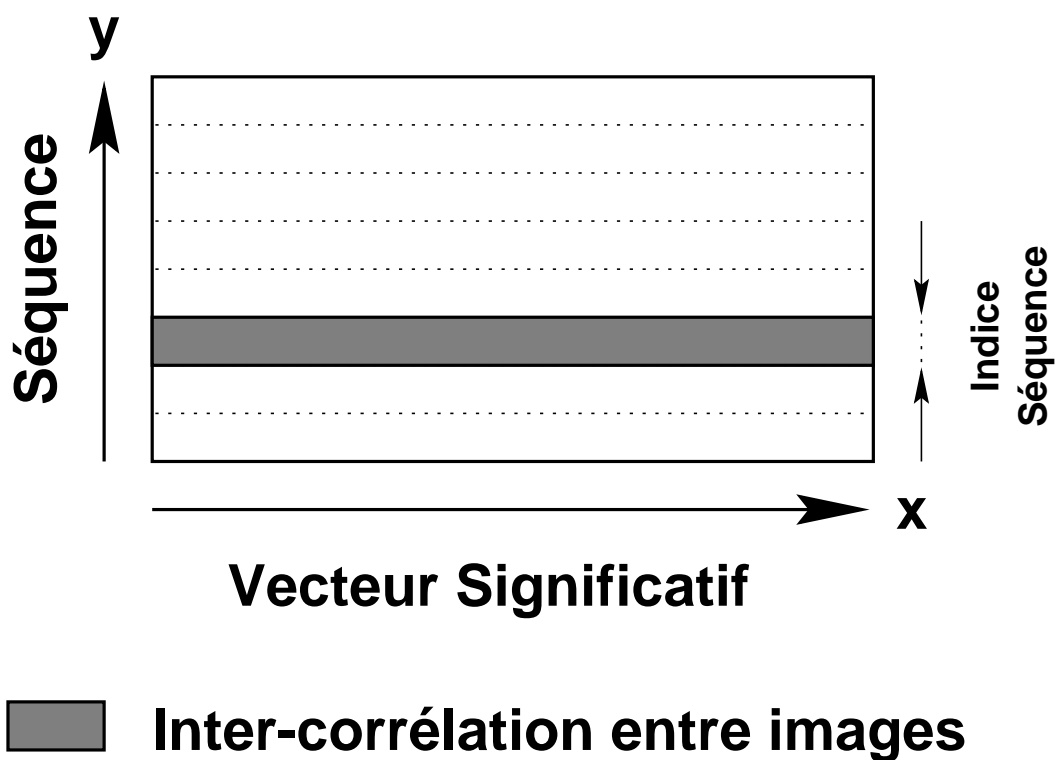


FIG. 2.7 – Second codage d'une signature à partir d'une séquence : L'évolution de la séquence est présentée suivant l'axe y . A chaque ligne j de la signature est associé un vecteur, résultat d'un filtrage puis d'une extraction sur l'élément i de la séquence.

$$\begin{aligned}
s(x, y) &= i_1 \left(x - \frac{d}{2}, y \right) + i_2 \left(x + \frac{d}{2}, y \right) \\
s(x, y) &= i_1 * \delta \left(x - \frac{d}{2}, y \right) + i_2 * \delta \left(x + \frac{d}{2}, y \right)
\end{aligned}$$

Si nous calculons la TF de l'expression 2.1, nous obtenons :

$$S(X, Y) = I_1(X, Y)e^{i\pi dX} + I_2(X, Y)e^{-i\pi dX} \quad (2.1)$$

Le calcul du carré du module de 2.1 établit donc des termes croisés entre les spectres X_1 et X_2 :

$$\begin{aligned}
|S(X, Y)|^2 &= (I_1e^{i\pi dX} + I_2e^{-i\pi dX})(I_1e^{i\pi dX} + I_2e^{-i\pi dX})^* \\
|S(X, Y)|^2 &= I_1^2 + I_2^2 + I_1I_2^*e^{2i\pi dX} + I_1^*I_2e^{-2i\pi dX}
\end{aligned}$$

Le passage dans l'espace direct fait alors naturellement apparaître la corrélation (au sens filtrage adapté) entre les images i_1 et i_2 :

$$\mathcal{F}_{|S|^2}(x, y) = (i_1 \otimes i_1) + (i_2 \otimes i_2) + (i_1 \otimes i_2) * \delta(x - d, y) + (i_2 \otimes i_1) * \delta(x + d, y) \quad (2.2)$$

Ainsi apparaissent les filtrages adaptés de i_1 par i_2 puis de i_2 par i_1 , déplacés par rapport au centre d'une distance double de celle qui séparait les deux images initialement, c'est-à-dire $2d$.

Toutefois, au traitement que nous venons d'évoquer, pour nous rapprocher du traitement générique présenté auparavant, il convient d'ajouter deux étapes, une en amont de la non-linéarité, une en aval :

- le produit du spectre joint par un filtre de prétraitement linéaire : son objectif est de permettre d'établir une sélection des fréquences spatiales caractéristiques de certains détails ;
- l'élévation à une puissance k du carré du module du spectre joint : cette non-linéarité permet de réaliser une opération proche de la puissance γ appliquée au dénominateur du filtre (égal à $1/2$ dans le cas d'un filtre de phase). De ce fait, elle permet le choix d'un compromis ajustable entre sélectivité et robustesse [18].

Ces deux opérations complémentaires permettront de se libérer du filtre adapté, trop peu sélectif.

Jusqu'à présent, nous n'avons évoqué, dans le traitement non-linéaire, que les changements d'espace de représentation, les opérations de filtrage linéaire (prétraitement intégré au filtre) ou non-linéaire (estimation du carré du module du spectre joint). Reste maintenant à développer le passage du plan d'inter-corrélation au vecteur à intégrer dans la signature 2.7.

Ce passage d'une image à un vecteur, tel que nous venons de l'évoquer, mérite que nous nous penchions sur ce que nous voulons extraire de l'image. Lorsque ce vecteur était formé d'un ensemble de PCE établissant la comparaison entre une image et une base de références, la transformation paraissait évidente : le PCE réduit le plan de corrélation

à un point unique, donc établit un passage direct de 2 dimensions à un scalaire. Nous devons passer de 2 dimensions vers une unique dimension.

Une technique classique consiste à parcourir l'image par un chemin particulier. Sa principale propriété doit être de conserver la proximité d'un point à un autre de l'image lors de la formation du vecteur. Le parcours de Hilbert B.1 semble satisfaire cette propriété. De plus, fractale auto-similaire, elle se construit à partir d'un processus récursif B.2. L'application d'un tel parcours sur une image pour en extraire un vecteur pose cependant le problème de la taille du vecteur : si une image dispose d'une dimension de 256×256 pixels, le parcours résultat sera de 65536 points. Pour réduire ce nombre de points, nous pouvons regrouper les points de ce vecteur par puissance de 4, donc 4^n . De par le parcours, tout se passe comme si l'image se segmentait en pavés de pixels de 2×2 , 4×4 , 8×8 , 16×16 , etc...

Regardons par exemple sur une image de 64×64 pixels (donc de 16384 points) le parcours fractal associé (figure 2.8), pour des regroupements de pixels par blocs de 4×4 , 8×8 , 16×16 pixels.

Toutefois, le problème qui survient par application d'une telle opération réside dans sa très faible robustesse à la translation de l'image. La figure 2.9 illustre cet élément. Si, par exemple, nous comparons par corrélation deux vecteurs issus d'images translatées l'une par rapport à l'autre, nous n'aurons qu'un pic de corrélation très émoissé...

Cette propriété du parcours de Hilbert, appliquée à une image, justifie la raison pour laquelle il n'est pas attractif au point de l'utiliser directement sur les images. En effet, pour que ce parcours ait un intérêt, il est indispensable que l'image ait une symétrie centrale. Cette symétrie peut naturellement être créée par application d'un filtrage non-linéaire, réduisant le spectre à son module (donc entraînant une perte de la phase). Bien que l'action de réduire le spectre à son module (ou à son module au carré) permette de «symétriser» l'image pour y effectuer la transformation $2D \rightarrow 1D$, elle doit être accompagnée. En effet, les images symétrisées – qui s'apparentent dans le cas d'une non-linéarité de type «module au carré» à une auto-corrélation – sont peu discriminantes pour des images au spectre centré, les visages par exemple. Le produit dans l'espace de Fourier par un filtre ne laissant passer que certaines fréquences spatiales – et en aucun cas les basses – semble adapté à cette tâche. L'utilisation des ondelettes constitue ici une solution des plus intéressantes : elle élimine systématiquement les basses fréquences, lesquelles sont communes dans de nombreux cas, et, pour des classes différentes, pour ne conserver que l'information la plus pertinente lors du filtrage.

L'emploi d'un filtrage par une ondelette avant l'introduction de la non-linéarité dans l'espace des fréquences renforce donc le choix de parcours transformant le résultat d'un filtrage non-linéaire en un vecteur. Reste toutefois à déterminer la nature des images sur lesquelles seront appliqués ces filtres non-linéaires. Lorsque nous avons proposé d'utiliser pareil traitement, nous nous plaçons en amont des images «assemblées». Cet assemblage sera développé, dans le chapitre des «Applications», suivant deux aspects complémentaires, en fonction de la nature des données que nous aurons à analyser :

- assemblage de données issues de différents capteurs synchrones :

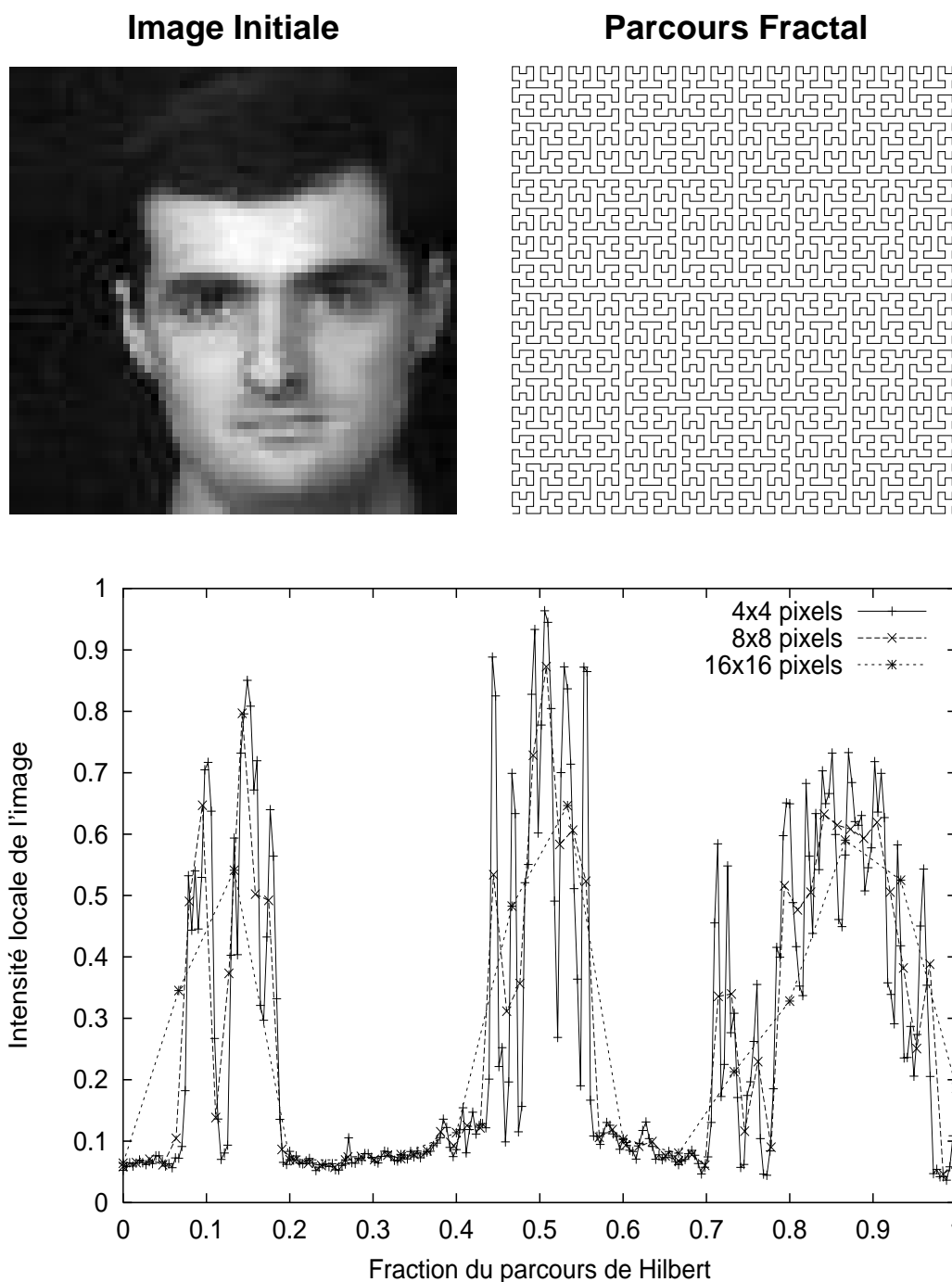


FIG. 2.8 – Exemple de transformation $2D \rightarrow 1D$ par un parcours de Hilbert : En haut figurent l'image originelle et le parcours de Hilbert associé (partant d'en bas à gauche). Dessous, les courbes représentent l'extraction des intensités des pixels de l'image suivant un parcours de Hilbert. L'image de 64×64 pixels est «sondée» par un parcours de 16384 points. Les points sont alors regroupés par blocs de 16, 64 et 256 pixels.

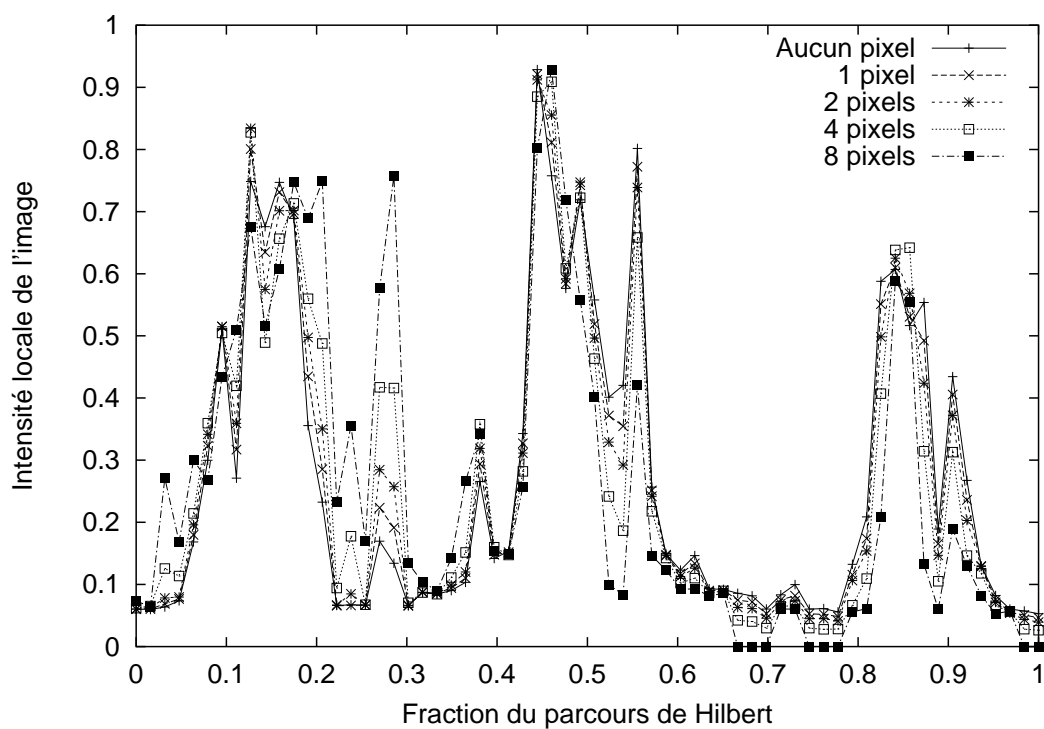


FIG. 2.9 – Influence d'un décalage sur la transformation 2D→1D : Les cinq courbes illustrent l'évolution de la courbe extraite d'un parcours fractal sur une image. L'image originelle est celle du visage présenté sur la figure 2.8. Chaque courbe comprend 256 points : chacun des points représente donc un pavé de 8×8 pixels.

les capteurs disposent de la même fréquence de rafraîchissement. Le «formatage» des images issues de ces capteurs est assuré par un recalage temporel. Ainsi, les images acquises simultanément par deux capteurs x et y formeront en amont une image $\{x_i, y_i\}$, où la variable i représente un indice désignant deux acquisitions séparées d'un instant Δt . Ainsi, $\{x_{i+1}, y_{i+1}\}(t) = \{x_i, y_i\}(t + \Delta t)$. Le principal avantage de cet assemblage réside dans la capacité du filtrage à établir les liens entre les images.

- assemblage de données issues d'un unique capteur, à des instants différents : le capteur x dispose d'un rafraîchissement de fréquence $1/\Delta t$. Nous avons ainsi $x_j(t) = x_i(t + n\Delta t)$ où n est entier. Le résultat de la jonction des deux images issues d'instant différents possède la forme $\{x_i, x_j\}$. Un tel assemblage trouve son intérêt lorsqu'il est nécessaire de renforcer les liens temporels. Dans le cas précédent, la signature établissait les liens, à un instant donné, entre deux capteurs aux mêmes instants, ici les liens sont établis entre les acquisitions d'un même capteur à des instants différents. De pareils liens entre images d'une même séquence peut se généraliser par un assemblage avec des données issues de séquences très antérieures, voire disposées dans la base de références. Dans ce cas, le lien serait «forcé» par rapports à des données existantes.

Ces deux «assemblages» forment donc les deux exemples que nous développerons dans le chapitre destiné à présenter les applications. Une généralisation est cependant possible : nous établissons ainsi la jonction de signaux issus de différents capteurs à des instants différents ; par exemple, une séquence d'un capteur liée avec une séquence d'un autre capteur déphasé.

2.6 Analyse de Signatures

La synthèse de signature étant terminée, notre tâche se tourne désormais vers son analyse. Comme nous l'avons très sommairement évoqué dans l'introduction, nous ne disposons, par hypothèse (ou plus pragmatiquement, par un choix issu des contraintes technologiques auxquelles nous étions confrontés), que du filtrage pour mener à son terme le processus d'identification de la forme ou de reconnaissance de l'attitude. Notre signature était un intermédiaire entre les données brutes et la décision finale.

Pour développer une telle analyse, nous allons utiliser nombre de résultats développés dans ce chapitre, notamment – et surtout – concernant les propositions que nous apportons au filtrage comme outil de reconnaissance de formes. En effet, la corrélation va être l'outil principal de décision dans cette seconde phase de traitement. Nous avons développé la corrélation comme un opérateur de comparaison. C'est effectivement dans ce rôle-là que nous allons généralement l'utiliser, qu'il s'agisse de signatures formées à partir de signaux 1D ou 2D en amont du traitement.

2.6.1 Signaux d'antenne synthétique

La formation d'une signature à partir de signaux d'antenne synthétique illustre son caractère bi-temporel : un temps de récurrence séparant deux acquisitions, donc séparant deux instants formant une ligne et un temps d'acquisition désignant la durée durant laquelle les signaux issus du capteurs sont numérisés.

Une nouvelle fois, cet exemple va nous servir à illustrer le problème de l'analyse de signatures.

Quels objectifs allons-nous poursuivre pour l'analyse d'une telle signature d'antenne ?

Nous nous sommes intéressés aux applications d'un filtrage comme outil de reconnaissance de formes. Nous allons en effet chercher, par l'analyse de signatures, à atteindre différents objectifs :

- la restitution de la forme ;
- l'identification de la forme ;
- la restitution de la forme par l'identification de ses composantes élémentaires.

Restituer la scène : mode «Imageur»

La première utilisation consiste, étant donné que chaque point forme une réponse d'antenne élémentaire, à restituer la scène observée. La formation de la réponse d'antenne d'une cible est linéaire, pour une distance d'un point de la cible donnée. Corréler la réponse d'antenne de la scène avec la réponse d'antenne d'une cible ponctuelle suffira à restituer la scène. Toutefois, un petit détail mérite d'être souligné : la réponse d'antenne est identique quel que soit l'azimut (la direction suivant laquelle se déplace l'antenne), mais pas suivant la distance (le lobe de l'antenne se dilate lorsque la distance augmente). Il faudrait donc théoriquement procéder à la comparaison de la totalité des réponses d'antenne de cible ponctuelle, quelle que soit la distance, pour restituer la scène. Heureusement, l'évolution en distance permet de s'affranchir de cette difficulté.

Le mode «imageur» va exploiter la ressemblance existant entre l'opération de corrélation utilisée et la déconvolution. La formation d'une signature d'antenne, pour une distance donnée, est un processus linéaire ; il est donc intéressant d'exploiter le filtrage, outil linéaire, pour assurer la restitution de la scène observée.

Dans cette optique, nous allons corréler la signature d'antenne de notre cible avec la signature d'antenne d'une cible ponctuelle. Le filtrage de phase, que nous utilisons, demeure toujours un intéressant compromis entre robustesse et discrimination.

Les cibles à reconstruire sont identiques à celles dont les signatures d'antenne sont préalablement établies : un point, une lettre «F» pixellisée et une mine ombrée. Ce dernier exemple est certainement le plus significatif puisqu'il va permettre de sonder la capacité de notre filtrage : opérer une reconnaissance utilisant l'absence de signal (l'ombre de la mine) et l'existence d'un autre signal (la rétrodiffusion directe de la mine) associées au même objet, lesquelles forment l'information la plus pertinente.

Les résultats présentés sur la figure 2.10 illustrent l'action de la corrélation entre les signatures d'antenne correspondant aux trois différentes cibles et la signature d'antenne d'une cible ponctuelle. Nous constatons que la forme générale de la cible est restaurée :

- la dimension du pic de corrélation indique, dans le cas de la restauration d'une cible ponctuelle, que les résolutions théoriques de compression d'impulsion sont bien

- atteintes ;
- la lettre «F» est restituée avec plus ou moins de finesse. Des pics secondaires, issus de la restitution des pics primaires correspondant eux-mêmes aux points originels de la lettre, s'additionnent, donnant non plus 5 points mais 9 points. Si une descente en bande de base a été faite lors de la synthèse de la signature, les points qui apparaissent de façon plus significative sont ceux de la barre verticale. Notons également que, dans ce cas, les lobes secondaires «bavent» suivant la direction de l'azimut ;
 - la mine ombrée est assez bien restituée. Un γ^3 de 3 sur l'image permet de s'assurer qu'il existe encore un large contraste entre l'ombre de la mine et le bruit l'environnant. Si toutefois les signaux ont fait l'objet d'une descente en bande de base, seule la moitié de la scène est restituée, de manière très sommaire, les modulations perturbant l'image reconstituée.

Les trois restitutions présentées sur la figure 2.10 l'ont été avec l'aide d'une seule et unique signature servant de référence. Il est alors intéressant de noter que si nous avons opté pour un filtre inverse – réalisant ainsi une déconvolution – la restitution aurait été plus sommaire : seuls quelques points auraient été extraits de la signature d'antenne.

Toutefois, la descente en bande de base, bien que généralement pratiquée dans l'analyse de signaux en télédétection, ne s'avère pas indispensable à la vue des quelques restitutions. L'origine d'une si mauvaise adaptation est à chercher, une fois encore, dans le spectre de la signature d'antenne. La conservation d'une porteuse génère nécessairement de hautes fréquences, lesquelles décentrent la modulation. Si nous ne conservons que la modulation, nous nous retrouvons dans une situation comparable à celle rencontrée pour des images à spectres centrés, les visages par exemple. C'est donc pour cette raison que nous n'appliquerons plus par la suite dans la synthèse de la réponse d'antenne la descente en bande de base.

Reconnaître un objet : mode «Pointeur»

Nous avons, dans le paragraphe précédent, cherché à restituer la scène par une corrélation entre signature d'antenne de la scène et signature d'antenne d'une cible ponctuelle. L'objectif recherché peut être désormais, non plus de restituer, mais de reconnaître si un objet se trouve dans la scène, et pointer l'endroit où il se situe.

La corrélation peut jouer les deux rôles de comparateur et de pointeur d'une référence présente dans une scène.

Nous avons donc construit, suivant le traitement décrit ci-dessus, les signatures d'antenne de différents objets déplacés par rapport au centre de l'image (un point, une lettre «F» pixellisée et une mine ombrée), puis les avons comparées, par corrélation, avec les signatures centrées présentées figure 2.10. La figure 2.11 restitue le pointage des différents objets dans la scène. Malgré la distorsion de la signature, due à une distance plus éloignée de l'antenne, la corrélation à partir d'un filtre de phase a été en mesure de pointer l'objet.

³La fonction γ ici est égale à : $V_{max} \exp [\log_{10} (V/V_{max}) / \gamma]$

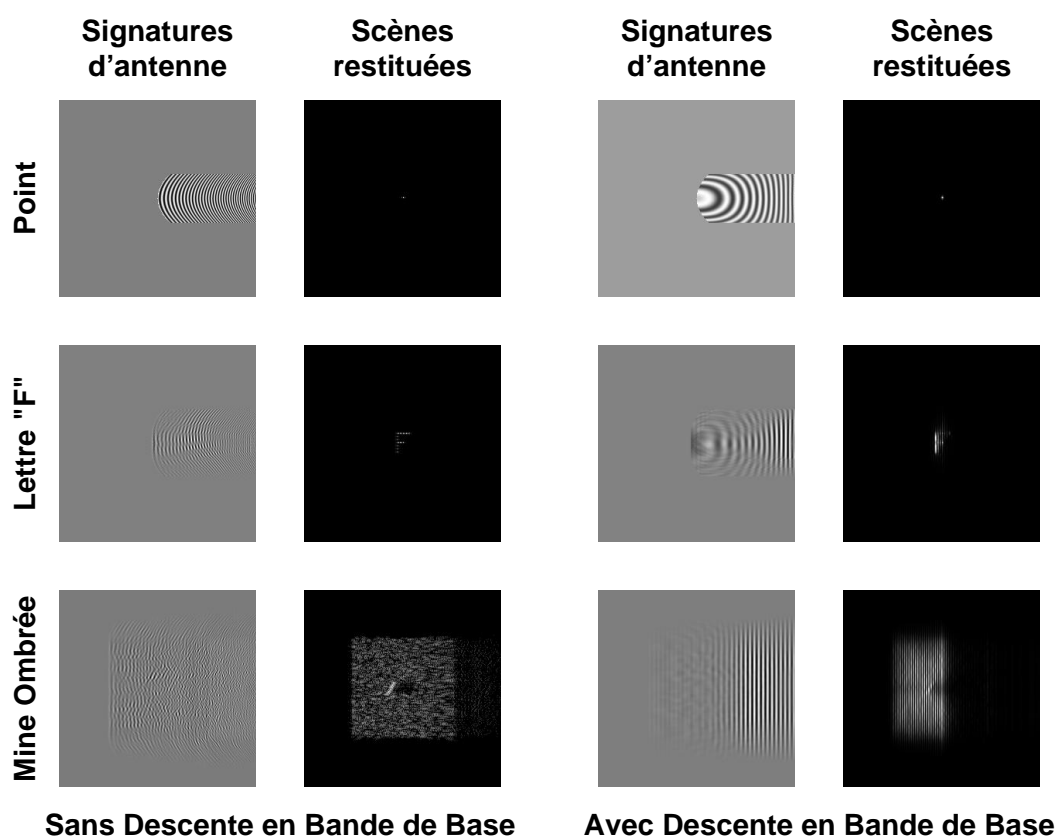


FIG. 2.10 – Restitution d'une cible par analyse de sa signature d'antenne : les signatures d'antenne sont corrélées avec la réponse d'une cible ponctuelle. Chaque cible est reconstruite plus ou moins fidèlement : le point est quasi-parfait, la lettre «F» lisible et la mine ombrée discernable.

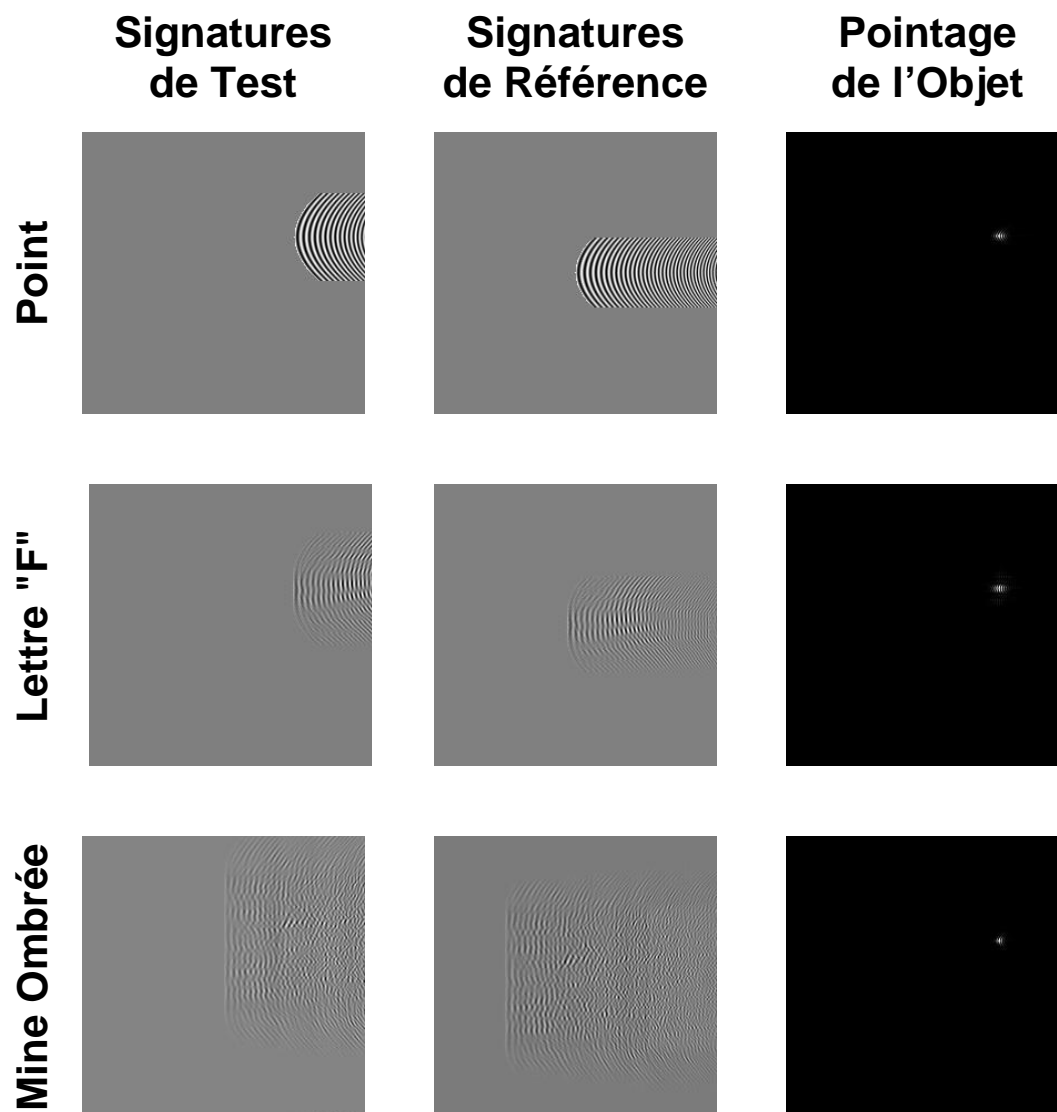


FIG. 2.11 – **Pointage d'une cible par analyse de sa signature d'antenne** : les signatures d'antenne sont corrélées avec des signatures d'antenne de référence. Un point indique la position de la cible dans le plan de corrélation.

Restituer la scène par ses objets : mode «Multi-pointeur»

Aux deux méthodes exposées s'ajoute une troisième, hybridation entre les précédentes : la scène peut être restituée, non à partir d'une signature d'antenne de cible ponctuelle, mais à partir de la réponse de cibles élémentaires rétrodiffusants le signal d'antenne. La scène n'est pas restituée mais reconstituée à partir des positions données par un traitement suivant le mode imageur.

2.6.2 Signatures de séquences d'images

La formation de signatures issues de signaux bidimensionnels, des images par exemple, suit le même protocole de synthèse que celui utilisé pour la formation d'une réponse d'antenne synthétique : l'évolution temporelle occupe un axe de la signature, l'autre étant destiné au codage d'une information, issue soit de la comparaison entre chaque image de la séquence initiale et une base de références, soit de l'image elle-même par une auto-corrélation au moyen d'un traitement non-linéaire.

La principale utilisation que nous allons faire de ces signatures consiste à les comparer avec une base de signatures que l'on suppose exhaustive, de manière à reconnaître la forme par son mouvement.

Dans cette optique une distinction est à opérer quant à l'origine des données ayant permis la formation des signatures à utiliser :

- le premier aspect cherchera à reconnaître l'attitude de la forme, pour ne conserver comme base de défilement de références que celles correspondant à cette attitude : une structure hiérarchique sera alors mise en place pour réduire à chaque étape le cardinal de références à présenter. Pour se faire, chaque ligne devra être la plus différente possible des autres, pour conclure sur l'attitude qu'il nous faut désormais considérer ;
- le deuxième aspect concerne la reconnaissance d'une forme soumise à de très fortes variations stochastiques, non pas à cause de mouvements intrinsèques (lesquels formeraient une attitude), mais par le canal de transmission. La signature devra, dans ce cas, rechercher à être la plus stable possible temporellement : elle ressemblera à un code barre. Cette propriété nécessitera l'implantation d'une fonction de convergence (le pic de corrélation étant nécessairement très étalé suivant la direction temporelle d'évolution).

2.7 Architecture générique de traitement

Nous avons, dans ce chapitre, développé la synthèse de signatures à partir de signaux d'antenne synthétique (de nature unidimensionnelle à l'origine), la synthèse de signatures à partir de séquences d'images. Une fois ces signatures établies, nous nous sommes penchés sur leur analyse, à partir d'une opération simple, le filtrage spatial [7].

Pour la synthèse d'une signature, en ne considérant que des signaux bidimensionnels, nous avons développé deux approches :

- la première utilisait une base de références. La formation de la signature se composait donc de pixels, où chaque élément présentait l'adéquation entre élément de la séquence et référence ;
- la seconde utilisait pour former chaque ligne un parcours fractal sur le résultat d'une inter-corrélation (par filtrage non-linéaire), dont nous avons forcé les liens par une association entre des images. Ces images provenaient de capteurs différents synchrones ou d'un même capteur, à des instants différents.

Ces deux approches ne sont pas, *a priori* exclusives : il est en effet possible de former des signatures dont l'extraction se réalise par un parcours fractal à partir d'un filtrage non-linéaire ayant utilisé une base de données. Toutefois, avec des filtres de phase correspondant à des références, la base de références est inutile. En effet, la non-linéarité supprime la phase du signal filtré, donc l'effet des filtres de phase. Aussi séparerons-nous ces deux approches par la suite.

Quant à l'opération d'analyse, nous avons, dans les cas évoqués ci-dessus, choisi de n'utiliser que l'outil-filtrage pour établir la reconnaissance. La décision finale se trouve donc sous-tendue à une analyse combinée de chacune des décisions élémentaires issues des *PCE* à chaque résolution.

Le schéma 2.12 illustre l'architecture générique pour la synthèse puis l'analyse de la signature. Il se compose donc de trois couches successives, que sont :

- la couche d'observations et de mesures ;
- la couche de traitement ;
- la couche de données et de connaissance.

Il se compose également (de gauche à droite) d'une association de deux blocs :

- un premier filtrage réalisant le passage entre les données brutes et la signature :

la synthèse

 ;
- un second filtrage permettant l'extraction d'une décision élémentaire à partir de la signature :

l'analyse

.

Ces deux blocs comprennent les mêmes opérations (à l'introduction de la non-linéarité près). Leur ressemblance permet, par exemple, d'envisager l'implantation d'un seul et unique processeur pour réaliser les deux opérations.

La formation intervenant de manière «naturelle» dans le cas d'une synthèse de signature d'antenne synthétique, seule la partie «analyse» du schéma subsiste.

La lecture du schéma 2.12 de haut en bas sera pour nous l'occasion de revenir sur les différentes couches constituant le traitement générique conduisant à la décision.

2.7.1 L'architecture générique, de haut en bas

La couche d'observation et de mesures

Cette couche comporte trois éléments. Les images issues des différents capteurs nécessitent une mise en forme de façon à former les données sur lesquelles le filtrage va intervenir. Elle consiste surtout en une agrégation de données, issues de plusieurs

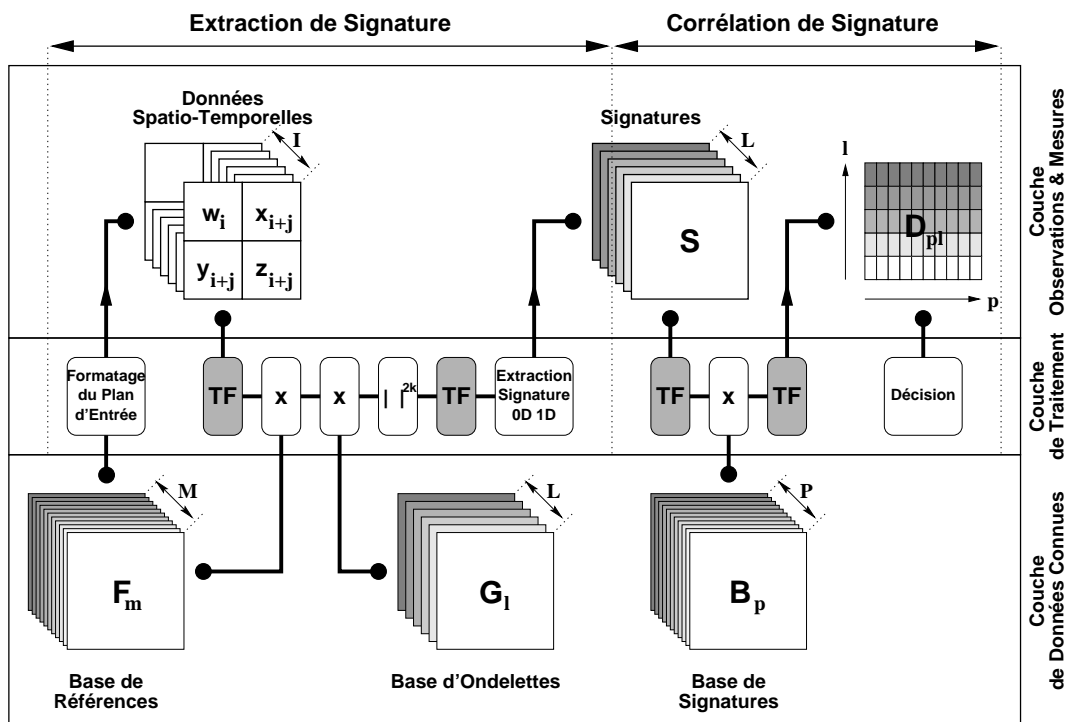


FIG. 2.12 – Schéma général de synthèse puis d'analyse par filtrage spatial : ce schéma général présente les différentes couches (mesures, traitement, bases de données) nécessaires à la synthèse puis l'analyse de la signature. A noter le filtrage illustré par les deux TF présents dans chaque phase de traitement.

capteurs synchrones ou asynchrones, ou d'un unique capteur à des instants différents.

Ainsi, sur la figure 2.12, lorsque les capteurs x , y , z sont rafraîchis I fois, le capteur w ne l'est qu'une seule fois. Les capteurs x , y , z sont synchrones, mais asynchrones avec le capteur w .

Sur ces données est donc appliqué un premier filtrage. Ce dernier permet la formation d'une signature, il s'agit donc de l'opération de synthèse.

L'analyse de la signature se développe ensuite par une corrélation qui forme une matrice de décision. C'est sur cette matrice de décision que se tranche la décision finale : la reconnaissance d'une attitude ou l'identification d'une forme.

La couche de traitement

La couche de traitement se scinde en deux groupes d'opérations : un premier réalise la synthèse et l'autre l'analyse :

- la synthèse se réalise par un passage dans l'espace des fréquences (TF). Deux bases (évoquées plus bas) sont utilisées : la première permet d'établir la comparaison entre plusieurs éléments et la seconde permet d'accéder à une nouvelle dimension, la résolution. Le filtrage utilise ces deux bases lors de deux produits dans l'espace des fréquences. Ensuite, suivant le choix d'un traitement linéaire (synthèse d'une signature par extraction de PCE) ou non-linéaire (synthèse d'une signature par extraction du plan d'inter-corrélation), l'élévation à une puissance $2k$ du module du spectre résultant est activée. Une seconde TF permet de revenir dans l'espace direct. Vient ensuite l'extraction, un simple PCE ou un parcours fractal du plan d'inter-corrélation selon qu'il s'agisse d'un traitement linéaire ou non-linéaire ;
- l'analyse se réalise par la comparaison entre les signatures obtenues et une base de signatures préalablement calculées. Cette comparaison est une corrélation réalisée dans l'espace des fréquences (d'où la présence de deux TF et d'un produit) ;
- la décision se réalise enfin par une analyse des différents résultats issus de la matrice de décisions élémentaires. Un simple polynôme semble faire l'affaire pour parvenir à un critère global de décision.

La couche de bases de données

Ces données suivent également la structure du traitement, et donc la structure des données intermédiaires que sont les images d'entrée ou la signature. Elles se composent de trois éléments distincts :

- une base de références ou de filtres permettant d'établir la comparaison directe entre des éléments de même nature. Il s'agira donc de visages lorsque des attitudes de visages seront à analyser, des représentations temps-fréquence lorsque des images temps-fréquence seront à analyser, etc. . . ;
- une base d'ondelettes, agissant dans le cadre du prétraitement linéaire. Elle permet de comparer la séquence aux références en introduisant la notion de résolution. Les comparaisons ainsi obtenues le sont à un certain niveau de résolution ;

- une base de signatures, disposant d'un ensemble d'attitudes et de formes. Elle permet, par exemple, à faible résolution, de reconnaître l'attitude d'un visage et à haute résolution, de conclure sur l'identité du visage.

La direction que nous avons choisie pour détailler ce traitement générique n'est pas innocente ; la «volatilité» des couches décroît suivant cette direction. En effet, la couche de mesures est perpétuellement rafraîchie. La couche de traitement peut être optimisée sans à aucun moment toucher à la couche de bases de données. Une fois cet ensemble de bases établi avec la meilleure finesse possible, nous n'y touchons plus. Elle demeure la mémoire de notre traitement et donc son principal atout.

De fait, la formation des bases, à l'origine, dépend nécessairement des traitements dans lesquels elles sont sollicitées. Elles peuvent également s'enrichir de résultats antérieurs. En ce sens, son cardinal ne fait que croître.

2.8 Conclusion

Ce chapitre a été l'occasion de développer la synthèse puis l'analyse de signatures temporelles. Nous avons, quelle que soit la nature des signaux, utilisé le filtrage spatial comme outil de reconnaissance.

Ce chapitre nous a permis de généraliser, par analogie, le principe de formation «naturelle» d'une signature temporelle à partir de signaux d'antenne synthétique. Nous avons suivi le canevas dressé par cette première application pour parvenir à deux approches suivantes : l'approche «PCE» et l'approche «parcours fractal».

De plus, les procédés de synthèse et d'analyse de signatures utilisent la même structure générique : un filtrage, donc deux transformations de Fourier et un produit dans l'espace de Fourier. L'introduction du filtrage non-linéaire dans une des deux synthèses possibles de signatures n'est toutefois pas à omettre. Toujours est-il que cette simplicité nous permet, d'envisager une seule et unique unité de traitement, qu'il s'agisse de la synthèse ou de l'analyse.

Ces méthodes décrites, leurs différences développées, nous aborderons dans le prochain chapitre les différentes formations possibles de signatures puis leurs utilisations pour des signaux classiques, des séquences d'images. Occasion nous sera ainsi donnée de comparer les deux variantes sur des cas «réels».

Chapitre 3

Applications

Le chapitre précédent nous a permis de nous familiariser avec la synthèse puis l'analyse de différents types de signaux, de nature unidimensionnelle ou bidimensionnelle. Nous avons également délimité quelques-uns des objectifs qu'une telle approche de synthèse puis d'analyse permettait de poursuivre. L'heure est maintenant aux applications.

- Pour cela, nous nous intéresserons successivement aux traitements possibles affectant :
- les attitudes de visages en visible et infrarouge (régime multi-senseurs) ;
 - l'observation temps-fréquence de bâtiments de surface (régime multi-référence).

3.1 Signaux Visible & Infrarouge couplés

Nous avons déjà utilisé l'outil «filtrage» pour extraire d'une signature d'antenne synthétique des informations de plusieurs natures, restituant successivement la scène ou pointant certains objets se trouvant dans cette scène. A l'origine, les signaux étaient ceux directement reçus par l'antenne. Leur nature unidimensionnelle devenait bidimensionnelle lorsque nous procédions à un «découpage» du signal reçu pour chacune des récurrences émises par l'antenne.

Les signaux avec lesquels nous allons désormais travailler sont de nature toute différente : s'agissant de séquences d'images, nous disposons d'un signal tridimensionnel. De plus, la nature de l'évolution temporelle que nous avons précédemment était déterministe : elle était «provoquée» par le mouvement relatif entre antenne et cible. Ici, le déterminisme est conservé, mais le mouvement ne concerne que la cible : un visage évoluant au cours du temps. Cette évolution se caractérise par des attitudes déterministes telles que celles développées dans l'annexe C. Le cadre de notre étude va donc être la réalisation puis l'étude d'une signature associée à une attitude de visage. Pour rendre les informations plus complètes, les attitudes de visages sont enregistrées par deux caméras travaillant dans les domaines visible et infrarouge (échelle linéaire de $3\mu m$ à $5\mu m$)¹. L'information issue du capteur infrarouge sur des images de 64×64 pixels, se réduit grossièrement à un contraste entre le visage et son fond. L'aspect multi-senseurs introduit donc ici une redondance d'information permettant d'améliorer la reconnaissance.

¹Les acquisitions suivant les canaux visible et infrarouge ont été menés dans les locaux de SFIM-ODS, Vélizy, France

Nous avons développé, dans le chapitre précédent, la synthèse de signatures suivant deux axes, selon que nous supposions ou non que les signaux étaient auto-suffisants pour former une signature. Nous allons donc illustrer ces deux approches sur les attitudes de différents visages.

3.1.1 Présentation des séquences utilisées

Afin de juger de la puissance des deux approches que nous proposons et de comparer leurs avantages et leurs inconvénients, nous n'allons extraire de la totalité des séquences que certaines d'entre elles.

Les séquences segmentées n'étant pas de qualité équivalente pour tous les visages et toutes les attitudes, nous nous limiterons aux visages «Emmanuel» et «Barbu». Parmi les dix séquences segmentées à partir des vidéo, trois seront sélectionnées pour former la base d'apprentissage, trois autres délimitant la base de test :

Base d'apprentissage	Base de test
bâillement	froncement & bâillement
hochement	hochement & bâillement
rotation	rotation & clignement

La base d'apprentissage permettra la formation d'une base de signatures. La seconde servira à la synthèse de signatures, lesquelles seront comparées à la base préalablement formée. La figure 3.1 illustre deux attitudes, une d'apprentissage et une de test, pour deux visages différents.

Pour décrire la structure de la base de données, nous désignons par x_i et y_i les images visible et infrarouge. Dans ce cas, les images sont synchrones.

3.1.2 Approche «PCE»

Présentation de la méthode

La première approche présentée utilisait une banque de références pour former la signature (figure 2.4). Le schéma général 2.12 peut se réduire au schéma 3.2.

Les images visible x_i et infrarouge y_i étant associées au même instant i , un premier filtrage leur est appliqué, par l'intervention de deux bases de filtres :

- une base de M filtres F_m (dont les références sont de même nature que les images de la séquence), ici les images des séquences de la base d'apprentissage ;
- une base de L ondelettes G_l , permettant d'accéder à un niveau de détail donné.

Du résultat de ce double filtrage est extrait le critère de corrélation PCE de $\{x_i, y_i\}$ filtré par F_m , à un niveau de contraction d'ondelette l . Si nous isolons les résultats issus des différents niveaux l , nous formons un pixel de la signature S indicé par i , indice temporel de la séquence et m indice du filtre présenté à $\{x_i, y_i\}$. Nous formons ainsi $S_{l(mi)}$:

$$S_{l(mi)} = PCE\{\{x_i, y_i\} \otimes \mathcal{F}_{F_m G_l}\} \quad (3.1)$$

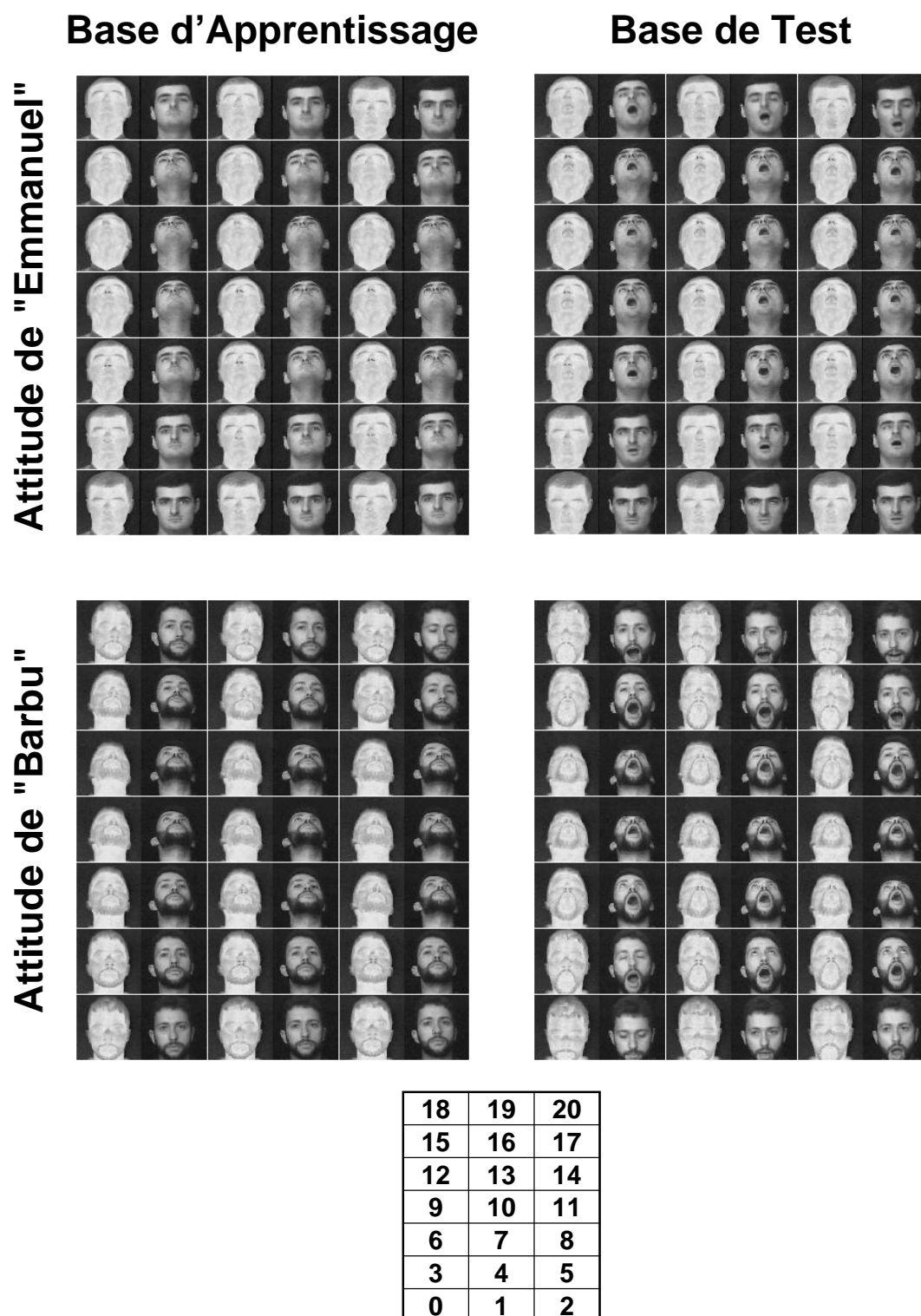


FIG. 3.1 – Exemples d'attitudes de visages : les séquences présentées pour les visages «Emmanuel» et «Barbu» sont celles d'un «hochement», pour la base d'apprentissage, et d'un «hochement & bâillement» pour la base de test.

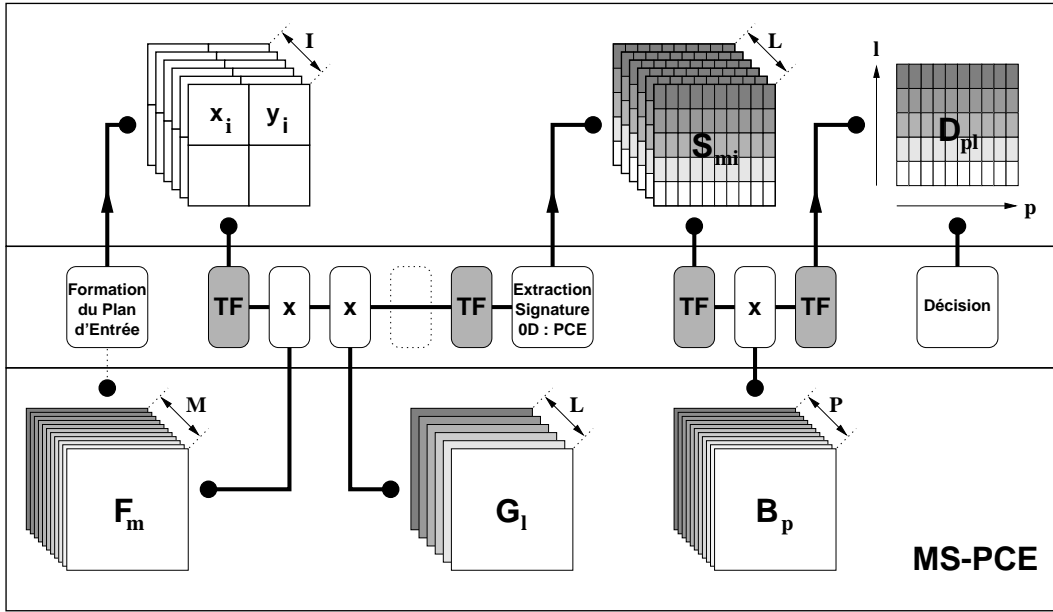


FIG. 3.2 – **Synthèse & Analyse de signatures à partir d’une Base de Références** : Chaque ligne correspond à un instant. Chaque pixel de la ligne porte le critère de corrélation PCE entre la référence et l’image jointe de la séquence.

Nous disposons, une fois la base F présentée dans sa totalité, d’un ensemble de L signatures S . Ces L signatures S_{mi} ainsi obtenues sont comparées par filtrage à un ensemble de P signatures B . Nous formons alors une matrice D constituée des PCE entre la référence B_p et la signature S_l .

$$D_{pl} = PCE\{S_l \otimes \mathcal{F}_{B_p}\} \quad (3.2)$$

Notons que nous ne comparerons, dans un premier temps, que des signatures ayant un niveau de résolution l commun. Cette simplification pourra, le cas échéant, être levée s’il s’avère intéressant de comparer des signatures de résolutions différentes, dans le cadre, par exemple, d’une structure auto-similaire au sein de la signature. Toujours est-il que c’est à partir de cette matrice de décision D_{pl} que sera donnée la décision finale : nous nous limiterons à considérer une forme polynomiale pour cette dernière décision :

$$D_p = \sum_{l=0}^{l=L-1} \alpha_l D_{pl}^{\beta_l} \quad (3.3)$$

Le choix des coefficients α_q et β_q sera établi par un apprentissage sur les différentes bases de signatures. Il est clair qu’une forte correspondance (*i.e.* un fort PCE) entre deux signatures à un niveau élevé l de résolution aura nécessairement plus de poids qu’à un niveau de résolution moindre.

Expérimentation

Nous allons pour l’attitude «hochement», laquelle sera ensuite perturbée par un «bâillement» (séquence «hochement & bâillement» de la base de test), établir les L

signatures correspondant aux deux visages «Emmanuel» et «Barbu».

La figure 3.3 illustre les différentes signatures obtenues pour des niveaux de résolutions, fixés par a , identiques à 2, 1, 1/2 et 1/4. D'un point de vue général, chaque signature dispose d'une symétrie axiale : à l'origine de cette symétrie, la symétrie de la séquence. Par ailleurs, nous remarquons que les différentes signatures présentent des «X» plus ou moins marqués. Cette structure en «X» a pour origine la disposition de la base de références à laquelle est présentée chaque image de la séquence : elle est, en effet, ordonnée temporellement, les séquences se succédant les unes aux autres. Aussi est-il cohérent que les images servant de références dans les séquences d'apprentissage «répondent» favorablement lorsqu'elles sont comparées entre elles. Ainsi, si nous nous bornons à observer la réponse d'une séquence à sa séquence associée, nous distinguons une symétrie plus centrale.

La position de la structure en «X» désigne la séquence la plus vraisemblable. Les signatures issues de la base de test (à droite) mettent en évidence le même caractère : un «X» indique que la séquence présentée se rapproche du hochement. La séquence présentée est en effet perturbée par un bâillement. Nous remarquons de plus que les références associées au bâillement «répondent» à la séquence de test. La structure n'est pas un «X» mais deux parenthèses tête-bêche couchées « \smile », signe que la séquence présentée, à son début et sa fin, se rapproche d'un bâillement.

L'influence du prétraitement est également très perceptible. En effet, à des valeurs de a importantes est associée une faible résolution. Un filtrage à de faibles niveaux de résolution induit une sélectivité moins grande du filtre de phase. Il apparaît ainsi, autant pour le visage «Emmanuel» que «Barbu», une croix désignant la séquence de la base d'apprentissage correspondant au hochement. Ainsi est-il possible de déterminer à quelle attitude correspond la signature, sans nécessairement connaître le visage associé. Cette propriété illustre la hiérarchisation qu'il nous sera possible d'établir pour restreindre le nombre de classes de la base d'apprentissage. A des niveaux faibles de résolution, nous n'aurons que quelques visages pour un ensemble d'attitudes important, l'attitude y étant discriminante. Une fois l'attitude connue avec une bonne certitude, seules des classes correspondant à des visages seront présentées à des résolutions plus fines.

Cette première étape de synthèse de signatures accomplie, nous nous devons maintenant de les analyser. Cette analyse de signatures se réalise grâce à une corrélation, opération nous permettant de comparer la signature obtenue aux signatures de la base B . Cette base B se construit donc à partir des filtres de phase correspondant à chaque signature. Un PCE établira la correspondance entre la signature et un élément de la base B_p . Notons que la comparaison aura lieu entre des signatures obtenues à des niveaux de résolution identiques.

La table 3.1 présente les résultats issus des corrélations entre les signatures de base d'apprentissage et celles obtenues par présentation de la séquence. Nous constatons que, généralement, la séquence présentée «hochement & bâillement» se retrouve reconnue, à chaque niveau de résolution, comme la séquence «hochement». A de faibles résolutions (a faible), une ambiguïté demeure entre l'attitude «hochement» et «bâillement». De plus, la

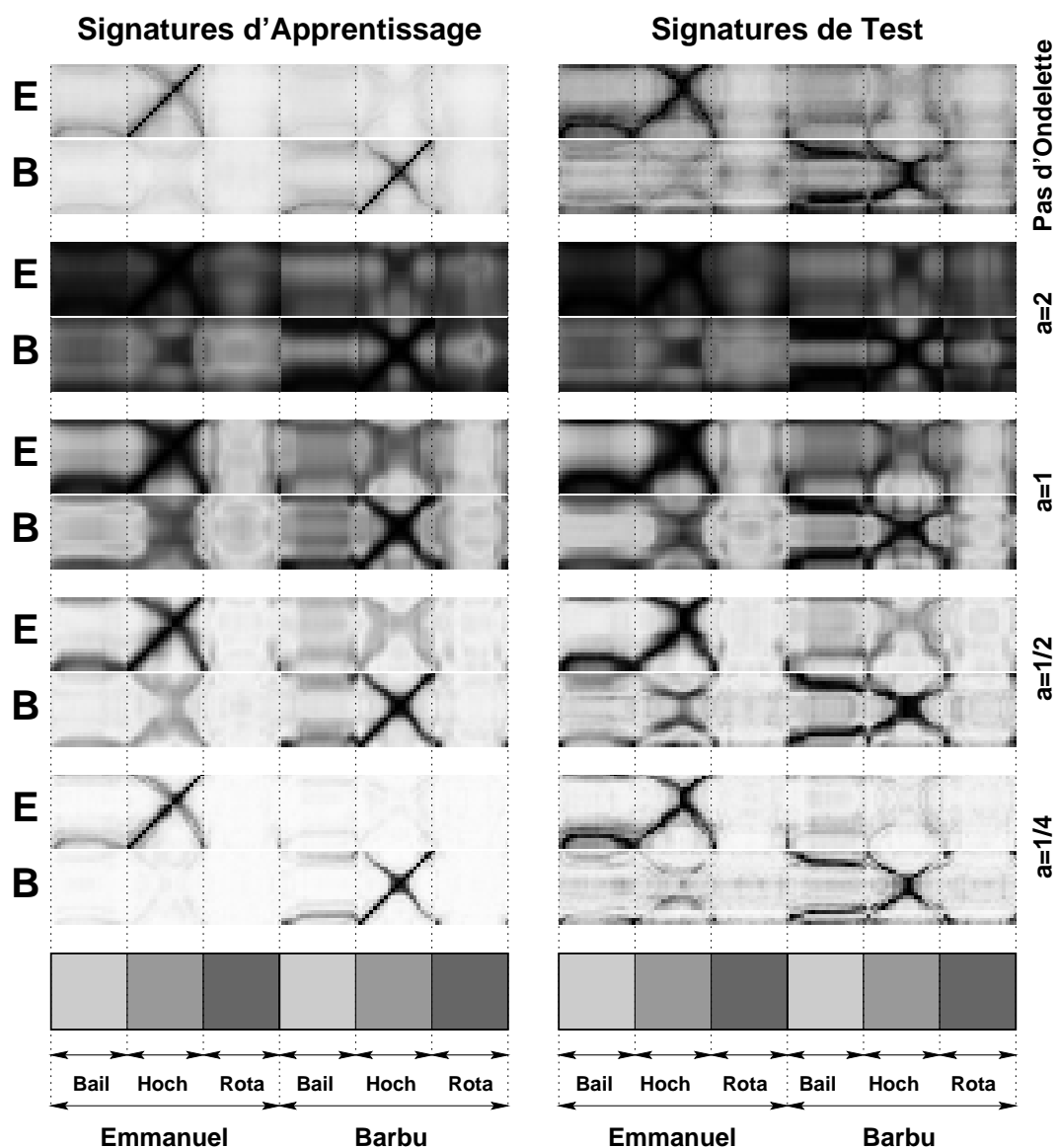


FIG. 3.3 – Exemples de signatures par la méthode «PCE» : la méthode «PCE» utilise une base de filtres F correspondant à des références d'autres séquences d'apprentissage. Sur la colonne de gauche apparaissent les signatures de la base d'apprentissage et sur la droite celles de la base de test. Pour les deux visages «Emmanuel» (noté **E**) et «Barbu» (noté **B**) et pour chaque niveau de contraction d'ondelette a se superposent les différentes signatures. La base de références présentée dispose de deux classes de visages, au sein desquelles se trouvent trois attitudes : «bail» pour «bâillement», «hoch» pour «hochement» et «rota» pour «rotation».

similitude des signatures à de faibles résolutions se retrouve dans les *PCE* présentés : les valeurs sont comparables pour $a = 1$. Il est d'ailleurs intéressant de noter que les *PCE* obtenus sur des visages différents, pour une même attitude, sont plus proches pour $a = 1$ que pour $a = 2$. De cette remarque nous pouvons supposer qu'il serait précipité, dans un processus de reconnaissance, d'éliminer directement les références les moins vraisemblables dès le premier niveau d'investigation.

		Emmanuel			Barbu		
Dilatation	Visage	Bail	Hoch	Rota	Bail	Hoch	Rota
Sans Ondelette	Emmanuel	14	23	8	8	6	5
	Barbu	5	5	4	9	14	10
$a = 2$	Emmanuel	13	14	11	8	9	8
	Barbu	10	11	8	14	17	11
$a = 1$	Emmanuel	17	34	12	17	13	6
	Barbu	15	24	11	24	20	12
$a = 1/2$	Emmanuel	18	65	13	15	9	3
	Barbu	10	10	6	17	17	9
$a = 1/4$	Emmanuel	20	50	11	2	1	3
	Barbu	5	6	5	7	15	15
Tous Niveaux	Emmanuel	68	163	47	42	32	20
	Barbu	40	51	30	62	69	47

TAB. 3.1 – **Table présentant les *PCE* issus des corrélations entre signatures** : les *PCE* ($\times 10^{-3}$) présentés sont obtenus à partir des corrélations entre les signatures obtenues à partir d'une séquence test et les signatures de la base d'apprentissage.

Décision finale

Nous avons proposé de prendre une décision finale en établissant un critère à partir de la totalité des critères précédents D_{pl} , p désignant l'indice de la signature et l la résolution. Nous devons donc établir un critère D_p à partir des différents D_{pl} . Pour cela, nous proposons une forme polynomiale D_p dépendant des décisions élémentaires D_{pl} auxquelles sont associés des coefficients ajustables (α_l, β_l) (équation 3.3).

Si nous fixons arbitrairement les couples (α_l, β_l) à l'unité ($(\alpha_l = 1, \beta_l = 1)$ pour $l = \{1, \dots, L\}$), les D_l apparaissent sur les deux dernières lignes de la table 3.1. La décision finale intervient donc en recherchant la valeur maximale de la ligne pour ensuite identifier la colonne correspondant à l'attitude et au visage.

Conclusion

La méthode «PCE» forme une signature dont chaque ligne est composée des *PCE* issus de la corrélation entre une image de la séquence temporelle et un ensemble de références.

Cette méthode permet de relever, directement par l'observation des signatures, certaines caractéristiques, comme la symétrie de la séquence, ou le fait que l'attitude présentée par la séquence soit une composition de plusieurs attitudes.

Le filtrage spatial permet la formation de chaque pixel de la signature. Il demeure également le lien actif établissant la comparaison, par corrélation, entre la signature obtenue et une base de signatures préalablement établies. Le critère de corrélation que nous avons jusqu'à présent utilisé, le *PCE*, sert autant à la formation de sa signature qu'à la conclusion sur la comparaison entre signatures. Une décision finale, enfin, établit la synthèse entre la matrice de décision D_{pl} , et le vecteur D_p , où p désigne autant la classe d'attitude que la classe de visage. Enfin, la recherche de la valeur maximale de D_p permet de conclure sur la nature de l'attitude et du visage. A une décision finale aussi sommaire peut se substituer une autre plus fine, laquelle fera, par exemple l'objet d'un apprentissage.

3.1.3 Approche «Parcours Fractal»

Présentation de la méthode

La méthode que nous venons de présenter nécessitait l'intervention d'une base de références indispensable à la formation de la signature. Cette nouvelle approche considère que les données issues des capteurs sont suffisantes à la formation d'une signature discriminante. Ainsi, la différence avec la méthode précédente repose-t-elle sur la formation d'une signature ayant la structure présentée figure 2.4. L'aspect temporel étant toujours codé verticalement, l'image, ou l'assemblage d'images provenant des capteurs visible et infrarouge devra être réduite à un vecteur s'insérant sur une ligne. Aussi, pour réaliser pareille opération, le schéma général 2.12 se réduit-il au schéma 3.4.

La procédure permettant la formation d'une ligne de la signature est illustrée figure 3.5. Sept étapes forment chaque ligne de la signature :

- un «*zero padding*», permettant d'éviter la formation des intercorrélations sur la périphérie de l'image ;
- une transformée de Fourier, permettant le passage dans l'espace des fréquences ;
- un produit avec un filtre d'ondelette permettant de ne sélectionner que les intercorrélations entre les images à un niveau de résolution donné ;
- une estimation du module puis son élévation à une puissance k . Cette opération forme la non-linéarité du traitement ;
- une transformée de Fourier pour revenir dans l'espace direct ;
- une extraction d'un des deux plans d'intercorrélations (lesquels sont symétriques) ;
- une transformation de l'intercorrélations extraite en un vecteur par application d'un parcours fractal (courbe de Hilbert).

Ces opérations sont à réitérer pour chacune des images formant la séquence. La figure 3.6 illustre la formation de la signature à partir d'un ensemble de plans d'intercorrélations.

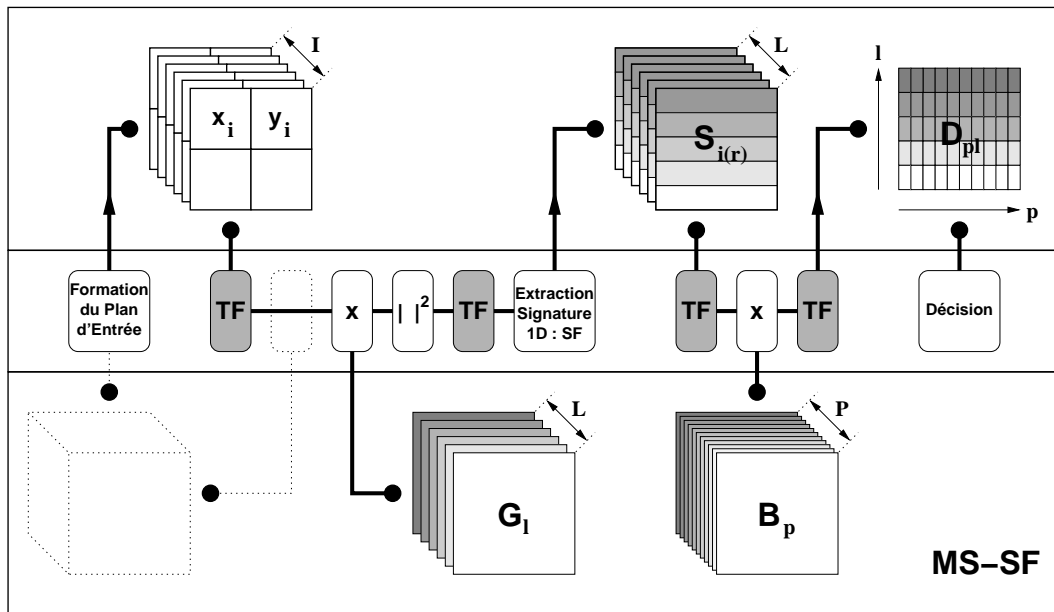


FIG. 3.4 – **Synthèse & Analyse de signatures avec extraction par parcours fractal** : chaque ligne correspondant à un instant, l'assemblage des images visible et infrarouge se réalise par application d'un filtrage non-linéaire suivie d'une réduction de l'intercorrélation à un vecteur par parcours fractal.

Expérimentation

Nous avons, pour les mêmes séquences d'apprentissage et la même séquence de test, recherché les différentes signatures, pour des niveaux de résolution a identiques à ceux explorés lors de l'approche «PCE».

Ainsi formons-nous les signatures des visages «Emmanuel» et «Barbu» pour trois attitudes d'apprentissage (bâillement, hochement et rotation) et trois attitudes de test (bâillement & froncement, hochement & bâillement et rotation & clignement), pour quatre résolutions différentes ($a = 2$, $a = 1$, $a = 1/2$ et $a = 1/4$).

La figure 3.7 présente les signatures de façon identique à la figure 3.3 : les signatures pour les visages «Emmanuel» et «Barbu» animés des attitudes «hochement» et «hochement & bâillement».

Nous constatons tout d'abord que les principaux signes distinctifs qui apparaissaient par la méthode de synthèse «PCE», ne se retrouvent pas sur les signatures : plus de «X» ni de « \sphericalangle ». Nous constatons également que les signatures conservent leur symétrie axiale, suivant la direction verticale, pour des raisons de symétrie liée à l'attitude elle-même.

L'exploitation des signatures est donc beaucoup plus complexe qu'elle ne l'était par la première approche : il n'est plus possible de conclure rapidement, uniquement par l'examen de la signature, sur la nature de l'attitude ou le visage associé. Le filtrage retrouve donc là encore tout son intérêt pour analyser globalement par une comparaison

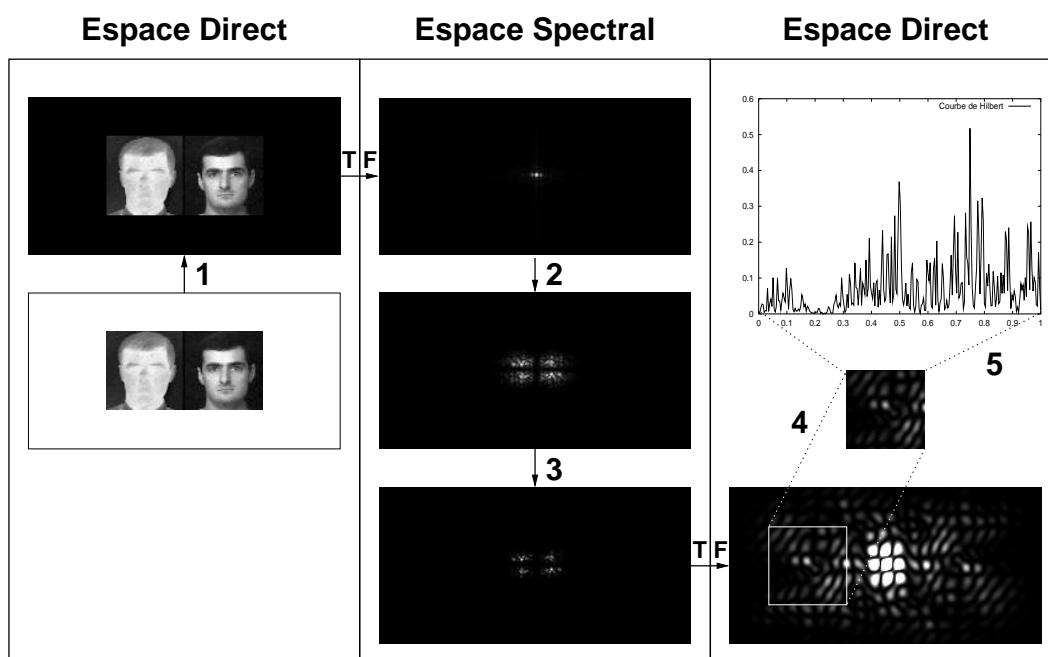


FIG. 3.5 – **Étapes dans la formation d'une ligne de signature** : La formation d'une ligne de signature à partir de la combinaison de deux images se décompose en 7 étapes : un «*zero padding*» [1], une *TF*, le produit par le filtre d'ondelette [2], le calcul du module et l'élévation à une puissance [3], une *TF*, l'extraction du plan d'intercorrélration [4] et la réduction du plan d'intercorrélration à un vecteur par application d'un parcours de Hilbert sur le plan [5].

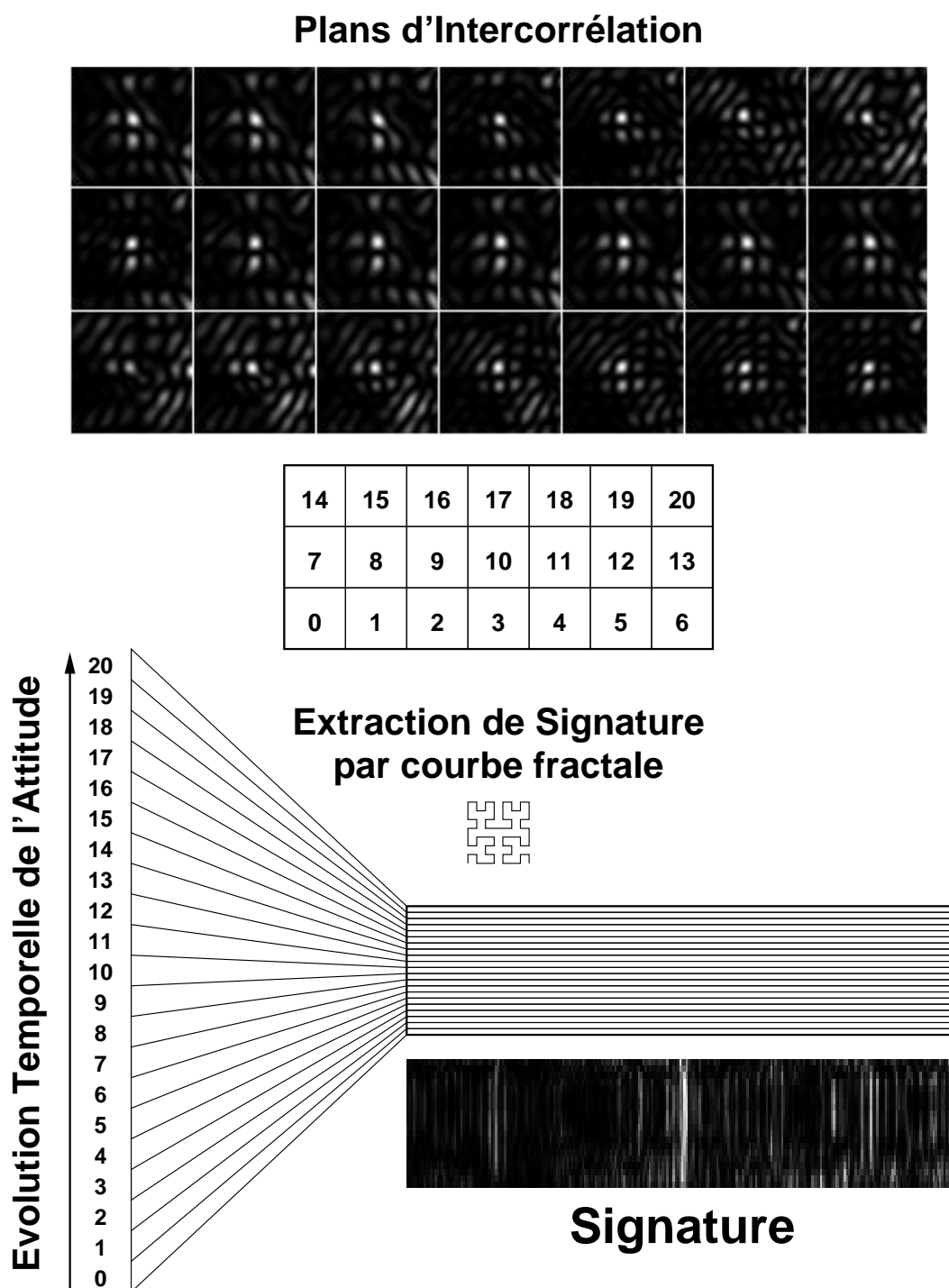
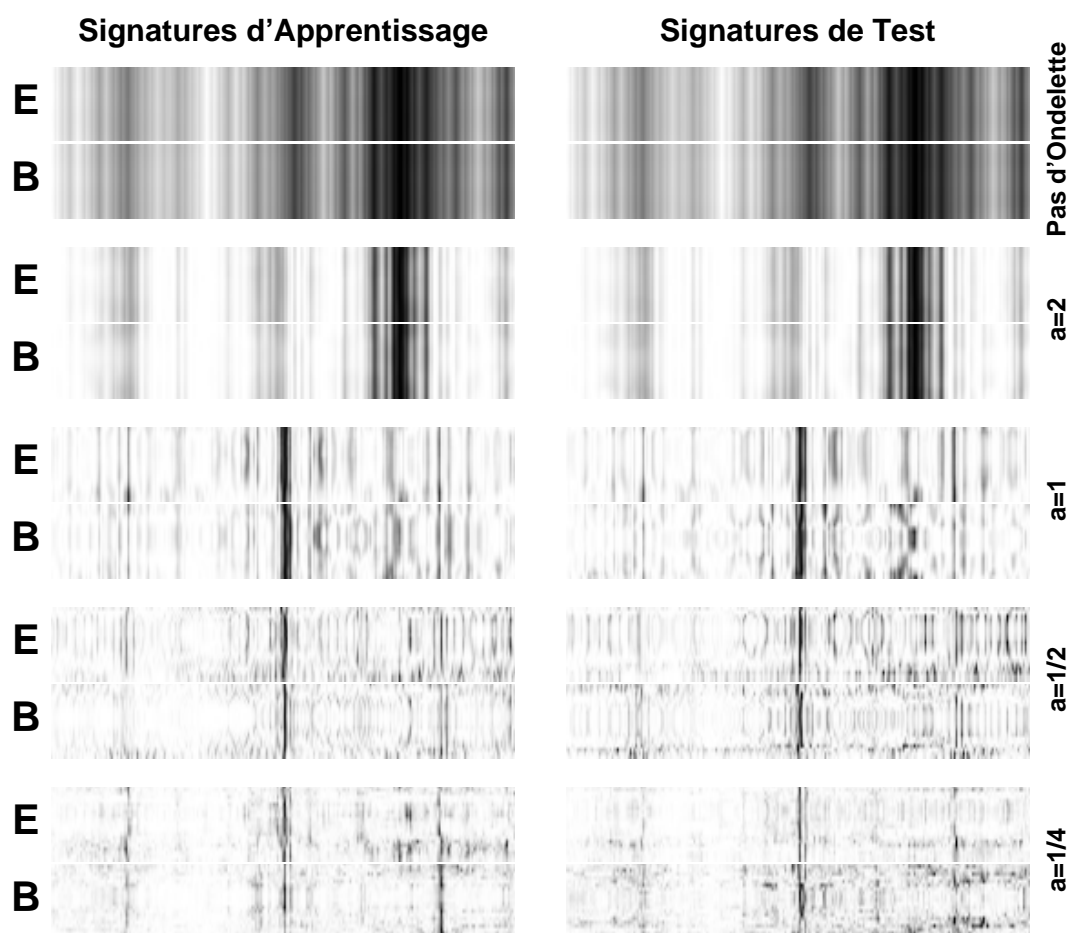


FIG. 3.6 – Formation d'une signature à partir des plans d'intercorrélation : le parcours fractal suivant la courbe de Hilbert établit le passage entre l'image (2D) et le vecteur (1D) s'insérant dans la signature.



E : Emmanuel & B : Barbu

FIG. 3.7 – Exemples de signatures d'attitudes de visages obtenues par parcours fractal : la synthèse de signatures par filtrage non-linéaire considère que les images formant les séquences sont suffisantes à la formation d'une signature discriminante de l'attitude et du visage. Sur la colonne de gauche apparaissent les signatures de la base d'apprentissage et sur la droite celle de la base de test. Pour les deux visages «Emmanuel» (noté **E**) et «Barbu» (noté **B**), les différentes signatures se superposent pour chaque niveau de contraction d'ondelette a .

quelle attitude correspond le mieux à la signature obtenue, puis quel visage présente une telle attitude.

La table 3.2 présente les *PCE* obtenus par la corrélation entre la signature test «hochement & bâillement» et les six signatures d'une base d'apprentissage, pour les visages «Emmanuel» et «Barbu». Sans introduction d'ondelette, aucune discrimination n'est possible. La discrimination va croissant pour des facteurs a décroissants, donc pour des niveaux de détail plus élevés. Nous constatons que, pour les deux visages, les deux attitudes les plus reconnues sont celles composant la signature de test, le hochement et le bâillement. Pour le visage «Emmanuel», il n'existe aucune ambiguïté, quelle que soit la résolution. Il n'en est cependant pas de même pour le visage «Barbu», lequel est confondu avec «Emmanuel» aux niveaux $a = 1/2$ et $a = 1/4$.

Dilatation	Visage	Emmanuel			Barbu		
		Bail	Hoch	Rota	Bail	Hoch	Rota
Sans Ondelette	Emmanuel	16	16	16	16	16	16
	Barbu	16	16	16	16	16	16
$a = 2$	Emmanuel	51	51	48	49	49	44
	Barbu	52	50	51	52	52	50
$a = 1$	Emmanuel	72	93	40	76	57	28
	Barbu	26	44	11	56	47	20
$a = 1/2$	Emmanuel	22	71	16	12	20	5
	Barbu	6	17	14	13	14	5
$a = 1/4$	Emmanuel	32	54	35	17	36	16
	Barbu	11	14	7	5	13	4
Tous Niveaux	Emmanuel	177	269	139	154	162	93
	Barbu	95	125	83	125	126	79

TAB. 3.2 – **Table présentant les *PCE* issus des corrélations entre signatures** : les *PCE* ($\times 10^{-3}$) présentés sont obtenus à partir des corrélations entre les signatures obtenues à partir d'une séquence test et les signatures de la base d'apprentissage.

Décision finale

Pour établir la décision finale à partir des décisions élémentaires D_{pl} , nous avons proposé une composition polynomiale, dont les coefficients seraient ajustés en fonction de l'application (équation 3.3).

Pour des coefficients α_l et puissances β_l mis à l'unité, les résultats se présentent sur la dernière ligne de la table 3.2. Pour «Emmanuel», aucune ambiguïté ne subsiste : les attitudes préférentiellement reconnues sont le hochement puis le bâillement. Il n'en est pas de même pour «Barbu», pour lequel les attitudes reconnues sont bien hochement et bâillement. Il est toutefois plus difficile de conclure sur l'identité du visage.

Pour pallier pareils inconvénients, nous proposons :

- soit de combiner les résultats des différentes attitudes pour un visage donné, ou de ne sélectionner que quelques attitudes pour lesquelles il existe la meilleure adéquation. Ainsi, dans le cas présent, puisque les attitudes de hochement et de bâillement sont les plus vraisemblables, nous les combinons pour ne former plus qu'une unique valeur par visage ;
- soit de changer d'attitude de test. Si nous prenons par exemple l'attitude de rotation, la discrimination serait meilleure.

Conclusion

Les résultats, par cette approche, pourraient être des plus mitigés, étant donnée l'ambiguïté concernant la reconnaissance du visage «Barbu». Toutefois, il apparaît que nous pouvons améliorer cette décision par une composition des résultats issus des différentes attitudes. De plus, l'attitude-test que nous avons considérée, une composition d'un hochement avec un bâillement, bouleverse singulièrement la structure de la séquence, donc de la signature associée.

Si nous devons par exemple reconnaître une personne par une attitude de visage, il serait préférable de prendre une attitude telle que la rotation. Notre reconnaissance n'en serait que meilleure.

3.1.4 Comparaison des deux approches

Avant d'apporter quelques remarques particulières, revenons sur la différence principale de formation de signatures entre les approches «PCE» et «Parcours Fractal».

Dans le premier cas, nous considérons la formation d'une «inter-signature», par comparaison entre nos signaux et une base de références de même nature. Il est donc possible d'injecter une information supplémentaire à la formation d'une signature. Cette approche nécessite cependant que nous disposions d'une information *a priori* sur les références auxquelles nous devons faire appel pour la synthèse de signatures, pour établir les meilleures auto-corrélations possibles entre séquences et références. Ainsi la performance de cette première méthode repose-t-elle implicitement sur la qualité de la base de références.

Dans le second cas, nous supposons qu'il n'est pas nécessaire de recourir à une base de références pour former une signature, laquelle se forme par elle-même. Ces «auto-signatures» sont donc nécessairement assez discriminantes entre elles pour que nous envisagions de les substituer à celles de la première méthode. Elles disposent d'autres avantages, nous y reviendrons, lorsque nous nous pencherons sur leur intégration.

De plus, une autre remarque entre les deux approches que nous venons de développer tient au fait que, dans l'approche «PCE», l'information pertinente était de nature locale : chaque point représentait la comparaison entre une image de la séquence et une référence de la base. Ainsi, les structures qui apparaissaient sur les signatures (les «X», les barres verticales, etc...) étaient autant d'informations pertinentes auxquelles nous pouvions accéder directement. Il n'en est pas de même pour l'approche par filtrage non-linéaire. Les

liaisons que nous avons forcées entre les images issues des capteurs visibles et infrarouge devenaient des plans de corrélation que nous réduisions à un vecteur. L'information, bien que conservée localement (entre pixels voisins de l'image), se répartissait sur toute la largeur de la signature. Elle était donc de nature plus globale.

Le fait que les approches développées forment des signatures plus locales ou globales dans leur structure justifie l'utilisation d'un corrélateur dans la seconde approche : seul un filtrage permet de traiter l'image dans sa totalité (sur toute sa surface). De fait, la corrélation entre une base de signatures et la signature obtenue peut-elle être appliquée à l'ensemble des signatures, quelle que soit l'approche utilisée pour les synthétiser.

La seconde remarque viendra des résultats des corrélations entre la signature et une base de signatures de référence. Même si l'ambiguïté demeure plus importante dans le cas d'un filtrage non-linéaire, il est intéressant de noter que les résultats présentés sur les tables 3.1 et 3.2 sont comparables ; en effet, les mêmes attitudes présentent systématiquement les *PCE* les plus importants, dans des proportions souvent comparables. Le fait que, dans certains cas, il faille «injecter» de l'information (la base de références) dans le traitement pour qu'il soit plus performant est même un bon signe...

Enfin, pour terminer, nous concluons par des considérations plus pragmatiques : le temps de calcul. En ce qui concerne la corrélation finale entre la signature et la banque de signatures, les cardinaux sont identiques. Toutefois, une drastique différence existe entre la le temps nécessaire à la formation d'une signature à partir d'une base de références et celui requis dans un filtrage non-linéaire : dans un cas, chaque ligne de la signature exige une corrélation, donc deux *TF*, deux produits et un calcul de *PCE* ; dans l'autre approche, deux *TF*, deux produits et un calcul de module seulement sont nécessaires pour former une ligne entière. Pour fixer un ordre de grandeur, 20 minutes sont nécessaires dans un cas, contre 30 secondes dans l'autre. Cette contrainte, liée au rafraîchissement de la signature, fixera certainement le choix d'une approche au détriment de l'autre.

3.2 Signaux Radar «Temps-Fréquence»

3.2.1 Importance de l'horizon temporel

La problématique exprimée dans le deuxième chapitre présentait les origines possibles de variabilité d'un signal. Ainsi, en télédétection, disposons-nous de trois composantes : cible, canal de transmission et détecteur. Cette variabilité avait une origine, déterministe ou stochastique. Entre ces deux évolutions temporelles possibles se place le problème de l'horizon. En effet, nous pouvons avoir des séquences disposant d'une évolution temporelle déterministe liée, par exemple, à un mouvement et d'une évolution temporelle stochastique, intrinsèque au canal de transmission. L'horizon temporel, ou l'échelle de temps dans laquelle s'insèrent les évolutions temporelles est des plus importantes selon la nature des informations que nous voulons relever sur notre signature. Aussi, deux cas sont possibles :

- si l'horizon temporel est long :
 - les évolutions temporelles peuvent alors caractériser une évolution intrinsèque à la cible, donc rendre compte d'un mouvement. Nous pourrions alors utiliser la synthèse

de signatures comme un outil permettant la visualisation du mouvement.

– si l’horizon temporel est court :

les seules évolutions temporelles sont associées à des phénomènes stochastiques.

Ainsi, la signature associée à de tels phénomènes devra être la plus «statique» (suivant l’axe temporel de formation de la signature) possible.

Les signaux «Temps-Fréquence» ou «*Range Gate Doppler*» (RGD), issus de radars aéroportés observant des bâtiments de surface, sont de cette dernière catégorie : sur un laps de temps court, les signaux sont soumis à un ensemble de perturbations liées aux mouvements intrinsèques du capteur, de la cible et aux perturbations électromagnétiques atmosphériques ou maritimes.

Sachant que les images issues d’une telle acquisition disposent d’une forte variabilité stochastique, nous allons donc chercher à réaliser la signature la plus stable possible. Cette stabilité temporelle transformera la signature en «code barre», lequel sera caractéristique d’un bâtiment.

3.2.2 Le choix du «Parcours Fractal»

Pour réaliser cette synthèse de signature, nous n’allons pas développer, en l’absence de données suffisantes, la méthode utilisant une base de références, mais celle réalisant un parcours fractal : un tel parcours est réalisé sur un plan d’intercorrélations, résultat d’un filtrage non-linéaire sur l’association de deux images (figure 3.8).

En effet, si nous utilisons un opérateur de comparaison pour former la signature à partir de *PCE*, les filtres que nous utiliserons seront par nature trop sélectifs pour que nous puissions établir une signature discriminante. De plus, si une référence est plus activée qu’une autre lors de la synthèse d’une ligne de la récurrence, il n’est pas du tout évident que cela se reproduise l’instant d’après. La figure 3.9 illustre la variabilité des signaux (deux images consécutives sont représentées pour les bases d’apprentissage et de test).

Le filtrage non-linéaire établit des liens entre les images de capteurs différents ou d’instantanés différents pour un unique capteur. Nous allons donc proposer de prendre une image correspondant à la fin d’une séquence et nous l’associons à I images consécutives qui la précèdent. Ainsi avons-nous les associations suivantes : $\{x_i, x_{i-I}\}$, $\{x_i, x_{i-I+1}\}$, $\{x_i, x_{i-I+2}\}$, . . . , $\{x_i, x_{i-2}\}$, $\{x_i, x_{i-1}\}$, $\{x_i, x_i\}$.

3.2.3 Expérimentation

Synthèse des signatures

A partir de ces I couples d’images, nous formons une signature de I lignes. Disposant de trois classes de bateaux différents (le remorqueur «Abeille-Frandre», un chalutier et un pétrolier), nous formons les signatures associées à plusieurs niveaux de détails (de $a = 2$ à $a = 1/4$), pour les classes d’apprentissage et de test. La figure 3.10 illustre les

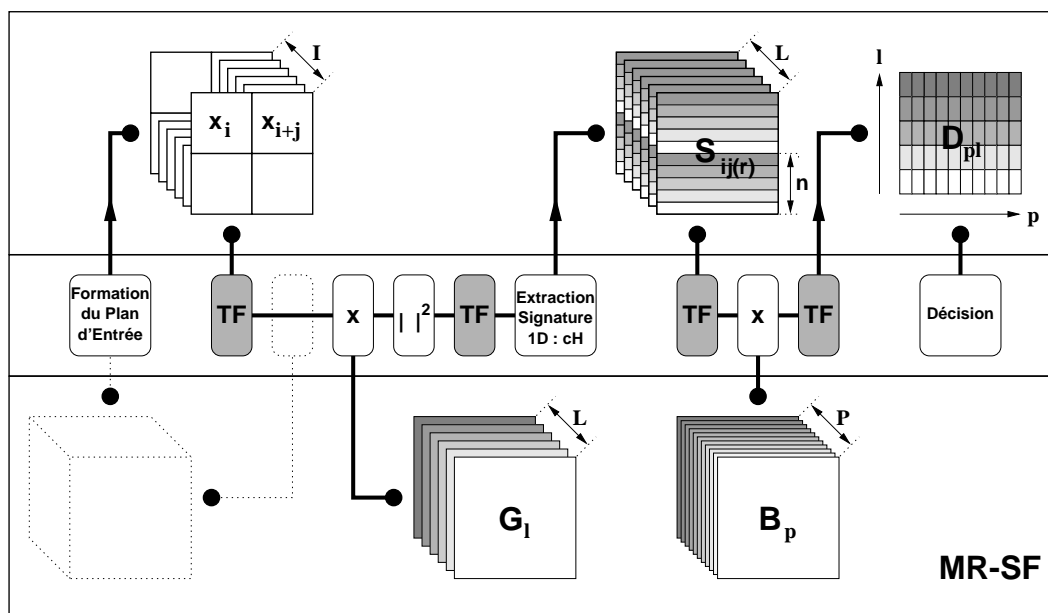


FIG. 3.8 – Synthèse & Analyse de signatures avec extraction par parcours fractal : les images, issues d'un même capteur, sont associées lors du formatage. Une image de cette séquence est jointe successivement à celles qui la suivent. Un filtrage non-linéaire assure l'intercorrélacion entre les images, laquelle est réduite à un vecteur par parcours fractal.

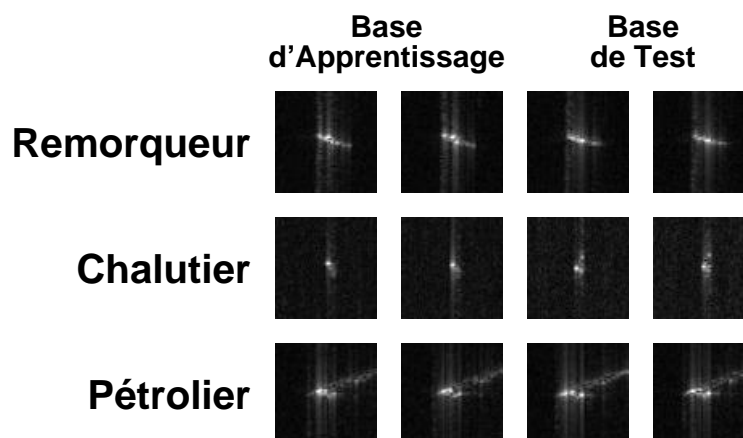


FIG. 3.9 – Exemples de signaux «Temps-Fréquence» de bâtiments de surface : ces signaux illustrent la variabilité de ces signaux au cours du temps. Pour les bases d'apprentissage puis de test, deux images consécutives sont présentées. Un examen rapide permet de noter la forte évolution temporelle stochastique d'un tel signal.

signatures obtenues à partir d'une filtrage non-linéaire.

Nous constatons tout d'abord que nous disposons bien, pour la majorité des signatures, d'une signature ressemblant à un code barre. La finesse du pic semble conditionnée par le niveau de détail a choisi. De plus, la répartition des barres est différente pour les trois bâtiments, ce qui permet de supposer que les signatures formées permettront bien, lors de leur analyse, d'identifier le bâtiment.

Nous allons donc réaliser l'analyse des signatures obtenues par une corrélation entre les signatures de la base de test et celles de la base d'apprentissage (table 3.3).

L'examen des résultats nous permet tout d'abord de tirer quelques conclusions :

- la reconnaissance n'est réalisée que si nous introduisons un prétraitement par une ondelette ;
- les bâtiments sont reconnus de manière inégale suivant le niveau de détail choisi. Cette constatation nous indique que, même s'il n'y a pas bonne adéquation entre signature présentée et signature de la base à un niveau de détail donné, il semble prématuré d'éliminer cet élément de la base de défilement pour des niveaux de détails supérieurs ;
- les identifications, en cas de reconnaissance, sont généralement plus contrastées que dans le cas de la reconnaissance de visage par l'attitude ;

Analyse des Signatures

Comme nous l'avons réalisé dans les applications précédentes, si nous associons toutes les décisions élémentaires pour ne former qu'une unique décision, par une sommation des PCE issus de chacun des niveaux de détail, nous obtenons un résultat très honorable de reconnaissance pour le pétrolier et le remorqueur. Pour le chalutier, une ambiguïté subsiste encore. Par ailleurs, comme nous pouvons le remarquer pour les signatures associées au chalutier, certaines traces apparaissent. Elles correspondent à des images très bruitées.

Toutefois, le cas que nous avons considéré était des plus particuliers. En effet, l'horizon temporel était si court que les seules évolutions que nous observions étaient de nature stochastique. Nous avons donc cherché à «gommer» ces évolutions dans les signatures.

Dans le cas général, nous allons justement rechercher à considérer un horizon temporel assez vaste pour établir sur la signature, une évolution intrinsèque de la cible, donc saisir, sur sa signature, son comportement. Si c'est un bâtiment de surface, il s'agira par exemple, de son attitude face à la houle, sa manœuvrabilité en opération. Le rôle de la signature et son interaction avec la base de signatures permettra de conclure sur l'identité du bâtiment, à la fois par sa «forme» (sa représentation par le senseur utilisé) mais aussi, et surtout, par son attitude.

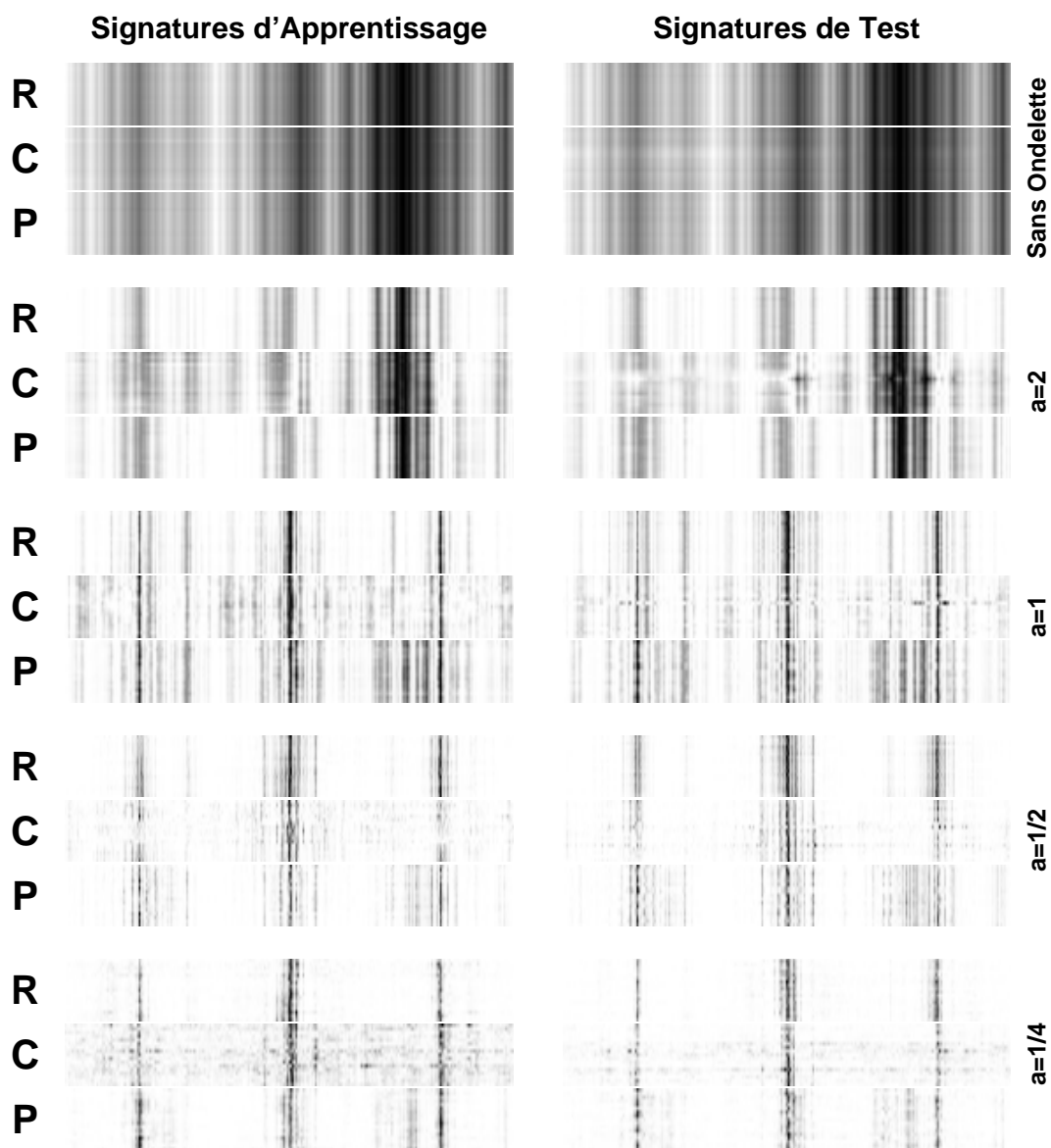


FIG. 3.10 – Exemples de signatures de bateaux obtenues par l’approche du filtrage non-linéaire : sur la colonne de gauche apparaissent les signatures de la base d’apprentissage et sur la droite celle de la base de test. Pour les trois bateaux «Remorqueur» (noté **R**), «Chalutier» (noté **C**) et «Pétrolier» (noté **P**) et pour chaque niveau de contraction d’ondelette a se superposent les différentes signatures.

Niveau d'Ondelette	Bateau	Remorqueur	Chalutier	Pétrolier
Sans Ondelette	Remorqueur	12	12	12
	Chalutier	16	16	16
	Pétrolier	18	18	18
$a = 2$	Remorqueur	45	19	34
	Chalutier	32	44	31
	Pétrolier	44	37	66
$a = 1$	Remorqueur	74	30	31
	Chalutier	67	65	42
	Pétrolier	32	21	93
$a = 1/2$	Remorqueur	54	20	20
	Chalutier	46	40	19
	Pétrolier	26	26	92
$a = 1/4$	Remorqueur	60	29	31
	Chalutier	42	43	20
	Pétrolier	55	36	44
Tous	Remorqueur	233	98	116
	Chalutier	187	194	112
	Pétrolier	157	139	295

TAB. 3.3 – Table présentant les *PCE* issus des corrélations entre signatures : les *PCE* ($\times 10^{-3}$) présentés sont obtenus à partir des corrélations entre les signatures obtenues à partir d'une séquence test pour chaque bateau et les signatures de la base d'apprentissage comprenant l'ensemble des bâtiments.

3.3 Conclusion

Nous avons donc, dans le cadre de données réelles, réalisé la synthèse puis l'analyse de signatures associées à des séquences d'images.

Parmi ces séquences d'images, nous avons des images synchrones issues de différents capteurs formant des séquences déterministes (attitudes de visage). Nous avons également des images issues d'un unique capteur à variabilité stochastique.

Ces signaux sur lesquels nous avons travaillé, étaient très différents, par leur nature intrinsèque ou l'assemblage que nous en faisons. Ils nous permettaient cependant de parvenir à la synthèse de signatures discriminantes. Une corrélation entre ces signatures et une base de signatures permettait enfin de revenir, dans le cas des visages, à l'attitude puis à l'identité du visage.

Dans un autre registre, la synthèse réalisée à partir de signaux stochastiques permettait la formation d'une signature stationnaire suffisamment caractéristique pour qu'il soit reconnu.

Toujours est-il que les applications que nous avons données ne sont que des cas particuliers des modes dérivés de l'architecture générale (figure 2.12), le mode «multi-senseurs» (figures 3.2 et 3.4) et le mode «multi-références» (figure 3.8).

Ces deux modes «multi-senseurs» et «multi-références» peuvent se scinder en deux sous-groupes, selon que, oui ou non, une base de références est utilisée pour «enrichir» la signature : les approches «PCE» et «Parcours Fractal». Le choix entre ces deux sous-modes se trouve intimement relié à la nature des signaux dont nous entendons former les signatures associées. De plus, à la nature des signaux à traiter s'ajoutent les contraintes liées à une implantation de tels traitements. Un aspect «temps-réel» exigerait par exemple un taux de formation de signatures tel que la méthode «PCE» (une corrélation par point) s'avèrerait incapable de soutenir un tel rythme.

Enfin, ces méthodes de synthèse et d'analyse, qui se retrouvent dans le schéma générique 2.12 ayant été illustrées, il nous reste désormais à exprimer par quels moyens, numériques ou optiques, nous comptons réaliser ces filtrages.

Chapitre 4

Implantations & Prospectives

4.1 Introduction

Les chapitres précédents ont été l'occasion de développer les opérations de synthèse et d'analyse de signatures, le rôle que le filtrage spatial comptait jouer dans ces deux opérations sur des signatures temporelles, caractéristiques d'un comportement. Une expérimentation, menée sur des attitudes de visages, observées conjointement en visible et infrarouge, a démontré les potentialités de notre approche.

Cette étape accomplie, il nous reste à présenter comment réaliser de telles opérations de filtrage, compte-tenu des données sur lesquelles nous sommes amenés à travailler.

Le cadre de notre travail s'insère comme une poursuite de travaux déjà entrepris par le département d'optique de l'ENST Bretagne sur les corrélateurs optiques pour la reconnaissance de formes, notamment la réalisation d'un démonstrateur compact *«temps réel»*.

Ainsi, le travail que nous présentons n'est qu'un des aspects du travail d'équipe entrepris depuis quelques années : conjointement aux architectures, lesquelles développent les algorithmes, les implantations se poursuivent. L'une d'elles illustrera (travail de A. Al Falou) une des approches présentées dans le premier chapitre.

Malgré tout, d'une vigilance particulière sur l'arrivée sur le marché de composants propres à chaque implantation, nous nous laisserons aller à quelques anticipations sur le devenir de notre traitement *«analyse & synthèse de signatures»*, proposant quelques applications originales.

4.2 Un traitement simple pour des implantations duales

Nous avons délibérément choisi de nous placer dans le cadre d'un traitement particulier : le filtrage spatial. Les différents développements que nous avons proposés pour ce dernier n'ont, à aucun moment, fait référence aux contraintes technologiques. Revenons dans un contexte plus pragmatique : l'implantation des deux filtrages spatiaux, tels qu'ils

apparaissaient dans le schéma général 2.12. Nous y décrivions deux filtrages successifs conduisant tout d'abord à la synthèse d'une signature temporelle, puis à son analyse.

Si nous nous penchons sur la réalisation «physique» de ces deux filtrages, nous constatons qu'ils peuvent être accomplis :

- soit par un unique dispositif de filtrage, qui, alternativement, passe d'un mode «synthèse» à un mode «analyse» ;
- soit par deux dispositifs disjoints, cette approche permettant de réaliser l'analyse pendant la synthèse.

Des contraintes, liées à l'application, imposent le premier choix ; si la masse de données est trop importante pour que nous puissions la confier à un seul «montage», deux «montages» sont indispensables. . . *A contrario*, un problème de place ou de bande passante (sur une carte numérique dédiée ou des flux de données associés, sur un montage optique) peut également nous contraindre à ne disposer que d'un seul montage permettant le filtrage optique. Les contraintes pécuniaires rentrent naturellement dans ce dernier cas.

Aussi supposons-nous que nous ne disposons que d'un seul et unique montage pour réaliser les opérations de synthèse et d'analyse de signatures.

Revenons aux opérations que nous rencontrons, toutes approches confondues, à la réalisation de nos filtrages :

1	un formatage des données
2	une transformation de Fourier
3	le produit du spectre par une référence appartenant à une base
4	le produit du spectre par un filtre d'ondelette
5	le calcul du module
6	l'élévation à une puissance k
7	une transformation de Fourier inverse
8	l'extraction du <i>PCE</i> d'une corrélation
9	l'extraction d'un parcours d'un résultat de corrélation

Parmi ces opérations, lors de la synthèse, seules certaines sont utilisées suivant les approches :

- approche «PCE» : opérations **1, 2, 3, 4, 7 et 8** ;
- approche «Parcours Fractal» : opérations **1, 2, 4, 5, 6, 7 et 9** ;

L'analyse, quant à elle, n'utilise que les opérations **1, 2, 3, 7 et 8**.

Nous allons donc proposer, pour chacune des implantations possibles, le composant permettant de réaliser pareilles opérations.

4.3 Implantation Optique

4.3.1 Généralités

Le filtrage spatial est une opération naturelle en optique depuis qu'Abbe et Porter ont proposé, dès le siècle dernier, de «filtrer» certaines fréquences spatiales d'une grille

bidimensionnelle afin de reformer par la suite une grille unidimensionnelle.

Ces montages ont été développés par la suite. En effet, disposant désormais d'une source de lumière cohérente puissante, le laser, et d'une possibilité d'enregistrer la phase par holographie, le filtrage spatial, initialement de nature «bloquante» (seules quelques fréquences étaient isolées), les applications gagnèrent le vaste domaine de la reconnaissance de formes.

Le montage de filtrage spatial tel que Abbe et Porter l'ont proposé n'a pas évolué dans sa structure (figure 4.1) : il est composé d'une source lumineuse monochromatique éclairant par une onde plane un plan d'entrée, sur lequel est disposé l'objet à filtrer. Cet objet est «codé» par sa transparence. Le fait d'éclairer l'objet par une onde plane permet de se placer dans le cadre de l'approximation de Fraunhofer, donc de la diffraction à l'infini. La diffraction d'une onde plane frappant un objet provoque la formation à l'infini de la transformation de Fourier de cet objet. Cette TF est rapportée à distance finie par une première lentille, agissant comme un transformateur de phase. Le spectre de l'objet formé dans le plan focal image de la première lentille est filtré par une fonction de transmittance : il s'établit ainsi le produit entre le spectre de l'objet et le filtre. Une seconde lentille est placée de façon à faire coïncider le plan focal objet avec le spectre filtré. Elle agit comme un second transformateur de phase, pour ramener finalement dans son plan focal image le résultat de l'image filtrée. A ce stade, il convient de préciser que la représentation de l'image filtrée que nous obtenons n'est que le carré de son module.

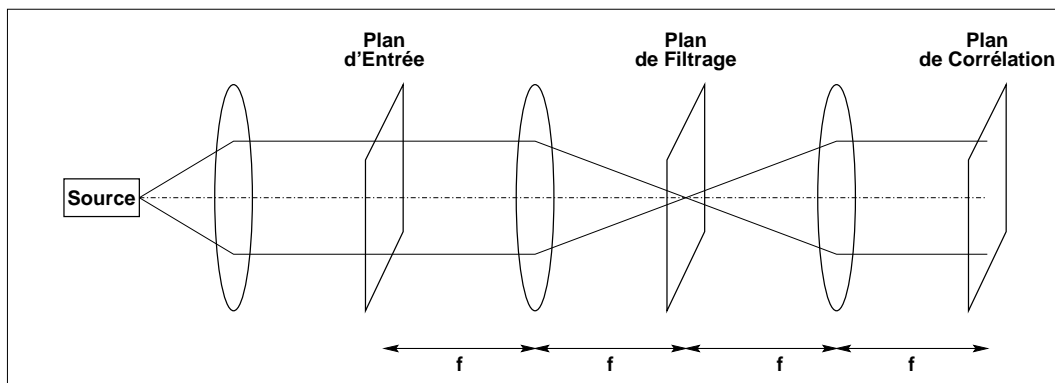


FIG. 4.1 – **Le montage «4f», base du filtrage optique** : le montage «4f» illustre l'action de la diffraction comme transformateur de Fourier. Cette transformation de Fourier naturelle est utilisée pour filtrer le spectre de l'objet placé dans le plan objet par une fonction de transmittance dans le plan de filtrage. Une nouvelle diffraction sur le spectre filtré opère une seconde transformation de Fourier, permettant de revenir dans l'espace direct.

Ainsi, la diffraction réalise, en nous plaçant dans le cadre de l'approximation de Fraunhofer, une opération de transformation de Fourier à l'infini, une lentille permettant de ramener cette diffraction à distance finie. Les opérations de transformation de Fourier ne semblent pas un problème dans l'accomplissement de notre filtrage spatial. Qu'en est-il des autres opérations que nous devons accomplir ?

4.3.2 Implantation des différentes opérations

Nous venons de préciser comment réaliser les opérations de transformation de Fourier. Nous avons également évoqué la mise dans le plan d'entrée de l'image à filtrer et la fonction de transmission qu'il nous était nécessaire de coder pour réaliser le produit entre le spectre et le filtre dans le plan de filtrage.

Les transformations de Fourier directe et inverse

L'approximation de Fraunhofer nous permet de considérer la diffraction d'un objet éclairé par un onde plane comme sa transformée de Fourier, la lentille jouant alors le rôle de transformateur de phase.

Dans notre filtrage, nous devons successivement passer de l'espace direct à l'espace des fréquences puis de l'espace des fréquences à l'espace direct. Ce passage est assuré par la TF , puis la TF inverse.

Pour la TF standard, nous n'avons aucune difficulté : la diffraction de Fraunhofer assure l'opération. En ce qui concerne maintenant la TF inverse, revenons à sa définition ; dans le chapitre destiné au filtrage, nous avons explicité la TF à deux dimensions par l'équation 1.2. La transformation inverse est alors donnée par :

$$f(x, y) = \overline{\mathcal{F}}_{F(\nu_x, \nu_y)}(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(\nu_x, \nu_y) e^{2i\pi(\nu_x x + \nu_y y)} d\nu_x d\nu_y$$

Si nous changeons x en $-x$ et y en $-y$, nous retrouvons l'expression d'une TF directe. Ainsi :

$$\overline{\mathcal{F}}_{F(\nu_x, \nu_y)}(x, y) = \mathcal{F}_{F(\nu_x, \nu_y)}(-x, -y)$$

Ainsi, la TF inverse se réduit-elle à la symétrique de la TF par rapport à l'origine. Nous n'avons donc pas besoin d'une transformée de Fourier inverse pour revenir dans l'espace direct. Une transformée de Fourier suffit, à condition qu'une symétrie centrale soit appliquée par la suite.

$$\boxed{TF \longrightarrow \text{Diffraction \& Lentille}}$$

Le formatage des images

Le plan d'entrée comporte donc les images sur lesquelles nous entendons réaliser un filtrage. Ces images sont rafraîchies à une certaine cadence (la fréquence video de 25 images par seconde semble suffisante), et doivent comporter la meilleure dynamique possible. Les modulateurs spatiaux de lumière électriquement adressés (souvent notés EASLM pour «*Electrically Addressed Spatial Light Modulator*»), à cristaux liquides nématiques, répondent à ces critères de rafraîchissement et de dynamique. Toutefois, bien qu'ils ne demeurent pas les seuls et uniques composants à proposer de telles opérations, ils sont les plus utilisés, malgré une fréquence de rafraîchissement ne pouvant excéder quelques dizaines de hertz.

$$\boxed{\text{Formatage des Images} \longrightarrow \text{EASLM nématique}}$$

Le produit par un filtre

Le produit n'est pas, comme d'autres opérations arithmétiques élémentaires (comme l'addition et la différence), difficile à implanter optiquement : la superposition du spectre à une fonction de transparence suffit à réaliser pareille opération.

Intéressons-nous tout d'abord à la nature de la transmittance nécessaire au filtrage. En effet, tout dépend des caractéristiques qui lui sont demandées. De fait, pour réaliser un blocage de certaines fréquences, un masque suffit. Si nous voulons par contre intégrer une fonction plus évoluée, un simple blocage est trop sommaire. Il est donc indispensable de procéder à un codage, autant en intensité (niveaux de gris) qu'en phase. Ce codage de la phase est plus problématique que les niveaux de gris, lesquels peuvent s'émuler par un contrôle de la densité des points formant les filtres (une photographie, par exemple, dont les grains d'argent bloqueront plus ou moins, selon leur densité, l'énergie qui traverse le film). Pour la phase, une première méthode peut consister à réaliser l'hologramme correspondant à la transformée de Fourier du filtre qui doit s'intégrer dans le plan de filtrage. Une autre solution propose encore de réaliser un composant optique diffractif avec un nombre de niveaux de phase suffisant. Une troisième utilise enfin des cristaux photoréfractifs [33].

Toutefois, au-delà de la nature et de la synthèse du filtre, un point apparaît crucial pour notre application : sa reconfigurabilité. Les filtrages successifs ont pour objectif de réaliser le filtrage, ou la corrélation, entre l'image présentée dans le plan d'entrée et un ensemble de filtres. Une solution serait d'intégrer les différentes transformées de Fourier correspondant aux filtres dans un même hologramme, mais le filtre resterait de toute façon statique. La mise à disposition d'un EASLM à cristaux liquides nématiques permettrait une certaine dynamique sur le codage des signaux, mais un choix serait à faire entre niveaux de gris et niveaux de phase. De plus, le rafraîchissement n'excédant pas quelques dizaines de hertz, nous ne pourrions pas réaliser un nombre de filtrages suffisant pour nos applications. Parmi les composants possédant de telles caractéristiques, nous avons des EASLM à cristaux liquides ferroélectriques qui permettent une fréquence de plusieurs kHz . Utilisables en amplitude ou en phase, ils ont toutefois le principal handicap d'être binaires.

Les filtres utilisés dans le cadre de la synthèse de la signature, étaient des filtres d'une base de références F_m et des filtres d'ondelette G_l . Les filtres issus d'une base de références sont des filtres de phase, choix établi dans le premier chapitre. Pour les filtres d'ondelette, nous constatons, sur la figure 1.20, leur nature blocante.

Ainsi, devons-nous réduire nos filtres de phase à des filtres binaires de phase pour la base F_m , mais également pour la base B_p . Par contre, nous devons réduire les filtres d'ondelette G_l à des filtres binaire d'amplitude. Le choix du chapeau mexicain est assez adapté à une telle restriction, son spectre étant réel pair.

En conclusion :

Filtre de référence \longrightarrow EASLM ferroélectrique en binaire de phase

Filtre d'ondelette \longrightarrow EASLM ferroélectrique en binaire d'amplitude

Si nous disposons d'EASLM à cristaux liquides antiferroélectriques, nous avons trois niveaux de codage : -1 , 0 et 1 . Nous possédons ainsi un filtre binaire de phase (-1 et 1) et un filtre binaire d'amplitude (0 et 1).

Filtres de référence & d'ondelette \longrightarrow EASLM antiferroélectrique ternaire

Module & élévation à une puissance k

La non-linéarité introduite dans le plan de filtrage consiste en une réduction du spectre à son module, puis son élévation à une puissance k . Une telle opération de calcul du module peut se réaliser en recueillant, dans le plan de filtrage, l'intensité lumineuse du spectre avec une caméra. L'énergie étant le carré du module du spectre, un traitement sur l'image obtenue est nécessaire avant d'utiliser un autre SLM pour réaliser optiquement la seconde transformation de Fourier. Une telle méthode a été expérimentée par Javadi dans le cadre d'un corrélateur à transformée conjointe. Toutefois cette approche demande deux conversions – analogique-numérique puis numérique-analogique – auxquelles s'ajoute un traitement numérique.

Un OASLM (Montage 4.3) – pour «*Optically Adressed Spatial Light Modulator*» – semble adéquat pour ce genre de tâche, puisqu'il réalise ces deux opérations de calcul du module puis d'élévation à une puissance k choisie [18]. Toutefois un OASLM demande à être lu ; en effet, lorsqu'un éclairage atteint l'OASLM, il provoque sur la face arrière un changement de réflectivité. Ce changement de réflectivité demande donc une seconde source lumineuse pour la lire.

Ainsi :

Module & puissance k \longrightarrow OASLM

Extraction de PCE ou de parcours fractal

Cette opération d'extraction est confiée à une caméra ou un PSD¹ (dans le cas où seul un pic est à extraire). Le signal recueilli est soit l'image filtrée, soit le plan de corrélation. Un balayage y est ensuite nécessaire pour déterminer le maximum et sa position, ainsi que l'énergie totale. Un simple rapport permet ensuite de conclure sur le PCE . Quant au parcours fractal, il est assuré par un balayage particulier de l'image.

Ainsi :

Extraction $0D$ ou $1D$ \longrightarrow Caméra

Conclusion

Nous avons donc, pour chacune des opérations formant le filtrage spatial, présenté quels étaient les outils «physiques» dont nous disposions : la diffraction à l'infini & la lentille pour les TF , l'EASLM nématique pour le plan d'entrée, l'EASLM ferroélectrique

¹Position Sensitive Device

pour les filtres de référence et d'ondelette, l'OASLM pour la non-linéarité et la caméra pour l'extraction finale du *PCE* ou du parcours de l'image filtrée.

Ainsi, l'approche «PCE» se réduit-elle au montage présenté sur le schéma 4.2 et l'approche «Parcours Fractal» à celui présenté schéma 4.3.

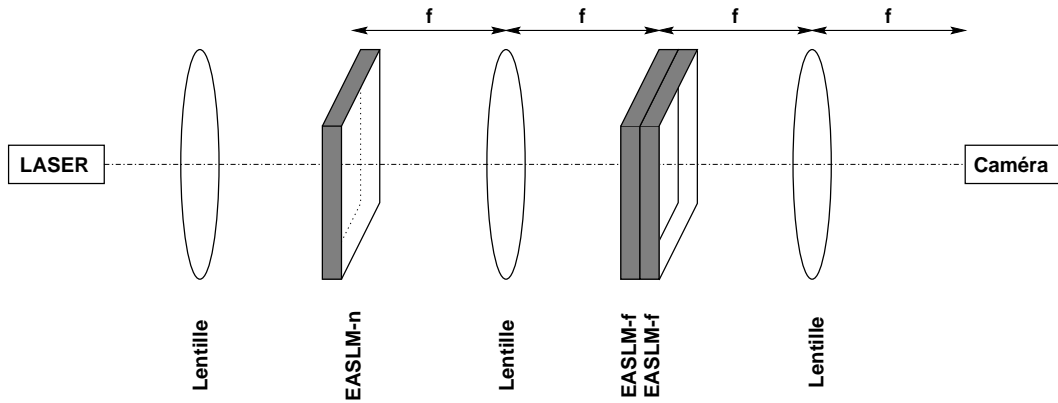


FIG. 4.2 – **Montage optique de l'approche «PCE»** : un premier EASLM nématique permet de placer l'image à filtrer dans le plan d'entrée; deux EASLM ferroélectriques sont utilisés pour présenter les filtres de références et les filtres d'ondelette; une caméra assure la conversion de l'image filtrée pour extraire le *PCE*.

4.3.3 Perte de dynamique & effets de binarisation

L'implantation optique du filtrage permettant la synthèse et l'analyse supposait le passage d'un filtre de phase à un filtre binaire de phase et d'un filtre d'ondelette réelle à un filtre binaire d'amplitude. De plus, l'image placée dans le plan d'entrée disposait désormais d'une dynamique limitée, fixée à 6 bits, comme le proposent certains composants commerciaux.

La réduction de la dynamique associée au plan d'entrée, lorsqu'elle est réduite à 6 bits ($2^6 = 64$ niveaux) est illustrée figure 4.4-a.

La binarisation des filtres de phase utilise une méthode des plus brutales, ne disposant que des niveaux de phase 0 et π , ou des valeurs 1 et -1 . Nous prenons -1 si la valeur réelle du pixel du filtre est négative et $+1$ dans le cas contraire. L'effet d'une telle binarisation est présentée figure 4.4-b.

La binarisation des filtres d'ondelette emploie par contre une méthode un peu plus fine. En effet, le filtre étant passif, ses valeurs sont nécessairement inférieures ou égales à l'unité. Nous disposons donc des valeurs de binarisation 0 et 1. Pour toute valeur supérieure à $1/2$, le pixel est mis à 1, 0 dans le sens contraire. L'erreur faite lors de la binarisation est reportée sur les pixels voisins. Evidemment, seuls les pixels n'ayant pas «subi» de binarisation voient l'erreur se reporter sur eux. Nous avons représenté sur la

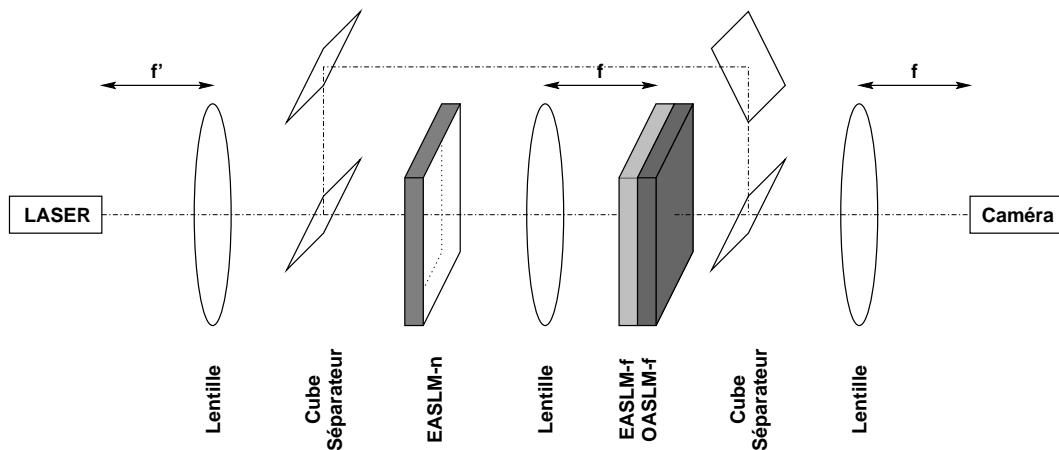


FIG. 4.3 – **Montage optique de l’approche «Parcours fractal»** : un premier EASLM nématique permet de placer l’image à filtrer dans le plan d’entrée ; un EASLM ferroélectrique est utilisé pour présenter les filtres d’ondelette ; l’OASLM réalise l’opération non-linéaire de réduction du spectre à son module, à une puissance k ; une partie du faisceau LASER, divisée par le premier cube séparateur, éclaire grâce au second cube la face arrière de l’OASLM. Le spectre filtré non-linéairement est diffracté avant de former sur une caméra le résultat duquel est extrait le parcours fractal.

figure 4.4-c une telle binarisation.

Le tableau 4.1-a présente les résultats de la corrélation entre le visage «Emmanuel» et quatre autres visages formant la base de référence, pour différents niveaux de détails, tenant compte des contraintes liées à l’implantation. Le tableau 4.1-b établit le rapport entre les résultats obtenus sans contraintes et ceux avec contraintes.

$PCE (\times 10^{-3})$					Rapports entre PCE				
Niveau	Référence				Niveau	Référence			
	1 ^e	2 ^e	3 ^e	4 ^e		1 ^e	2 ^e	3 ^e	4 ^e
(a) 4	0,08	0,13	0,08	0,08	4	1,8	1,3	1,7	1,7
2	0,12	0,17	0,10	0,07	2	2,4	2,2	1,8	3,7
1	0,22	0,38	0,13	0,26	1	1,2	3,2	1,4	1,8
1/2	0,17	1,4	0,15	0,35	1/2	1,6	4,4	2,5	2,7
1/4	0,24	5,3	0,19	0,48	1/4	2,3	4,3	1,4	2,2

TAB. 4.1 – **Influence des réductions de dynamique sur les PCE** : les deux tables représentent les $PCE (\times 10^{-3})$ et les rapports entre les PCE obtenus ici et ceux du tableau 1.1. Quatre images forment une base de références, lesquelles sont corrélées avec une image correspondant à la seconde référence, pour différents niveaux de contraction.

Ainsi la réduction de dynamique due à l’implantation provoque-t-elle une baisse très sensible du PCE : un facteur 4 est généralement constaté. Toutefois le contraste reste

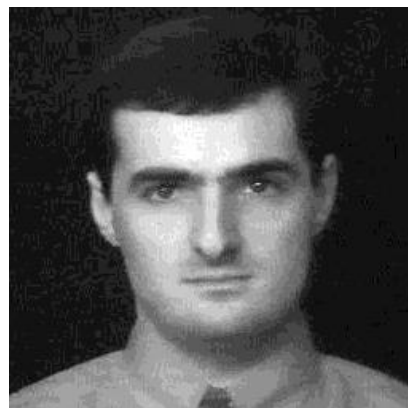
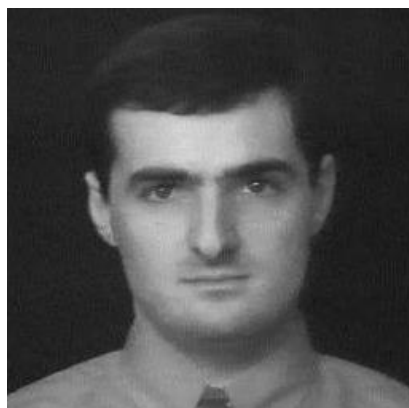
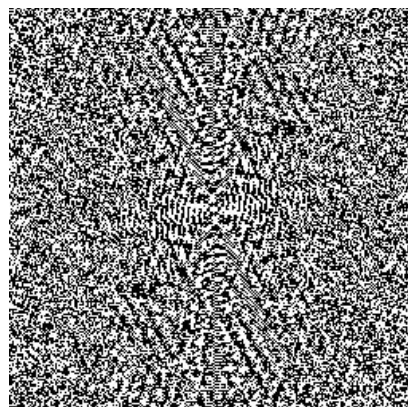
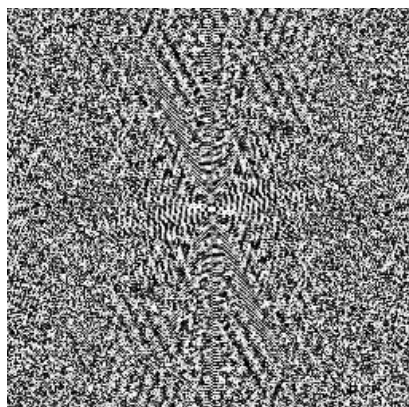
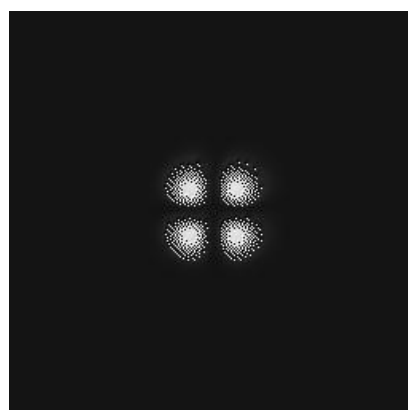
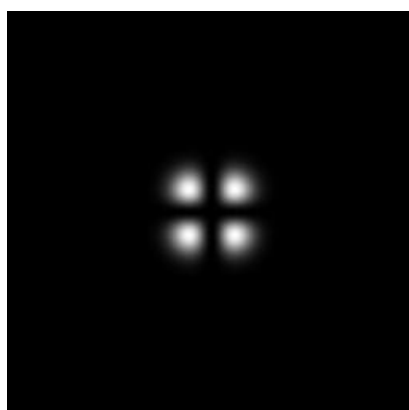
(a) Réduction à 64 niveaux de gris**EASLM-n****(b) Réduction à 2 niveaux de phase****EASLM-f****Binarisation par Seuillage****(c) Réduction à 2 niveaux d'amplitude****EASLM-f****Binarisation par Diffusion d'Erreur**

FIG. 4.4 – Réduction de dynamique sur les éléments : le codage des images placées dans le plan d'entrée est réduit à 6 bits, soit 64 niveaux (a), des filtres de phase (b) et des filtres d'ondelette (c) à un unique bit. La binarisation des filtres de phase est réalisée par seuillage, tandis que celle des filtres d'ondelette l'est par diffusion d'erreur.

encore assez important pour que nous puissions conclure sur la référence activée sans ambiguïté.

4.3.4 Quelques éléments d'implantation

La performance des montages optiques de filtrage spatial est, nous venons de le rappeler, intimement liée à celles des éléments qui la composent [16]. En effet, le filtrage optique ne peut se suffire à lui-même : des interfaces électro-optiques sont indispensables à la mise en forme des données électroniques sur un support physique, tandis que d'autres, opto-électroniques, assurent le retour des résultats afin de réaliser une prise de décision.

Si nous nous plaçons à l'instant présent, nous disposons d'un plan d'entrée formé d'un EASLM nématique d'une taille de 640×480 pixels, avec un facteur de remplissage de 61% et d'un contraste de 100. Nous pouvons donc espérer une dynamique identique à celle que nous avons citée, c'est-à-dire de 6 bits. Sa fréquence de rafraîchissement est la fréquence video, soit 25 Hz . Pour le codage des filtres, nous disposons d'EASLM ferroélectriques de 256×256 pixels, cadencés à la fréquence maximale de 3 kHz , lesquels nous servent à coder le binaire de phase et le binaire d'amplitude [24] [36].

La figure 4.5 présente le mécanisme permettant d'utiliser deux EASLM identiques binaires, l'un en phase, l'autre en amplitude à la suite : le premier EASLM (EASLM-bp) possède ses axes de cristal liquide de part et d'autre de la direction de polarisation de la lumière incidente ($22, 5^\circ$ de chaque côté de l'axe \mathbf{x}). La commutation du cristal liquide réalise une rotation du plan de polarisation de 45° . Le cube polarisant opère une projection suivant l'axe \mathbf{y} des directions $\mathbf{1}$ et $\mathbf{2}$, lesquelles forment ensuite les directions $\mathbf{1}'$ et $\mathbf{2}'$. Ces deux directions se retrouvent suivant l'un des axes neutres du cristal liquide. Dans la position \mathbf{a} du second EASLM, les directions restent constantes. La position \mathbf{b} provoque par contre une rotation des polarisations de 90° . Seules les polarisations $\mathbf{1}'\mathbf{b}$ et $\mathbf{2}'\mathbf{b}$ subissent la réflexion du cube séparateur. Ainsi est-il possible, par deux EASLM identiques, d'introduire le filtrage de phase et le filtrage d'amplitude correspondant respectivement aux références et aux ondelettes. Le schéma 4.5 présentant la structure du plan de filtrage nécessite donc un rayonnement polarisé dans son plan de filtrage (une telle condition est réalisée par les lasers classiques, lesquels sont polarisés rectilignement), deux EASLM identiques, un cube séparateur polarisant.

Enfin, dernière étape afin de réaliser ce montage «4f», sachant désormais quoi placer dans le plan de filtrage, il nous faut savoir quelle focale de lentille mettre entre le plan d'entrée et le plan de filtrage, pour faire correspondre le lobe principal de diffraction d'un pixel de l'EASLM nématique du plan d'entrée à la taille du EASLM ferroélectrique du plan de filtrage. Pour une taille inter-pixel de l'EASLM nématique a , une longueur d'onde $\lambda = 6328 \text{ \AA}$ (laser Hélium-Néon) et une focale de lentille f , la taille du lobe principal de diffraction est D_a (équation 4.1).

$$D_a = \frac{2\lambda f}{a} \quad (4.1)$$

Cette taille D_a doit correspondre à la dimension de l'EASLM ferroélectrique servant de

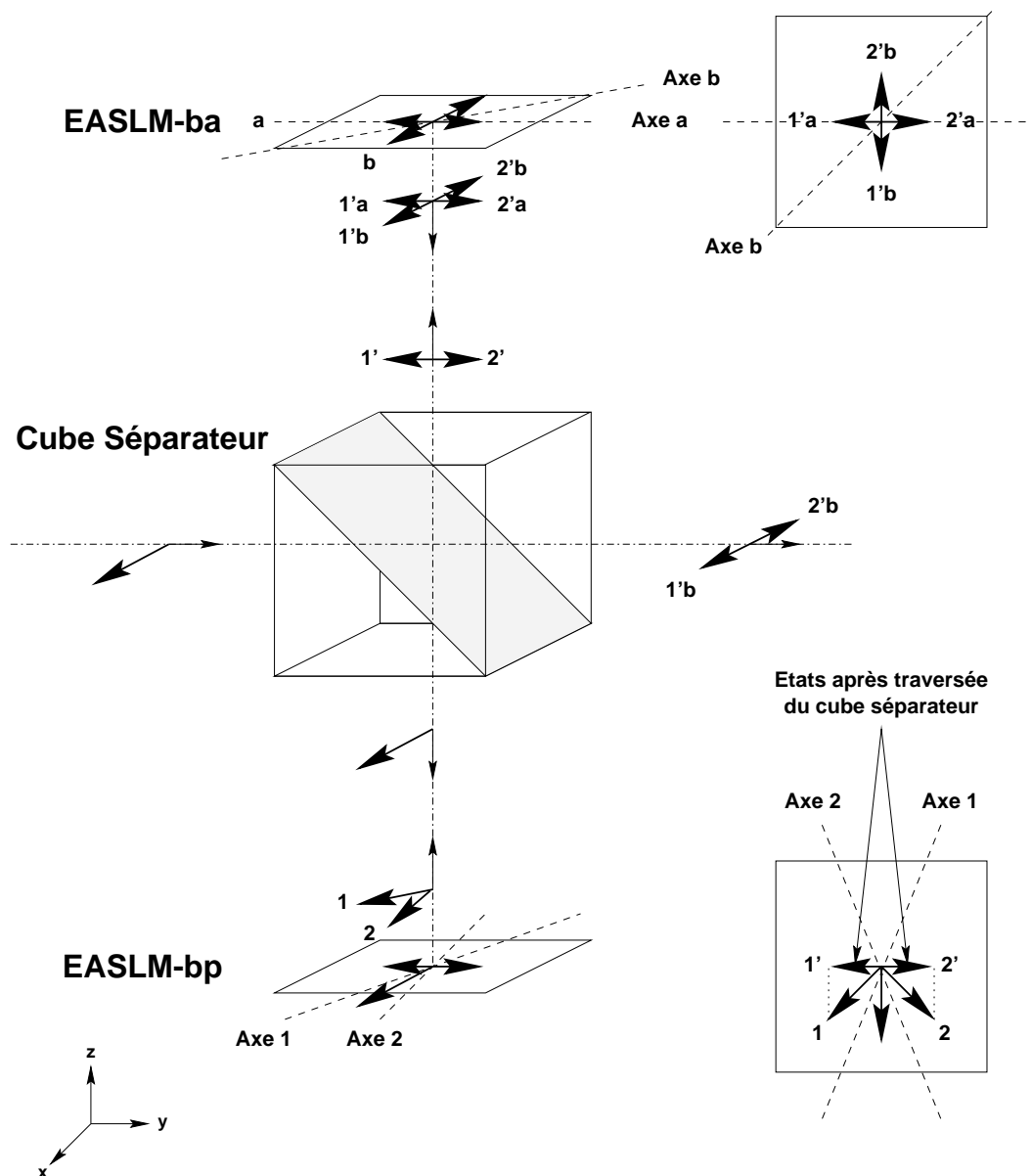


FIG. 4.5 – **Couplage de deux EASLM pour un double filtrage** : les deux EASLM (EASLM-bp et EASLM-ba) servent au codage des filtres de références et d'ondelettes. Un cube séparateur polarisant est nécessaire pour réfléchir ou transmettre les ondes suivant leur polarisation. La commutation de l'EASLM-bp permet la création des états **1** et **2**, puis **1'** et **2'** après la traversée du cube séparateur. La commutation de l'EASLM-ba propose ensuite les états **a** et **b**. Après passage au travers du cube polarisant, seuls les états **1'b** et **2'b** sont transmis.

filtre. Ainsi, si la dimension inter-pixel de l'EASLM de filtrage est d , sa taille 256×256 pixels, nous avons :

$$\frac{2\lambda f}{a} = 256 d \quad \text{ou} \quad f = \frac{128 d a}{\lambda} \quad (4.2)$$

Dans la pratique, pour adapter le grandissement de la tache principale de diffraction à la dimension de l'EASLM servant de filtre, nous utiliserons deux lentilles convergentes.

Quant à la focale de la seconde lentille, réalisant la transformation de la phase de la TF à l'infini du spectre filtré, elle se calcule en considérant de la même manière la distance inter-pixel d de l'EASLM.

Enfin, la réalisation du double filtrage de phase et d'amplitude nécessitait un certain espace entre les deux EASLM et le cube séparateur. Entre ces pièces des effets de diffraction se produisent. Bien que sur l'EASLM en amplitude, les effets soient moins préjudiciables que sur l'EASLM en phase, de par la nature blocante de son action, il serait peut-être nécessaire de procéder à une imagerie entre les deux EASLM.

Pour conclure, avant de revenir à quelques résultats expérimentaux, notons que de tels corrélateurs ont été réalisés [35] et utilisés dans quelques applications [22].

4.3.5 Éléments d'une validation expérimentale

L'ensemble des filtres développés dans le premier chapitre a fait l'objet d'implantations réelles. Ce travail fait partie de la thèse de doctorat d'A. Al Falou.

Les résultats obtenus confirment les prévisions des simulations. Nous ne présenterons ici qu'un exemple dans un des cas les plus complexes de niveau de codage, c'est-à-dire un filtre composite intégrant des porteuses.

Quatre références ont été intégrées dans le filtre (figure 4.7-c), chacune disposant de sa propre porteuse, correspondant chacune au quart du plan de corrélation. Les plans de corrélation 4.7-d de deux empreintes 4.7-a de spectres 4.7-b (l'une appartenant aux références intégrées au filtre, l'autre «intruse») dénotent la sélectivité du filtrage et confirment les résultats obtenus lors des simulations numériques.

Les matériels utilisés (figure 4.6) pour cette expérimentation sont :

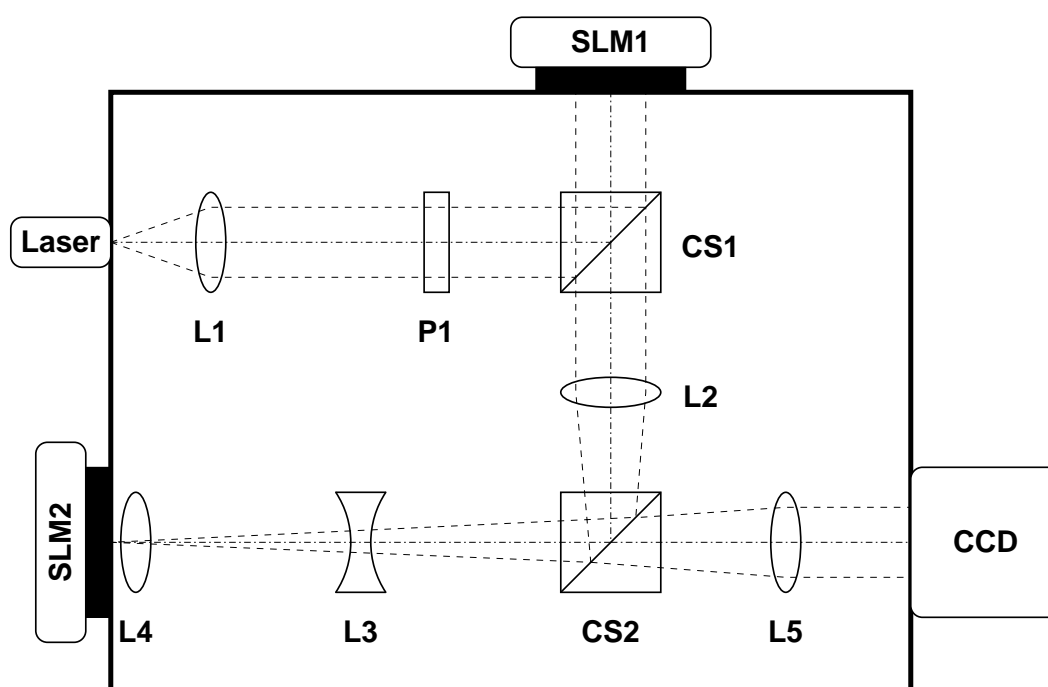


FIG. 4.6 – Montage expérimental de réalisation d'un corrélateur compact : deux EASLM binaires sont utilisés, le premier en binaire de phase, le second en binaire d'amplitude.

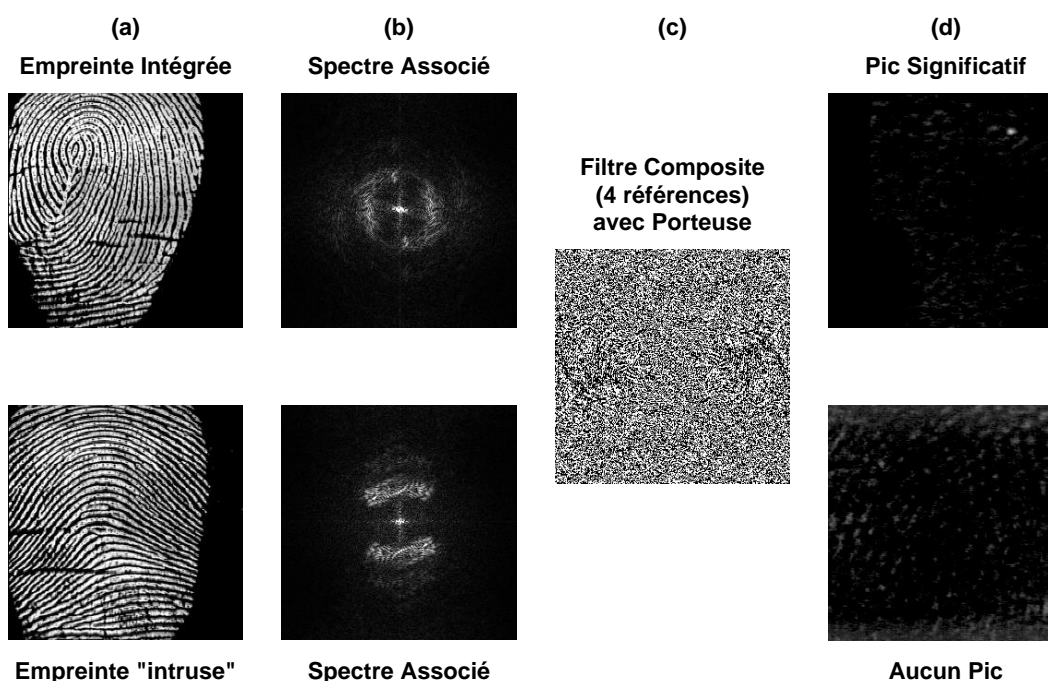


FIG. 4.7 – Implantation d'un filtre composite avec contrôle de positionnement : deux empreintes sont présentées à un filtre composite (chacune des références intégrées au filtre dispose d'une porteuse différente). Le plan de corrélation répond favorablement uniquement pour la référence que le filtre intègre.

Élément	Nature	Caractéristiques
Source	Laser	Hélium-Néon (6328Å)
Plan d'entrée SLM1	FLC-SLM binaire d'amplitude	de marque Displaytech Dimension de 256×256 pixels Taille du pixel de $14 \times 14 \mu m$ Facteur de remplissage de 85%
Plan de filtrage SLM2	FLC-SLM binaire de phase	de marque Displaytech Dimension de 256×256 pixels Taille du pixel de $14 \times 14 \mu m$ Facteur de remplissage de 85%
Plan de Corrélation CCD	Caméra CCD	Analogique Taille du pixel de $13 \times 13 \mu m$

La lentille **L1** assure la mise en forme du faisceau issu du laser. Le premier cube séparateur polarisant (**CS1**) permet l'utilisation du premier EASLM (**SLM1**) en binaire d'amplitude, tout comme le second cube séparateur (**CS2**) avec le second EASLM (**SLM2**) en binaire de phase. Les lentilles **L2**, **L3** et **L4** permettent la formation de la TF de l'image de **SLM1** sur **SLM2**. La dernière lentille **L5** image sur la caméra (**CCD**) le résultat final de la corrélation.

4.4 Implantation Numérique

La majorité des résultats que nous avons présentés sont donc des simulations, la plupart étant réalisée sur des données réelles.

Nous avons montré dans les précédents paragraphes que les interfaces opto-électroniques disponibles permettaient de réaliser les opérations de filtrage conduisant à la synthèse puis à l'analyse des signatures, par filtrage spatial.

Nous allons désormais nous intéresser aux opportunités que le numérique offre au filtrage spatial, au travers de composants spécifiques que sont les DSP (pour «*Digital Signal Processing*»).

En effet, nous manipulons des images, lesquelles sont des matrices réelles dont chaque point, indicé (i, j) , portait l'information sur l'image. Les manipulations présentées 4.2 étaient à accomplir pixel à pixel, par un balayage de toutes les valeurs possibles de i et j , telles que les produits entre spectre et filtre, l'élévation à une puissance et le calcul du PCE . Les autres nécessitaient d'introduire une relation moins sommaire entre les pixels, lors du formatage ou de l'extraction par parcours fractal. Toutefois, la totalité des pixels est concernée par le calcul de la TF . Cette opération est, de loin, la plus coûteuse en terme de temps de calcul. Pour la réaliser numériquement, Cooley et Tukey ont proposé un algorithme, la FFT (pour «*Fast Fourier Transform*», éliminant les redondances de calcul. Ainsi, d'une loi en N^2 sommes-nous passés en une loi en $N \log N$. Cet algorithme forme le cheval de trait de tout filtrage, de par les économies de calcul qu'il propose. Il forme donc l'épine dorsale de notre traitement par double filtrage : par synthèse puis par analyse de signatures.

4.4.1 Un coût de calcul non-négligeable

Ainsi, le filtrage dans l'espace de Fourier reporte les contraintes de temps de calcul liées à la convolution sur les transformations de Fourier.

Il nous faudra donc, si nous voulons réaliser nos opérations de filtrage, considérer :

- soit des images de tailles réduites :
le temps de calcul nécessaire évolue en $N \log N$, N étant le nombre total de pixels. Si nous réduisons la taille de nos images, par exemple d'un facteur 2 suivant les deux dimensions, notre temps de calcul se réduit d'un facteur $4 \log 4$, c'est-à-dire plus de 5 fois ;
- soit des processeurs numériques dédiés permettant de réaliser un nombre de *FFT* très supérieur à celui dont nous disposons sur un ordinateur standard.

Les processeurs de traitement numérique (traduction de «*Digital Signal Processor*»), ou DSP, sont en mesure de réaliser de telles opérations. Dans ce secteur pourtant très concurrentiel, la firme *Texas Instruments* a provoqué une petite révolution, voici deux ans, par la sortie d'une famille de DSP d'une puissance dix fois supérieure à celle de ses concurrents : la série des *TMS320C6x*. C'est en nous basant sur les performances d'un tel produit, disponible et d'un prix très raisonnable, que nous allons rechercher comment implanter notre double filtrage sur un dispositif numérique.

Pour fixer quelques idées sur les performances d'un DSP de la famille des *TMS320C6x*, nous présentons dans l'annexe D quelques estimations théoriques de la performance d'un tel DSP au calcul d'une FFT.

4.4.2 Une implantation sur DSP

Les performances du DSP *TMS320C6202*, permettent, en théorie, d'atteindre plusieurs centaines de corrélations par seconde sur des images dont la taille (128×64) correspond aux nôtres.

Cependant, si un unique DSP n'est pas suffisant pour réaliser la totalité de ces opérations, notamment si nous l'utilisons dans un mode «PCE» (très «gourmand» en FFT), nous avons la possibilité d'en cascader 4 à la suite, auxquels sont confiées les tâches suivantes :

- le 1^e : la FFT ;
- le 2^e : le produit avec filtres et/ou la non-linéarité ;
- le 3^e : la FFT ;
- le 4^e : l'estimation du *PCE* ou du parcours fractal.

Une structure à base de *pipeline* pourra, le cas échéant, être proposée (table 4.2). A partir du quatrième cycle, une corrélation complète est réalisée par cycle.

4.5 Comparaison entre implantations

Le traitement du signal a souvent opposé l'optique et le numérique. Pourquoi, dans notre étude, avons-nous voulu conserver un traitement le plus dual possible ?

N° DSP	Opération	Cycle					
		1	2	3	4	5	...
1 ^e	FFT						
2 ^e	\times ou $ ^{2k}$		1	2	3	4	...
3 ^e	FFT			1	2	3	...
4 ^e	<i>PCE</i> ou <i>Hc</i>				1	2	...

TAB. 4.2 – «**Pipelining**» de DSP : plusieurs DSP utilisés conjointement suivant une architecture de «*pipeline*» permettent de diviser le temps de calcul final.

La réponse à cette question se situe dans les données elles-mêmes :

- dans leur dimension ;
- dans leur dynamique ;
- dans leur rafraîchissement.

En effet, la principale force de l'optique réside dans sa capacité à réaliser une *TF* quasi-instantanée sur de grandes dimensions. Seul le rafraîchissement du codage du filtre fixe le délai séparant deux corrélations. Toutefois, les dynamiques sur lesquelles l'image, et surtout le filtre sont codés, s'avèrent, jusqu'à présent, des plus sommaires souvent binaires.

Quant au numérique, sa souplesse, sa facilité de mise en œuvre en font un outil très performant. Les DSP modernes ont permis, ces toutes dernières années, de gagner plusieurs ordres de grandeur en puissance brute. Exploités au maximum, les DSP pèchent face à l'optique sur les FFT, mais proposent des filtrages plus fins.

Ainsi, ces deux outils disposent d'avantages (flexibilité, dynamique pour le numérique, *TF* pour l'optique) et d'inconvénients (coût d'une *FFT* en numérique, limitation des interfaces en optique).

L'optique sera donc favorisée si de larges images sont à considérer ; dans le cas contraire, le numérique primera par sa flexibilité.

Enfin, existe un problème commun à nos deux implantations : l'interaction entre la couche de traitement et la couche de bases de données (figure 2.12). Devant des bases de références exhaustives, les interfaces électro-optiques et les bus (interactions avec les mémoires) constituent de véritables goulots d'étranglement. Une gestion de cet ensemble de bases doit être la plus judicieuse possible et nécessite une «hiérarchisation» des priorités.

Toutefois, dans le cas d'une implantation optique, si les mémoires sont également optiques [27], il est possible d'envisager une réduction partielle de ce goulot, renforçant par voie de conséquence l'intérêt d'une intégration physique du filtrage spatial.

4.6 Prospectives

Les seules prospectives que nous pouvons donner concernent les outils permettant l'implantation de notre traitement.

En optique, les valves optiques pallieraient le principal inconvénient du filtrage : une dynamique quasi-continue en phase et des fréquences de fonctionnement de plusieurs centaines de kHz de fonctionnement. Plusieurs ordres de grandeur, autant en vitesse qu'en finesse, seraient atteints.

En numérique, la loi de Moore² existe depuis plus de trente années et reste toujours valable. Les progrès récents sur les gravures de cuivre permettraient d'atteindre des fréquences d'horloge supérieures au GHz .

4.7 Conclusion

Ce chapitre nous a sensibilisés aux problèmes d'implantations, suivant les axes optique et numérique. Il apparaît que chacune d'elles dispose d'avantages (puissance brute de filtrage pour l'optique et flexibilité pour le numérique) et d'inconvénients.

Toutefois, ne pouvant conclure sans tomber dans une prospective dangereuse, nous entendons conserver notre traitement tel quel. Sa simplicité demeure sa force, les bases de données qui lui sont associées, son efficacité. Dans ce sens, il convient de s'intéresser préférentiellement aux données sur lesquelles nous devons travailler plutôt que cliver notre traitement.

²L'intégration des composants électroniques double tous les 18 mois

Conclusion

Nous avons donc, grâce au filtrage spatial, réalisé la **synthèse** de signatures temporelles. Ces dernières étaient caractéristiques de l'objet, mais surtout de son comportement. Un second filtrage spatial permettait l'**analyse** de ces signatures. Ainsi, un seul et unique outil, très simple, composé de deux TF , permettait, si des bases adéquates lui étaient associées, de réaliser la reconnaissance, autant de l'attitude que de l'objet, par cette étape intermédiaire que constitue la signature.

Au filtrage spatial traditionnel utilisé en reconnaissance de formes, a été associé un prétraitement linéaire. Ce dernier intégré au filtre, permettait d'extraire l'information à un niveau de détail déterminé. Les signatures ainsi formées intégraient donc, par leur nature, ce niveau de détail.

De plus, lorsque l'introduction d'une base de références n'était plus indispensable à la synthèse de la signature, un filtrage, non-linéaire cette fois, permettait d'établir une liaison entre des signaux différents. Ces liens, inscrits dans le plan d'intercorrélacion ainsi créé, étaient extraits par un parcours dont la nature – une fractale auto-similaire – conservait les liens de proximité lors de la transformation $2D \rightarrow 1D$.

Nous avons appliqué cette approche par double filtrages successifs à deux domaines différents : la reconnaissance d'attitudes de visages, puis la reconnaissance de bâtiments de surface. La signature, dans chacun des cas, et suivant l'horizon temporel dont elle disposait, était assez discriminante pour conclure sur la nature de l'attitude, l'identité de l'objet.

La simplicité du traitement que nous avons envisagé, le filtrage spatial, était délibérée; les rares contraintes qui nous étaient imposées nous permettent donc d'envisager des implantations numériques (à base de DSP) ou optiques (à base de FLC-SLM). Malgré les inconvénients respectifs de ces approches (traitement des FFT ou contraintes liées aux FLC-SLM), nous ne sommes pas en mesure, de façon objective, de nous projeter sur ce que nous réservent les futurs progrès technologiques des disciplines respectives.

En ce sens, le traitement que nous avons proposé, une synthèse, procédée d'une analyse de signature, par un outil simple, le filtrage spatial, trouverait son implantation *hic et nunc* si nous connaissions avec certitude la nature (dimensions, dynamique et fréquence) des données qu'il nous faut traiter. Seules, de telles contraintes posées par l'application seraient susceptibles de nous faire trancher en faveur d'une implantation.

Toujours est-il que, si nous revenons au traitement générique (figure 2.12), nous distinguons nos trois couches. Les deux premières, les couches d'observation et de traitement, viennent d'être évoquées. Quant à la couche de données, elle n'est pas à négliger, bien au contraire. Formant la couche la moins «volatile» de notre traitement, elle donne au filtrage spatial toute sa puissance. C'est sur son exhaustivité, sa qualité de conception que repose en partie le succès de notre approche. Il convenait donc de développer, comme nous nous y sommes attardés dans le premier chapitre, la manière de concevoir ces filtres formant ces bases.

Annexe A

Les ondelettes continues

A.1 Définition

Prenons une fonction ψ à valeurs réelles, de spectre Ψ , dont le «critère d'admissibilité» C_Ψ satisfait :

$$C_\Psi = \int_{-\infty}^{+\infty} \frac{|\Psi(X)|^2}{|X|} dX < \infty \quad (\text{A.1})$$

Le dénominateur $|X|$ de l'expression A.1 implique que le spectre $\Psi(X)$ soit nul en 0. De ce fait, la valeur moyenne de ψ sur $] -\infty, +\infty[$ est nulle (expression A.2).

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = 0 \quad (\text{A.2})$$

Il est donc possible de définir une famille d'ondelettes $\{\psi_{a,b}(x)\}$ à partir de $\psi(x)$ en posant :

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (\text{A.3})$$

Les coefficients a et b sont réels (a est positif) : a agit comme un facteur de contraction de l'ondelette et b comme une translation de cette ondelette suivant l'axe x .

Supposons une fonction f de carré sommable (donc appartenant à L^2), la «transformée en ondelette continue de f » s'exprime de la manière suivante :

$$W_f(a, b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{+\infty} f(x) \psi_{a,b}(x) dx \quad (\text{A.4})$$

A partir des $W_f(a, b)$, il est possible de revenir à la fonction $f(x)$ par la «transformée en ondelette inverse continue» :

$$f(x) = \frac{1}{C_\Psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} W_f(a, b) \psi_{a,b}(x) db \frac{da}{a^2} \quad (\text{A.5})$$

A deux dimensions, les expressions précédentes se généralisent d'elles-mêmes. Nous avons comme transformée continue en ondelette :

$$W_f(a, b_x, b_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \psi_{a,b_x,b_y}(x, y) dx dy \quad (\text{A.6})$$

et comme transformée inverse :

$$f(x, y) = \frac{1}{C_\Psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W_f(a, b_x, b_y) \psi_{a, b_x, b_y}(x, y) db_x db_y \frac{da}{a^3} \quad (\text{A.7})$$

avec :

$$\psi_{a, b_x, b_y}(x) = \frac{1}{|a|} \psi\left(\frac{x - b_x}{a}, \frac{y - b_y}{a}\right) \quad (\text{A.8})$$

La transformation d'ondelette continue est un filtrage ; nous retrouvons dans l'équation A.4 l'expression d'une corrélation entre $f(x)$ et ψ_a :

$$W_f(a, b) = f \otimes \psi_a(b) = f * \tilde{\psi}_a(b) \quad (\text{A.9})$$

où $\tilde{\psi}_a(x) = \psi_a(-x)$.

Ainsi, pour estimer la décomposition d'une fonction $f(x)$, il suffit de réaliser, pour un «niveau» a donné, le filtrage adapté entre la fonction f et l'ondelette ψ_a .

A.2 Illustrations

Comme ondelette simple, adaptée à une implantation autant optique que numérique, nous avons le «chapeau mexicain» ou «sombbrero». L'expression de cette ondelette $\psi(x)$ est la suivante :

$$\psi(x) = \frac{2}{\sqrt{3\sqrt{\pi}}} (1 - x^2) e^{-\frac{x^2}{2}} \quad (\text{A.10})$$

$$\Psi(X) = 4\pi^2 X^2 \sqrt{\frac{8\sqrt{\pi}}{3}} e^{-2\pi^2 X^2} \quad (\text{A.11})$$

Le coefficient $2/\sqrt{3\sqrt{\pi}}$ permet la normalisation à l'unité de l'«énergie» de l'ondelette.

La figure A.1 illustre pour quelques valeurs de a l'allure des $\psi_a(x)$ et de leurs spectres associés $\Psi_a(X)$.

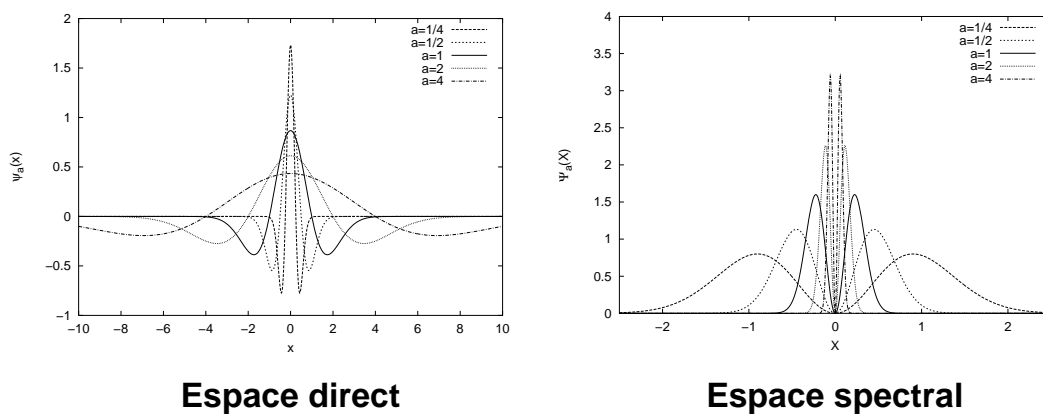


FIG. A.1 – **Ondelette du «Chapeau Mexicain»** : La fonction Ψ_a possède les propriétés d'une ondelette (moyenne nulle). Elle est représentée dans l'espace direct (à gauche) puis l'espace des fréquences (à droite), pour plusieurs valeurs de a .

Annexe B

Un Parcours Fractal : la Courbe de Hilbert

Parmi les transformations permettant de convertir une image en un vecteur, certaines possèdent des propriétés qu'il semblait intéressant d'introduire dans nos applications. La principale propriété de ces courbes en faisait d'intéressants outils pour «balayer» un plan suivant des motifs géométriques à caractère fractal. C'est pour cette raison que ces courbes sont appelées «*Space Filling Curve*».

B.1 La courbe de Hilbert

Hilbert[14], en 1890, présente un article sur un parcours disposant d'une structure auto-similaire, permettant de parcourir un réseau de points sur une surface carrée.

La figure B.1 illustre, pour trois images de taille 2×2 , 4×4 et 8×8 pixels, les parcours de Hilbert permettant de passer une fois et une seule sur chaque pixel. Nous retrouvons le motif élémentaire (un «U») se répétant sur chaque pavé de 2×2 pixels. Entre ces pavés, les liens assurent la continuité.

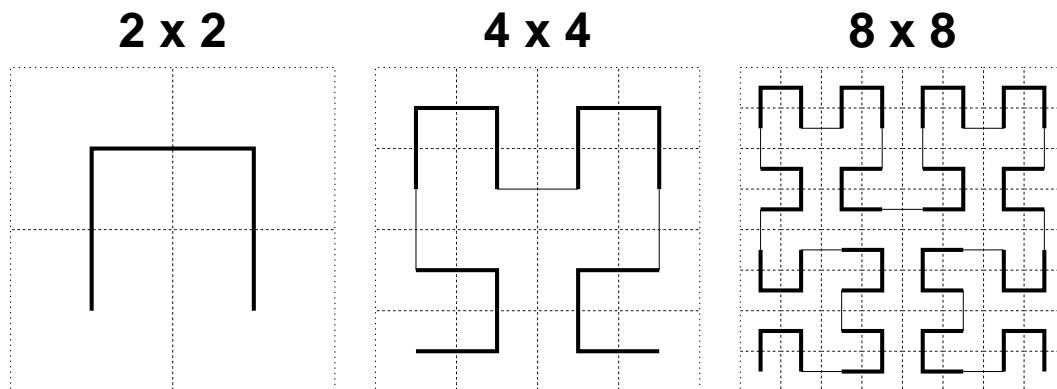


FIG. B.1 – **Parcours de Hilbert pour trois images** : sont représentés successivement les parcours de Hilbert pour trois images de taille 2×2 , 4×4 et 8×8

Nous imaginons donc que, pour appliquer un parcours fractal sur une image, il faut que cette dernière soit de $2^n \times 2^n$ pixels.

B.2 Un algorithme de synthèse récursif

La figure B.1 montre donc le «noyau» élémentaire du parcours. Reste maintenant à déterminer, pour une image carrée de taille $N \times N$ quelconque ($N = 2^n$), l'algorithme permettant de passer une fois et une seule sur ce maillage.

Comme le «noyau» élémentaire se retrouve quelle que soit la taille, nous allons utiliser un algorithme récursif (donc s'appelant lui-même), à partir de la procédure citée par Andrew Cumming sur son site WWW¹.

```

procedure hilbert(x0,y0,xis,xjs,yis,yjs,n)

if (n>=tooBig) then
  {
    LineTo(x0+(xis+yis)/2,y0+(xjs+yjs)/2);
  }
else
  {
    hilbert(x0,y0,yis/2,yjs/2,xis/2,xjs/2,n+1)
    hilbert(x0+xis/2,y0+xjs/2,xis/2,xjs/2,yis/2,yjs/2,n+1)
    hilbert(x0+xis/2+yis/2,y0+xjs/2+yjs/2,xis/2,xjs/2,yis/2,yjs/2,n+1)
    hilbert(x0+xis/2+yis,y0+xjs/2+yjs,-yis/2,-yjs/2,-xis/2,-xjs/2,n+1)
  }
end procedure;

```

Dans cette procédure récursive, les différentes variables sont :

- (x0,y0) : coordonnées du coin en bas à gauche;
- (xis,xjs) : coordonnées du vecteur directeur x ;
- (yis,yjs) : coordonnées du vecteur directeur y ;
- n : la taille de l'image

Ainsi, pour parcourir un domaine de 64×64 pixels, nous appelons la procédure `hilbert` avec les paramètres :

```
hilbert(0,0,0,64,64,0,64)
```

B.3 Application

Jusqu'à présent, la principale application d'un tel parcours fractal était de modifier l'image. La binarisation d'une image pouvait être accomplie par un procédé de ce type, en diffusant l'erreur faite sur la binarisation au pixel suivant du parcours. Aucune direction n'étant privilégiée, l'erreur était diffusée dans tout le plan[26].

¹<http://www.dcs.napier.ac.uk/a.cumming/hilbert.html>

Annexe C

Réalisation d'une Base de Données Multi-Senseurs

La validation de notre dispositif de traitement d'attitudes issues de plusieurs capteurs nécessitait la conception d'une base de données présentant plusieurs visages sous plusieurs attitudes différentes.

C.1 Protocole de prise de vue

Les scénarii, exécutés par 4 personnes, sont à diviser en deux catégories :

- ceux de la base d'apprentissage ;
- ceux de la phase de test.

Voilà les scénarii de la phase d'apprentissage :

- rotation de la tête : de 90° vers la droite à 90° vers la gauche, puis de cette dernière position à une position normale à la caméra ;
- hochement de tête : de 45° vers le haut à 45° vers le haut, puis de cette dernière position à une position normale à la caméra ;
- sourire (dents invisibles) ;
- sourire (dents visibles) ;
- bâillement ;
- froncement de sourcils ;
- clignement d'yeux ;
- lecture d'une citation «*Regarder un atome le change, regarder un homme le transforme, regarder l'avenir le bouleverse*».

Les scénarii de la phase de test sont au nombre de 8 :

- rotation de la tête & clignement d'yeux ;
- hochement de tête & bâillement ;
- sourire (dents visibles) & clignement d'yeux ;
- froncement de sourcils & bâillement ;
- rotation de la tête & citation ;
- hochement de tête & citation ;
- clignement d'yeux & citation ;

- froncement de sourcils & citation.

Les classes à établir à partir des prises de vue visible et infrarouge sont de deux ordres :

- la classe «Visage» ;
- la classe «Attitude».

C.2 Matériel de prise de vue & de numérisation

Matériel utilisé pour la prise de vue :

- un camescope pour la prise de vue visible ;
- une caméra à infrarouge sensible à la bande de 3 à 5 μm ;
- un fond noir (la veste d'un costume de couleur anthracite) ;
- deux magnétoscopes VHS.

Matériel utilisé pour la numérisation des séquences video :

- un magnétoscope Akai VHS ;
- un ordinateur de type PC ;
- une carte d'acquisition video «Crunch It 2000» de «Diamond Multimedia».

Le tableau C.1 référence le nom des séquences numérisées pour la création des bases de données d'apprentissage et de test. Chacune des séquences est au format AVI dont chaque image est aux normes JPG. Les images, dont la chrominance a été supprimée lors de la numérisation, sont d'une taille de 352×272 pixels avec une trichromie RVB d'une profondeur de 3×8 bits.

Nom	Séquences	
	Infrarouge	Visible
Emmanuel	<i>ir_1.avi</i>	<i>vi_1.avi</i>
Dominique	<i>ir_2.avi</i>	<i>vi_2.avi</i>
Olivier	<i>ir_3.avi</i>	<i>vi_3.avi</i>
Thesard	<i>ir_4.avi</i>	<i>vi_4.avi</i>

TAB. C.1 – **Séquences numérisées** : Les séquences de visages, stockées sur bande VHS, sont numérisées séparément, pour chaque visage, avec une carte d'acquisition video.

Le tableau C.2 donne les conclusions sur l'exploitation des séquences video de la base d'apprentissage.

Le tableau C.3 donne les conclusions sur l'exploitation des séquences video de la base de test.

Il est procédé manuellement, à l'aide d'un utilitaire écrit en *Tcl/Tk* à la segmentation des animations en imagettes. Ces imagettes sont formées par juxtaposition des deux images infrarouge et visible de taille 64×64 pixels (Figure C.1). Le script exécute l'utilitaire de visualisation *XAnim*, ainsi que les programmes *import*, *convert* et *montage* du logiciel *ImageMagick*.

Attitudes	Visages				Base
	Emmanuel	Dominique	Olivier	Thesard	
Rotation	●	○	○	●	21
Hochement	●	○	○	●	21
Sourire	●	○	●	●	11
Dents	●	○	●	●	11
Baillement	●	○	●	●	21
Froncement	●	●	●	●	11
Clignement	●	●	●	●	21

- : séquence correcte
- : séquence inexploitable

TAB. C.2 – **Base d’apprentissage** : les attitudes de la base d’apprentissage forment la base de références servant à établir les signatures d’apprentissage.

Attitudes	Visages				Base
	Emmanuel	Dominique	Olivier	Thesard	
Rotation & Clignement	●	●	○	●	21
Hochement & Baillement	●	○	○	●	21
Froncement & Baillement	●	○	●	●	11

- : séquence correcte
- : séquence inexploitable

TAB. C.3 – **Base de test** : les attitudes de la base de test sont des attitudes combinant plusieurs attitudes de la base d’apprentissage.

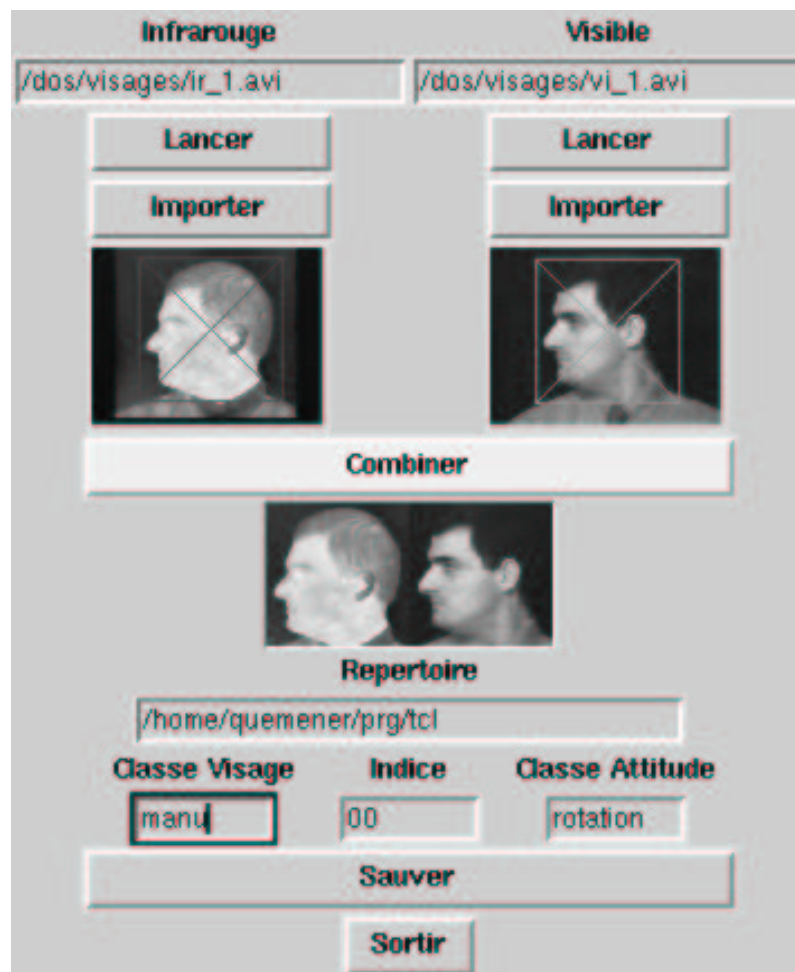


FIG. C.1 – Utilitaire Tcl/Tk : cet utilitaire a été conçu pour segmenter, calibrer et former les imagettes de la base de données formant chacune des séquences d'images.

La figure C.2 présente une attitude formée de 21 images : «rotation & clignement» du visage «Emmanuel».



FIG. C.2 – **Exemple d'attitude** : l'attitude, composée de 21 images, est une «rotation» du visage «Emmanuel». Le visage part de face, tourne sur sa droite puis revient de face. L'ordre de lecture est indiqué par la matrice dont les cadrants sont numérotés.

Annexe D

Performances d'un DSP Moderne

D.1 Un exemple de DSP : la famille *TMS320C6x*

Les informations présentées ici sont issues de la vitrine que présente *Texas Instruments* sur sa vitrine du réseau des réseaux¹. La firme américaine y présente ses composants récents, notamment en matière de DSP. Ses petits derniers (la famille *TMS320C6x*) y sont abondamment décrits, notamment sur le temps nécessaire à de tels processeurs pour réaliser des opérations de traitement de signal.

Le dernier-né de cette famille de DSP, le *TMS320C6202* travaillant sur des nombres de 32 bits, en virgule fixe («*fixed point*»), dispose d'une mémoire de 384 Ko, laquelle se divise en deux blocs de 256 Ko (2×128 Ko) pour les programmes et 128 Ko (2×64 Ko) pour les données. Le bus est cadencé à une fréquence de 250 MHz, ce qui lui confère une puissance brute de 2000 MIPS.

Diverses applications de ce DSP, ainsi que les performances qui lui sont associées, sont présentées sur leur site². Les calculs de FFT y figurent en bonne place. Ainsi, le temps nécessaire (en nombre de cycles d'horloge) à un DSP pour réaliser la transformation de Fourier rapide d'un vecteur de N points (N étant une puissance de 4), est :

$$c(N) = \log_4(N) \left(10 \times \frac{N}{4} + 33 \right) + 7 + \frac{N}{4} \quad (\text{D.1})$$

Ainsi, pour le *TMS320C6202*, le temps de calcul nécessaire pour assurer une FFT d'une image 2^n pixels est $t(n)$, tel que :

$$\begin{aligned} t(N) &= \frac{50 \times 10^{-6}}{13\,228} \left[\frac{\ln(2^n)}{\ln(4)} \left(10 \times \frac{2^n}{4} + 33 \right) + 7 + \frac{2^n}{4} \right] \\ &= \frac{50 \times 10^{-6}}{13\,228} \left[\frac{n}{2} (5 \times 2^{n-1} + 33) + 7 + 2^{n-2} \right] \end{aligned}$$

La figure D.1 présente le temps nécessaire à ce DSP pour réaliser une FFT sur des images de tailles croissantes. Notons que les valeurs élevées de n sont à prendre avec précaution. En effet, pour des images de taille supérieure à 64×64 pixels, la mémoire interne

¹<http://www.ti.com/sc/docs/dsps/products/c6x/index.htm>

²<http://www.ti.com/sc/docs/dsps/products/c6000/c62x/benchmk.htm>

ne peut stocker de telles quantités d'information. Aussi, les performances se dégradent-elles au-delà.

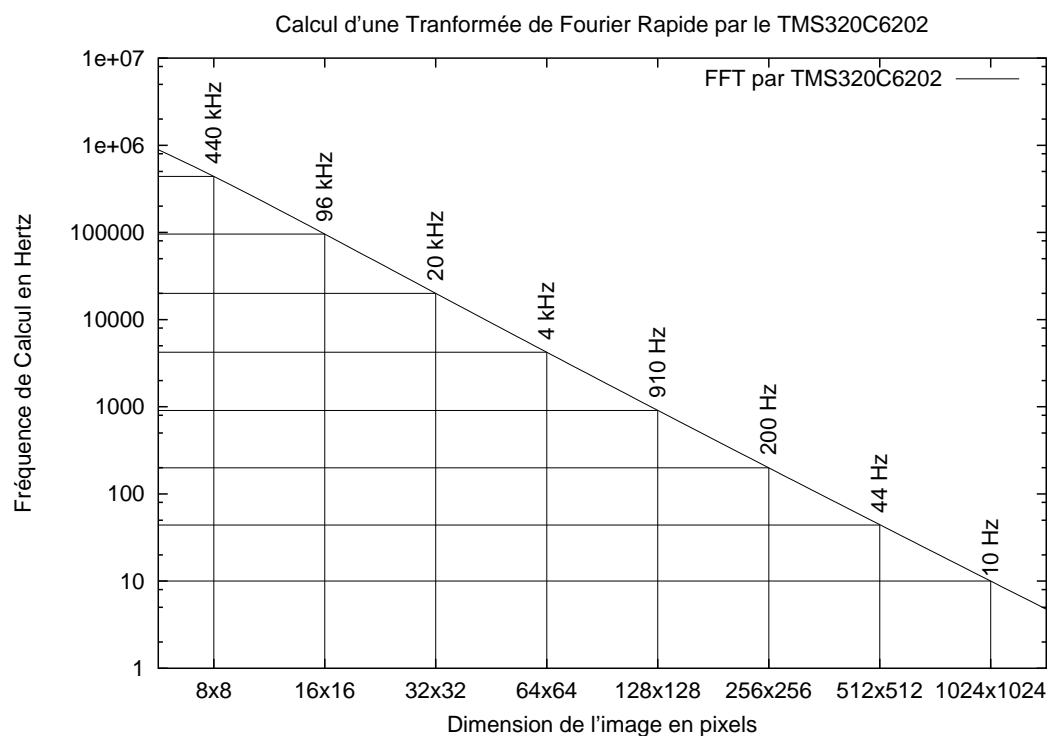


FIG. D.1 – Performances d'un DSP pour le calcul d'une FFT : cette courbe présente la fréquence de traitement du DSP *TMS320C6202*.

Bibliographie

- [1] B.D. Bock, T.A. Crow, and M.K. Giles. Design considerations for miniature optical correlation systems that use pixelated input and filter transducers. In *Proc SPIE Optical Information-Processing Systems and Architectures II*, volume 1347, pages 297–309, 1990.
- [2] R. Brunelli and T. Poggio. Face recognition : Features versus templates. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 15, page 1042, October 1993.
- [3] K. R. Castleman. *Digital Image Processing*. Prentice Hall, New Jersey, 1996.
- [4] H.J. Caulfield and R. Haimes. Generalized matched filtering. *Applied Optics*, 19 :181–183, 1980.
- [5] N. Collings. *Optical pattern recognition using holographic techniques*. Addison-Wesley, New-york, 1988.
- [6] J.-L. de Bougrenet de la Tocnaye, Quémener E., and Pétilot Y. Composite versus multichannel binary phase only filtering. *Applied Optics*, 36 :6646–6653, 1997.
- [7] J.-L. de Bougrenet de la Tocnaye, E. Quémener, and G. Keryer. Fourier optics and its use in multi-sensor and multi-decision processing environments. *Applied Optics*, 1999. submitted october 1998.
- [8] C. Elachi, T. Bicknell, R. L. Jordan, and Chialin Wu. Spaceborne synthetic-aperture imaging radars : Applications, techniques, and technology. *Proceedings of the IEEE*, 70(10) :1174–1209, October 1982.
- [9] A.V. Goncharsky, I.V. Kochikov, and V.B. Shikalov. Optical correlators in pattern recognition. *Pattern Recognition and Image Analysis*, 1 :440–449, 1991.
- [10] J. W. Goodman. *Introduction to Fourier Optics*. Mc Graw-Hill, New-York, 1968.
- [11] L. Guibert, G. Keryer, A. Serval, H.S. MacKenzie M. Attia, P. Pellat-Finet, and J.L. de Bougrenet de la Tocnaye. On-board optical joint transform correlator for real-time road sign recognition. *Optical Engineering*, 34 :135–143, 1995.
- [12] J.-P. Hardange, Ph. Lacomme, and J.-C. Marchais. *Radars aéroportés et spatiaux*. Masson, 1995.
- [13] E. Hecht. *Optics (Second edition)*. Addison-Wesley, New-york, 1987.
- [14] D. Hilbert. Über die stetige abbildung einer linie auf ein flächenstück. *Mathematische Annalen*, 38(459-460), 1890.
- [15] J. L. Horner and P. Gianino. Phase-only matched filtering. *Applied Optics*, 23(6) :812–816, March 84.

- [16] G. Keryer. Traitement Optique des Images par Ondelettes. Technical Report XVI, Ecole Navale, rapport interne, 1997.
- [17] Gilles Keryer. *Étude de corrélateurs optiques à corrélation jointe mono ou multicanaux : application à la reconnaissance de formes*. Thèse de doctorat, Université d'Orsay (Paris XI), Département Optique, ENST Bretagne, BP 832, 29285 BREST Cedex, France, janvier 1996.
- [18] A. Kohler, B. Fracasso, M. Killinger, J.-L. de Bougrenet, and P. Ambs. Joint transform correlator using nonlinear ferroelectric liquid crystal spatial light modulators. In *Proceedings Optical Applied Sciences and Engineering SPIE*, pages 236–243, San Diego, 1991.
- [19] B. V. K. Vijaya Kumar and L. Hassebrook. Performance measures for correlation filters. *Applied Optics*, 29(20) :2997–3006, July 1990.
- [20] B.V.K. Vijaya Kumar. Tutorial survey of composite filter designs for optical correlators. *Applied Optics*, 31 :4773–4801, 1992.
- [21] B.V.K. Vijaya Kumar, R.D. Juday, and P.K. Rajan. Maximization of snr of filters for optical pattern recognition under the condition of input and detector noise. In *Proceedings IEEE of the ISCAS91*, volume 5, pages 2637–2640, 1991.
- [22] S.C. Lindel. Transfer of Optical Processing to Systems (TOPS) Optical Correlation Program. In *Proceedings of the IEEE*, volume 1958, pages 7–18, 1994.
- [23] A. Mahalanobis, B.V.K. Vijaya Kumar, and D. Casasent. Minimum average correlation energy filters. *Applied Optics*, 26 :3633–3640, 1987.
- [24] D.J. McKnight, K.M. Johnson, and R.A. Serati. 256x256 liquid-crystal-on-silicon spatial light modulator. *Applied Optics*, 39 :2775–2783, 1994.
- [25] Y. Meyer. *Les ondelettes*. Armand Colin, Paris, 1994.
- [26] H.-O. Peitgen, H. Jürgens, and D. Saupe. *Chaos and Fractals*, chapter 2, page 100. SpringerVerlag, 1992.
- [27] D. Psaltis, F. Mok, S. Li, and K. Curtis. 3d holographic workshop storage in image recognition. *Proceedings on Optical Pattern Recognition*, pages 419–428, 1994.
- [28] Y. Pétillet, L. Guibert, and J.L. de Bougrenet de la Tocnaye. Fingerprint recognition using a partially rotation invariant composite filter in a FLC joint transform correlator. *Optics communication*, 126 :213–219, 1996.
- [29] Yvan Pétillet. *Vers une implantation optique de corrélateurs optiques temps réel – Application à divers problèmes de reconnaissance de formes*. Thèse de doctorat, Université de Bretagne Occidentale, Département Optique, ENST Bretagne, BP 832, 29285 BREST Cedex, France, mars 1996.
- [30] E. Quémener. De l'ouverture synthétique en général au traitement de sonar à ouverture synthétique en particulier. Technical Report XI, Ecole Navale, rapport interne, 1996.
- [31] P. Réfrégier. Filter design for optical pattern recognition : multicriteria optimization approach. *Optics Letters*, 15 :854–856, 1990.
- [32] D. Roberge and Y. Sheng. Optical wavelet matched filters. *Applied Optics*, 33 :5287–5293, 1994.

- [33] J. Rodolfo, H. Rajbenbach, and J.P. Huignard. Performance of a photorefractive joint transform correlator for fringerprint identification. *Optical engineering*, 34 :1166–1171, 1995.
- [34] A. Rosenfeld. *Multiresolution Image Processing and Analysis : Some Useful Properties of Pyramids*. Springer-Verlag, Berlin, 1984.
- [35] S.A. Serati, T.K. Ewing, R.A. Serati, K.M. Johnson, and D.M. Simon. Programmable 128x128 ferroelectric-liquid-crystal spatial-light-modulator compact correlator. In *Proceedings of the SPIE Optical Pattern Recognition IV*, volume 1959, pages 55–68, 1993.
- [36] I. Underwood, D.G. Vass, A. O’Hara, D.C. Burns, P.W. McOwan, and J. Gourlay. Improving the performance of liquid-crystal-over-silicon spatial light modulators : issues and achievements. *Applied Optics*, 33 :2768, 1994.
- [37] A. Vander-Lugt. Signal detection by complex spatial filtering. *IEEE Trans.Inf. theory IT-1*, page 139, 1964.
- [38] S. Weaver and K. Wagner. Nonlinear techniques in optical synthetic aperture radar image generation and target recognition. *Applied Optics*, 34(20) :3981–3995, July 1995.

Résumé

Les techniques modernes de télédétection offrent sur une cible une masse croissante de signaux de natures très différentes : spectrale (différents capteurs), spatiale (images) et temporelle (observation sur un laps de temps). L'analyse globale, rendue plus difficile par cette nécessité d'associer une masse d'informations hétérogènes, devient critique lorsque des contraintes opérationnelles sont posées, notamment lors d'applications dites «*temps réel*». Utilisant le modèle de formation d'une réponse d'antenne synthétique, une signature temporelle sera créée à partir de séquences d'images. La création puis l'utilisation d'une signature temporelle, image caractéristique de la cible et de son attitude, seront successivement abordées pour parvenir à la reconnaissance de la cible par l'attitude. Ces opérations de synthèse et d'analyse sont réalisées à l'aide d'un outil simple, autant dans sa structure que par ses implantations : le filtrage spatial. Il sera dans cette optique associé à un prétraitement linéaire, lequel permettra d'accéder, lors de la synthèse, à un niveau de détail donné. La validation de cette approche «synthèse & analyse» sera tout d'abord menée sur des attitudes de visages, puis sur des signaux radar «temps-fréquence». Cette dernière démontrera alors que la puissance de cette technique repose en majorité sur la qualité des bases de références adjointes au filtrage spatial. Ainsi, face à des traitements puissants, compliqués et donc nécessairement coûteux en temps de calcul, une approche simple voire sommaire conserve son intérêt. Associé à des bases de données ou de références, le filtrage spatial reste encore et demeure toujours un bon outil de reconnaissance de formes, mais aussi de formes par l'attitude. Sur cette base, une discussion des implantations optique et numérique sera développée, fixant les avantages respectifs de chacune dans le cadre considéré.

Mots clés :

Reconnaissance de formes, filtrage spatial, signature temporelle,
traitement optique de l'information, corrélateur optique ;

Abstract

Modern techniques in teledetection provide, for a given target, an ever-growing variety of data of a widely differing nature : spectral (different sensors), spatial (pictures) and temporal (observations over time) The global analysis is made more difficult by the need for an association between heterogeneous data. This requirement becomes critical when *real time* applications are considered. Extending the design of the synthetic aperture antenna response, a temporal signature is created from image sequences. The construction and use of this temporal signature, a defining characteristic of the target and its attitude, is successively developed for pattern recognition by attitude. This synthesis and analysis is performed with a basic tool : spatial filtering. The filtering is aided by the use of a linear pre-processing operation, which enables us to compare different patterns with a selected precision. The validation of this “synthesis & analysis” approach is tested with face attitudes and “time-frequency” radar signals. It is shown that the power of this technique is based mainly on the quality of filter references database. In this way, a simple filtering-based approach is shown to be competitive with more sophisticated and powerful but high time-cost techniques. When associated with quality filters, filtering is and remains a useful tool for pattern recognition and pattern recognition by attitude. Starting from this encouraging result, optical and numerical implementations are described, underlining the advantages of each according to the application considered.

Key Words

Pattern recognition, spatial filtering, temporal signature, optical processing, optical correlator ;