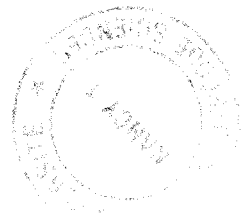
Université Henri Poincaré Nancy 1
UFR Sciences Techniques Mathématiques Informatique Automatique
Ecole Doctorale IAE+M
DFD Automatique et Production Automatisée



CNRS URA DO 821

THESE

présentée pour l'obtention du

DOCTORAT DE L'UNIVERSITE HENRI POINCARE - NANCY I

Spécialité : Automatique

par

Christophe SIMON

Vérification automatique de l'identité par perception multi-échelle et discrimination intégrée floues de l'image de la signature manuscrite

soutenue le 8 juillet 1996 devant la commission d'examen composée de :

Président :	Monsieur Claude HUMBERT, Professeur à l'UHP-Nancy 1 (ESSTIN)
Rapporteurs :	Monsieur Hubert EMPTOZ, Professeur à l'INSA de Lyon
	Monsieur Jack-Gérard POSTAIRE, Professeur à l'UST de Lille 1
Examinateurs :	Monsieur Jacques BREMONT, Professeur à l'UHP-Nancy 1
	Monsieur Eric LEVRAT, Maître de Conférences à l'UHP-Nancy 1
	Monsieur Robert SABOURIN, Professeur à l'ETS de Montréal, Canada

Avant-propos

Je tiens à remercier Monsieur Claude HUMBERT, Professeur à l'Université Henri Poincaré Nancy 1, Directeur de l'Ecole Supérieure des Sciences et Technologies de l'Ingénieur de Nancy et du Département de Formation Doctorale, de m'avoir fait l'immense honneur de présider ce jury.

Je remercie chaleureusement Messieurs les Professeurs Hubert EMPTOZ, de l'INSA de Lyon, et Jack-Gérard POSTAIRE, de L'Université des Sciences et Technologie de Lille, d'avoir jugé et apprécié ce travail.

Je voudrais tout particulièrement remercier Monsieur le Professeur Robert SABOURIN, de l'Ecole de Technologie Supérieure de Montréal Canada, pour avoir traversé l'Atlantique afin de participer à ce jury, ainsi que pour ses nombreux conseils durant ces années de recherche, qui ont participé à la qualité de cette thèse.

Je remercie également Monsieur le Professeur Jacques BREMONT de l'Université Henri Poincaré Nancy 1, de m'avoir accueilli au sein de son équipe et pour ses conseils avisés et les nombreuses discussions que nous avons eues.

Je tiens également à remercier Monsieur Eric LEVRAT, Maître de Conférence à l'Université Henri Poincaré Nancy 1, pour son aide, les nombreuses heures qu'il m'a accordées et cette complicité que nous avons établie durant les longues périodes de recherche. Que son épouse reçoive également mes plus sincères remerciements pour sa patience, ainsi que ses enfants lorsque leur papa travaillait avec moi.

Enfin, je remercie Frédérique BICKING pour son aide efficace et son soutien dans les moments difficiles, ainsi que toute l'équipe Perception et RAISonnement dans les Systèmes Industriels et Humains pour cette bonne ambiance et l'amitié mutuelle que nous nous sommes portée. Merci, à Jocelyne, Laurent, Alexandre, Claude, Ruben et tous ceux que j'ai oubliés sur ces quelques lignes.

Table des matières

Introduction générale ____ *1*

Chapitre 1 De la reconnaissance de formes à la vérification de signatures manuscrites ____ *6*

1. La reconnaissance de formes et l'écrit ____ **7**
1.1. Introduction ____ 7
1.2. La chaîne de reconnaissance de formes ____ 8
1.2.1. Introduction ____ 8
1.2.2. La perception ____ 10
1.2.3. Le raisonnement ____ 12
1.3. L'analyse de l'écrit ____ 26
1.3.1. Introduction ____ 26
1.3.2. Imprimé ou manuscrit ____ 27
1.3.3. Les sources de complexité de l'écriture manuscrite ____ 27
1.3.4. Le système de reconnaissance de l'écriture manuscrite ____ 29
1.3.5. Structure d'une application de reconnaissance hors-ligne de l'écrit ____ 31

2. La vérification hors-ligne de signatures manuscrites ____ **34**
2.1. Le contexte ____ 34
2.2. La vérification automatique de signatures manuscrites ____ 35
2.2.1. Les caractéristiques propres aux signatures ____ 36
2.2.2. Le système de vérification de signatures ____ 42
2.3. L'approche proposée par R. Sabourin ____ 44
2.3.1. Introduction ____ 44
2.3.2. La base de données ____ 45
2.3.3. La perception ____ 45
2.3.4. Le raisonnement ____ 52
2.3.5. Le protocole de test ____ 52
2.3.6. Coopération de classificateurs ____ 53
2.4. Les améliorations possibles ____ 55

3. Conclusion ____ **55**

Chapitre 2 Perception des signatures, une approche statique et pseudo-dynamique ____ *57*

1. Introduction ____ **58**

2. Le facteur de formes pour une approche statique de la vérification ____ **59**
2.1. Les prétraitements ____ 59
2.2. Les motifs multi-échelle ____ 59
2.2.1. Modification de la méthode de projection ____ 59
2.2.2. Adaptation de la scrutation aux signatures ____ 60
2.3. Introduction du codage de la distance ____ 62
2.4. La description des codages ____ 63
2.4.1. Corrélation de signaux ____ 63
2.4.2. Description statistique ____ 66
2.5. Evaluation du choix du descripteur en correspondance avec la technique de codage ____ 67
2.5.1. Compléments d'analyse pour la confirmation du choix du descripteur des bâtons ____ 75
2.6. Quantification de l'apport du codage de la distance ____ 77
2.7. Quantification pour la méthode de codage de la distance avec modification de la scrutation ____ 80
2.8. Conclusion ____ 81

3. Un facteur de formes pour une approche pseudo-dynamique de la vérification 82
3.1. Introduction du niveau de gris dans l'Extended Shadow Code avec codage de la distance 82
3.2. Codage simultané de la pseudo-dynamique et de la statique 83
3.3. Définition d'un opérateur flou de codage de signatures 85
3.3.1. Fuzzification 85
3.3.2. Règles d'inférence 87
3.3.3. Défuzzification 88
3.4. Pertinence du FESC pour la discrimination des faux aléatoires 89
3.5. Amélioration du FESC 91
3.6. Pertinence du nouveau facteur de formes 93
3.7. Application à la discrimination des faux par transfert 94
3.7.1. Simulation d'un faux par calque 95
3.7.2. Pertinence pour le traitement des faux par photocopies 98
3.8. Conclusion 102

4. Conclusion 104

Chapitre 3 Raisonnement : Discrimination intégrée *106*

1. Introduction 107

2. Le Raisonnement pour la vérification de signatures manuscrites 107

3. La base de nos travaux 108

4. Les améliorations possibles 111
4.1. L'apprentissage 111
4.2. La coopération des classificateurs 112
4.3. Prise en compte de l'approche multi-échelles 112

5. L'algorithme de discrimination locale 114
5.1. L'algorithme des k plus proches voisins flous de Béreau 116
5.1.1. La procédure de classification automatique 116
5.1.2. La procédure de discrimination 117
5.1.3. Discussion 120
5.1.4. Le flou : assouplissement ou rigidification de la contrainte 121

6. Adaptation de la discrimination pour la vérification automatique de l'identité 122
6.1. Particularisation de la fonction de discrimination 123
6.1.1. Incertitude d'affectation 123
6.1.2. Formulation de b_i 124
6.1.3. Intégration de la pertinence de la perception 126
6.2. Applications de la fonction de discrimination 127
6.2.1. Influence de b_i 128
6.2.2. Influence du nombre de voisins 129
6.2.3. Influence du facteur de formes λ_{vraies} 131
6.2.4. Influence du facteur $\lambda_{fausses}$ 132
6.2.5. Synthèse des pouvoirs discriminants 133
6.2.6. Conclusion 134

7. Une discrimination intégrée pour la vérification 134
7.1. Introduction 134
7.2. Intégration des discriminations 136
7.2.1. La coopération extra-classificateur 136
7.2.2. Expérimentations 137
7.2.3. La coopération intra-classificateur 138

7.3. La structure de la discrimination intégrée ____ 140
7.4. Expérimentations ____ 142
7.5. Conclusion ____ 143

8. Conclusion ____ 144

Conclusion et perspectives ____ 146

Liste des figures

Figure 1 : Décomposition fonctionnelle d'une chaîne de reconnaissance de formes 9
Figure 2 : Proportions des étapes de la chaîne 11
Figure 3 : Les approches possibles pour une reconnaissance vectorielle 15
Figure 4 : Principe d'enchaînement séquentiel de classificateur 22
Figure 5 : Bornes sur le rapport séquentiel de vraisemblance 22
Figure 6 : Principe de la coopération parallèle [Xu 92] 24
Figure 7 : Exemple de coopération hybride dans un problème de reconnaissance de caractères 25
Figure 8 : Différentes présentations de texte 28
Figure 9 : Espace situant la complexité d'un système de reconnaissance de l'écrit 29
Figure 10 : Différents types d'analyse de la signature suivant l'attribut choisi et la connaissance des classes 43
Figure 11 : Histogramme d'une image de signature 46
Figure 12 : Translation de la signature au centre du cadre de zoning 47
Figure 13 : Motif de base 47
Figure 14 : Motifs multi-échelle 49
Figure 15 : Illustration du mécanisme de projection de l'Extended Shadow Code 50
Figure 16 : Système complet de vérification de signatures selon l'approche de R. SABOURIN 54
Figure 17 : Différence de mécanisme de projection 60
Figure 18 : Illustration graphique du problème d'aliasage 61
Figure 19 : Résolution du problème d'aliasage 61
Figure 20 : Différence de codage d'information 62
Figure 21 : Transformée de Walsh-Hadamard du 1^{er} bâton horizontal du motif a pour 3 scripteurs, méthode de l'Extended Shadow Code. 64
Figure 22 : Transformée de Fourier du 1er bâton horizontal du motif a pour 2 scripteurs méthode Extended Shadow Code 65
Figure 23 : Approches signal ou distribution de la description du codage 66
Figure 24 : Plan des observations selon la méthode de référence projection binaire/moyenne du signal (D1) 68
Figure 25 : Plan des observations selon la méthode de référence projection binaire/3 moments du signal (D2) 69
Figure 26 : Plan des observations selon la méthode de référence 70
Figure 27 : Plan des observations selon la méthode de référence, projection binaire/3 moments de la distribution (D4) 70
Figure 28 : Plan des observations selon la méthode de codage de la distance, 71
Figure 29 : Plan des observations selon la méthode de codage de la distance, 71
Figure 30 : Plan des observations selon la méthode de codage de la distance, 72
Figure 31 : Plan des observations selon la méthode de codage de la distance, 72
Figure 32 : Plan des observations selon la méthode de codage de la distance, 73
Figure 33 : Plan des observations selon la méthode de codage de la distance, 73
Figure 34 : Plan des observations selon la méthode de codage de la distance, 74
Figure 35 : Plan des observations selon la méthode de codage de la distance, 74
Figure 36 : Comparaison des performances des descripteurs de bâton pour tous les motifs 76
Figure 37 : Performances du codage de la distance selon le mécanisme de projection de R. Sabourin/D1 78
Figure 38 : Performances comparées selon la résolution des motifs 78
Figure 39 : Représentation de ε_i versus les différents motifs pour K voisins 80
Figure 40 : Performances comparées selon la résolution 81
Figure 41 : Fuzzification de la pseudo-dynamique 86
Figure 42 : Fuzzification de la distance 87
Figure 43 : Table d'inférence pour la fusion de la distance et du niveau de gris 88
Figure 44 : Défuzzification du codage par le centre de gravité 88
Figure 45 : Performances du codage FESC/D1 89
Figure 46 : Performances comparées de l'ESC, de l'ESC avec codage de la distance et nouvelle scrutation et le Fuzzy ESC. 90
Figure 47 : Nouvelle fuzzification de la pseudo-dynamique 91
Figure 48 : Table d'inférence pour une approche image de signatures 92
Figure 49 : Surface de codage en fonction des entrées 92

Figure 50 : Comparaison des performances des facteurs de formes 93
Figure 51 : LUT pour la simulation d'une photocopie 95
Figure 52 : Image initiale et ses photocopies avec δ = 75%, 50% et 25% 96
Figure 53 : Correspondance des termes de fuzzification pour une signature authentique 97
Figure 54 : Position des signatures fausses et photocopiées dans le plan des observations pour le motif a 100
Figure 55 : Exemple de définition et d'application du seuil de discrimination 109
Figure 56 : Exemples d'erreurs possibles de vérification 111
Figure 57 : Evolution de la notion de ressemblance selon le motif de perception 113
Figure 58 : Comportement de $\mu_i(y)$ en fonction du nombre de voisins et selon b_i 118
Figure 59 : Evolution de la zone d'attraction d'un prototype selon la valeur de λ 119
Figure 60 : Modification de la contrainte selon b_i 122
Figure 61 : Comportement de f_i selon n et m 125
Figure 62 : Comportement de f_i en fonction de m pour un espace à 6 dimensions 125
Figure 63 : Comportement de b_i 126
Figure 64 : Détermination des voisins au sens multi-échelles 139
Figure 65 : Structure de la discrimination intégrée 141

Liste des tableaux

Tableau 1 : Résumé des divergences fausses/vraies signatures. 39
Tableau 2 : Erreur totale en % avec k plus proches voisins pour le couple codage de R. Sabourin/D1 75
Tableau 3 : Erreur totale en % avec k plus proches voisins pour le couple codage de R. Sabourin/D2 75
Tableau 4 : Erreur totale en % avec k plus proches voisins pour le couple codage de R. Sabourin/D3 76
Tableau 5 : Erreur totale en % avec k plus proches voisins pour le couple codage de R. Sabourin/D4 76
Tableau 6 : Erreur totale en % avec k plus proches voisins pour le couple codage de la distance/D1 77
Tableau 7 : Erreur totale ε_t, en %, obtenue par modification de la scrutation et codage de la distance pour les différents motifs, par la règle des k plus proches voisins 80
Tableau 8 : Erreur totale en % avec k plus proches voisins pour le couple codage Fuzzy Extended Shadow Code/D1 89
Tableau 9 : Variances des erreurs totales et pour les méthodes ESC et FESC 90
Tableau 10 : Erreur totale en % avec k plus proches voisins pour le couple codage Fuzzy Extended Shadow Code/D1 93
Tableau 11 : Variances des méthodes ESC, ESC avec distance et scrutation FESC et FESC global. 94
Tableau 12 : Erreur totale et erreur de type I en %, avec la règle des k plus proches voisins, sans référence de photocopies dans l'ensemble de référence 99
Tableau 13 : Erreur totale et erreur de type I en % avec 1 plus proche voisin, avec 5 références de photocopies 101
Tableau 14 : Influence du nombre de voisins k sur le taux d'erreur de vérification en % 129
Tableau 15 : Influence du poids λ_{vraies}, pour 5 voisins, sur le taux d'erreur de vérification en % 131
Tableau 16 : Influence du poids $\lambda_{fausses}$, pour 5 voisins, sur le taux d'erreur de vérification en % 132
Tableau 17 : Récapitulatif des performances par motif 133
Tableau 18 : Coopération extra-classificateurs, motifs a, f et g 137
Tableau 19 : Coopération extra-classificateur, motifs m, n et o 138
Tableau 20 : Performances de vérification du système dégradé, motifs a, f et g 142
Tableau 21 : Performances de vérification, motifs m, n et o 143

Introduction générale

Les travaux exposés dans ce mémoire s'inscrivent dans le cadre d'une collaboration de recherche franco-québécoise établie en 1994 entre l'Université Henri Poincaré-Nancy 1 et l'Ecole de Technologie Supérieure de Montréal. Une partie de cette collaboration repose sur l'équipe Perception et RAISonnement dans les Systèmes Industriels et Humains du Centre de Recherche en Automatique de Nancy, équipe dirigée par le Professeur J. Brémont, et sur le Laboratoire d'Imagerie de Vision et d'Intelligence Artificielle au travers du Professeur R. Sabourin.

Le sujet traité est la vérification hors-ligne de l'identité par analyse de l'image de la signature manuscrite. Cette problématique s'insère dans le cadre général de la reconnaissance de formes, et plus particulièrement dans la reconnaissance de l'écrit par l'image. Le contexte d'utilisation d'un système de vérification se situe dans le milieu bancaire et judiciaire, c'est-à-dire privilégiant le niveau de sécurité de la décision par rapport au confort. C'est en cela un problème particulier, puisque l'on rencontre ici le concept d'imitation et de falsification que l'on ne retrouve pas dans les problèmes de reconnaissance de formes ne traitant pas de l'identité.

Il s'agit de déterminer automatiquement si l'image saisie hors-ligne, après le tracé de la signature, correspond à une signature authentique émise par le scripteur annoncé. La vérification de l'identité est un problème différent de l'identification de scripteur, mais non moins difficile.

Les principales difficultés de cette problématique résident dans :

- la variabilité des caractéristiques des signatures authentiques, provoquée par des facteurs tels que le contexte et les moyens techniques d'apposition,
- la multiplicité des types de fausses signatures, allant jusqu'à une qualité telle qu'il ne soit plus possible de faire la différence avec une signature vraie,
- la définition d'un ensemble pertinent de signatures d'apprentissage.

De plus, l'acquisition hors-ligne de l'image de la signature élimine une grande partie des connaissances de la dynamique du tracé, et ne conserve principalement que les caractéristiques statiques de la signature.

La restriction du problème à un type de fausses signatures, dont la qualité d'imitation est faible, a permis à R. Sabourin de proposer un système automatique capable de vérifier l'identité. Cette catégorie, que l'on appelle faux aléatoires, correspond à l'ensemble des signatures authentiques des autres abonnés au système de vérification et dont le libellé est différent des vraies. Ce problème lié aux bases de données paraît trivial mais reste un problème ouvert, car il exprime déjà toute la complexité de la vérification.

Ce système a un fonctionnement qui s'apparente à celui d'un jury de 15 experts graphologues, munis chacun d'une loupe avec un grossissement spécifique, et d'un ensemble d'exemples de vraies et fausses signatures. Ils seraient chargés de fournir une réponse sur la validité de la signature par rapport à leurs connaissances et leur point de vue, la décision finale étant prise à la majorité. Le système est basé sur une perception globale de l'image de la signature, c'est-à-dire sans chercher à la caractériser en termes de segments, de lettres ou de

mots. La perception s'opère suivant différents points de vue (15 au maximum) qui mettent en évidence des caractéristiques de l'image de la signature selon différents degrés de finesse dans l'observation. La perception est basée sur l'Extended Shadow Code, une extension proposée par R. Sabourin d'un facteur de formes développé par D.J. Burr pour la reconnaissance de caractères. Chacune des échelles de perception fournit des caractéristiques de l'image de la signature sous une forme vectorielle, dans des espaces de représentation de dimension 6 à 276.

Cette approche du processus de perception présente un certain nombre d'intérêts dont le premier est d'être insensible au texte. Elle peut donc s'appliquer aux différents types de signatures que l'on risque de rencontrer, les signatures nord-américaines, et les signatures européennes sous forme de parafes. De plus, la multiplication de points de vue différents sur la signature, autorise l'application du système à de nombreuses formes de signatures et participe également à la robustesse du système de vérification. Enfin, la représentation vectorielle de l'image de la signature ne force pas à l'utilisation de méthodes de reconnaissance structurelle ou directe, dont plusieurs auteurs ont montré l'inadéquation avec la complexité des signatures.

La décision vraie/fausse signature est prise par un classificateur intégré, constitué d'au plus 15 classificateurs à distance minimale, opérant chacun dans un des différents espaces de représentation. Leurs réponses sont ensuite exploitées par un module de coopération parallèle à vote majoritaire qui fournit la décision finale. Les règles de discrimination locales sont obtenues par apprentissage sur un ensemble de signatures vraies et fausses (aléatoires), en début de procédure, et sont ensuite appliquées sur les signatures à vérifier.

R. Sabourin a montré les qualités de son approche lorsque les faux rencontrés se limitent aux faux aléatoires. Il souhaitait généraliser ce système afin de répondre plus complètement au problème de vérification de l'identité, en présence d'autres types de faux tels que les faux par calque ou par transfert optique.

Depuis les travaux de J. Brémont en reconnaissance de la parole en 1975, notre équipe s'appuie sur les concepts de la théorie des ensembles flous pour résoudre avec succès, des problèmes de commande, de modélisation, de vision artificielle et de reconnaissance de formes. L'aptitude de cette théorie à intégrer le recouvrement de classes, à prendre en compte l'imprécision et l'évolution de la connaissance, ainsi que ses capacités à supporter des raisonnements qualitatifs, ont suscité de nombreux développements tant théoriques qu'applicatifs en reconnaissance de formes.

Le potentiel de cette théorie en reconnaissance de formes, ainsi que notre expérience dans son application, la difficulté de la problématique de vérification de l'identité, la pertinence de l'approche classificateur intégré multi-échelle, la volonté de généralisation du système ont suscité un intérêt réciproque dans une collaboration et ont motivé notre recherche.

Cette collaboration s'est matérialisée notamment par la mise à notre disposition d'une base de données d'images de signatures et de protocoles d'évaluation des performances, indispensables à un positionnement de nos résultats.

L'objectif de notre recherche est donc la généralisation de l'approche proposée par R. Sabourin à la vérification automatique de l'identité en présence de faux tels que les faux par

calque ou par transfert optique. Cette généralisation ne vise pas uniquement de bonnes performances numériques de vérification, mais également une prise en compte globale de la problématique. Nous chercherons à doter le système de capacités en adéquation avec la complexité du problème.

Nos contributions couvrent les deux constituants du système de vérification, la perception et le raisonnement et notre mémoire s'articule selon trois chapitres.

Le premier présente la chaîne de reconnaissance de formes dans le cadre particulier d'une représentation vectorielle de formes. Nous n'avons pas détaillé la reconnaissance de formes basée sur des représentations structurelles ou syntaxiques, qui ne s'insère pas directement dans notre problématique. Nous précisons ensuite les caractéristiques et les difficultés de la reconnaissance de l'écrit, afin de situer notre problème. Puis, nous définissons la vérification automatique de l'identité et précisons la nature des informations influençant le tracé de la signature ainsi que les caractéristiques des faux rencontrés. La fin de ce chapitre est dédiée à la présentation de l'approche proposée par R. Sabourin.

Le second chapitre est consacré à nos travaux sur l'étape de perception, il se divise en deux parties, consacrées respectivement à l'approche statique et à l'approche pseudo-dynamique de la perception.

Dans la première partie, nous présentons nos améliorations du facteur de formes proposé par R. Sabourin, au niveau de la méthode de projection et de la scrutation des images de signatures. Nous introduisons ensuite la distance séparant le tracé de la signature du maillage superposé comme information améliorant le pouvoir descriptif du facteur de formes. Les performances de description de ces différents facteurs de formes sont évaluées et comparées tout au long de cette partie.

La deuxième partie est consacrée à l'élaboration d'un facteur de formes répondant simultanément aux problèmes de discrimination des faux aléatoires et des faux par transfert optique. Après avoir établi les liens existants entre la dynamique du tracé de la signature et les niveaux de gris de son image, nous présentons le nouveau principe de codage basé sur un système de règles floues. Ces règles s'appuient sur les caractérisations floues de l'échelle des niveaux de gris de l'image et de la distance de la signature au maillage. Le pouvoir discriminant de ce facteur de formes est évalué et comparé aux précédents résultats en fin de chapitre.

Dans le dernier chapitre, nous présentons nos contributions à l'étape de raisonnement. Après avoir rappelé les contraintes et difficultés de la discrimination pour la vérification de l'identité, nous introduisons nos différentes propositions d'amélioration. Nous explicitons ensuite les différents critères qui nous ont conduits à retenir la fonction de discrimination des k plus proches voisins flous de M. Béreau [Béreau 86]. Nous analysons ses caractéristiques et

proposons les adaptations qu'elle nécessite pour son application à chacune des échelles de perception. Nous proposons enfin d'évaluer les performances de vérification de cette fonction pour les deux types de faux rencontrés, pour chacun des motifs de perception.

Après avoir présenté les deux principes de coopération parallèle intra et extra classificateurs, nous définissons le système de discrimination intégrée, constitué de n motifs de perception, de leur fonction de discrimination associée et d'une fonction de coopération externe. La dernière partie de ce chapitre est consacrée à la présentation des performances de ce système de vérification, ainsi qu'à leur analyse.

La conclusion de ce mémoire présentera la synthèse de nos travaux et dégagera les perspectives d'application envisagées.

Chapitre 1

De la reconnaissance de formes à la vérification de signatures manuscrites

1. La reconnaissance de formes et l'écrit

1.1. Introduction

Dans [Belaïd 92], l'auteur propose la définition suivante de la reconnaissance de formes :

« C'est la première étape d'un long processus de compréhension de notre univers, pour lequel la reconnaissance de formes doit résoudre les problèmes de ***codage****, de* ***paramétrisation*** *et de* ***discrimination*** *des formes. C'est une tâche difficile car les formes appartiennent à un* ***monde physique continu****, et la transcription dans* ***le monde numérique discret*** *est complexe. La reconnaissance est d'autant plus difficile que la nature des formes et leur apparence varient d'un échantillon à l'autre au sein d'une même famille. »*

Cette définition présente la reconnaissance de formes sous l'angle d'une fonction évoluée propre à un système automatique, lui permettant de comprendre son environnement. Elle laisse également supposer que les informations produites par la fonction de reconnaissance des formes sont consommées par une autre fonction du système. Ainsi, la reconnaissance de formes n'est pas une finalité, c'est une fonction de transformation de la nature de l'information depuis le monde physique vers le monde des symboles, des signes.

Si l'on se réfère à une vue systémique d'un système automatisé, la reconnaissance de formes est donc la fonction d'observation du système, qui fournit les informations symboliques nécessaires et suffisantes à la fonction de réflexion reliée à l'action. La reconnaissance de formes est donc une perception finalisée de l'environnement d'un système automatisé, la finalité se situant dans la fonction réflexion.

Une autre définition est proposée par Dubuisson [Dubuisson 90] :

« La reconnaissance de formes est une science de définition de méthodes permettant de classer des objets, dont l'aspect a varié par rapport à un objet type. Il s'agit de définir à quelle forme type une forme observée ressemble le plus. »

Dans cette définition, la reconnaissance de formes n'est plus vue comme une fonction d'un système mais comme un domaine de recherche.

Dans la suite de notre mémoire, nous traiterons la reconnaissance de formes comme une fonction et nous parlerons plus volontiers de la chaîne de reconnaissance de formes.

La notion de formes concerne une entité concrète ou abstraite, appréhendée par des sens, des capteurs ou fournie par un raisonnement. Nous considérons ici une forme sous une représentation vectorielle, c'est-à-dire comme un point dans un espace de représentation multi-dimensions. Chacune des dimensions de cet espace porte un caractère ou attribut particulier de la forme.

Les autres représentations possibles d'une forme, structurelle ou syntaxique ne sont pas abordées dans ce mémoire. Le lecteur trouvera dans [Heutte 94], [Miclet 84], [Hall 80], des développements concernant ces deux représentations.

L'objectif de la reconnaissance de formes est d'étiqueter, de nommer par un symbole, des formes du monde physique continu. Les symboles sont issus d'un vocabulaire fini caractérisant des ensembles de formes connues (prototypes), aux aspects approximativement analogues. L'allure des formes inconnues à nommer s'en rapproche plus ou moins. La difficulté est d'en apprécier la proximité. Il est donc nécessaire que nous analysions et recherchions une relation entre les formes à reconnaître et des symboles représentatifs de ces formes.

La reconnaissance de formes automatique doit passer par la **perception** du monde physique continu sous une forme numérique discrète afin d'imiter le **raisonnement** humain et de déterminer naturellement les symboles les plus adaptés pour caractériser ces formes.

Dans le cas naturel comme dans la situation automatique, l'objectif est de nommer ce qui est perçu par rapport à ce qui est connu.

Ainsi, pour assurer la fonction de reconnaissance, on se doit de construire des représentations numériques de formes (perception), des méthodes de regroupement de formes selon leurs caractères (coalescence, classification) et des méthodes d'affectation (discrimination) afin d'attribuer à ces formes, des étiquettes absolues ou relatives (symboles).

Dans ce chapitre, nous décrivons d'abord les grandes étapes rencontrées dans une chaîne de reconnaissance de formes, puis nous spécifions la particularité de la reconnaissance de l'écrit auquel la vérification hors-ligne de signatures manuscrites se rattache. Enfin, nous présentons le problème particulier de la vérification de signatures, et notamment l'approche proposée par R. Sabourin, base des développements ultérieurs de nos travaux.

1.2. La chaîne de reconnaissance de formes

1.2.1. Introduction

La spécificité des problèmes de reconnaissance de formes rencontrés et les multiples possibilités de résolution de tout ou partie de ces problèmes, empêchent d'envisager une solution unique à un problème particulier. Cette spécificité a pour principale conséquence de laisser l'expérimentateur libre de la conception orientée d'une chaîne de reconnaissance adaptée au problème qu'il se propose de résoudre.

Cependant, il est communément admis que cette chaîne peut être décomposée selon une structure en 5 blocs fonctionnels. La solution proposée par l'expérimentateur est développée suivant 5 étapes présentées dans la Figure 1 [Postaire 87]. Chaque bloc réalise une

action, assurant la transformation de la nature de l'information à transmettre aux blocs suivants, sans préciser les moyens pour le faire. Les actions des différents blocs concourent à l'affectation d'une étiquette ou d'un symbole à une forme.

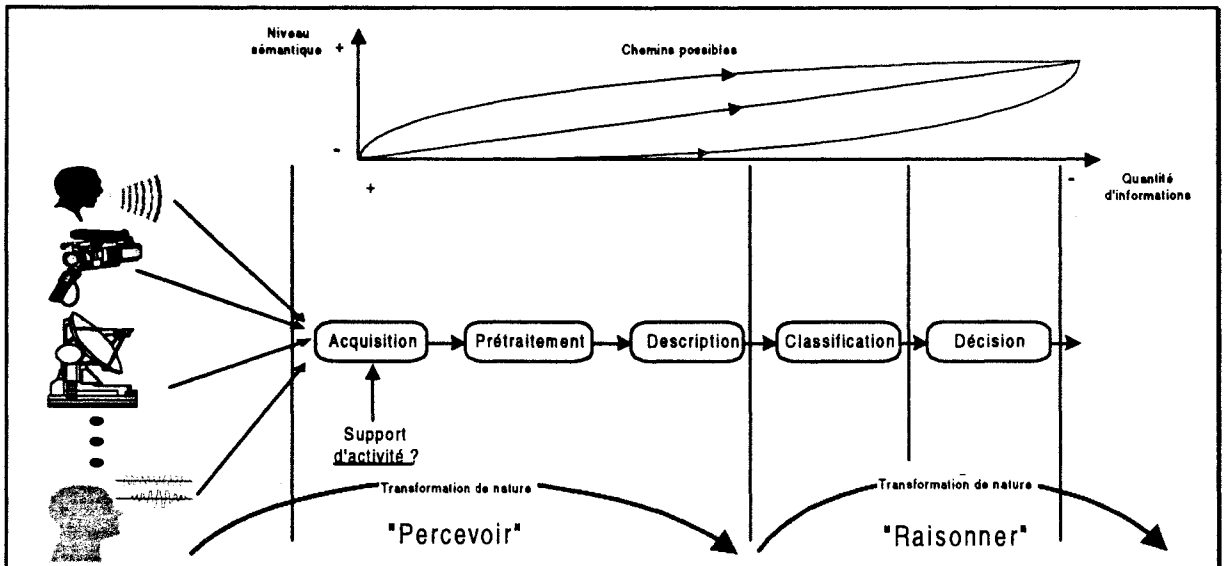


Figure 1 : Décomposition fonctionnelle d'une chaîne de reconnaissance de formes

Les fonctions réalisées par ces différents blocs ne relèvent pas toutes, à proprement parler, de la reconnaissance de formes, au sens des définitions, mais de domaines connexes. La perception est d'une importance capitale pour la reconnaissance [Milgram 93], [Postaire 87], de par sa position en aval de la chaîne Perception Raisonnement. La perception réalise la transformation de nature de l'information, depuis le monde analogique continu vers le monde numérique discret. Le raisonnement permet de passer du monde numérique discret vers le monde symbolique.

La maîtrise de la conception d'une chaîne complète est difficile car elle passe par la maîtrise des différents domaines impliqués. L'interaction des différentes étapes est tout aussi primordiale, puisque toute information perdue à un niveau de la chaîne le reste pour la suite des opérations. Toute information, dont le sens est diminué par une étape, nécessite un nombre d'opérations importantes, pour l'extraire à nouveau dans les étapes ultérieures. Comme l'a exprimé Lorette [Lorette 92], le niveau de performance des différentes opérations est à définir lors de la conception, puisque chacun des chemins possibles peut mener plus ou moins facilement à une solution de reconnaissance (Figure 1). La connaissance du problème à résoudre, ainsi que l'expérience du concepteur peuvent conduire à un développement plus poussé de l'étape de Perception, ou de l'étape de Raisonnement. Cependant, les performances globales de la chaîne ne peuvent être qu'estimées a priori, et vérifiées a posteriori. L'élaboration d'une chaîne complète est très expérimentale, il n'existe pas de définition automatique des supports des différentes activités présentées dans la Figure 1. Les performances de la reconnaissance dépendent donc de l'expérience acquise dans les domaines mis en jeu, de la connaissance générale du procédé étudié et des moyens mis en oeuvre.

1.2.2. La perception

1.2.2.1. L'acquisition

L'étape d'acquisition est la première de tout système de traitement automatique de l'information, et particulièrement en reconnaissance de formes. Elle s'appuie sur des connaissances en physique, en électronique. C'est l'instrumentation de la chaîne. Elle convertit des valeurs physiques analogiques en valeurs numériques. Le choix approprié et crucial des capteurs oriente le type d'information que l'on veut obtenir. Malheureusement, les contraintes physiques interdisent parfois l'utilisation d'un capteur pour une transduction optimale de l'information. On rencontre notamment cette situation lorsque l'on veut qualifier ou quantifier des phénomènes subjectifs tels que le confort thermique ou postural. Des capteurs flous ou symboliques sont actuellement le point de mire de recherches pour résoudre ces problèmes d'ordre subjectif [Benoît 93], [Bouchon-Meunier 94], [Levrat 96], [Hubert 95].

L'efficacité de l'acquisition, outre le choix du capteur approprié, tient aussi à l'environnement physique de cette acquisition. La vision par ordinateur illustre concrètement ce propos, où le choix de l'éclairage sur la scène est souvent plus important que le choix du capteur. L'analyse et la prise en compte sérieuses de cette étape permettent de simplifier à l'extrême les tâches aval de la chaîne de reconnaissance de formes, où les algorithmes les plus astucieux ne viennent pas toujours à bout de problèmes sous contraintes, rencontrés souvent dans l'industrie.

1.2.2.2. Prétraitements de l'information

L'opération de prétraitement est engagée dans le but précis de dégager au mieux l'information utile, en se donnant les moyens de la rendre la plus présente dans la représentation numérique fournie par le capteur. A un degré plus ou moins fort, elle est indispensable.

Le prétraitement est destiné à préparer les phases ultérieures de la reconnaissance, c'est une phase d'épuration qui élimine certaines informations, soit inutiles, soit pouvant conduire à la génération d'ambiguïtés de reconnaissance. Il relève du domaine du traitement du signal mono ou bi-dimensionnel selon le capteur utilisé.

Cette étape préparatoire des données est généralement constituée d'opérations de conditionnement, d'amélioration (filtrage), ou de restauration en relation avec le contexte d'acquisition (bruit de mesure, mélange de signaux, ...). Les algorithmes de traitement du signal qui sont engagés dans cette étape peuvent être déportés, dans certains cas, vers une réalisation électronique. Dans des applications sur des données images, de nombreuses opérations sont implémentées sur circuits électroniques libérant les tâches d'un calculateur (e.g. imagerie satellite).

1.2.2.3. Description

La description, dernière étape de la partie Perception, donne la représentation numérique finale de l'objet analogique. Elle joue un rôle de préparation des données,

d'extraction et de synthèse de l'information. Elle vise la mise en évidence des caractéristiques invariantes des formes observées, par rapport à la décision. Le bloc fonctionnel de description des données relève des mathématiques et du traitement du signal.

La description fournit un vecteur caractéristique de la forme, d'une dimension inférieure ou égale à la dimension originale de la forme numérique. Ce vecteur est obtenu par une phase de segmentation en entités élémentaires portant de façon pertinente l'information, puis d'extraction de caractéristiques des segments. Ces deux opérations sont réalisées par ce que nous appellerons dans la suite de notre mémoire, le facteur de formes. L'espace de représentation des formes doit être celui où la séparabilité des classes est la plus forte possible. Plus cette étape réalise le compromis entre la qualité et la quantité d'information, plus l'étape de classification/discrimination en aval peut être simplifiée relativement au contexte d'application. La Figure 2 illustre cette remarque.

L'étape de perception joue le rôle d'un filtre sur l'information, puisque les formes analogiques connues et à reconnaître sont vues au travers d'elle. Aussi, la relation réelle entre les objets du monde analogique et les classes de formes est transformée selon la complexité de la description. La détermination de cette relation ne peut donc être réalisée qu'entre les caractéristiques extraites par la description et les classes de formes associées à la décision.

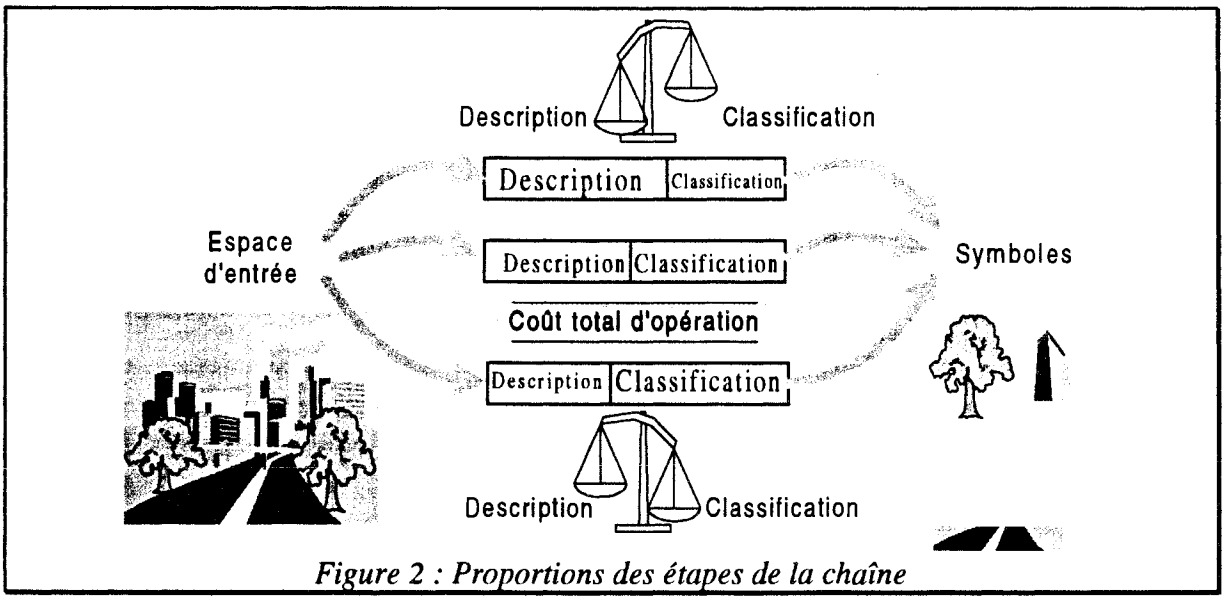


Figure 2 : Proportions des étapes de la chaîne

La perception joue un rôle primordial pour la qualité du raisonnement puisqu'elle définit un point de vue particulier sur les données, une transformation supposée simplificatrice.

Pour clore cette présentation de l'étape de Perception, nous pouvons remarquer que l'ensemble des opérations engagées dans cette étape (acquisition, prétraitements, description) bien que visant à augmenter la qualité et la pertinence de l'information, sont également des sources d'imprécision, qui induisent une ambiguïté de distinction des formes, pouvant augmenter la difficulté de l'étape de Raisonnement.

1.2.3. Le raisonnement

1.2.3.1. Introduction

Cette étape relève du domaine de l'informatique, de la reconnaissance de formes, de l'intelligence artificielle. Son rôle pourrait se résumer à « apprendre pour connaître, puis reconnaître » justifiant à notre avis l'appellation d'étape de Raisonnement. Elle définit une relation entre les formes perçues et les classes de formes liées à la décision. Dans l'espace de représentation des formes, fourni par la Perception, cette étape doit déterminer les classes de formes de caractères invariants, la classification, pour ensuite prendre une décision d'affectation, la discrimination [Lorette 92].

La recherche d'une relation entre les formes perçues et les classes de formes est conduite selon la connaissance que l'on possède du problème posé. Cette connaissance est en partie contenue dans un lot de formes d'apprentissage, dans lequel on va chercher à « connaître son problème ». L'apprentissage est donc le départ obligé de la phase de Raisonnement pour déterminer la relation intrinsèquement exprimée par les données. Le raisonnement consistant à déterminer les règles de décision à partir des données, est un mode de raisonnement des plus classiques dit inductif.

La maîtrise de l'apprentissage, dans la recherche d'une relation, est extrêmement délicate. Puisque la relation est déterminée à partir d'exemples, seul ce qui est exprimé peut être trouvé. Cela soulève la question de la partialité de la vue de la relation au travers des données, associée à l'exhaustivité des données d'apprentissage par rapport au problème posé. La relation n'est alors qu'approximative, sa qualité et sa robustesse dépendent du nombre de données d'apprentissage et de leur représentativité du problème de reconnaissance [Fukunaga 90].

Aussi, la relation n'est figée définitivement que dans l'hypothèse où toutes les données exprimant "clairement" la nature du problème sont présentes. Cependant, la relation peut être considérée comme satisfaisante à un instant donné, avec pourtant un nombre restreint d'informations. Cette situation laisse tout de même la possibilité, par l'introduction de nouvelles données, d'atteindre ou de converger vers une solution générale et définitive, qui est le but final de la reconnaissance. On fait apparaître ici la notion d'apprentissage permanent pour lequel on peut faire le parallèle avec l'apprentissage humain dont les capacités apporteraient de nombreux atouts à un système de reconnaissance de formes [Dubuisson 90].

1.2.3.2. Apprentissage

Une phase d'apprentissage démarre sur la base d'informations qui définissent l'acquis. En cours d'apprentissage, les nouvelles informations permettent de confirmer ou d'infirmer ce que l'on sait déjà, représenté par des regroupements de formes similaires. On cherche également à affiner la définition des classes de formes, en cernant au mieux leurs limites.

Une classe peut être formée de sous-classes, (le concept de méta-classe proposé par Dubuisson [Dubuisson 90]). Il faut établir alors les relations entre les cas généraux et ceux moins généraux constituant une méta-classe que l'on peut voir comme un symbole de sémantique élevée. Dans cet agrégat de sous-classes sémantiquement liées, disjointes ou non-disjointes, il est nécessaire d'interpoler entre les différentes sous-classes, pour obtenir une

définition des objets situés à l'intérieur de la méta-classe. Aux abords de ses limites extérieures, une légère extrapolation rend la décision plus souple, et permet également d'absorber des phénomènes d'évolution des classes. Simultanément, lorsque l'extrapolation n'est pas trop souple, elle autorisera la création de nouvelles sous-classes définissant de nouvelles connaissances. L'extrapolation, l'interpolation, associées à la réintégration de données dans la définition des sous-classes, permettent d'augmenter la connaissance du système.

La possibilité d'évolution de la représentation des connaissances pose le problème de comportement de l'apprentissage dont on doit ici se préoccuper. Harnad [Harnad 87] présente un point de vue général sur l'apprentissage et les comportements que l'on risque de rencontrer. L'apprentissage doit converger vers une formulation stable des connaissances. Ainsi, une situation d'ordre est atteinte et il n'y a plus d'apprentissage possible. Inversement, le désordre est une situation où l'apprentissage doit s'exercer, pour obtenir à nouveau une situation d'ordre. Dans un contexte d'apprentissage permanent, on doit retrouver ces deux situations successivement et la situation terminale doit être une situation d'ordre, stable.

Pour la prise de décision durant l'apprentissage, situation obligatoire lors d'un apprentissage permanent, la relation est déterminée, mais provisoire. On peut alors considérer que la décision est satisfaisante à un moment donné, aux vues de la connaissance possédée à cet instant. L'introduction de nouvelles données relance l'apprentissage et la décision doit être continuellement prise par rapport à la représentation instantanée des connaissances.

On distingue classiquement deux types d'apprentissage :

- L'apprentissage supervisé ou avec professeur, dans lequel le nombre de classes et la classe de chaque objet sont connus a priori. La classification recherche la règle qui minimise l'erreur d'étiquetage sur les données d'apprentissage. Il est donc nécessaire de posséder en même temps formes et étiquettes définies par le professeur, qui est en général un expert du domaine d'application. La règle trouvée est ensuite employée pour la discrimination de nouvelles formes.

- L'apprentissage non supervisé, ou sans professeur, où il n'y a pas spécification des classes, ni des objets attribués aux classes. L'opération de raisonnement doit structurer l'ensemble d'apprentissage en classes, puis définir la règle de discrimination relative aux classes formées. La définition des classes est réalisée par une méthode de coalescence, méthode algorithmique, qui itère jusqu'à l'optimisation d'un critère qui lui est propre (inertie intra-classe minimum, ...), afin d'obtenir les classes et les étiquettes associées.

Cette situation d'apprentissage se trouve couramment lorsqu'il n'y a pas accès à la connaissance d'un expert, ou lorsqu'il est trop coûteux de posséder les étiquettes. L'apprentissage non supervisé représente la plus grosse difficulté pour la reconnaissance de formes.

Enfin, on peut citer l'apprentissage semi-supervisé qui est une composition des deux notions précédentes. En effet, il existe parfois des situations où l'expert ne possède qu'une partie de l'information. Selon Pedrycz [Pedrycz 90], même une faible proportion de connaissance améliore significativement les résultats, aussi dès que cela est possible, il est conseillé de faire intervenir la connaissance extérieure.

1.2.3.3. Les méthodes

1.2.3.3.1. Introduction

En reconnaissance de formes, la résolution d'un problème est guidée par l'information caractéristique qui le spécifie. La reconnaissance de formes se divise en deux grandes familles de méthodes qui sont la reconnaissance globale [Milgram 93], [Postaire 87] ou vectorielle [Heutte 94], et la reconnaissance structurelle [Miclet 84]. Chacune des familles est liée à la nature de la représentation des formes.

Les méthodes relevant de la reconnaissance structurelle exploitent la structure des objets connus et à reconnaître. Ces méthodes se basent sur des outils adaptés à la représentation de structures comme les graphes [Rocha 94], les arbres ou encore à partir de représentations syntaxiques. Ces méthodes peuvent cependant présenter des inconvénients lorsque les représentations des formes sont complexes et sensibles à des influences extérieures [Schneider 92]. Nous pouvons remarquer dans certaines applications que la théorie des ensembles flous renforce l'efficacité des méthodes de reconnaissance structurelle [Miclet 84], en raison de sa souplesse. La littérature contient des travaux dans ce sens [Lan 95], [Craine 94].

A la différence des méthodes structurelles, les méthodes globales exploitent des caractéristiques ou attributs pris sur les formes. Ces méthodes utilisent une représentation vectorielle qui place la forme comme un point dans un espace de représentation. Certains auteurs, [Heutte 94], [Milgram 93], intègrent dans la reconnaissance globale, la reconnaissance directe où la perception des formes est « rétinienne ». Une telle approche peut convenir lorsque les modèles de formes sont peu complexes, de taille raisonnable, et les formes peu variables, sinon, le nombre de modèles croît très rapidement. Ce type d'approche convient assez bien à la reconnaissance de caractères [Crettez 89] [Lebourgeois 91] et des développements particuliers ont permis de s'affranchir de la complexité des modèles [Krattenthaler 94], [Epstein 94]. La théorie des ensembles flous permet de s'affranchir de l'inconvénient de la multiplication des modèles, par l'utilisation d'opérateurs appropriés [Wang 95].

Cependant, plusieurs travaux ont montré que la reconnaissance structurelle [Tappert 90] ainsi que les méthodes directes [Gaillat 79] sont moins adaptées à la reconnaissance hors-ligne de l'écriture manuscrite. Aussi, nous ne détaillerons pas plus ces méthodes, pour nous focaliser sur la reconnaissance globale.

1.2.3.3.2. La reconnaissance globale

Dans le cadre d'une approche globale, on doit se poser la question : « Connaît-on le nombre de classes présentes, et possède-t-on des prototypes, c'est-à-dire des formes étiquetées à ces classes ? » La réponse conditionne le séquencement des opérations de raisonnement à réaliser (Figure 3).

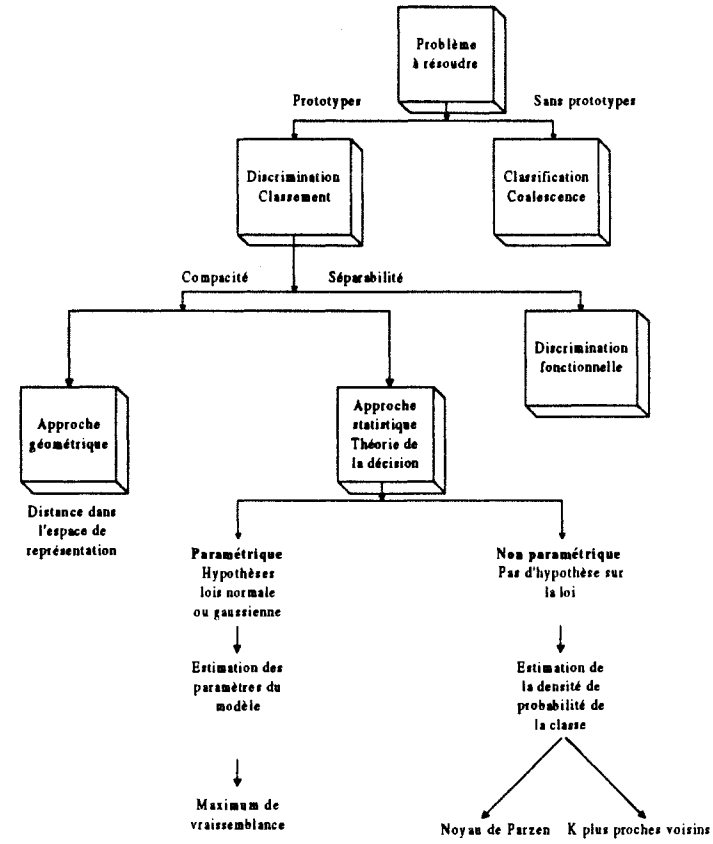


Figure 3 : Les approches possibles pour une reconnaissance vectorielle

1.2.3.3.2.1. Les méthodes de classification automatique

Si on ne possède aucune connaissance sur les classes en présence, et sur les étiquettes des données, on utilise des méthodes de coalescence [Dubuisson 90] ou de classification automatique [Postaire 87] qui vont structurer les données en classe selon un critère propre à l'algorithme. On cherchera à optimiser ce critère, en restructurant de manière itérative les données. La classification automatique est un problème difficile puisque l'on cherche à extraire l'information exprimée par les données sans disposer de prototypes, sans connaître le nombre et les caractéristiques (statistiques ou géométriques) des classes. Postaire [Postaire 87] précise que dans ce cas, la seule information accessible est la fonction de densité de probabilité sous-jacente à l'ensemble des observations à classer. Si une hypothèse de loi normale est prise, alors la classification est un problème paramétrique, et revient à identifier un mélange gaussien. Postaire [Postaire 87] donne une méthode générale sans connaissance a priori pour résoudre ce problème. Il existe d'autres méthodes, mais nécessitant des connaissances a priori comme l'égalité des matrices de covariance ou le nombre de classes [Celeux 88], [Geva 89]. Si aucun modèle probabiliste ne peut être défini, l'approche est alors non paramétrique. On peut trouver de nombreux développements de ces méthodes de classification automatique dans [Postaire 87], [Dubuisson 90], ainsi que des approches floues de la coalescence [Bezdek 81]. Certains travaux utilisent une approche floue de la coalescence pour définir une structure particulière des classes. Dans [Jajuga 91] et [Bobrowski 91], les classes formées sont de type hyper-cubique grâce à l'utilisation de la distance de Minkowski dans L^1 ou L^∞. On peut citer

également les travaux de Gustafson [Gustafson 79] dans lesquels les classes recherchées sont de forme hyper-elliptique. En outre, la complexité des classes formées a évolué jusqu'à des formes très particulières par l'utilisation des Fuzzy C-shells [Dave 92a], [Dave 92b] et [Krishnapura 92].

1.2.3.3.2.2. Les méthodes de discrimination

Dans l'alternative, nous nous dirigeons vers une méthode de classification/discrimination ou de classement. Dans ce cas une condition de séparabilité ou de compacité des classes conduit à choisir des méthodes de classement différentes.

L'hypothèse de séparabilité suppose qu'il est possible de séparer les prototypes par une frontière de décision qui minimise l'erreur de classement, c'est la discrimination fonctionnelle. On choisit en général des fonctions simples, de type polynomial, linéaires ou quadratiques. La nature de ces fonctions est à relier à des hypothèses statistiques des classes pour lesquelles la décision optimale conduit à ce type de frontières [Duda 73]. Dans la discrimination fonctionnelle, les problèmes multi-cas nécessitent la détermination de nombreuses surfaces pour former des zones de l'espace de représentation où se situent les classes. Mais, l'accroissement du nombre de frontières entraîne une augmentation du nombre de zones, dont l'affectation à une classe unique devient plus difficile voire impossible [Heutte 94].

L'hypothèse de compacité conduit à deux approches. Une approche géométrique, où l'on cherche à former des classes compactes, sur la base des distances séparant les prototypes, ou une approche statistique se référant aux fonctions de densité de probabilité des classes.

♦ Approche géométrique

Dans l'approche géométrique, le classement se base sur la notion de distance comme évaluation du concept de ressemblance. Ainsi, les distances séparant les observations d'une même classe sont normalement plus petites que celles séparant les observations de classes différentes. Cette approche consiste donc en la définition de paramètres de réglages comme des seuils [Sabourin 94a] ou des pondérations, appliqués aux distances pour réaliser le classement. Ces méthodes sont simples à mettre en oeuvre, mais on ne peut garantir un taux d'erreur minimum [Postaire 87].

♦ Approche statistique

L'autre approche possible dans le cas de la compacité des classes est l'approche statistique, c'est-à-dire utilisant explicitement des fonctions de densité de probabilité. Les observations sont alors la réalisation d'un vecteur aléatoire continu dont la distribution est complètement définie par la fonction de densité de probabilité multi-variables, caractéristique de la classe à laquelle les observations appartiennent [Postaire 87].

♦ La théorie de la décision

Dans ce cadre, la théorie de la décision constitue une approche fondamentale pour la reconnaissance de formes. Le point de vue statistique utilisé lui donne son caractère de méthode de référence. En effet, cette théorie se base sur des lois de probabilités générales. Les lois complètement définies et l'hypothèse d'un nombre infini d'échantillons considérée

impliquent une connaissance parfaite du problème. Ainsi, la théorie de la décision donne directement la règle de décision optimale de Bayes au sens de l'erreur de discrimination, et détermine la borne minimum par la probabilité d'erreur.

Si on considère un problème à deux classes avec deux densités de probabilités connues, la règle de Bayes permet de déterminer la décision d'affectation d'un objet à la classe C_1 ou à la classe C_2 à partir de la probabilité a posteriori $P(C_i|X)$ de l'objet.

L'objet X appartient à la classe C_1 si $P(C_1|X) > P(C_2|X)$ *(Eq. 1)*

X appartient à la classe C_2 sinon

où $P(C_i|X) = \dfrac{p(X|C_i)P(C_i)}{\sum_{j=1}^{2} p(X|C_j)P(C_j)} = \dfrac{p(X|C_i)P(C_i)}{p(X)}$

$p(X|C_i)$ est la densité de probabilité conditionnelle de X sachant C_i,

$P(C_i)$ est la probabilité a priori de C_i, et $P(C_i|X)$ est la probabilité a posteriori d'appartenance à la classe C_i si on se trouve au point X.

Ainsi, la règle de décision de Bayes (Eq. 1) permet d'attribuer pour chaque point X l'appartenance à la classe C_i si $P(C_i|X) = p(X|C_i)P(C_i)$ est maximum. Le mode d'affectation posé en terme de probabilités minimise la probabilité d'erreur (Eq. 2)

$$P(erreur) = \int_{-\infty}^{+\infty} P(erreur|X)p(X)dX \qquad \textit{(Eq. 2)}$$

La règle de Bayes définit le seuil de décision quelles que soient les lois de probabilités, lorsque celles-ci sont parfaitement connues. La probabilité d'erreur est donc la mesure idéale pour déterminer les performances d'un système de reconnaissance de formes.

L'application de la règle optimale de Bayes présente cependant un inconvénient majeur. L'hypothèse probabiliste qui s'y rattache demande un nombre infini de données ainsi que des informations précises sur les probabilités mises en jeu (probabilités conditionnelles et probabilités a priori) dont on ne dispose que très rarement dans un problème pratique. Aussi, en suivant une approche statistique, deux cas peuvent se présenter selon l'hypothèse de l'existence d'un modèle probabiliste ou non. Ces approches sont respectivement paramétrique et non paramétrique.

◊ Approche paramétrique

Dans le cas paramétrique, le problème se résume à l'estimation des paramètres d'une loi de densité de probabilité des classes. Cependant, l'estimation des paramètres statistiques, à partir d'échantillons, ne s'applique de façon réaliste, que dans le cas de lois normales. Il suffit alors d'estimer la moyenne et la matrice de variance covariance des lois associées aux classes.

La méthode la plus courante est l'estimation du maximum de vraisemblance qui consiste à rendre maximum la probabilité d'observer les données [Milgram 93], [Dubuisson 90]. On obtient ainsi les estimées du vecteur moyenne et de la matrice de variance covariance qui définissent complètement la loi. On peut alors directement appliquer la règle de Bayes.

A partir de ces estimations, on peut utiliser plusieurs classificateurs. Un classificateur couramment utilisé est le classificateur à distance minimum de Mahalanobis où la distance d'un point aux moyennes des classes tient compte des matrices de variance covariance des classes. Ce classificateur affecte une observation à la classe pour laquelle la distance l'en séparant est minimum [Fukunaga 90], [Raudys 91] et [Basseville 88]. On trouve également des classificateurs plus complexes comme le classificateur quadratique par morceaux [Smolarz 87].

L'hypothèse de normalité, bien que souple et simplificatrice, peut conduire à des performances moins intéressantes si la loi réelle régissant les données s'en écarte.

◊ Approche non paramétrique

Ici, on ne pose aucune hypothèse de loi précise, il s'agit donc d'estimer la densité de probabilité à l'aide de méthodes non paramétriques.

Le principe d'estimation relève d'une hypothèse simple. Si l'on admet qu'une fonction de densité de probabilité ne présente pas de variations importantes à l'intérieur d'un domaine limité de l'espace de représentation, alors on peut considérer que la valeur de la fonction de densité, en tout point de ce domaine, est la valeur moyenne de la fonction de densité sur ce domaine [Postaire 87], [Dubuisson 90]. Si N prototypes sur les M existants sont dans ce domaine, alors le rapport N/M constitue une estimation simple de la fonction de densité de probabilité sur le domaine considéré.

Pour réaliser cette estimation, il est nécessaire de définir le domaine de l'espace de représentation. Deux estimations sont alors possibles. Dans la première, on considère un domaine de taille fixe, lié au nombre de prototypes, de façon à pouvoir réaliser l'estimation. La méthode des noyaux de Parzen fonctionne sur ce principe [Parzen 62]. Dans la seconde, on fixe un nombre de prototypes, et on modifie la taille du domaine pour les trouver, c'est la méthode des k plus proches voisins [Cover 67].

Dans la méthode des noyaux, le domaine d'estimation locale de la densité de probabilité est posé a priori. L'inconvénient d'une telle procédure est que l'estimateur n'a pas une résolution suffisante si le domaine est trop grand. Inversement, un domaine trop petit donne une grande sensibilité de l'estimateur. Il peut même arriver qu'en réduisant la taille du domaine, l'estimation soit nulle car aucune donnée n'est présente dans cette zone. Le choix de la taille du domaine est donc crucial, et résulte d'un compromis entre une estimation locale précise mais variable, et une estimation peu variable où on ne voit pas les variations importantes de la fonction de densité.

Pour ne pas fixer la taille du domaine, on utilise la méthode des k plus proches voisins. Cette approche consiste en l'ajustement de la taille du domaine d'estimation en fonction de la densité des prototypes dans le voisinage d'un point. Le domaine va croître jusqu'à ce qu'un nombre fixé de prototypes soit considéré. Si la densité des prototypes est élevée, le domaine

sera automatiquement petit et l'estimateur sera précis. En revanche, si la densité des prototypes est faible, alors le domaine d'estimation sera grand, et l'estimation moins précise.

Il est important de préciser que les deux estimateurs convergent vers la densité de probabilité réelle lorsque le nombre de données est infini. La démonstration de la convergence de la règle du plus proche voisin vers la règle de décision Bayésienne, dans l'hypothèse de lois de probabilités normales ou gaussiennes ainsi qu'un nombre infini d'échantillons, en fait une référence dans le cadre non paramétrique [Cover 67]. Cependant, sur un nombre limité d'exemples, on ne sait pas situer la performance du classificateur. La seule certitude vient sur la convergence asymptotique vers une fenêtre de performances ($P_{Bayes} < P_{ppv} < 2P_{Bayes}$) quand le nombre d'échantillons tend vers l'infini. Dans le cas d'un nombre limité d'échantillons, cette performance n'est plus assurée, et il arrive qu'avec un nombre insuffisant d'exemples les performances soient meilleures avec une règle des k plus proches voisins ($k \neq 1$), notamment avec un recouvrement des classes pour lequel une règle du plus proche voisin est sensible au bruit.

La performance est donc liée au nombre de voisins, ainsi qu'au type de distance utilisé, puisque cette combinaison modifie directement le comportement du classificateur. Généralement, le choix de la distance se porte sur la distance euclidienne car les autres distances ont un effet plus difficile à appréhender dans ce classificateur. Néanmoins, Cover a démontré la propriété de convergence quelle que soit la distance, et Raudys [Raudys 91] signale que la distance de Mahalanobis peut parfois conduire à de meilleurs résultats de discrimination, en utilisant comme pondération la variance naturelle des données selon chacun des attributs des formes.

Le fait que la règle du plus proche voisin amène une bonne performance suggère que la plupart de l'information soit contenue dans le plus proche voisin. Cependant, l'expérience montre qu'on peut atteindre, avec un nombre limité d'exemples, des performances supérieures avec plus de voisins. Il faut adapter la valeur de k à l'application et Bezdek [Bezdek 86] propose de procéder par itérations pour définir la meilleure valeur de k.

La règle des k plus proches voisins avec $k \neq 1$ impose une attribution de classe sur la base d'un vote majoritaire. Or, on peut retrouver des configurations de voisin où seul un voisin provoque l'étiquetage, ce qui rend la décision peu certaine. La méthode de (k,l) plus proches voisins [Devijver 82] [Dubuisson 90], qui est la généralisation de la méthode des k plus proches voisins, renforce la certitude sur la décision. Son fonctionnement est peu différent puisque le principe consiste à attribuer l'étiquette à un point si au moins l voisins sur les k sont de la même classe, sinon on attribue une étiquette de rejet. Cette méthode permet également l'utilisation d'un l différent par classe. Cependant, cela pose le problème du choix de l qui ne doit pas être trop élevé, car il implique aucune classification, ni trop bas car la règle peut proposer plusieurs choix [Bezdek 86].

Malgré l'intérêt majeur que peut présenter la méthode générale des (k,l) plus proches voisins, cette méthode souffre d'un inconvénient qui est propre aux méthodes sans compilation, la complexité. Les méthodes non paramétriques impliquent la conservation de tous les points de référence, le nombre de calcul pour la recherche des k plus proches voisins est donc important, ce qui rend la méthode lourde en temps de calcul et en place mémoire. De plus, si on ne conserve pas suffisamment de points, la convergence asymptotique vers la performance maximale ne peut être obtenue. Ces dernières remarques font que la règle des k plus proches

voisins doit être appliquée avec précaution. Cependant, les extensions floues de ces méthodes permettent d'améliorer notablement les performances des k plus proches voisins ordinaires lorsque le nombre de prototypes est réduit [Kissiov 92], [Denoeux 95]. Béreau [Béreau 86] propose une règle floue des k plus proches voisins où un rejet de distance est intégré, ce qui conduit également à l'amélioration des performances.

1.2.3.3.2.3. Estimation de l'erreur de discrimination

Dans la pratique, l'expression analytique de la probabilité d'erreur n'est pas accessible. L'erreur ne peut alors être qu'estimée à partir d'un nombre fini d'échantillons. Un certain nombre d'estimateurs de cette erreur existent [Raudys 91], mais le seul estimateur non biaisé est le pourcentage d'observations mal classées [Fukunaga 90].

Pour pouvoir quantifier cette erreur, il faut établir la loi de discrimination à partir d'un lot d'apprentissage, puis estimer l'erreur à partir d'un lot de test. Ces deux lots de données sont pris statistiquement indépendants, ou au moins différents. Fukunaga [Fukunaga 90] précise que le biais de l'estimation de l'erreur est provoqué par la taille finie de l'ensemble d'apprentissage, et que la variance de l'estimation provient de la taille finie de l'ensemble de test.

L'estimation de l'erreur est donc dépendante de la technique de construction des ensembles d'apprentissage et de test. La méthode de resubstitution utilise le même ensemble pour l'apprentissage et le test. L'estimation est donc biaisée, optimiste avec une variance réduite.

La méthode « holdout » consiste à partager aléatoirement les échantillons en deux lots, un pour l'apprentissage, l'autre pour le test. La difficulté de cette méthode est de savoir comment diviser les échantillons en deux groupes, et fixer la taille de ces groupes. Cette technique présente l'inconvénient de réduire la taille des lots, et donne une estimation pessimiste biaisée.

La méthode de validation croisée consiste à former C_n^k classificateurs utilisant k des n échantillons pour l'ensemble d'apprentissage et $(n\text{-}k)$ échantillons pour l'ensemble de test. L'opération est répétée pour tous les k et le taux moyen d'erreur est estimé. Lorsque $k=1$, cette méthode est celle du « leave-one-out » donnant une estimation non biaisée à grande variance. Cette méthode est intéressante mais coûteuse en temps de calcul.

La méthode « bootstrap » échantillonne les données avec remplacement pour donner de nouveaux ensembles. Cette technique est très coûteuse en temps de calcul. L'estimation de l'erreur est plus précise que la méthode « leave-one-out » uniquement lorsque l'erreur de discrimination est élevée.

Finalement, l'estimée de la probabilité d'erreur de discrimination est la moyenne de l'estimée par la méthode « leave-one-out » et de l'estimée par resubstitution [Raudys 91], mais on préfère donner un intervalle de l'estimation de l'erreur dont les bornes sont les deux estimations précédentes [Fukunaga 90].

1.2.3.3.3. Coopération de classificateurs

En reconnaissance de formes, le choix de l'utilisation d'un classificateur dépend essentiellement du contexte applicatif et de l'expérience de l'expérimentateur. De plus, chacune des méthodes possède des spécificités de fonctionnement, et donc les performances dépendent de l'application et des caractéristiques choisies. Comme aucune méthode n'a de suprématie dans un contexte universel et que chacune possède des qualités particulières adaptées à certaines catégories de problèmes, on cherche à exploiter les différents apports par la coopération des classificateurs.

Selon l'avis de Lorette [Lorette 92], il faut faire collaborer au sein d'un même système des sources de connaissances multiples représentant plusieurs couches de niveaux contextuels (morphologique, linguistiques, pragmatiques, ...). Cependant, l'utilisation de données de nature hétérogène (continues, discrètes, binaires, structurelles) dans le même classificateur est délicate. Pour l'instant, on trouve beaucoup d'approches linéaires ou hiérarchiques du problème (analyse par le flux d'informations, approche cartésienne des systèmes compliqués [Durand 94]). Lorette souligne qu'il faudrait aborder le problème par un système hétérarchique (complexe au sens systémique du terme) composé de systèmes bouclés avec interactions multiples entre modules ou avec modules indépendants collaborant entre eux pour réaliser la tâche commune. Il faut définir la stratégie générale, gérer les interactions entre modules, utiliser des méta-connaissances et résoudre les conflits entre systèmes [Lorette 92].

Heutte [Heutte 94] relève trois méthodes pour la coopération de classificateurs. L'approche séquentielle où le principe consiste à fournir à une méthode de reconnaissance de formes les résultats d'une reconnaissance préalable et ainsi de suite dans un enchaînement de méthodes similaires ou différentes. L'enchaînement en niveaux hiérarchiques séquentiels permet de renforcer une décision, rejeter des objets entachés d'une incertitude de décision, ou de condenser le nombre potentiel de décisions vers une solution finale. La combinaison parallèle place les méthodes de reconnaissance à un niveau hiérarchique équivalent et toutes les propositions des méthodes sont combinées dans une décision finale par un vote démocratique, ou par un vote dirigé si un poids est attribué aux différentes décisions. Enfin, la méthode hybride opère dans le sens d'une composition selon une architecture à la fois séquentielle et parallèle des propositions initiales.

Le dilemme de la coopération de méthodes, pouvant d'ailleurs être perçue comme une coopération d'experts, est qu'aucun schéma n'est pré-définissable pour atteindre la solution optimale. Encore une fois, la performance apportée par la coopération dépend de l'expérience de l'expérimentateur et de sa connaissance sur la procédure de reconnaissance qu'il introduit par la conception de la structure de coopération. En ce sens, la coopération hybride est la plus délicate à exercer par la maîtrise des interactions et la gestion de la structure.

1.2.3.3.3.1. Coopération séquentielle ordinaire, et coopération séquentielle floue.

- Principe

La coopération séquentielle propose l'enchaînement d'opérations de reconnaissance de formes différentes ayant des propriétés différentes. On envisage également dans la coopération séquentielle d'utiliser une méthode de reconnaissance fixe appliquée successivement à des paramètres extraits des objets à reconnaître.

♦ Approche ordinaire

Une telle coopération séquentielle a été proposée par Fu [Fu 68] où la classification est réalisée paramètre par paramètre pour assurer la discrimination des formes. Le principe général est d'évaluer un critère qui permet de classer les méthodes sur les paramètres, par ordre croissant de leur pouvoir discriminant, afin de définir l'ordre d'enchaînement des méthodes (Figure 4).

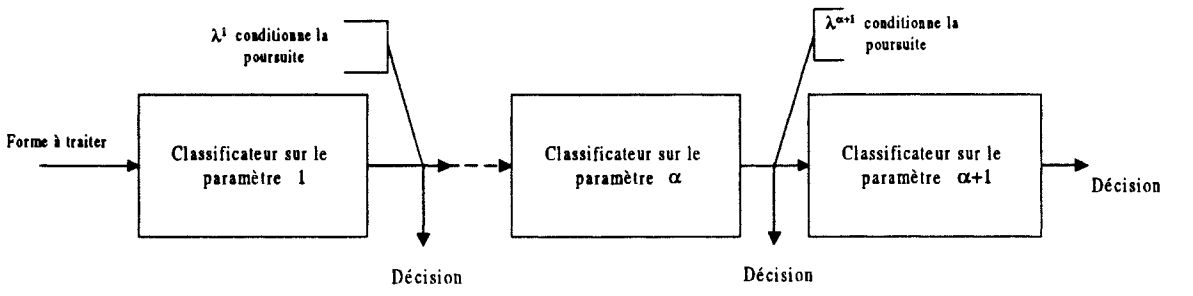


Figure 4 : Principe d'enchaînement séquentiel de classificateur

Dans cet objectif, Fu propose d'évaluer le risque d'erreur de classification par le rapport de vraisemblance séquentiel λ^{α}, calculé sur les caractéristiques α extraites des formes :

$$\lambda^{\alpha} = \frac{P\left(X^{\alpha}|C_i\right)}{P\left(X^{\alpha}|Cj\right)} \qquad \textit{(Eq. 3)}$$

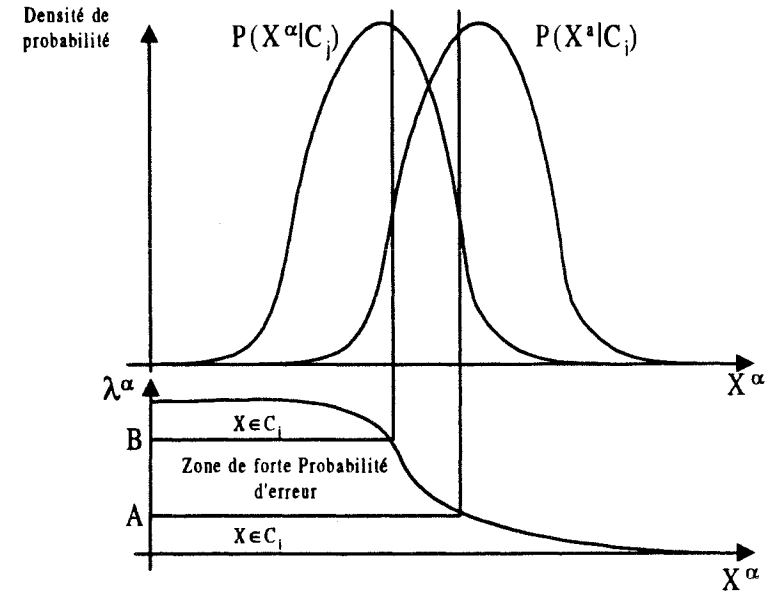


Figure 5 : Bornes sur le rapport séquentiel de vraisemblance

En posant deux bornes A et B sur le rapport λ^{α} (SPRT), on peut cerner la probabilité d'erreur ou de confusion (Figure 5). Ainsi, si à une étape α, correspondant à une caractéristique particulière, le rapport de vraisemblance est compris entre A et B, alors la probabilité de confusion est grande et il est nécessaire d'utiliser une caractéristique supplémentaire pour affiner la discrimination. Dans le cas contraire, la donnée de test X appartient à la classe C_i si :

$$\lambda^{\alpha} \geq A = \frac{1 - e_{ji}}{e_{ij}}$$

ou à C_j si : (Eq. 4)

$$\lambda^{\alpha} \leq B = \frac{e_{ji}}{1 - e_{ij}}$$

avec e_{ij} et e_{ji} les probabilités que l'on décide $X \in C_i$ (respectivement C_j) quand actuellement $X \in C_j$ (respectivement C_i) est vrai.

L'utilisation d'un seuil fixe λ^{α} pour engager une reconnaissance sur le paramètre suivant peut poser problème. En effet, des seuils trop forts imposent le passage de beaucoup de paramètres, et réduisent l'intérêt de la méthode. A contrario, des seuils trop faibles ne permettent aucune décision. C'est pourquoi Fu a proposé une version modifiée de la procédure qui consiste à prendre des valeurs différentes pour A et B à chaque étape. Cette modification des seuils est assurée soit par une fonction algébrique, soit par programmation dynamique établissant une valeur optimale de chaque seuil.

L'établissement du rapport séquentiel de probabilité de Wald (SPRT) se fait dans le domaine de la statistique paramétrique, c'est-à-dire avec la connaissance a priori des lois de probabilité, et limite son application au domaine paramétrique, ce qui n'est pas toujours possible. Cela induit deux problèmes. Si seul le modèle des lois de probabilités est connu, Fu propose une estimation de leurs paramètres par les facteurs d'expansion de Karhunen-Loève pendant le calcul du SPRT. Par contre, si on ne connaît pas les lois de probabilités, alors on se reporte sur une version non paramétrique de la méthode.

La structure proposée par Fu s'organise sur une base de reconnaissance réalisée paramètre par paramètre extrait des objets à reconnaître. Pour une exploitation plus riche de cette démarche, Lorette [Lorette 92] précise qu'une collaboration de sources de connaissances multiples (lexicale, syntaxique, sémantique, morphologique, ...) doit fournir un apport substantiel. Il existe des méthodes qui coopèrent dans le sens donné par Lorette, c'est-à-dire des méthodes dont les propriétés sont différentes. Pascual [Pascual 92] utilise ce principe d'approche séquentielle multi-connaissances pour le contrôle de la décomposition de documents en archivage informatisé. Avec l'enchaînement d'une reconnaissance syntaxique d'informations linguistiques et d'une étape de reconnaissance par règles des propriétés morpho-dispositionnelles, il détermine la structure d'un document.

- Approche floue

La procédure d'utilisation séquentielle des paramètres pour obtenir une décision "optimale" peut être abordée par la théorie des ensembles flous plutôt que par une approche probabiliste. La stratégie reste basée sur l'évaluation des pouvoirs discriminants des différentes étapes de discrimination réalisées, puis à engager une reconnaissance efficace sur les caractères extraits des objets. La principale différence réside dans la méthode d'évaluation de ce pouvoir discriminant.

Strackeljan [Strackeljan 95] montre une utilisation concrète d'une détermination des paramètres pertinents pour une reconnaissance, par la définition d'un degré d'efficacité calculé sur les degrés d'appartenance des formes à classer. L'algorithme détermine itérativement les

degrés d'efficacité, qu'il classe par ordre croissant et qu'il combine pour obtenir l'ordre des caractères servant à la reconnaissance. Le critère de présélection de paramètre est basé sur la correspondance entre le degré d'efficacité de la sous-sélection proposée et le degré d'efficacité de tous les paramètres.

1.2.3.3.3.2. Coopération parallèle, ordinaire et floue.

♦ Approche ordinaire

La coopération parallèle est une alternative logique à la combinaison séquentielle. Une coopération parallèle type est le vote majoritaire. Le principe général est de combiner les réponses de différents classificateurs et de combiner les résultats soit à vote démocratique, *i.e.* en donnant la même importance à toutes les réponses, soit à vote dirigé, i.e. en pondérant les réponses des votants (Figure 6). Le point crucial réside dans la définition optimale des coefficients de pondération conduisant à la décision finale.

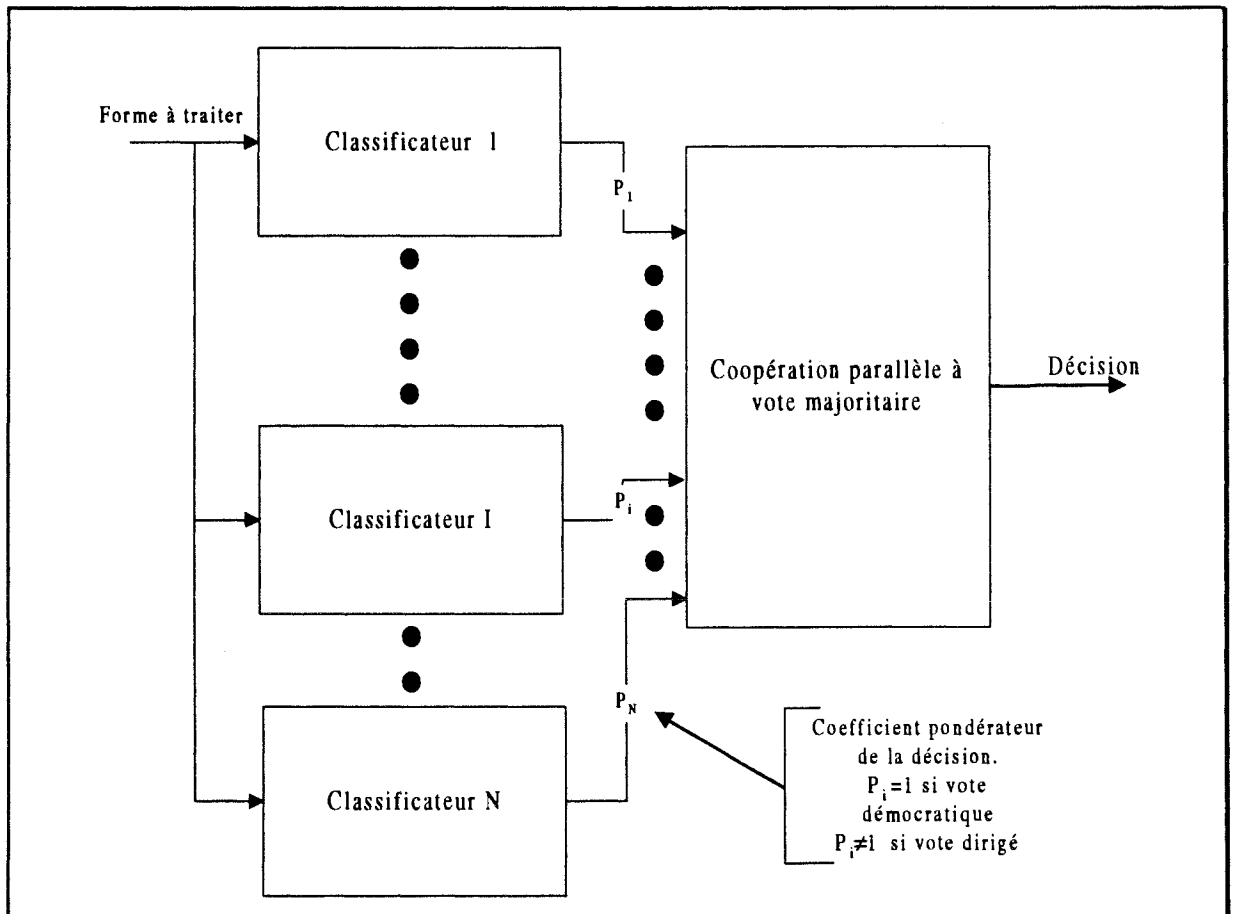


Figure 6 : Principe de la coopération parallèle [Xu 92]

Dans une assemblée de votants, chacun fait son choix, a priori de manière indépendante. La décision finale correspond à la proposition majoritaire. La décision a normalement plus de valeur (normalement la meilleure). Ce principe est intéressant lorsque les votants ou experts procèdent par approches différentes, ce qui amène de la complémentarité, de la richesse d'information, parfois de la synergie, et permet l'élimination d'erreurs individuelles. Cependant, cette méthode ne change rien au fait que les données entrées soient

mauvaises. Si les experts se basent sur les mêmes primitives, alors il y a plus de possibilités de se tromper suite à la corrélation initiale des informations [Anigbogu 92].

♦ Approche floue

La coopération floue de classificateur est extrêmement courante car l'une des applications de la théorie des ensembles flous est l'agrégation multi-critère, ce qui correspond à une combinaison de réponses. Cette agrégation proposée est vue comme un vote dirigé dont les coefficients de pondération sont liés à la pertinence des classificateurs. C'est la détermination de ces coefficients qui amène la coopération parallèle floue, basée sur la théorie des ensembles flous.

Grabisch [Grabisch 95] propose l'utilisation des intégrales floues pour combiner les différentes réponses des classificateurs en déterminant une agrégation floue par les opérateurs flous de conjonction sur la base de la minimisation d'un critère global, calculé à partir des critères d'erreurs des sous classificateurs.

La coopération parallèle à vote dirigé est donc propre à l'agrégation floue de réponses grâce à la notion de conjonction qui finalise les inférences floues. Kuncheva [Kuncheva 95] utilise l'opérateur maximum ou un opérateur moins restrictif pour calculer un degré d'appartenance issu de l'agrégation de degrés d'appartenance provenant de systèmes flous de reconnaissance.

1.2.3.3.3.3. Coopération hybride ordinaire et floue.

Le principe d'une coopération hybride est de combiner les approches parallèle et séquentielle. Pour pouvoir définir la structure d'une telle coopération, il est nécessaire de maîtriser les différents classificateurs utilisés pour profiter des performances spécifiques de chacun d'eux. Cohen [Cohen 91] a proposé une coopération hybride en faisant coopérer quatre classificateurs en parallèle dont les décisions sont exploitées par un arbre de décision. Les quatre classificateurs reposent sur des principes différents de reconnaissance globale et structurelle (Figure 7).

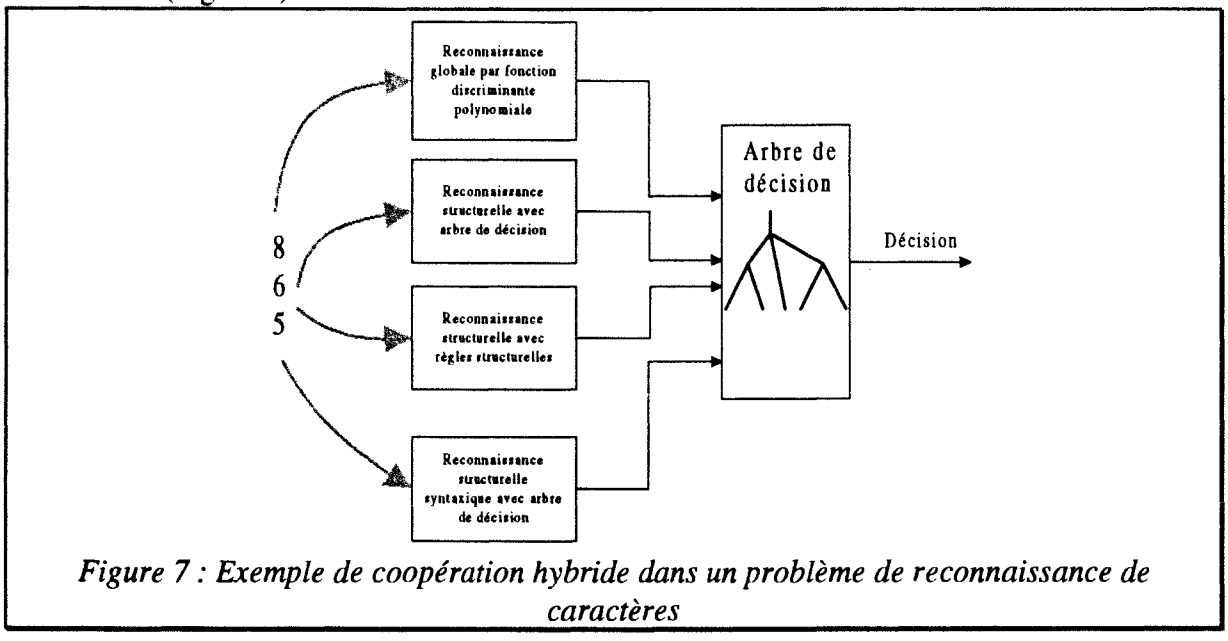


Figure 7 : Exemple de coopération hybride dans un problème de reconnaissance de caractères

1.2.3.3.4. Conclusion

La plupart des méthodes que nous avons vues présentent une alternative floue. L'apport de la théorie des ensembles flous dans ces algorithmes se ressent principalement dans les performances finales, mais concerne des applications particulières. En effet, les algorithmes ordinaires sont toujours performants lorsqu'ils sont utilisés dans les conditions adéquates. L'une des premières conditions est d'avoir un nombre suffisant de données. De ce fait, les algorithmes flous sont à utiliser dans les autres conditions. Ils permettent d'extraire une plus grande connaissance des données, notamment lorsque celles ci sont en nombre insuffisant. La théorie des ensembles flous permet également de minimiser l'influence de l'imprécision des données et prend en compte l'ambiguïté des classes.

Pedrycz souligne que les approches ordinaires et floues doivent coopérer plutôt que combattre. Un certain nombre d'articles [Sudkamp 92], [Stanislaw 93], [Dubois 92], [Guan 93] discutent de ce rapport flou et probabilité. Les probabilités sont un cas restreint des possibilités qui, elles-mêmes, sont un cas restreint de la théorie de l'évidence. "*La théorie des possibilités et la théorie des probabilités sont des descendantes de la théorie de l'évidence*" [Bouchon-Meunier 94]. Les deux approches sont donc fondées sur les mêmes bases théoriques mais leurs développements permettent des raisonnements différents. Les probabilités s'attacheront à un problème fréquentiel alors que la théorie des ensembles flous permettra de s'affranchir d'une connaissance complète pour développer un raisonnement plus souple.

Cette présentation de la chaîne de reconnaissance de formes et de ses grands principes nous a permis de placer la problématique générale dans laquelle s'inscrivent nos travaux. Nous présentons dans les paragraphes suivants la spécificité de cette chaîne lorsque l'on aborde le problème de l'écrit.

1.3. L'analyse de l'écrit

1.3.1. Introduction

Les travaux sur la reconnaissance de l'écrit sont un exemple typique des applications de reconnaissance de formes. Ils ont longtemps activé la recherche, puis se sont quelque peu épuisés. Le lecteur trouvera dans [Harmon 72], [Mantas 86], [Tappert 90] et [Suen 93], des synthèses des travaux dans ce domaine. Ils connaissent un regain d'intérêt, suite aux avancées technologiques de ces dernières années, portant notamment sur les capteurs (tablettes à digitaliser, scanners,...) ainsi que sur l'analyse et la conception des interfaces utilisateurs [Heutte 94]. Comme nous l'avons précisé dans le paragraphe précédent, l'étape d'acquisition, en entrée de chaîne, est un élément prépondérant dans la résolution générale du problème. Aussi, la mise en place de capteurs adaptés et performants permet la réalisation de nombreuses applications qui remettent ce thème au goût du jour. Néanmoins, tous les problèmes ne peuvent être résolus avec le progrès technologique et les travaux dans ce thème ne sont pas sur le point de disparaître. Ils sont au contraire stimulés par l'attrait économique de l'exploitation automatique de « l'écrit ».

Dans le contexte général de la reconnaissance de formes, l'analyse de l'écrit n'est qu'un problème particulier. La démarche entreprise pour sa résolution suit le schéma classique de la reconnaissance de formes et seule sa spécificité engage des opérations particulières. Ces opérations seront dépendantes des deux approches générales de la reconnaissance de l'écrit, c'est-à-dire l'approche en-ligne et l'approche hors-ligne. Ces approches conditionnent, sous le même scénario, des opérations différentes.

Le traitement de l'écrit est un problème simple pour l'homme et quasi universel dans les catégories de problèmes que l'on peut rencontrer, mais il est pourtant difficile à résoudre de manière automatique. Il n'existe pas de systèmes véritablement fiables lorsque la sécurité nécessite une reconnaissance parfaite. Les problèmes de reconnaissance automatique de l'écrit sont complexes ; un certain nombre de critères définissent le degré de cette complexité. La suite de ce paragraphe présente les différentes sources de complexité de l'écrit.

1.3.2. Imprimé ou manuscrit

Il est tout d'abord nécessaire de faire la distinction entre imprimé et manuscrit. L'imprimé ne pose plus réellement de problèmes, on peut s'en rendre compte avec les nombreuses applications bureautiques de reconnaissance optique de caractères. Néanmoins, la reconnaissance sur des supports particuliers (métal, bois, ...) présente encore des difficultés, qui ne se situent pas exclusivement sur la reconnaissance, mais plutôt sur les prétraitements. On peut définir comme degré de difficulté la notion de caractères multi-fontes qui rend le problème plus compliqué, mais qui peut être résolu par la séparation des caractères.

En revanche, le cas de l'écriture manuscrite est loin d'être complètement résolu. Les problèmes sont beaucoup plus complexes, particulièrement, lorsque l'ensemble des lettres à reconnaître est constitué de lettres liées. L'approche par caractères isolés n'est plus valide et il faut envisager une reconnaissance globale de mots ou de fraction de mots dont les modèles sont plus nombreux et différents. C'est dans ce cadre particulier que nos travaux se placent.

1.3.3. Les sources de complexité de l'écriture manuscrite

Lorette propose dans [Lorette 92] une typologie des problèmes de la reconnaissance de textes manuscrits en définissant des degrés de la complexité, correspondant à des situations de reconnaissance plus ou moins contraintes. Deux natures de contraintes sont à prendre en compte, les contraintes externes qui interviennent sur la présentation du texte [Tappert 90] et des contraintes internes liées à la particularité de l'écriture des scripteurs.

1.3.3.1. Présentation du texte

Le cas le moins contraignant, et donc le plus difficile, concerne le cas général sans contraintes de position du texte à reconnaître. En effet, dans ce cas, toutes les étapes évoquées dans les paragraphes précédents sont présentes, et la chaîne est forcément plus lourde [Viard-Gaudin 92]. L'introduction d'une contrainte externe de disposition, par l'utilisation d'une ligne guide (le guidé) permet de décharger la phase de description pour la localisation de la ligne d'écriture. Le zoné, par la définition d'un cadre, réduit encore la difficulté de localisation de

l'écriture, et enfin le précasé où l'écriture est présegmentée par l'utilisateur, est le problème le plus simple. La Figure 8 présente les différents types de présentation de texte.

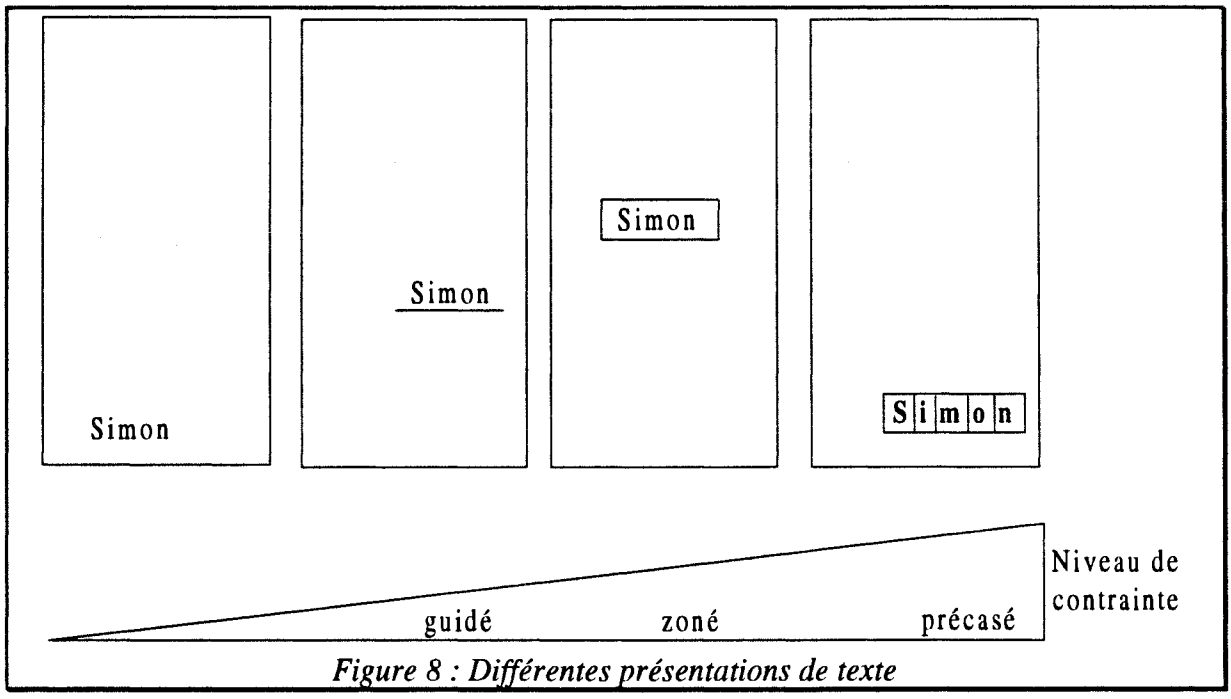


Figure 8 : Différentes présentations de texte

1.3.3.2. L'écriture

Le style d'écriture dont la difficulté s'évalue en termes de contraintes internes peut se décomposer en trois catégories :

- les lettres séparées (lettres bâtons) est le cas le plus simple,
- les groupes de lettres liées, présentés sous la dénomination de graphèmes est un cas plus complexe [Heutte 94],
- les lettres entièrement liées, ou écriture cursive, représentent la plus grosse difficulté.

1.3.3.3. Nombre de scripteurs

La complexité générale du problème de reconnaissance de l'écrit est également liée au nombre de scripteurs, elle croît avec ce nombre. Une application mono-scripteur est la situation la plus simple où les applications fonctionnent bien. Une sous-classe de ce problème est le multi-mono-scripteurs où l'approche revient à spécialiser la reconnaissance par scripteur. Les applications multi-scripteurs, où une reconnaissance préalable détermine le scripteur puis s'adapte à son écriture, sont de complexité plus grande. Enfin, les applications omni-scripteur, sans conteste les moins contraintes, sont plus difficiles, le système doit s'adapter à n'importe quel scripteur, inconnu a priori. Cette dernière approche générale n'a pas encore trouvé de solution et la particularisation de l'approche de reconnaissance va à l'encontre de cette généralisation.

1.3.3.4. La taille du vocabulaire

Un vocabulaire limité (inférieur à 100 mots) définit une application de complexité raisonnable, a contrario un vocabulaire très étendu devient délicat à traiter, en raison du volume des références.

1.3.3.5. La morphologie de l'écriture

Finalement, les contraintes morphologiques définissent la généralité du système de reconnaissance. Celle-ci augmente en fonction de la capacité du système à prendre en compte différents types de modèles d'écriture, pour imposer moins de contraintes au scripteur. Trois niveaux sont possibles :

- nombre limité d'entités isolées (lettre, chiffre, symbole),
- entités plus globales (mots ou fraction de mots),
- reconnaissance globale du texte (types d'écriture différents, annotations, retouches).

1.3.4. Le système de reconnaissance de l'écriture manuscrite

La performance d'un système de reconnaissance de l'écriture manuscrite est dépendante des contraintes imposées pour la reconnaissance. Le niveau de généralité et de complexité du système peut être représenté dans un espace Nombre de scripteurs/Invariance aux variations/Taille du vocabulaire, comme l'a proposé Lorette dans [Lorette 92]. La Figure 9 présente cet espace.

Ainsi, plus le système est général, plus il est complexe. L'utilisation d'un modèle fixe implique beaucoup de contraintes, mais simplifie le problème, tout en diminuant la généralité. Si on veut peu contraindre le système, il faut le concevoir "souple et intelligent" pour accepter les variations.

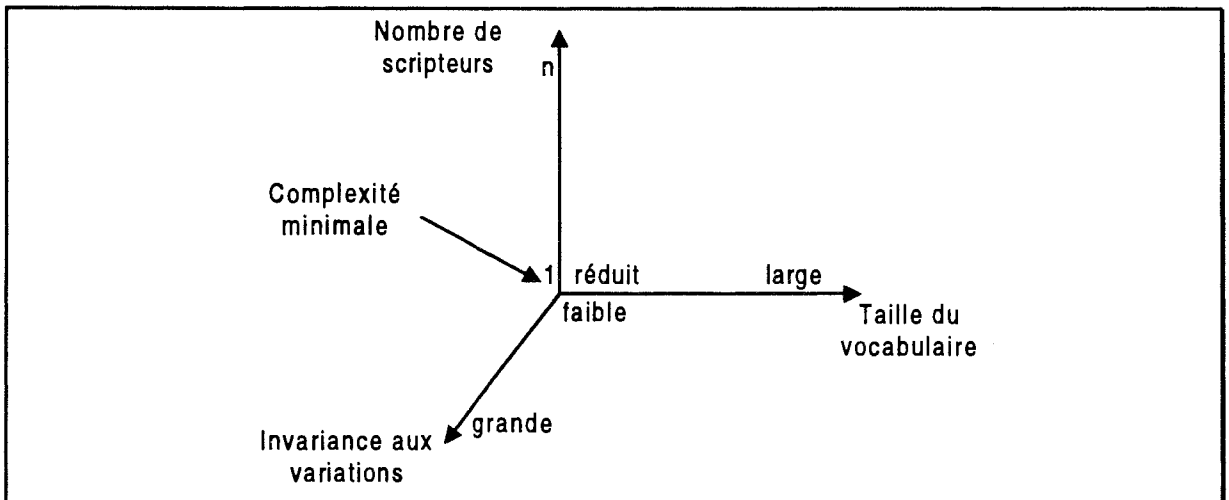


Figure 9 : Espace situant la complexité d'un système de reconnaissance de l'écrit

En outre, si le système est doté d'un apprentissage permanent pour s'adapter aux évolutions successives d'écritures, il devrait tendre vers une universalité que chacun cherche à mettre en oeuvre.

Les méthodes de reconnaissance de l'écrit utilisent implicitement ou explicitement un modèle de l'écrit [Lorette 92], et le scripteur impose un modèle fonctionnant avec des règles propres. Ces modèles portent sur l'une ou l'autre des deux composantes fondamentales de l'écrit :

- l'aspect ou la forme, le signifiant qui caractérise l'auteur [Plamondon 89], [Plamondon 90], [Sabourin 90],
- le contenu sémantique, signifié ou libellé qui caractérise le sens de l'écrit.

Le choix d'une composante particulière (signifiant ou signifié) conduit à l'utilisation de systèmes de reconnaissance différents. Si le modèle porte sur la forme, la connaissance du signifié apporte une information contextuelle supplémentaire qui lève les ambiguïtés de reconnaissance et conduit à de meilleures performances. Notre application porte sur ce dernier cas de figure.

Le choix des éléments de la chaîne est orienté selon les caractéristiques accessibles et désirées de l'écrit. A chaque type de problème, un ou plusieurs scénarios de résolution peuvent être proposés. Le choix d'une stratégie de résolution du problème de l'écrit demande au préalable de s'orienter vers une approche en-ligne, c'est-à-dire dans le cas où l'utilisateur est en train d'écrire, ou hors-ligne, c'est-à-dire après écriture. Ce choix oriente les méthodes d'acquisition et les opérations à mettre en oeuvre. En effet, une approche en-ligne, plus performante, est souvent traitée par une approche signal guidée par une dimension temporelle. L'approche hors-ligne est traitée par une approche image sans référence temporelle.

1.3.4.1. La reconnaissance en-ligne

Cette approche a beaucoup été étudiée et a donné naissance à de nombreuses applications (e.g. NEWTON™ de chez Macintosh™). L'utilisation d'une reconnaissance en-ligne permet le traitement immédiat de la forme et de la dynamique de tracé des symboles, grâce à un système d'acquisition qui capte ces deux informations de l'écriture. Notons que le traitement en temps différé de la reconnaissance est toujours considéré comme en-ligne, lorsque la dynamique de la signature est utilisée.

L'aspect dynamique de l'écriture est traduit en un ensemble d'informations comme le nombre de traits, leurs ordres, la direction et la vitesse pour chaque trait. Le trait est la portion de tracé délimité par un posé et un levé de stylo. Cet ensemble d'informations complète l'information globale définie par la forme des symboles (statique), et amène de nombreux avantages à la reconnaissance.

Les performances obtenues par une analyse en-ligne sont très bonnes, car cette analyse en temps réel permet une interactivité avec l'utilisateur qui corrige les erreurs produites par le système. Il y a donc un phénomène d'adaptation mutuelle entre l'utilisateur et la machine. Pour

la reconnaissance en-ligne réalisée en temps différé, on perd l'interactivité mais les performances restent très bonnes [Tappert 90].

Ce phénomène d'adaptation est également le talon d'Achille de ces outils puisque l'utilisateur doit s'adapter, l'équipement n'est plus naturel, ni confortable comme l'est le traditionnel papier crayon.

1.3.4.2. La reconnaissance hors-ligne

La reconnaissance hors-ligne est différente de l'approche précédente, principalement sur le plan de l'acquisition. Elle ne peut fournir les mêmes informations, puisque la notion de dynamique est perdue. Les systèmes d'acquisition pour le traitement hors-ligne sont les scanners ou caméras linéaire ou matricielle, couleurs ou niveaux de gris. L'information est présentée sous la forme d'images. Actuellement, ces systèmes d'acquisition sont assez rapides, ils ont également une définition spatiale grandissante puisque que l'on atteint des définitions courantes de l'ordre de 600 à 1200 points par pouce en millions de couleurs.

Cependant, ces systèmes entraînent classiquement des prétraitements coûteux et imparfaits (extraction de contours, amincissement ou squelettisation) pour restituer l'information de tracé, sans épaisseur propre à la statique de l'écrit. Ces traitements permettent également de régénérer une information de séquence du tracé, à partir des informations obtenues hors-ligne (traits) et ramènent la résolution vers une utilisation de la stratégie de reconnaissance en-ligne [Sabourin 90]. Cependant, cette conversion n'apporte pas une compensation suffisante pour combler la différence de performances avec l'acquisition en-ligne des informations.

La reconnaissance hors-ligne est clairement un problème plus difficile, dont les performances sont bien inférieures à celles de la reconnaissance en-ligne, puisque seule la notion de forme (statique) peut être utilisée [Tappert 90]. C'est cette dernière approche qui nous préoccupe.

1.3.5. Structure d'une application de reconnaissance hors-ligne de l'écrit

Nous développons, uniquement dans les paragraphes suivants, les spécificités de la reconnaissance de l'écrit pour une représentation vectorielle des formes. Les méthodes de raisonnement sont les mêmes que celles présentées dans le paragraphe 1.2.3.3.

1.3.5.1. Acquisition

Les développements actuels des moyens s'orientent vers des acquisitions multi-niveaux de gris, ou couleurs, avec une grande résolution afin d'obtenir une bonne source initiale et limiter les pertes possibles. Si les étapes suivantes sont inhibitrices de l'information pertinente, l'apport en quantité d'information sera absorbé et n'aura plus d'intérêt.

Comme nous l'avons précisé auparavant, le problème d'acquisition est plus complexe que la seule définition du matériel. Outre le choix technologique qui doit être adapté à

l'information que l'on désire véhiculer, il est nécessaire de définir et de résoudre le problème posé par la scène. L'acquisition est un ensemble constitué par le système matériel d'acquisition, l'éclairage et la scène elle-même. Bien souvent, un travail au niveau de l'éclairage, de la position des différents éléments et la simplification de la scène conduisent à une chaîne de traitement simplifiée. Cette démarche qui consiste à déporter une partie de la pertinence de la chaîne dans l'acquisition et son environnement n'est pas toujours réalisée ou réalisable.

Dans de nombreuses applications, soit par contraintes technologiques, soit par choix, l'information numérique véhiculée reste binaire. Lorsque la binarisation est un choix, elle implique une injection de connaissance de la part du concepteur, puisqu'elle sert de filtre dissociant l'information véhiculée de celle qui ne le sera pas. Cette séparation de l'information est irrémédiable et doit être réalisée correctement. A titre d'exemple, les systèmes de reconnaissance basés sur une extraction de composantes connexes sont perturbés lorsqu'une information portée par une composante se retrouve scindée trop durement en information pertinente et en information qui ne l'est pas.

Cette procédure simplifie les traitements automatiques, mais on devra souvent compenser ultérieurement ses effets néfastes.

Il existe cependant peu de systèmes qui remettent réellement en cause la binarisation, et peu d'études portent sur l'apport d'une analyse de l'écrit en niveaux de gris [Belaïd 92]. Notons cependant, que cette remarque ne concerne que le cas de l'écrit où l'on s'intéresse principalement à la sémantique. Dans le cas de la vérification de l'identité, nous verrons dans la suite de ce mémoire, que le niveau de gris peut être une information discriminante.

1.3.5.2. Prétraitements

Comme nous l'avons précisé au paragraphe 1.2.2.2, l'opération de prétraitement est engagée dans le but de résoudre un problème particulier provoqué par l'ensemble électronique d'acquisition, mais également par des imperfections liées au geste de l'écriture. Ce sont, pour l'approche hors-ligne, des prétraitements de type traitement d'images, puisque l'écriture est abordée comme un problème d'imagerie. Notons que la reconnaissance de l'écriture en-ligne est un problème qui peut être vu comme un problème direct de traitement du signal mono-dimensionnel.

Le bruit introduit par la discrétisation ou la transmission de données (images satellite), est réduit par des opérations de filtrage, de réduction de points non pertinents, d'élimination de points doubles, ou de points isolés.

Les mouvements de main erratiques (écriture cursive) ou une tenue particulière du stylo provoquent la diffusion du point d'attaque et l'étalement du tracé qui peuvent être résolus par des techniques de dehooking, de deskewing, de la squelettisation.

On peut trouver également des traitements liés à la scène, tels que l'élimination du fond lorsque l'écriture est sur fond texturé. C'est un problème typique d'une approche hors-ligne. La texture est courante surtout pour la reconnaissance de caractères sur des pièces métalliques [Girod 93], [Lachaux 94], ou des signatures sur des chèques. Le redressement d'écriture

penchée, le redressement de la ligne de base, la normalisation de la taille, qui sont également liés à l'écriture cursive, sont courants.

La multiplicité des problèmes entraîne celle des opérations, mais le prétraitement le plus courant reste la binarisation, soit par seuil global, soit par seuil local adaptatif. La réduction d'information qu'elle engendre est appropriée lorsque la reconnaissance se focalise sur la forme de l'objet. Elle permet une certaine sélection de l'information, mais sa rudesse est parfois mal adaptée. Tout le problème est d'en définir la validité par rapport à l'objectif de reconnaissance. La binarisation a souvent été une étape a priori, parfois même intégrée directement dans l'acquisition. La plupart des travaux partent de cet état de fait pour mettre en place une stratégie de reconnaissance.

La définition d'un seuil global de binarisation s'appuie sur l'hypothèse d'un histogramme d'image bimodal (fond/objet). Il est délicat à appliquer lorsque la densité de pixels de l'objet est faible, ou lorsque le fond est tramé ou variable, se traduisant par un histogramme non bimodal ou dont la bimodalité n'est pas explicite. L'usage d'un seuil local assouplit la vérification de l'hypothèse de bimodalité en la reportant à une bimodalité locale.

Cependant, la validité générale de la binarisation reste assez incertaine surtout par rapport à la variabilité des données (les niveaux de gris). L'élimination logique du problème de la binarisation, passe par l'extraction des primitives de l'image en niveaux de gris, ou en couleurs. Cette approche rend l'information plus riche, mais plus complexe. Parfois, l'avantage rend la tâche plus ardue, notamment lorsque cela intervient sur la forme du tracé dont les caractéristiques se trouvent modifiées.

1.3.5.3. Description

Dans la reconnaissance de l'écriture manuscrite, la description se trouve généralement décomposée explicitement en deux phases de segmentation et d'extraction de caractéristiques.

1.3.5.3.1. Segmentation

L'objectif de cette étape est d'obtenir des entités élémentaires qui seront analysées par la méthode de raisonnement mise en oeuvre. On différencie deux types d'entités, l'entité physique qui est le pixel, et l'entité logique constituée d'un groupement de pixels (trait, graphème).

Une approche très générale basée sur les données brutes issues du prétraitement d'une image de document, consiste en une décomposition en entités logiques élémentaires (mots, lettres), liée à la structure de l'écrit. Ces entités servent alors de base à la reconnaissance sous une forme appropriée [Tappert 90].

Une approche plus simple consiste à obtenir directement ces entités sous la forme adéquate, c'est ce que l'on retrouve lorsque des caractères sont séparés par le scripteur (Boxed Discrete Characters), si la reconnaissance se pose en termes de caractères, ou sous la forme de l'entité de base.

1.3.5.3.2. Extraction de caractéristiques

Lorsque l'on possède l'entité fournie par l'étape de segmentation, si elle a été nécessaire, ou par l'étape de prétraitement si les entités logiques étaient directement accessibles, alors l'extraction de caractéristiques est envisagée. Le terme envisagé précise en effet que cette étape peut ne pas exister dans la mesure où la reconnaissance travaille directement sur l'entité, c'est le cas des méthodes de reconnaissance directe qui sont courantes en reconnaissance de caractères, mais que nous ne développons pas ici.

Si une extraction de caractéristiques est effectuée, elle oriente obligatoirement la reconnaissance vers la reconnaissance globale. En effet, l'extraction de primitives donne accès à une représentation vectorielle des formes.

Il existe de nombreuses techniques d'extraction de primitives, elles ont été largement utilisées dans le domaine de la reconnaissance de formes, et dans bien d'autres domaines également. L'approche description de signal dans des espaces transformés, comme la transformée de Fourrier, Hough ou Hadamard, est une forme classique de description de signaux. D'autres techniques telles que les descripteurs statistiques, les moments ou des techniques de vision (morphologie mathématique), sont également très utilisées en reconnaissance de caractères.

Le choix du descripteur est délicat à faire, car il est prépondérant pour l'efficacité de la méthode de reconnaissance. Le concepteur doit l'orienter en fonction de son but et des informations accessibles et recherchées. De nombreux travaux sur l'extraction de primitives, soit locales [Ammar 90], [Brockelhurst 85], [Nagel 77], soit globales [Simon 93], [Nouboud 88], [Phelps 82], [Forre 86], [Chuang 77], [Ammar 86], ont été réalisés.

Cependant, il est extrêmement difficile de définir quelle est la meilleure description a priori. Encore une fois, la qualité de ce choix est liée à l'expérience du concepteur et à la spécificité du problème, une validation a posteriori entérine le choix. Des méthodes telles que l'analyse en composantes principales ou la coalescence [Dubuisson 90] permettent de valider le caractère discriminant d'un descripteur.

2. La vérification hors-ligne de signatures manuscrites

2.1. Le contexte

Le problème auquel s'attellent les travaux exposés dans ce mémoire, concerne la vérification automatique de signatures manuscrites. L'intérêt des travaux se trouve dans le besoin des sociétés à recourir à des systèmes automatiques de vérification de l'identité. En effet, le besoin d'identifier se ressent lors de transactions bancaires, immobilières, notariales. Les taux de communication actuels sont tels que l'identification nécessite une grande rapidité de vérification. La modification des outils de communication contribue à accentuer ce besoin. Les échanges au travers de réseaux, les transactions à distance, dépersonnalisent les communications et renforcent la nécessité d'une identification rapide et fiable.

Le choix de la signature manuscrite comme véhicule de l'identité vient naturellement, car c'est un moyen couramment utilisé et facile à réaliser. Cependant, d'autres supports d'identification sont utilisables. Parmi les représentations de l'identité, il est possible d'utiliser un numéro d'identification personnel, un mot de passe, une carte, une clef, ou des empreintes digitales, des vascularisations rétiniennes, ... Les premières propositions ne sont personnelles que dans l'attribution, elles ne représentent pas des attributs propres. De ce fait, elles sont facilement dérobables. Le choix d'un attribut personnel et physique rend l'usurpation d'identité plus difficile, et parmi les attributs personnels, la signature est le véhicule de l'identité le plus accessible.

D'autres critères contribuent au choix de la signature manuscrite. Elle est le résultat d'un geste réflexe inconscient et traduit des caractéristiques propres au scripteur [Mathyer 61], [Qi 95]. De plus, c'est le seul mot que la population peu scolarisée ou analphabète sait écrire sans effort. Sa réalisation est indépendante du niveau de connaissance, ce qui fait son universalité. Bien qu'elle présente une certaine fiabilité, elle n'apporte pas une certitude absolue quant à l'élimination de l'usurpation [Shaw 80].

La vérification automatique de signatures manuscrites dans un environnement bancaire s'attache plus particulièrement au traitement des chèques, et plus généralement à la vérification sur support papier. Le support papier reste le plus universel, et le plus communément utilisé.

L'émission de la signature peut être réalisée dans de multiples situations, orientant ainsi les recherches vers une approche hors-ligne. Néanmoins, une approche en-ligne en temps différé n'est pas écartée, bien qu'elle nécessite un équipement particulier qui peut être coûteux.

Cette situation place la vérification de signatures dans la catégorie des problèmes d'expertise de document par l'analyse hors-ligne de l'écrit [Harrisson 81]. Elle sera naturellement abordée comme un problème basé sur l'image, avec la panoplie des traitements associés.

2.2. *La vérification automatique de signatures manuscrites*

La vérification automatique de l'identité par l'analyse de l'image de signatures manuscrites consiste à déterminer si l'image de la signature présentée au système est celle du scripteur annoncé. Cette signature est déclarée vraie ou fausse. Ce problème est à différencier d'un contexte général de reconnaissance car le terme vérification suppose une hypothèse sur l'identité réelle du scripteur. Le système doit répondre à la validité de la proposition et non à la détermination du scripteur. C'est donc un problème de reconnaissance de formes restreint à deux classes, les vraies et les fausses signatures. Cependant, cette simplification n'est qu'apparente car il existe plusieurs types de faux. Dans les paragraphes suivants, nous présentons les caractéristiques des signatures authentiques et celles des différents types de fausses signatures.

2.2.1. Les caractéristiques propres aux signatures

2.2.1.1. Caractéristiques des authentiques

Les authentiques possèdent dans leur tracé différentes caractéristiques qui sont difficilement imitables [Sabourin 90]. L'harmonie de la signature apposée librement et sans contrainte est liée à un geste naturel. L'utilisation de cette caractéristique d'harmonie est possible par un expert, mais peu abordable par un système automatique sans quantification de son niveau. Elle permet à un expert graphologue de différencier plusieurs types de faux.

Les variations naturelles globales et locales des proportions hauteur/largeur, sans pourtant de corrélation hauteur/largeur, et sans notion de temps, constituent un critère de forme global propre à chaque scripteur qui fait partie de son essence [Evett 85]. Il faut cependant noter que cette proportion est variable sur une base quotidienne aussi bien que mensuelle.

Le rythme du geste qui traduit la dynamique est une notion physique qui, en association avec la forme du tracé, se présente également comme une caractéristique propre au scripteur. Le rythme est lié à l'harmonie dans l'exécution du geste, qui est un prolongement de l'état physique et psychique du scripteur. La stabilité du rythme sur une base de temps quotidienne ou mensuelle reste donc liée à cet état.

La signature authentique est un ensemble harmonieux dans l'espace (la forme) et dans le rythme (la dynamique), sur un laps de temps court. Cette harmonie s'exprime par les ombrages (pleins et déliés) où se mêlent effectivement les notions de statique et de dynamique. Selon Gupta [Gupta 79], c'est ce mélange qui représente le cas le plus difficile à imiter. De plus, l'authentique s'affranchit des contraintes spatiales homothétiques, ce qui rajoute de la difficulté à l'usurpation.

2.2.1.2. Informations influentes du tracé

La classe des authentiques est difficile à cerner et à qualifier. Elle présente des variations naturelles dans le tracé et dans le temps, ce qui la rend complexe et difficilement contrôlable par un système automatique. Un certain nombre de paramètres modifie l'information présente sur le tracé, en intervenant sur sa forme et entraînant le fait qu'un scripteur est incapable de reproduire exactement sa signature.

2.2.1.2.1. Le contexte

Les circonstances dans lesquelles la signature est apposée influencent son tracé. Le facteur téléologique est un facteur d'influence prépondérant de la signature. Il divise les signatures en catégories distinctes, elles-mêmes soumises à d'autres facteurs de modification. Ces catégories, au nombre de trois, sont définies en fonction de l'intensité du contexte :

- la signature formelle pour les documents importants tel qu'un acte notarié,
- la signature informelle pour un document de routine, une note de service,
- la signature fugitive concernant les tickets de paiement réalisés au bord d'une caisse de magasin.

L'intensité du contexte se définit donc par rapport à l'impact que peut avoir la signature, et intervient sur l'attention et la rigueur dans l'acte du scripteur.

Il existe également d'autres facteurs circonstanciels tels que :

- la maladie,
- les blessures,
- les contraintes temporelles,
- la drogue,
- la température,
- la contrainte spatiale,
- la régionalisation qui implique un type de calligraphie,
- l'apprentissage [Moon 77].

Dans les facteurs d'influence du tracé de la signature, on trouve également le niveau d'habileté du scripteur. Un scripteur habile signe avec un geste dynamique pour lequel la plume est en mouvement avant de toucher le papier. Le geste est un geste réflexe, avec une importante structure dynamique. A l'opposé, un scripteur novice a une plume stationnaire sur le papier avant mouvement et le tracé a une dynamique complètement différente. Il faut tenir compte de ce paramètre, car un scripteur reste rarement novice, sa situation va migrer vers une position de scripteur habile, dont la signature aura évolué et affirmera son identité.

Toujours en termes de paramètres liés au scripteur, on relève des facteurs d'influence individuels sur la forme de la signature tels que :

- l'âge,
- le sexe,
- la main directrice (gaucher/droitier) [Totty 83],
- le niveau scolaire.

Les trois derniers facteurs sont constants pour l'état de la personne, mais l'âge modifie la signature au cours du temps.

2.2.1.2.2. La technique

Pour le tracé de la signature, on trouve également des facteurs influents indépendants du scripteur. Il s'agit de l'outil et du substrat dont Masson [Masson 85] précise l'importance du choix. La nature de l'outil a une influence non négligeable sur le tracé :

- le crayon, dont l'utilisation n'est pas recommandée pour les utilisations légales, n'est pas traité dans ce contexte ;
- la plume renforce les pleins et déliés, ce qui implique des difficultés pour l'imitation. Elle renforce souvent le début de la signature par un écoulement d'encre, notamment lorsqu'il n'y a pas de point d'attaque rapide. C'est le cas d'un scripteur débutant, mais ce renforcement

doit disparaître avec l'habitude, et cette disparition sera perturbatrice pour un système automatique ;

- le stylo bille est un outil dur qui ne révèle quasiment pas d'ombrage et de dynamique de la signature. Il ne génère pas d'écoulement d'encre lors d'un arrêt du scripteur ;
- le plume feutre a une terminaison généralement carrée, qui tend à déformer le tracé. En revanche, le débit d'encre délivrée, lié à la vitesse d'écriture, présente une caractéristique intéressante pour retrouver une information dynamique dans un contexte hors-ligne. La largeur du trait, suite à des variations de vitesse et de pression, fournit également une information pertinente.

Le second facteur technique de modification du tracé concerne la nature du substrat. Le substrat revêt une grande importance, puisque sa qualité doit être en relation avec le type d'outil de tracé. A titre d'exemple, un écoulement d'encre important force le choix d'un substrat peu capillaire, ne diffusant pas trop l'encre, afin de limiter la formation de cils de diffusion autour du tracé.

2.2.1.3. Caractéristiques des différents types de faux

Comme les vraies signatures, les fausses présentent un certain nombre de caractéristiques particulières qu'un expert graphologue exploite pour la vérification de l'identité.

Le faux par déguisement est très ressemblant à l'authentique, mais avec une déformation des lettres, de la pente générale [Locard 59] et des changements de vitesse. Il est réalisé par le scripteur annoncé, qui modifie volontairement sa signature dans un but de fraude.

Le faux servile est une imitation pour laquelle la graphie est assez ressemblante, mais diverge par les proportions, les espacements, les alignements, l'inclinaison relative des lettres et des grammas (éléments de composition d'une lettre). L'inclinaison moyenne est souvent différente, il y a des variations de la ligne de base, un tracé lent et hésitant, avec reprises et retouches.

Le faux par imitation libre est réalisé de mémoire par un faussaire, après de nombreux essais jusqu'à satisfaction. Cette capacité du faussaire rend le faux très difficile à détecter, mais peu de personnes sont à même de réaliser de tels faux. Le faux présente néanmoins quelques discordances avec le vrai, notamment sur les proportions relatives des lettres, des espacements, des alignements, où il manque l'alternance des pleins et déliés.

Le faux par calque ou par transfert optique, manque de spontanéité et montre des imprécisions dans le tracé, une apparence de mouvement lent, de pression uniforme et élevée. Il y a perte d'ombrages naturels, et les retouches y sont visibles.

Le faux aléatoire est un cas lié au contexte des systèmes de vérification. La forme et le libellé sont différents de l'authentique.

Le faux simple présente une copie du libellé, mais la forme de la signature est fondamentalement différente. Le cas est donc proche de celui du faux aléatoire et permet en

l'absence de faux simples d'utiliser des faux aléatoires comme représentation restreinte des faux simples.

2.2.1.4. Les caractéristiques divergentes

Le problème de vérification de signatures manuscrites n'exprime que deux classes, définissant les signatures vraies et signatures fausses. Cependant, les classes renferment une complexité importante due, dans la classe des vraies à une impossibilité du scripteur à reproduire exactement sa signature, et dans la classe des faux aux différents types de faux subissant également de nombreuses variations.

Malgré ces difficultés, nous pouvons focaliser notre attention sur l'étude des deux catégories d'information susceptibles de traduire la différence entre signatures vraies et fausses. Les caractéristiques d'ordre statique sont traduites par la forme générale du graphique et celles d'ordre dynamique sont exprimées par la rythmique du geste.

Dans notre application, la vérification est posée comme un problème de reconnaissance hors-ligne. Ainsi, l'information dynamique n'est jamais accessible, seule une information dite pseudo-dynamique peut être utilisée. Cette information n'est qu'un succédané de la dynamique, que l'on extrait du niveau de gris. En effet, le niveau de gris traduit simultanément une information de vitesse et de pression sans notion d'ordre dans les segments du tracé. Le Tableau 1 résume, pour les caractéristiques statiques et dynamiques des signatures, les principales différences entre les vraies et les fausses signatures.

Caractéristiques dynamiques	Servile	Libre	Calque	Simple	Aléatoire
Vitesse du tracé	1	3	1	3	3
Hésitation, tremblement	2	4	1	4	4
Retouche	3	4	1	4	4
Perte d'ombrage	2	3	1	3	4
Pression uniforme	3	4	1	4	4
Caractéristiques statiques					
Forme des lettres	3	3	4	1	1
Proportions relatives des lettres	2	3	4	1	1
Espacements	2	3	4	1	1
Alignements	2	3	4	1	1
pente relative	2	4	4	1	1
Libellé	4	4	4	4	1

1 : très grande différence, 2 : grande différence, 3 : petite différence, 4 : différence négligeable

Tableau 1 : Résumé des divergences fausses/vraies signatures.

Ce tableau, extrait de [Sabourin 90], montre que la détermination des faux aléatoires et des faux simples s'exerce sur des informations statiques.

La discrimination des faux par calque ou par transfert optique se place exclusivement sur le niveau dynamique, seul à même de révéler la différence de signatures.

Pour la détection des faux libres, le Tableau 1 montre que les possibilités de discrimination sont très faibles, quelles que soient les informations utilisées. Dans l'approche hors-ligne, les divergences statiques, bien que faibles sont le plus à même de dégager la différence avec l'authentique. L'information dynamique ne présente que peu de possibilités de discrimination, essentiellement contenues dans la vitesse et l'ombrage.

Le faux servile présente une difficulté de discrimination moindre, puisque la différenciation se place moyennement sur la forme générale (statique) et assez fortement sur la vitesse du tracé et sur l'ombrage (dynamique).

2.2.1.5. Les faux rencontrés dans notre application

La vérification dans son contexte applicatif demande une procédure particulière. Lors de l'émission d'un chèque, l'abonné présente un Numéro d'Identification Personnel (NIP) et produit une signature manuscrite. La vérification automatique de l'identité de la personne associée au NIP consiste à évaluer la validité de la signature, en fonction de références fournies au système de vérification, pour porter un jugement de type vrai/faux.

Dans notre application, nous serons confrontés à des faux aléatoires, des faux simples, mais également à des faux par transfert et par imitation libre.

Le problème des faux aléatoires est lié au contexte de bases de données, c'est-à-dire à la vérification des signatures des autres scripteurs abonnés au système de vérification. Ce type de faux ne devrait normalement pas être rencontré, dans la mesure où on ne soumet pas une signature à la vérification par rapport à tous les scripteurs de la base. Cependant, pour nos travaux, une extension au problème "signatures vraies / signatures quelconques" peut être faite et permet d'évaluer le système dans un contexte général où l'usurpateur utiliserait une sémantique différente. Aussi, nous utiliserons les faux aléatoires en lieu et place de faux simples qui ne sont pas disponibles dans notre base de données.

Les faux par calque sont obtenus par des techniques de transfert. Le niveau sémantique y est présent, mais de nombreuses irrégularités sur l'aspect existent. Néanmoins, il faut tenir compte des nouvelles méthodes de transfert comme les photocopieuses qui vont rendre la tâche d'un système de vérification beaucoup plus difficile.

Les faux par imitation libre sont réalisés en mémorisation visuelle par un faussaire. Cela représente la plus grande difficulté que l'on peut relier à la performance du faussaire.

Quelle que soit la démarche utilisée, le système de vérification ne sera réalisable, utilisable et viable que s'il est plus sensible aux variations provenant des incriminés qu'à celles

des authentiques. Ceci traduit le dilemme classique en reconnaissance de formes, de la minimisation de la distance intra-classe et la maximisation de la distance inter-classes.

Pour traiter l'ensemble des problèmes de fausses signatures, tout en évitant la confusion avec des vraies, il est nécessaire d'exploiter simultanément le potentiel discriminatoire des informations statiques et pseudo-dynamiques. Pour cerner plus précisément le comportement que doit avoir le système de vérification, nous pouvons nous inspirer de la démarche d'un expert graphologue explicitée par Harrisson [Harrisson 81].

2.2.1.6. La démarche du graphologue

Dans un premier temps, il faut savoir que l'erreur est toujours possible. Seul le scripteur peut dire si la signature présentée est vraie et il ne peut l'affirmer qu'au moment où il signe. Après, sa position revient à celle d'un expert vis-à-vis de sa propre signature et l'erreur est toujours possible.

Aussi, il peut exister plusieurs points de similarité entre incriminés et authentiques, notamment dans les aspects les plus évidents. Une bonne imitation d'une signature est globalement ressemblante. Cette remarque nous indique que dans une approche globale de la Perception de la signature, une vue à une échelle grossière sera moins discriminante qu'une observation à une échelle fine. De nombreux articles sur la vérification de signatures manuscrites rapportent cet état de fait. Qi [Qi 95] précise en ce sens, que lors d'une observation à une échelle grossière, les signatures se ressemblent fortement et à une échelle très fine (le pixel), elles sont toutes différentes. Aussi, deux signatures authentiques ne se ressembleront pas dans le détail, car il n'y a pas de réplique exacte entre les authentiques. Dès lors, une trop grande ressemblance implique que l'incriminé est un faux réalisé par une technique de transfert performante et qu'il doit être rejeté.

Puisque les vraies signatures présentent une variation naturelle dans leur forme, il est nécessaire de l'intégrer en disposant d'un nombre suffisant d'échantillons de signatures, représentant les états possibles de la signature. Gayet [Gayet 61] affirme qu'en pratique un expert utilise une vingtaine de signatures pour obtenir un panel complet des habitudes du scripteur. Notons que l'expert a besoin de moins de références pour la comparaison de deux spécimens authentiques que pour deux spécimens venant de deux individus différents.

Selon [Hilton 71], dont l'étude a porté sur l'évaluation du nombre de références nécessaires en fonction du niveau de difficulté de la vérification, il faut en moyenne 10 à 12 références d'authentiques pour assurer une vérification satisfaisante par un expert.

Pour des signatures simplifiées, c'est-à-dire sans personnalisation profonde de la graphie, l'authentification se fait sur de petites différences répétées. La même remarque s'applique à des faux par calque, par imitation servile, ou par imitation libre de la signature authentique. Pour ces cas, il faut 40 à 50 signatures sélectionnées en termes de date, buts et circonstances de l'émission pour réussir une bonne discrimination. En dessous de 25 signatures, l'expert doit utiliser l'expérience pour décider, ce qui met en jeu une forme d'apprentissage des mécanismes de falsification.

Les caractéristiques principales utilisées par l'expert sont :

- la dynamique,
- les terminaisons,
- les tremblements,
- l'ombrage.

Ensuite, pour renforcer sa décision, l'expert travaille sur l'information statique avec :

- les alignements,
- les proportions relatives lettres espacements,
- les pentes relatives entre caractères.

Ces points sont de première importance pour la détermination des faux. Mais, l'étude des proportions relatives n'est pas généralisable, notamment sur les graphies personnalisées comme nous avons en Europe, et la généralisation que nous cherchons à réaliser bute sur ce point. Le système à mettre en place doit éviter l'utilisation de telles informations, d'autant que la recherche de ces primitives n'est pas triviale [Sabourin 90].

La décision finale d'expertise repose sur une étude de ces caractéristiques et sur l'expérience de l'expert puisque notre objectif est de remplacer globalement ou partiellement le jugement "subjectif" de l'expert. Le système de vérification doit réaliser un apprentissage pour acquérir une "expérience" dans les mêmes termes que l'expert. Cet apprentissage représente la tâche principale de l'étape de raisonnement et sûrement la plus grande difficulté qu'elle doit résoudre.

Cependant, l'instabilité temporelle des signatures rend complexe cette phase d'apprentissage, car les caractéristiques d'un scripteur se modifient au cours du temps. Kapoor et Sharma [Kapoor 85] confirment la variabilité des caractéristiques, et précisent que les références d'authentiques sont à prendre sur une même période et doivent être renouvelées régulièrement.

2.2.2. Le système de vérification de signatures

Dans le domaine de l'écrit, la signature possède moins d'informations que le texte. Les signatures de type européennes sous la forme d'un parafe, auxquelles nous voudrions que nos travaux puissent s'appliquer, ne portent pas nécessairement de sens. Aussi, le choix d'une vérification s'appuyant sur une base sémantique est à exclure puisque nous désirons nous affranchir du sens véhiculé par la signature. Une approche basée sur la forme de la signature est donc plus appropriée. Elle nous permet de nous focaliser sur l'aspect de la signature, et donc d'être insensible au texte.

La Figure 10 présente différentes situations de reconnaissance suivant l'attribut de la signature choisi et la connaissance a priori des classes [Sabourin 90]. Si l'attribut de la signature est d'ordre sémantique, son libellé est connu, alors on fera de la reconnaissance de libellé si l'on connaît les classes a priori, sinon de l'apprentissage de libellé. En revanche, si l'attribut est de nature morphologique (singulière), la connaissance des classes orientera vers la

vérification de signatures, l'analyse constituera l'identification de scripteur dans le cas contraire.

Le système de vérification peut être placé dans l'espace situant le niveau de complexité d'un système de reconnaissance de l'écrit proposé par Lorette dans la Figure 9 du paragraphe 1.3.4. Ainsi, le problème de vérification se réduit à la question, "la signature proposée est-elle vraie ou fausse ?", nous nous plaçons dans une approche multi-mono scripteur. Chaque scripteur a son propre système de vérification. La taille du vocabulaire n'a aucun sens, puisque nous voulons offrir la possibilité de traiter des parafes sans contexte sémantique, c'est-à-dire par une reconnaissance globale du mot. Sur le plan des contraintes extérieures, le choix se porte sur une contrainte assez forte par l'utilisation d'une zone de signatures. Cette contrainte est cohérente dans le contexte applicatif de la vérification sur des chèques ou documents papier.

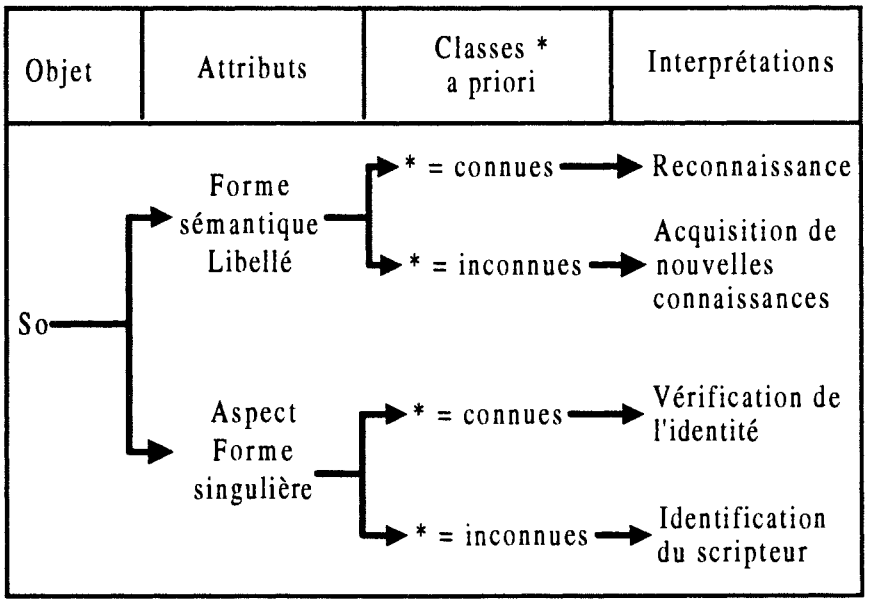


Figure 10 : Différents types d'analyse de la signature suivant l'attribut choisi et la connaissance des classes

Une approche globale nous semble la plus adaptée pour résoudre le problème de vérification de signatures manuscrites, cependant d'autres techniques de reconnaissance de formes ont été envisagées.

R. Sabourin [Sabourin 90] a proposé dans son mémoire de Philosophiae Doctor, une approche de type compréhension de scène, basée sur une analyse structurelle de l'image de la signature manuscrite. Son approche visait à traiter différents problèmes de faux, et notamment les faux libres réalisés par des faussaires "habiles", par l'utilisation d'informations statiques et pseudo-dynamiques. Sa démarche exploitait les différences de caractéristiques entre vraies et fausses signatures dans une approche locale des caractéristiques de la forme, de leur importance en termes de surface, de leur position et de leurs proportions, ainsi que des relations entre les primitives. L'exploitation de cette technique a montré une grande complexité notamment dans l'établissement des relations entre primitives caractérisant la structure. Dans

un contexte hors-ligne, le choix d'une approche structurelle est moins intéressant que dans un contexte en-ligne. En effet, dans une méthode en-ligne, nous possédons la notion d'ordre des primitives qui sert de base à la reconnaissance et obtenons de meilleures performances par rapport à la méthode hors-ligne. Dans l'approche hors-ligne, cette notion d'ordre des primitives est perdue. Aussi, le choix d'une méthode structurelle demande la recherche d'un pseudo-ordre dont la complexité et le résultat ont conduit R. Sabourin à réfuter cette approche pour le problème particulier de la vérification de signature.

Tappert [Tappert 90] et Forre [Forre 86] ont montré que des approches de reconnaissance directe n'étaient pas adaptées à la complexité de la signature.

Le rejet des méthodes structurelles et directes conforte notre choix d'une approche globale basée sur une représentation vectorielle des signatures.

Cette approche impose au concepteur d'élaborer une perception adaptée au problème de vérification de signatures manuscrites.

Nous présentons dans les paragraphes suivants l'approche proposée par R. Sabourin, qui servira de base à nos développements ultérieurs.

Nous présentons également le cadre méthodologique et applicatif de notre étude. Nous définirons notamment les bases qui nous permettent d'évaluer nos apports, en présentant la base de données, ainsi que les différentes méthodes d'évaluation utilisées par R. Sabourin.

2.3. L'approche proposée par R. Sabourin

2.3.1. Introduction

R. Sabourin [Sabourin 95a] a proposé une solution au problème de la vérification automatique de signatures répondant à l'énoncé des différentes caractéristiques de la vérification. Ce système a donné des résultats satisfaisants et nous a semblé pertinent par rapport à la problématique posée, notamment en perception. Nous avons cherché à améliorer les performances du système complet, sur les aspects Perception et Raisonnement, en collaboration avec R. Sabourin.

Dans le domaine de la vérification de signatures manuscrites, nous avons vu que différentes catégories de faux, dont la complexité est plus ou moins grande, existaient. Un problème encore non résolu concerne l'élimination des faux aléatoires, c'est-à-dire les signatures des autres scripteurs de la base. On peut également remarquer que la résolution du problème des faux aléatoires nous rapproche de celle des faux simples, notamment si nous nous attachons aux caractéristiques permettant leur discrimination. Ces caractéristiques sont résumées dans le Tableau 1. Cette remarque est d'autant plus vraie que l'approche à adopter doit être insensible au texte puisque le libellé de la signature est une caractéristique négligeable pour la discrimination. R. Sabourin s'est attaché à l'élimination des faux aléatoires et par conséquent des faux simples.

Ses travaux se situent dans une approche globale de la reconnaissance de formes. Ils sont basés sur la coopération parallèle de plusieurs blocs de discrimination associés à des facteurs de formes permettant une perception multi-échelle de l'image de la signature. Ce système forme ce que R. Sabourin appelle un classificateur intégré.

Avant de présenter en détail les travaux de R. Sabourin, nous définissons la base de données sur laquelle s'appuieront nos travaux.

2.3.2. La base de données

Dans le cadre d'une approche hors-ligne de la vérification de signatures, R. Sabourin a établi une base de données sous forme d'images. Cette base de données est le point commun aux différents travaux engagés sur la problématique. Pour une bonne compréhension des explications ultérieures, nous précisons les informations relatives à la constitution de cette base d'images de signatures.

Vingt scripteurs ont participé à l'expérience en fournissant chacun 40 spécimens de signatures selon un protocole d'apposition fixe et défini [Sabourin 92]. Ce protocole d'apposition précise l'outil d'écriture et le substrat, de façon à limiter leur influence sur les signatures, mais également pour fournir au mieux les informations pertinentes, par rapport aux différentes catégories de faux que l'on peut rencontrer. L'outil est un stylo de type *pilot fine liner* à encre noire, choisi pour son aptitude à exprimer les variations de vitesse, tout en ne déformant pas la géométrie du tracé. Le substrat est un papier blanc réduisant au minimum une possibilité de texture et sa porosité est choisie en adéquation avec l'outil d'écriture.

En termes de contraintes extérieures, des contraintes spatiales sont imposées en fixant un cadre de 3x12cm (le zoné), afin de réduire le problème de localisation segmentation et d'éviter une dérive spatiale des proportions des signatures. Toutefois, ce cadre est peu influant sur les proportions hauteur/largeur et sur l'orientation naturelle du tracé.

L'acquisition est réalisée par une caméra vidicon et une carte d'acquisition fournissant des images d'une dimension de 128x512 pixels, quantifiées en 256 niveaux de gris. L'évolution actuelle des matériels d'acquisition remet en cause cette procédure de numérisation des images dont la qualité pourrait être améliorée par l'utilisation d'un scanner. Cependant, les conditions d'acquisition ne changent en rien la démarche proposée par R. Sabourin. Leur amélioration ne peut qu'entraîner de meilleures performances finales ainsi qu'une certaine robustesse du système de vérification.

2.3.3. La perception

2.3.3.1. Les prétraitements

L'image de la signature est soumise à un prétraitement de type image dans un double objectif, éliminer les perturbations telles que le bruit ou les variations de luminance du fond de la scène et restreindre les informations à la géométrie de la signature. A cette fin, un opérateur de binarisation est appliqué. La binarisation fournit une image où seule la position des pixels de

la signature devient exploitable. Les informations d'ordre pseudo-dynamiques, traduites par les niveaux de gris, ont disparu. Les informations restantes conviennent alors parfaitement à l'élimination des faux aléatoires et faux simples.

Le choix de l'opérateur de binarisation doit être fait dans le cadre de son application aux images de signatures. L'outil d'écriture et le substrat utilisés favorisent une dilatation des densités de niveaux de gris. En outre, le tracé et le substrat ne peuvent être d'une homogénéité parfaite. Le tracé d'une signature ne peut représenter qu'une faible partie de l'histogramme de l'image. L'allure de l'histogramme est logiquement bimodale puisque le tracé se place dans la gamme des niveaux gris de faibles luminances et le fond dans celui des fortes luminances. Le tracé et le substrat sont caractérisés par des densités de niveaux de gris relativement larges, entraînant un recouvrement entre le mode correspondant au fond de la scène et celui de l'objet signature comme le montre la Figure 11.

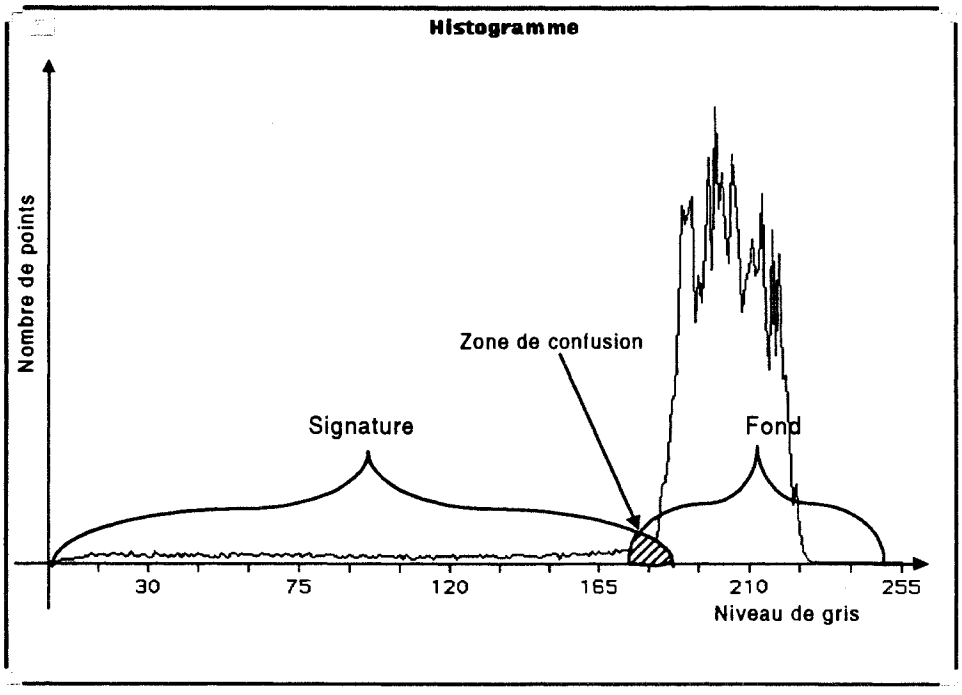


Figure 11 : Histogramme d'une image de signature

Le choix de l'opérateur de binarisation doit tenir compte de la bimodalité et du recouvrement des modes. R. Sabourin [Sabourin 90] a étudié différents opérateurs de binarisation et a conclu que l'opérateur d'Otsu [Otsu 79] était le plus adapté aux images de signatures.

Parallèlement à l'étape de binarisation, l'image de la signature est placée au centre de la zone d'écriture par translation, en faisant coïncider le centre d'inertie de la signature avec le point central du cadre (Figure 12). Cette translation ne modifie pas les contraintes spatiales définies par la technique de « zoning », ni les caractéristiques de la signature. En outre, elle réduit de façon homogène les variations locales naturelles du tracé par rapport au centre de gravité.

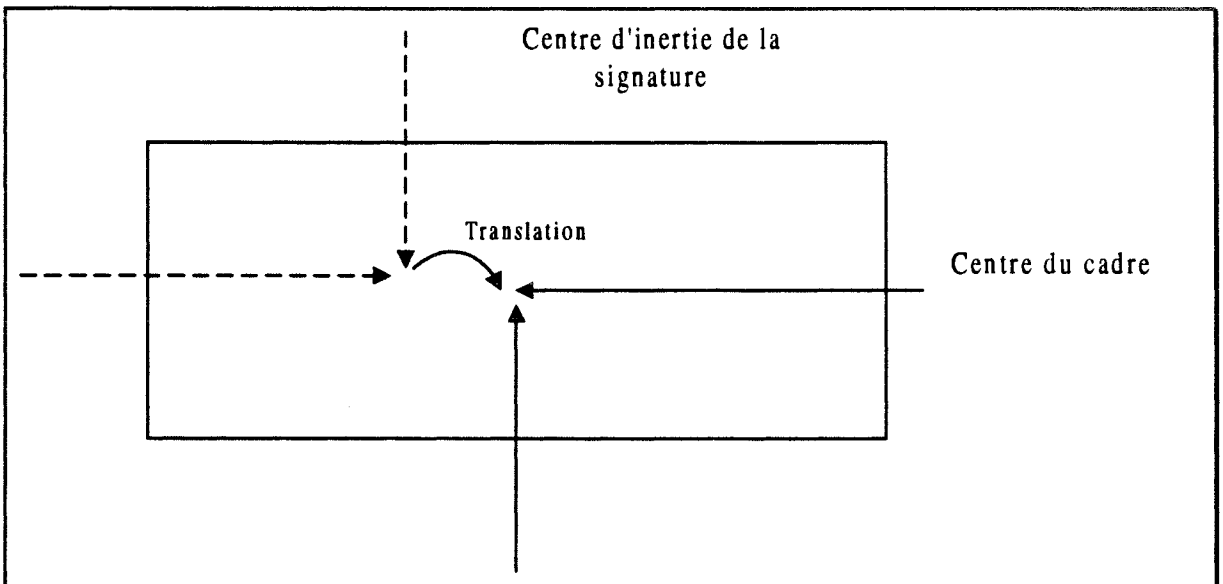


Figure 12 : Translation de la signature au centre du cadre de zoning

2.3.3.2. L'extraction de caractéristiques

Sabourin et al. [Sabourin 92] rapportent deux approches principales entreprises par les chercheurs pour la définition d'un facteur de formes dans le cadre de l'élimination des faux aléatoires. Une première approche s'attache à des mesures locales sur les signatures, nécessitant une segmentation en lettres, constituant un problème non résolu en pratique. Une telle approche s'attache au libellé des signatures, elle est dite sensible au texte. De plus, des graphies particulières comme celles de parafes Européens la rendent inutilisable. Une seconde approche est basée sur des mesures globales de la forme des signatures amenant une insensibilité au texte, et donc un caractère plus général que la première approche.

Afin de réaliser le compromis local/global, R. Sabourin a proposé une approche multi-échelle en reconnaissance globale offrant simultanément une vue globale des signatures à différents points de vue, liés aux échelles utilisées.

Le facteur de formes proposé est issu d'une méthode d'extraction de caractéristiques pour la reconnaissance de caractères manuscrits isolés, appelée Shadow Code [Burr 87], [Burr 88]. Son application directe n'est pas appropriée aux images de signatures, car une signature est une entité plus complexe qu'un caractère. R. Sabourin a donc étendu le principe du Shadow Code à la signature [Sabourin 93], [Sabourin 94a], [Sabourin 94b], et a défini la technique de l'Extended Shadow Code.

L'Extended Shadow Code consiste, à partir d'une image de signature binarisée, à projeter l'information de la signature sur un maillage superposé à l'image. Ce maillage est composé de la répétition d'un motif de base présenté Figure 13 :

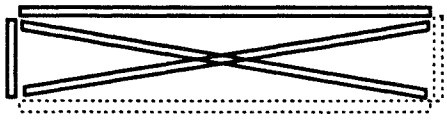


Figure 13 : Motif de base

Ce motif est composé de trois types d'axes ou bâtons, les bâtons verticaux, horizontaux et diagonaux, pour lesquels les deux directions diagonales sont considérées.

Le principe de projection sur des axes de références est assez classique en reconnaissance de caractères. On obtient, sur ces axes, des signaux de projection caractérisant l'objet signature. Les travaux de Al-Yousefi et Huang [Al-Yousefi 92], [Huang 92] utilisent cette approche pour la reconnaissance optique de caractères, par analyse des projections orthogonales des points de la signature sur un motif carré exinscrit.

L'Extended Shadow Code est plus complet, puisqu'il est construit par la répétition du motif de base (Figure 13), en formant une série de motifs globaux conduisant à une approche multi-échelle (Figure 14). Une telle approche est très intéressante du point de vue de la caractérisation obtenue, qui peut être grossière et plus ou moins fine selon le motif choisi. Ainsi, des détails d'importance relative peuvent être perçus.

Conjointement, un second intérêt provient du fait qu'une ressemblance existe entre toutes les signatures à une échelle très globale, et toutes les signatures sont différentes à une échelle très locale. Comme toujours, il y a un juste compromis d'échelle susceptible de répondre correctement aux problèmes posés. La définition des différents motifs permet de proposer différents types d'examens des signatures.

Cependant, R. Sabourin n'a pas encore résolu le problème du choix automatique d'une configuration particulière de motifs, pour un scripteur particulier. Cette particularisation impose de disposer d'un lot de signatures conséquent, pour un scripteur donné, si l'on veut s'assurer de la pertinence de ce choix. Bien qu'il ait conduit des études sur la pertinence des différents motifs [Sabourin 94a], [Sabourin 95a], R. Sabourin n'a pu définir une méthodologie générale de sélection.

En ce qui nous concerne, nous n'avons pas approfondi ce problème, qui mériterait des études plus poussées.

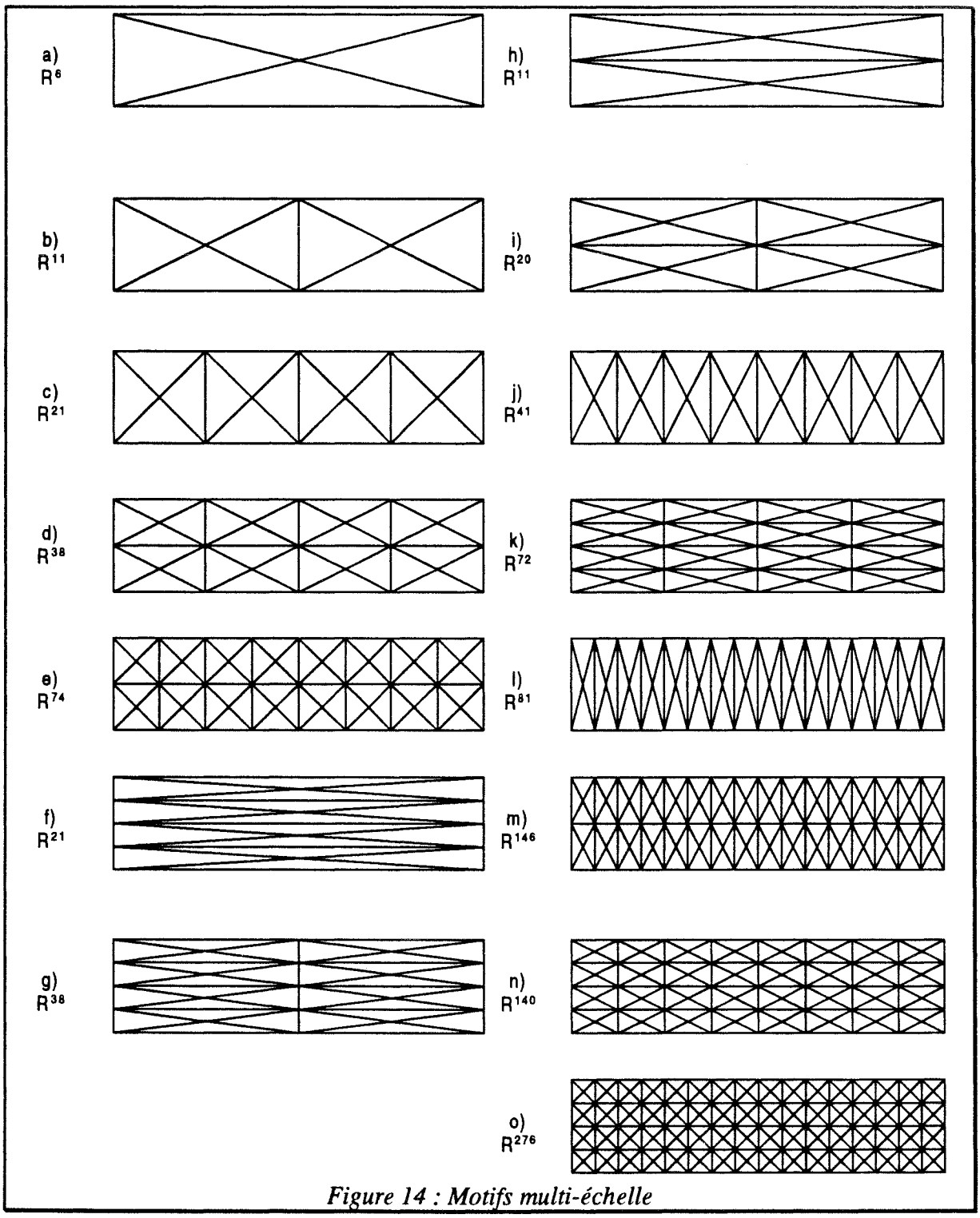


Figure 14 : Motifs multi-échelle

La projection orthogonale concerne le motif de base et n'est réalisée que pour les points inscrits dans ce motif. Même si le maillage est devenu plus complexe, par la répétition du motif de base pour former les différentes échelles, la projection ne concerne que les zones de l'image cernées par les motifs de base (Figure 15). Chaque bâton du motif porte la silhouette numérisée du tracé obtenu, par la projection des pixels de l'écriture manuscrite. Ce principe de projection détermine d'abord le constituant horizontal et respectivement vertical le plus près

des pixels du tracé de la signature, en terme de distance Euclidienne. Puis, ces pixels sont projetés sur les détecteurs diagonaux.

Chaque bâton des motifs est constitué d'un nombre fini de cases élémentaires en correspondance avec la quantification spatiale de l'image. A chaque case élémentaire est attribuée une valeur correspondant à la caractéristique de la signature codée. La projection initialement proposée par R. Sabourin reste dans la logique de Burr, elle traduit la notion de présence de l'objet dans la zone de projection liée au motif (Figure 15). La signature objet est alors exprimée par une série de signaux résultant de la projection de la présence de pixels de la signature.

L'intérêt de cette méthode est de permettre le codage de plusieurs caractéristiques différentes de la signature comme la position, l'orientation, en la percevant sous plusieurs aspects, pour répondre à différents problèmes de faux.

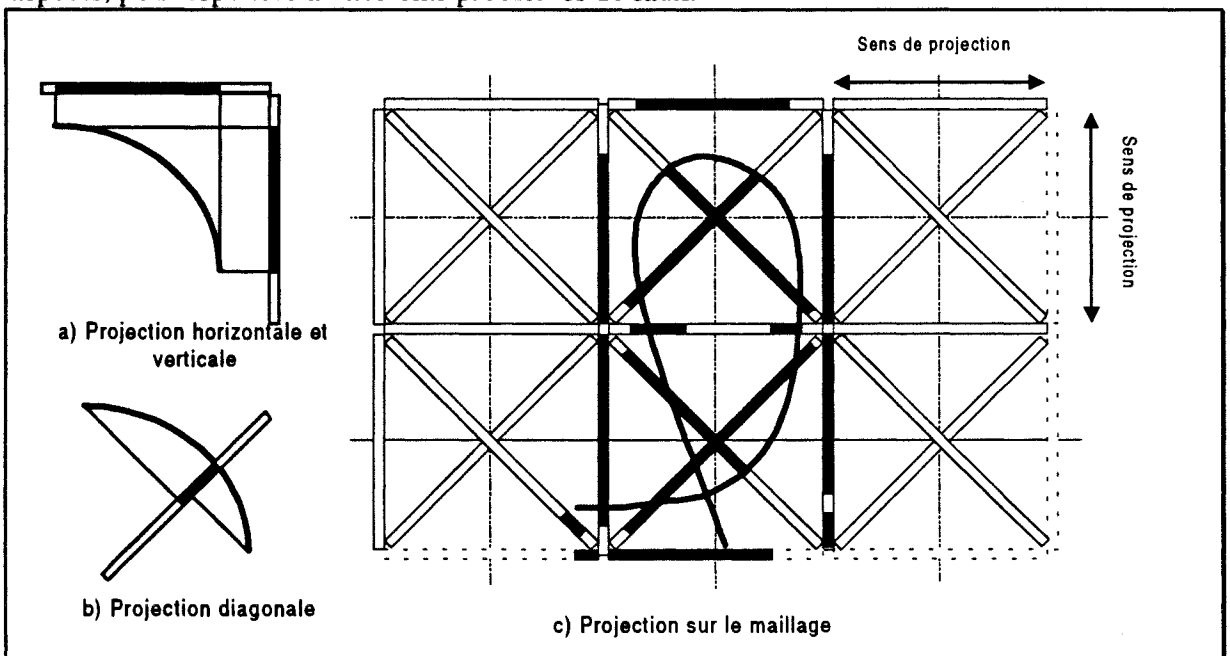


Figure 15 : Illustration du mécanisme de projection de l'Extended Shadow Code

L'utilisation des bâtons codés, comme vecteurs caractéristiques dans une reconnaissance par approche globale représente une masse d'information trop importante. En prenant l'exemple des bâtons du motif *a*, on obtient un espace de paramètres dans $\Re^{2336}$. Cette réduction importante par rapport à l'espace image dans $\Re^{16777216}$, n'est pas suffisante pour la base d'images de signatures en notre possession. Elle devient encore plus faible lorsque le motif a une résolution plus grande. L'exemple du motif *o* donnerait une réduction dans $\Re^{8832}$. Aussi, l'information portée par chaque bâton doit être synthétisée afin d'en réduire la masse. La qualité du descripteur de bâton est indissociable de celle de la technique de codage utilisée. Elle se traduit par le pouvoir discriminant et pertinent du couple codage/descripteur et se quantifie a posteriori par un taux d'erreur de classification.

Lorsque tous les pixels de la signature ont été projetés, le nombre de cellules activées sur chaque bâton est obtenu, puis décrit par le ratio des cellules activées sur le nombre total de cellules du bâton. Ainsi, la dimension du vecteur caractéristique correspond au nombre de bâtons présents dans le motif.

L'utilisation de cette projection locale du tracé de la signature, en relation avec les échelles des motifs, réalise un compromis entre des approches globales, où la silhouette de la signature est considérée comme un tout, et des approches locales où les projections caractérisent des portions locales du tracé. L'Extended Shadow Code permet d'obtenir des mesures globales ou locales, sans segmenter au préalable l'image, réduisant ainsi l'influence pernicieuse d'une étape de segmentation.

2.3.3.3. Le protocole d'évaluation du pouvoir discriminant du facteur de formes

Pour évaluer la pertinence du facteur de formes constitué de la méthode de codage et du descripteur, il est nécessaire de construire un protocole d'évaluation de son pouvoir discriminant. L'établissement de ce protocole passe par le choix de la méthode de constitution des ensembles de référence (§ 1.2.3.3.2.3) et de test, ainsi que du classificateur utilisé.

La pertinence du couple codage/description et, plus globalement de l'ensemble de perception, ne reste souvent qu'une estimation plus ou moins fiable. En effet, la représentativité des lots de données par rapport à l'ensemble total des données est indéterminée. Les performances obtenues doivent donc être relativisées.

La constitution des lots de données est effectuée par une méthode de type « holdout » avec choix aléatoire multiple des données. Le classificateur choisi pour l'évaluation de la pertinence est la règle des k plus proches voisins. Cette règle permet, dans le cas d'une approche non paramétrique, d'obtenir une évaluation de la borne minimale de l'erreur de vérification avec un seul voisin, sous l'hypothèse d'un nombre suffisant de données.

Puisque l'étude porte sur un ensemble de 20 scripteurs fournissant chacun 40 signatures, soit une base de données de 800 images, on constitue un ensemble de référence E_{ref} par scripteur, à partir de ses 20 premières signatures, et de 5 signatures choisies aléatoirement parmi les 20 premières signatures des 19 autres scripteurs. Un ensemble de test E_{gen} est formé selon les mêmes considérations : 20 dernières signatures du scripteur, puis 5 signatures choisies aléatoirement parmi les 20 dernières des 19 autres scripteurs. Les ensembles E_{ref} et E_{gen} sont donc composés de 115 signatures, et le choix aléatoire de ces ensembles permet d'obtenir une indépendance statistique de l'évaluation.

L'évaluation de la probabilité d'erreur qualifiant la pertinence du facteur de formes est alors donnée par le taux d'erreur de classification d'un k plus proches voisins, à partir du taux d'erreur ε_1, c'est-à-dire le taux de refus de signatures authentiques et ε_2 le taux d'acceptation de faux. Le taux d'erreur totale ε_t est donné à partir des taux ε_1 et ε_2 selon (Eq. 1).

$$\varepsilon_t = \frac{\varepsilon_1 + \varepsilon_2}{2} \qquad \textit{(Eq. 1)}$$

La pertinence exprimée par ε_t est liée à la technique de perception, mais également au choix des faux de référence et de test. Pour en relativiser l'influence et pour obtenir une évaluation générale de la pertinence, une simulation de type Monte-Carlo est réalisée selon le protocole défini dans [Sabourin 93], [Sabourin 94a], [Sabourin 94b]. Ce protocole tient

également compte des différentes résolutions, de la technique de codage pour donner une évaluation générale de l'ensemble de l'approche multi-échelle.

Protocole d'évaluation:

POUR motif = a à o
POUR itr =1 à 25
    POUR individu = 1 à 20
    Choisir aléatoirement les 5x19 échantillons de ω_2
        Constituer les ensembles de référence ω_1 et de test ω_2
        Calculer les distances et évaluer les erreurs ε_1 et ε_2 pour 1, 3, 5, 7, 9 voisins
    FIN POUR
    Evaluer les erreurs du système (pour chaque itération)
FIN POUR
FIN POUR
Evaluer l'erreur moyenne totale du système ainsi que la dispersion, pour 1, 3, 5, 7, 9 voisins

2.3.4. Le raisonnement

Dans l'objectif d'une implémentation très simple, et d'une économie de place mémoire, R. Sabourin [Sabourin 94b] [Sabourin 95a] utilise des classificateurs à seuil opérant sur les caractéristiques extraites des différents facteurs de formes de l'approche de perception multi-échelle. De fait, 15 classificateurs sont utilisés simultanément, pour vérifier l'identité annoncée d'un scripteur. L'ensemble des classificateurs est particularisé pour un scripteur particulier au travers du seuil.

Le principe du classificateur à seuil est de déterminer dans l'espace des caractéristiques, à partir d'un lot d'apprentissage, un seuil de comparaison de distance τ qui minimise l'erreur totale de classification en mémorisation. A partir d'un nouveau lot de signatures de référence, ce seuil est appliqué ensuite pour déterminer l'appartenance de nouvelles données, par rapport à un ensemble réduit de prototypes de comparaison mémorisé.

2.3.5. Le protocole de test

Pour évaluer les performances de l'ensemble facteur de formes et classificateur à seuil, un protocole spécifique de test a été défini. D'une part, un ensemble E_{comp} est formé de N_{comp} signatures authentiques du scripteur considéré, prises parmi ces 20 premières signatures. Il sert de comparaison lors de l'évaluation des performances en généralisation du classificateur à seuil. D'autre part, pour la détermination du seuil τ, N_{app} signatures, c'est-à-dire 20- N_{comp}, sont choisies parmi les signatures restantes des 20 premières signatures du scripteur considéré, et 5x19 signatures authentiques sont choisies aléatoirement parmi les 20 premières signatures des autres scripteurs de la base. Ils forment l'ensemble des faux aléatoires pour l'ensemble

d'apprentissage E_{app}. Enfin, un ensemble de test E_{test} est formé à partir des 20 dernières signatures du scripteur et 5x19 signatures choisies aléatoirement parmi les 20 dernières signatures des autres scripteurs. Il sert à évaluer les performances en généralisation du classificateur à seuil pour le scripteur considéré. Comme pour l'évaluation de la pertinence du facteur de formes, un certain nombre d'itérations redéfinit les différents ensembles, relativisant les résultats par rapport à leur définition.

La règle de discrimination utilisée consiste à étiqueter la signature testée respectivement comme vraie ou fausse si la distance euclidienne minimum qui la sépare des signatures de comparaison est inférieure, ou si la distance est supérieure à τ. L'évaluation totale est obtenue à partir du protocole suivant :

Protocole d'évaluation :
POUR motif = a **à** o
POUR itr =1 **à** 25
N_{comp}=6
POUR scripteur = 1 **à** 20
Choisir N_{comp} signatures de comparaison parmi les signatures de référence du scripteur

Constituer l'ensemble d'apprentissage E_{app}

Evaluer le seuil de comparaison t qui minimise l'erreur totale en mémorisation

POUR toutes les signatures de test choisies dans l'ensemble de généralisation

Evaluer la distance Dmin entre chaque signature et les signatures de comparaison

SI Dmin > τ et signature de test est élément de la classe des vrais
ALORS une erreur de type I

SI Dmin < τ et signature de test est élément de la classe des faux
ALORS une erreur de type II

FIN POUR

FIN POUR
Evaluer les erreurs du système (pour chaque itération)
FIN POUR
Evaluer l'erreur moyenne du système ainsi que la dispersion, pour 1, 3, 5, 7, 9 voisins
FIN POUR

2.3.6. Coopération de classificateurs

L'objectif de cette étape est de finaliser la décision prise par chaque classificateur associé à chaque niveau de perception, afin de former un classificateur intégré adapté à chaque scripteur. En effet, la coopération de classificateur doit permettre la mise en oeuvre d'un système complet de vérification, adapté aux différents styles d'écriture. En particulier, certains

motifs de perception ont un potentiel de discrimination plus important pour la signature d'un scripteur que pour un autre. Ainsi, l'objectif final est de singulariser le système complet pour un scripteur particulier. On peut remarquer que cette démarche permet également de sécuriser la décision en multipliant les points de vue sur une même signature.

A cette fin, R. Sabourin a proposé une coopération de classificateurs par vote majoritaire. Chacun des classificateurs à seuil prend une décision quant à l'appartenance d'une signature à la classe des vraies ou des fausses, en spécifiant l'étiquette de la classe concernée (ω_1 ou ω_2). Cet étiquetage est défini à partir de la règle de discrimination établie au cours de l'apprentissage des classificateurs. Le vote majoritaire consiste alors, à affecter une étiquette définitive à la signature incriminée, à l'aide d'une règle classique de combinaison par vote. Ainsi, la signature est définie comme vraie si la somme des propositions d'appartenance à la classe vraies est supérieure ou égale à $\alpha \times K$ ou étiquetée fausse si la somme des appartenances à la classe fausses ne vérifie pas cette condition. K correspond au nombre de classificateurs mis en jeu avec $K \in \{3,5,7,9,11,13,15\}$ et $\alpha = 0.5$ pour un vote à majorité simple. Cette règle permet la définition d'une classe de rejet lorsqu'aucune des conditions précédentes n'est atteinte.

La Figure 16 présente la structure de ce classificateur. Par un choix moins critique du facteur de formes, ce système complet a un comportement plus robuste. Le classificateur intégré s'affranchit ainsi du style d'écriture du scripteur, qui n'entre plus en considération dans l'utilisation globale. Dans [Sabourin 94b], l'auteur a montré l'efficacité de ce classificateur intégré.

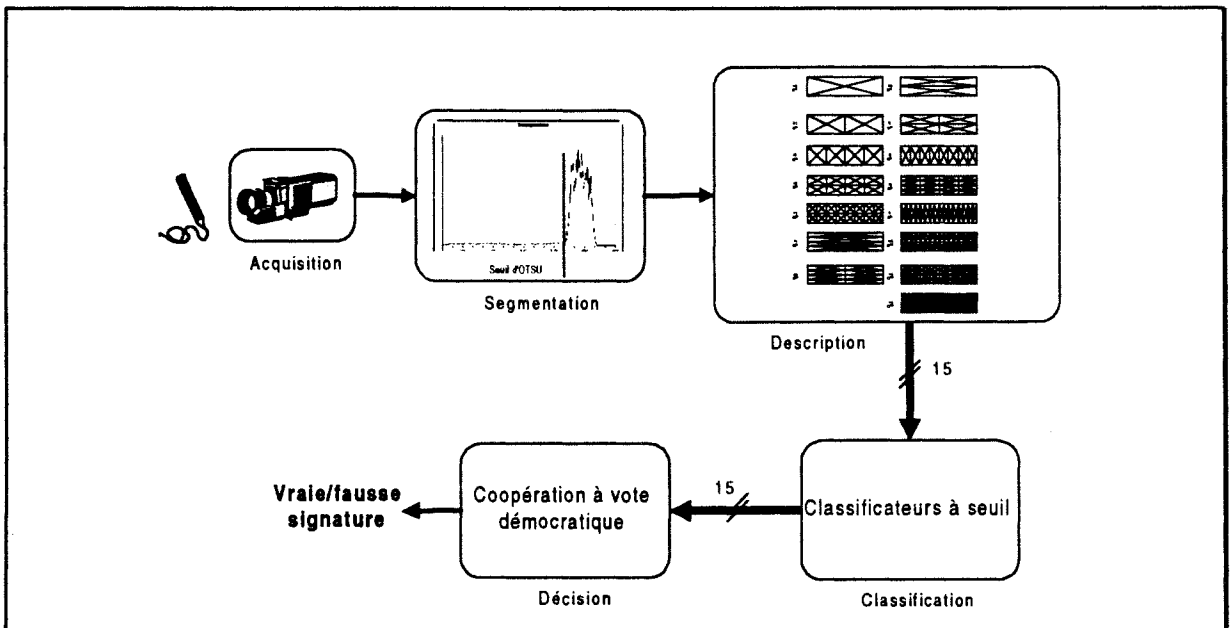


Figure 16 : Système complet de vérification de signatures selon l'approche de R. SABOURIN

2.4. Les améliorations possibles

Cette description du système de vérification de signatures manuscrites défini par R. Sabourin nous amène à faire plusieurs remarques. La dichotomie en 2 étapes de perception et de raisonnement précise bien les deux fonctions essentielles du système.

Pour l'étape de perception, le facteur de formes proposé et sa structure sont appropriés au problème de discrimination des faux aléatoires. En outre, il ouvre la voie au traitement d'autres faux, si l'information codée est susceptible de montrer la différence les séparant des signatures authentiques. L'approche multi-échelle offre également une richesse d'information qui permet l'utilisation de la méthode, quel que soit le style de signatures, c'est-à-dire européen ou nord-américain. Simultanément, on conçoit parfaitement qu'une amélioration du facteur de formes augmentera la performance du système complet.

La conception du classificateur intégré permet de finaliser le système, et autorise une décision plus performante et robuste pour la vérification de signatures. La définition du nombre de classificateurs impliqués dans la prise de décision contribue à la singularisation d'un système pour un scripteur particulier. On peut cependant reprocher à l'étape de raisonnement de ne pas être complètement intégrée. En effet, chaque classificateur autonome effectue une décision sur la signature incriminée. De ce fait, une erreur locale de décision est transmise au système de coopération qui la prend en compte. Cette erreur locale peut être due à une décision binaire, prise à partir d'une mauvaise connaissance de ce classificateur à l'échelle locale. On note également que les règles de discrimination des classificateurs sont établies une fois pour toutes sur la base d'un ensemble d'apprentissage dont la représentativité du problème de vérification est indéterminée.

La coopération par vote majoritaire est un choix logique, mais reste trop séparée et dépendante des classificateurs en amont. Elle traite de façon abstraite les sous-décisions données par les classificateurs. En outre, sa singularisation n'est actuellement définie qu'a posteriori en relation avec les erreurs finales. Par ailleurs, elle n'est pas associée à la sécurité que l'on peut exiger d'un système de vérification de signatures, et seul le facteur α permet de sécuriser a posteriori et manuellement la décision finale.

Il paraît donc légitime de renforcer la pertinence du facteur de formes pour son application au traitement des faux aléatoires, mais également de l'adapter pour ouvrir les capacités du système vers le traitement de faux plus complexes comme les faux par calque ou transfert optique.

Simultanément, l'étape de raisonnement doit évoluer vers la définition d'un système le plus intégré possible, offrant des possibilités d'évolution et d'adaptation à un scripteur particulier tout en amenant un surcroît de sécurité sur la décision, en relation avec le problème spécifique posé par la graphie du scripteur.

3. Conclusion

En termes de complexité, la vérification de signatures manuscrites est un problème difficile. Un certain nombre de facteurs apportent quelques simplifications. L'utilisation d'un cadre réduit les problèmes de localisation de la signature et permet de concentrer l'étude sur la

signature elle-même. Le vocabulaire est réduit puisque l'approche est multi-mono-scripteur. De plus, la solution proposée s'oriente vers une reconnaissance globale. Le choix d'une telle approche de reconnaissance est justifié par les études menées par R. Sabourin et d'autres chercheurs, tant sur les approches directe que structurelle. En outre, les différentes revues de l'écriture manuscrite montre qu'une approche hors-ligne conduit au choix d'une approche globale dès que la complexité des objets à reconnaître est importante.

Dans l'approche globale choisie, le facteur de formes constitue un potentiel déterminant de l'amélioration des résultats pour le problème posé. L'approche multi-échelle étend la discrimination à n'importe quelle graphie ou style d'écriture, ce qui lui donne un caractère général. Certaines échelles, plus appropriées à un scripteur qu'à un autre, autorisent la singularisation d'un système de vérification à un scripteur particulier, même si cette particularisation exige un lot de formes de référence assez conséquent.

Le mécanisme de projection ouvre le codage d'un éventail d'informations différentes selon le problème posé. En effet, le codage de la présence de point amène à un certain degré de performances, dont l'amélioration ne peut être que profitable pour le système complet de vérification. En outre, le codage d'informations relatives à d'autres type de faux peut ouvrir de nouvelles perspectives.

L'étape de Raisonnement est structurée par l'approche multi-échelle impliquée par le facteur de formes. L'objectif de former une étape de raisonnement intégrée contribue à la finalisation du système de vérification. Chaque classificateur, associé à une échelle particulière de perception, synthétise la connaissance exprimée par les données sous une forme qui lui est propre. Le problème de vérification est complexe et chaque classificateur subit cette complexité. En effet, le problème, bien qu'à deux classes, montre en réalité que les classes de vraies et de fausses sont composées de sous classes relatives aux variations de tracé, aux habitudes du scripteur, ou encore à l'aspect téléologique de l'acte d'écriture. Simultanément, on ne possède qu'une connaissance restreinte des signatures au travers d'un nombre limité de données. Un apprentissage permanent est alors peut-être plus réaliste? En outre, les signatures, perçues plus ou moins grossièrement par les motifs de perception, sont des informations plus ou moins imprécises. Le raisonnement qui lie ces informations ne peut être qu'incertain. En conséquence, on doit conduire un renforcement de la sécurité de la décision.

La prise en compte de l'imprécision des informations et de l'incertitude du raisonnement dans l'approche multi-classificateurs traitée par R. Sabourin, ainsi que le recouvrement des classes de vraies et fausses signatures, induisent notre proposition : une solution basée sur le concept des ensembles flous, en adéquation avec ces caractéristiques.

Chapitre 2

Perception des signatures, une approche statique et pseudo-dynamique

1. Introduction

La conception du système de vérification passe par celle d'un facteur de formes qui permette le passage de l'espace image à celui des vecteurs caractéristiques, exploités par l'approche globale de la reconnaissance de formes. Il faut définir un mécanisme de perception dont le fonctionnement soit orienté vers un transfert approprié des informations pertinentes en relation avec l'expression des caractéristiques discriminantes vis-à-vis du problème à résoudre. Parallèlement, il faut viser l'obtention de performances finales, exprimées en terme d'erreur de reconnaissance, dépendantes de la pertinence du facteur de formes par rapport aux invariants des classes.

Ce chapitre concerne l'amélioration du facteur de formes initialement proposé par R. Sabourin. L'approche multi-échelle est conservée eu égard aux possibilités qu'elle offre. Comme nous l'avons vu, le principe de l'Extended Shadow Code est dérivé d'une méthode dédiée à la reconnaissance de caractères manuscrits isolés. Son application au codage de signatures manuscrites demande quelques adaptations. Par ailleurs, l'information statique utilisée, trop dénudée, mérite d'être enrichie pour conduire à de meilleures performances.

L'information extraite des signatures relève du problème initialement posé, pour la discrimination des faux aléatoires. Cependant, le principe de l'Extended Shadow Code, indépendant de la nature de l'information extraite des signatures, nous permet de faire évoluer la nature de cette information en portant notre étude sur le traitement simultané de faux aléatoires et de faux plus complexes comme les faux par photocopie.

2. *Le facteur de formes pour une approche statique de la vérification*

2.1. Les prétraitements

Dans le cadre de l'évolution de l'approche de perception, les étapes de prétraitement établies par R. Sabourin, pour la base de données à l'étude, ne demandent pas de modification du fait de leur bonne adéquation au problème. Ainsi, l'ajustement de la signature sur le centre de gravité du cadre d'écriture est conservé. Le seuil d'Otsu, approprié à une problématique d'image de signatures, est conservé dans la suite des travaux.

2.2. Les motifs multi-échelle

Nous avons vu au chapitre précédent, que l'approche de perception multi-échelle avec le principe de l'Extended Shadow Code était appropriée à la vérification de signatures manuscrites. En effet, l'apport de plusieurs sources d'informations, fournies sous des points de vue différents à des échelles de perception de résolutions grossières à fines, permet de renforcer la décision de vérification et de s'adapter aux graphies particulières des scripteurs. Aussi, est-il judicieux de profiter de cet apport en conservant l'approche multi-échelle et les différents motifs utilisés. En outre, le principe général de l'Extended Shadow Code, qui autorise une évolution possible vers un facteur de formes plus performant, peut être conservé.

2.2.1. Modification de la méthode de projection

Le principe de projection orthogonale, l'Extended Shadow Code (ESC), où les projections ne concernent que les points inscrits dans chaque motif, est adapté à la reconnaissance optique de caractères manuscrits isolés car l'objet caractère dans son ensemble est inscrit dans le motif. Or, pour la vérification de signatures, il n'y a pas d'objets dissociés inscrits dans chaque motif. La signature est une entité unique dont nous ne recherchons pas le sens véhiculé, ce qui n'oblige pas à sa segmentation en lettres. Dans ce cas, l'ESC entraîne une troncature de la projection de la signature par le motif, dans le cadre d'une projection exinscrite au motif. Cette troncature modifie l'information projetée et amplifie les déviations spatiales naturelles du tracé, propres au scripteur. Cette troncature ne s'effectue que lors de la projection sur les bâtons diagonaux du motif de base, mais le descripteur, en raison de son effet globalisant, va en minimiser l'impact. Néanmoins, pour éviter cette influence du motif et adapter réellement l'Extended Shadow Code aux images de signatures, il faut adopter la projection effectuée sur les bâtons horizontaux et verticaux comme mécanisme de projection sur les bâtons diagonaux, c'est-à-dire que l'on "code la signature sur le bâton le plus proche".

Nous appliquons ce principe quelle que soit la direction de codage, sans tenir compte de l'appartenance des points de la signature à un motif de base ou à un autre motif voisin. Ainsi, l'opération de projection est rendue homogène pour toutes les directions. Un exemple montrant la différence de concept est présenté sur la Figure 17.

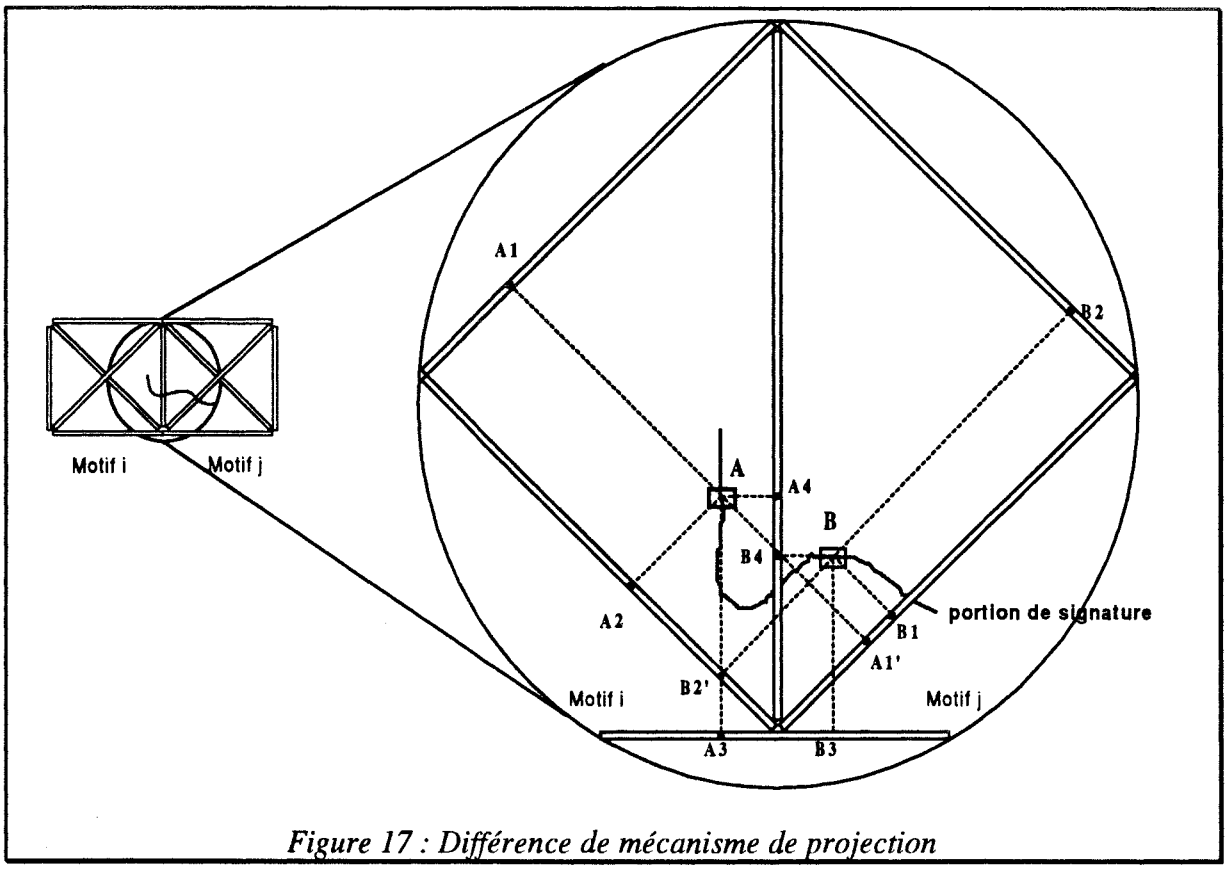


Figure 17 : Différence de mécanisme de projection

Le codage de deux pixels A et B de la signature, proches l'un de l'autre, donne les projections (A1, A2, A3, A4) et (B1, B2, B3, B4) par la méthode de R. Sabourin. Lorsque l'on utilise le maillage comme un motif et que l'on code au bâton le plus proche, les projections deviennent (A1', A2, A3, A4) et (B1, B2', B3, B4). Ainsi, le codage de deux points proches, donne deux codages proches, quel que soit le bâton concerné par la projection. On obtient ainsi une application homogène de la notion de projection aux plus proches bâtons.

2.2.2. Adaptation de la scrutation aux signatures

Indépendamment du mécanisme de projection, la scrutation proposée par R. Sabourin présente des difficultés d'ordre pratique. En effet, la scrutation de la signature pour la projection sur les bâtons diagonaux introduit des problèmes d'aliasage, bien connus en image et provoqués par la numérisation de l'image de la signature. La difficulté du codage des cases élémentaires sur les bâtons diagonaux s'exprime en termes de pixels dans le motif choisi, selon sa forme rectangulaire.

Cette représentation géométrique produit une perte d'information du fait de l'inadéquation entre les droites analytiques du motif et la position physique des pixels. De ce fait, lors du balayage d'une zone de l'image, il se peut que des pixels ne soient pas scrutés ou le soient plusieurs fois. La qualité de représentation du codage dépend alors de la forme du motif de base. Plus celui-ci sera proche d'un carré, moins le codage traduira le tracé réel de la signature puisque dans ce cas, seul un point sur deux est codé.

Cette perte d'information peut être capitale, notamment dans le cas de motifs carrés tels que les motifs *c*, *e*, *o* et selon le type de signatures. En effet, les signatures de notre base sont de type nord-américaines et leur graphie s'étend principalement dans le sens horizontal. Cependant, une graphie européenne, plus stylée, peut être plus sensible à ce problème car plus orientée dans le sens diagonal (Figure 18). Afin de compenser ce problème de codage dont l'importance peut être capitale selon le tracé de la signature, une série de points est ajoutée sur le bâton. La méthode de scrutation de Sabourin ainsi modifiée permet de coder toute l'information contenue dans l'image.

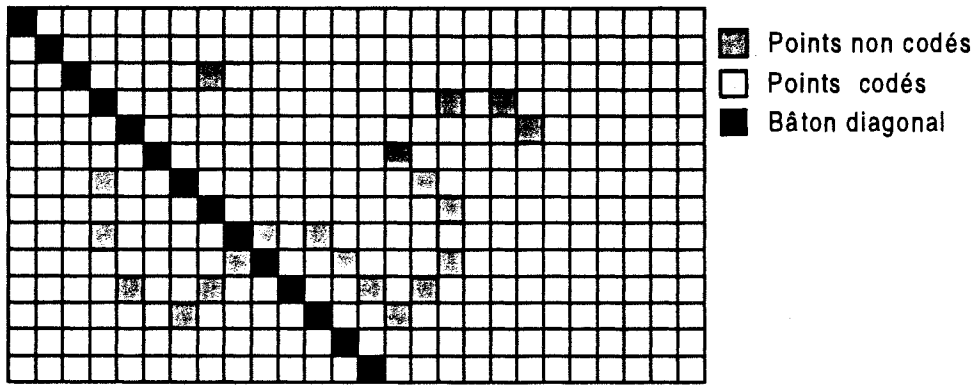


Figure 18 : Illustration graphique du problème d'aliasage

L'ajout du point origine de la scrutation permet d'éliminer toutes les portions d'image non scrutée précédemment et de coder de façon plus systématique l'image de la signature (Figure 19).

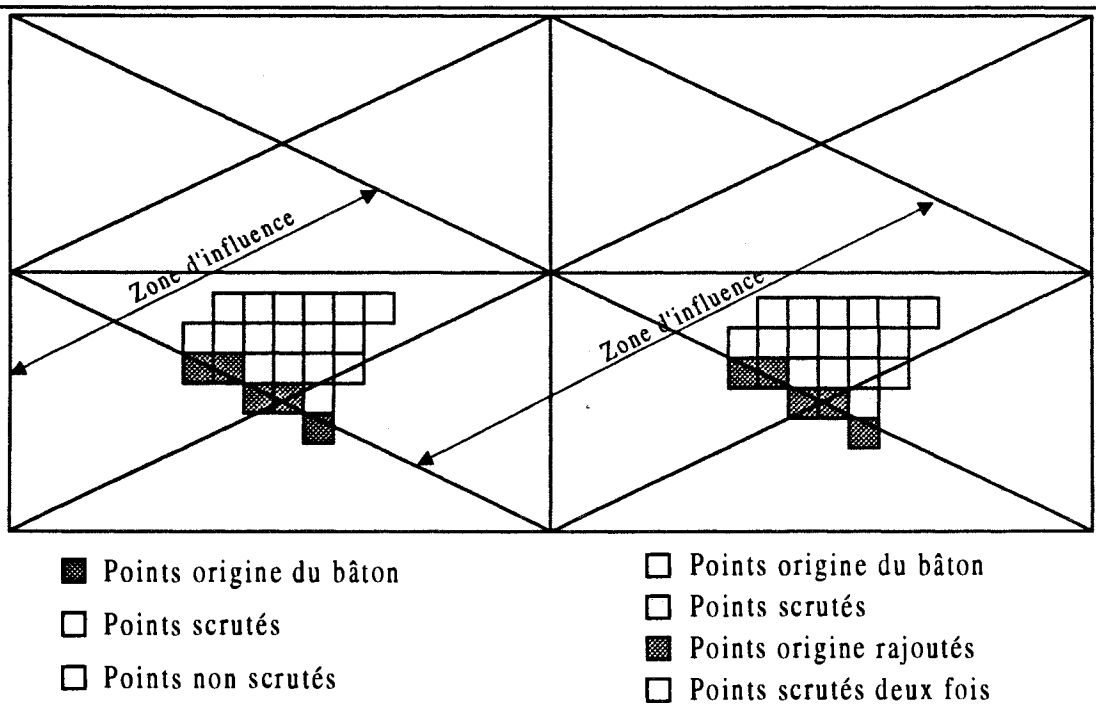


Figure 19 : Résolution du problème d'aliasage

2.3. Introduction du codage de la distance

Le problème de l'élimination des faux aléatoires se pose principalement en termes de géométrie de la signature exprimée par ses caractéristiques statiques. Le codage proposé par R. Sabourin (Figure 20a), ainsi que la méthode de reconnaissance de caractères de Burr utilisent la présence de points dans la portion de l'image scrutée. Cependant, cela ne reflète pas fidèlement l'expression géométrique de l'image de la signature.

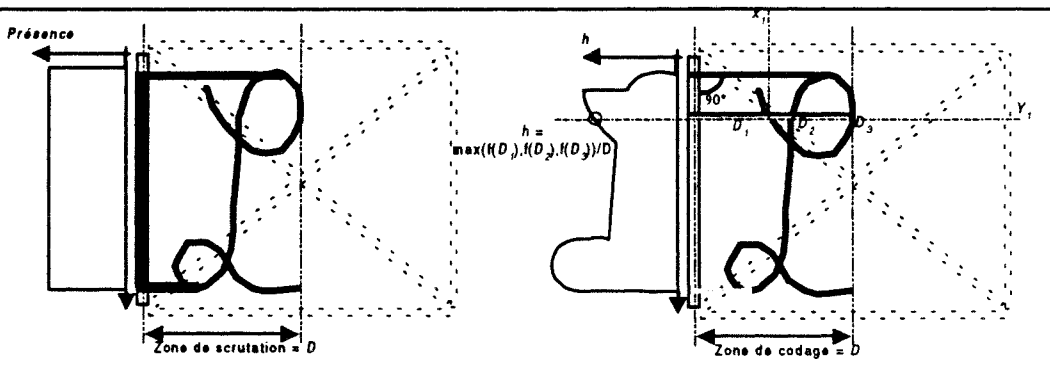
L'introduction dans la méthode de codage d'une information de distance séparant les points de la signature du bâton sur lequel l'image est codée semble alors plus représentative. La distance est une information de nature géométrique qui traduit plus explicitement la forme de la signature plutôt que sa présence. En effet, la notion de distance implique celle de présence de point, aussi la présence est-elle incluse dans la distance. La distance apporte donc plus d'information, pouvant ainsi limiter le recouvrement entre les classes de vraies et de fausses signatures et améliorant ainsi le taux de vérification.

L'introduction des distances dans le codage permet donc de coder la proximité des pixels de la signature par rapport aux bâtons, quelle que soit la méthode de scrutation utilisée. Ainsi, plus un pixel est éloigné d'un bâton, plus son influence sur ce bâton diminue (Figure 20b). Pour définir la valeur du codage de la case élémentaire du bâton concerné, on calcule la distance h d'un pixel de la signature de coordonnées (x,y) par rapport au bâton. Puis, h est normalisée par rapport à la distance maximum que l'on peut coder, située entre le bâton et la limite de la zone de codage. Elle peut être exprimée pour une droite définie par : $ax + by + c = 0$ selon l'équation (Eq. 5).

$$\mu_d = 1 - \frac{h}{D}$$
$$h = \frac{|ax_1 + by_1 + c|}{\sqrt{a^2 + b^2}} \qquad \textit{(Eq. 5)}$$

où μ : le niveau de proximité du pixel,
h : la distance du pixel au bâton,
D : la distance entre le bâton et la limite de la zone de codage,
et (x_1, y_1) : le couple de coordonnées.

On choisit le point le plus proche du bâton pour déterminer la valeur effective du codage dans le cas où plusieurs points sont à coder sur la même case élémentaire (Figure 20b).



a) Codage de la présence b) Codage de la distance

Figure 20 : Différence de codage d'information

Cette opération fournit l'enveloppe des portions de signature présentes dans le motif. Pour le motif *a*, le plus grossier, le signal de codage traduit l'enveloppe extérieure de la signature, ce qui se rapproche du facteur de formes proposé par Nouboud [Nouboud 88]. Cette approche du codage bénéficie d'une perception multi-échelle grâce aux différents motifs, ce qui nous donne un facteur de formes plus descriptif avec une approche grossière à fine.

2.4. La description des codages

Il est délicat d'intégrer directement la quantité d'informations fournies par les techniques de projection dans un système de discrimination. En effet, la taille de l'espace des caractéristiques est très importante. Bien qu'en-dessous de l'espace naturel que constitue l'image, la réduction apportée par la projection reste faible. Bellman, le père de la programmation dynamique, introduit l'expression "malédiction de la dimensionnalité" [Bellman 58] pour exprimer le fait qu'utiliser des vecteurs formes de grande taille est source de problèmes. Aussi, il est nécessaire de limiter cette quantité d'information par l'utilisation d'un descripteur de bâton.

R. Sabourin a proposé de décrire les codages sur chaque bâton par le taux de présence exprimé en pourcentage. Cette technique donne le niveau moyen de présence à un facteur 100 près, ce qui correspond à une approche statistique au travers de la description par la moyenne du signal.

Pour décrire les projections de la signature, une approche logique est de concevoir le mécanisme de projection comme une transformation de signaux. Le signal bi-dimensionnel image se transforme en une série de signaux mono-dimensionnels. Dès lors, on peut effectuer des opérations classiques de caractérisation de signaux mono-dimensionnels comme des analyses fréquentielles, des corrélations de signaux, ou des caractérisations statistiques. Nous avons donc envisagé ces techniques.

2.4.1. Corrélation de signaux

Pour la caractérisation de signaux, une approche fréquentielle est envisageable puisque le classificateur se retrouve comme corrélateur de données. Ainsi, nous pouvons proposer deux méthodes fréquentielles, l'analyse par transformée de Fourier ou par transformée de Walsh-Hadamard, selon la nature de l'information codée sur les bâtons.

La projection de type binaire introduit de fortes réponses lors de l'analyse fréquentielle par la transformée de Fourier, car les signaux portés par les bâtons des motifs sont de forme carrée. Une transformée de Walsh-Hadamard proposant une corrélation avec des signaux carrés donne l'équivalent binaire de la transformée de Fourier et convient mieux à la projection binaire [Ahmed 75].

2.4.1.1. Transformée de Walsh-Hadamard

La transformée de Walsh-Hadamard a été appliquée sur les premiers bâtons horizontaux de trois scripteurs pour le motif le plus grossier dans le cas du codage de la présence. L'objectif

est de déterminer si cette analyse conduit à la détermination d'invariants caractéristiques sous forme de fréquences pour les différents scripteurs.

Les résultats de la transformée de Walsh-Hadamard sur les signaux issus de la méthode Extended Shadow Code, montrent que le rang 0 est un bon point de comparaison des signatures. Ce point caractéristique confirme la proposition initiale de description de R. Sabourin puisque cela correspond à la valeur moyenne du bâton. Par ailleurs, certaines fréquences, représentées par les différents pics cerclés (Figure 21), sont potentiellement intéressantes pour une possible discrimination.

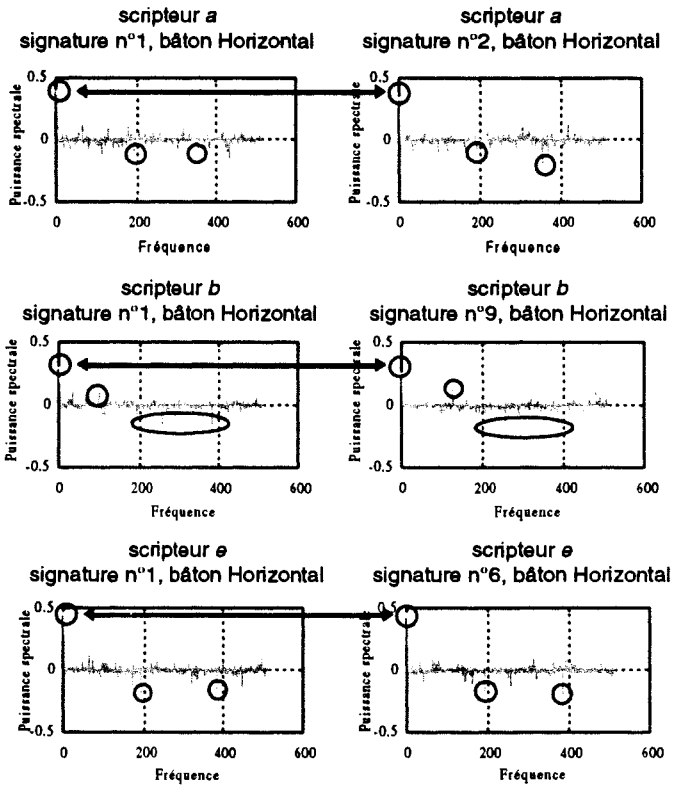


Figure 21 : Transformée de Walsh-Hadamard du 1er bâton horizontal du motif a pour 3 scripteurs, méthode de l'Extended Shadow Code.

Cette étude montre un certain potentiel de corrélation des signaux mais la détermination des corrélations discriminantes demande une analyse très poussée pour déterminer lesquelles sont représentatives. L'utilisation de toute l'analyse par la transformée de Walsh-Hadamard repose le problème de la masse d'information à traiter par un classificateur. En effet, la dimension de la transformée est de même taille que celle du signal analysé. En outre, cette analyse ne convient pas lorsque les signaux sont issus du codage de la distance.

Le résultat de l'analyse ne peut pas être considéré entièrement par le système de vérification.

2.4.1.2. Transformée de Fourier

La transformée de Fourier sur un signal binaire montre une forte densité spectrale de puissance dans les fréquences supérieures suite aux nombreuses transitions (Figure 22); cela

n'est donc pas d'un grand intérêt. Cependant, elle est adaptée aux signaux issus du codage de la distance. Selon la même analyse que celle réalisée pour la transformée de Walsh-Hadamard, la transformée de Fourier montre que le fondamental fournit un invariant permettant la discrimination. D'autres fréquences de cette analyse pourraient être utilisées pour caractériser les différents scripteurs, mais l'ensemble des combinaisons possibles est aussi vaste que celui des signaux de départ. Le choix d'invariants fréquentiels n'est donc encore pas le meilleur.

L'exploration des analyses fréquentielles montre que le choix de la moyenne du signal projeté, exprimée par la fréquence 0, est susceptible de caractériser correctement les ensembles de signatures similaires. En effet, sur la Figure 22, on remarque que le scripteur *a* a des puissances spectrales similaires, autour de 70 pour la fréquence 0, alors que pour le scripteur *b*, ses puissances se situent autour de 35. Mais, l'objectif de cette analyse est de mettre en évidence si l'ensemble de fréquences est spécifique d'un scripteur, c'est-à-dire de dégager une propriété fréquentielle caractéristique d'un groupe de signatures. La détermination d'une telle propriété est difficile, en regard de la masse d'informations donnée par les transformées. Cependant, au delà du fondamental, les puissances sont très faibles, et l'association avec la moyenne ne serait intéressante qu'avec une pondération adéquate. Par ailleurs, il est nécessaire d'utiliser une transformée appropriée à la nature des signaux projetés (binaires → Hadamard; analogiques → Fourier). Il faudrait également tenir compte des fréquences selon l'orientation du bâton.

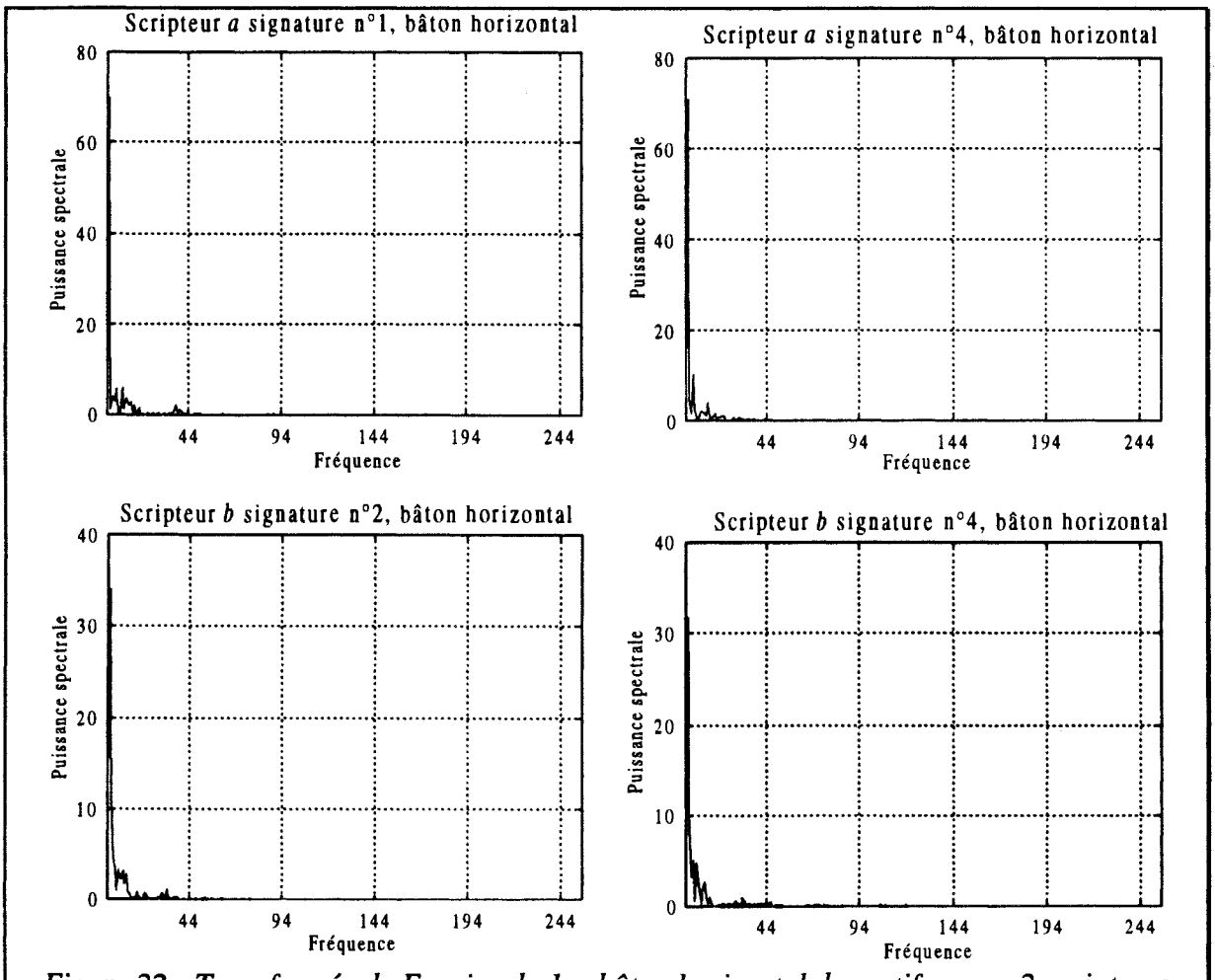


Figure 22 : Transformée de Fourier du 1er bâton horizontal du motif a pour 2 scripteurs méthode Extended Shadow Code

A partir de ces remarques, il est donc clair que la détermination de fréquences caractéristiques ne peut se faire que par une méthode automatique. Dès lors, il n'y a pas assez de données pour réaliser une telle sélection de fréquences. En conclusion, sans une étude très poussée sur la recherche de fréquences caractéristiques, seule la moyenne du signal apparaît discriminante.

2.4.2. Description statistique

La caractérisation des signaux de projection peut se faire par une approche statistique contenue dans les analyses de corrélation précédemment faites, notamment au niveau du fondamental qui correspond à la valeur moyenne du signal. De ce fait, une proposition conventionnelle et légitime de description statistique de signaux est une caractérisation en termes de statistiques du signal, mais aussi en termes de statistiques de la distribution [Monjallon 63]. Quatre descriptions peuvent être proposées (Figure 23).

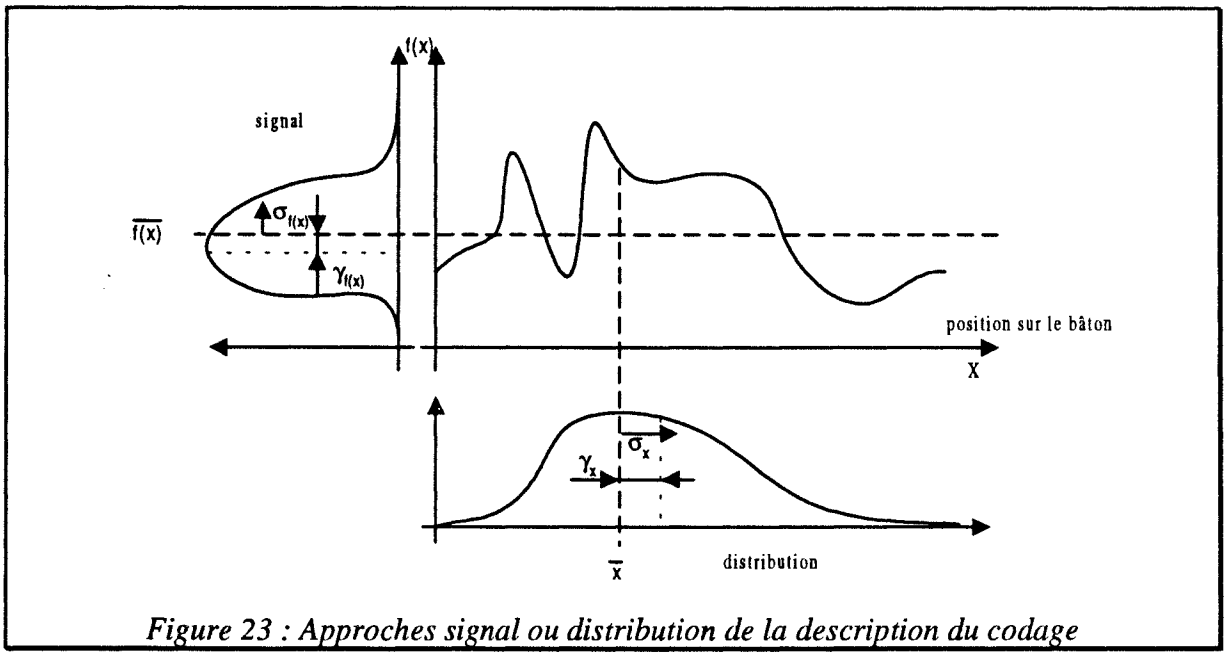


Figure 23 : Approches signal ou distribution de la description du codage

La moyenne du signal doit absolument être évaluée au vu des différentes analyses de corrélation de signaux. Elle correspond au descripteur D1. Une description statistique plus complète du signal peut être intéressante, par l'évaluation de la moyenne, la variance et la dissymétrie du signal. Elle est proposée sous la dénomination D2. Une description par la statistique de la distribution en considérant les signaux comme une fonction de distribution peut également apporter une caractérisation intéressante. Le descripteur D3 concerne la moyenne de la distribution et le descripteur D4, la moyenne, la variance et la dissymétrie de la distribution. Les relations permettant les calculs de ces différents descripteurs sont présentées respectivement par les équations (Eq. 6) (Eq. 7) (Eq. 8) et (Eq. 9).

$$D1 : \overline{f(x)} = \frac{\sum_x f(x)}{card(f(x))} \qquad \text{(Eq. 6)}$$

$$\text{D2} : \left[\overline{f(x)} = \frac{\sum_x f(x)}{card(f(x))} \quad \sigma^2_{f(x)} = \frac{\sum_x \left(f(x) - \overline{f(x)}\right)^2}{card(f(x)) - 1} \quad \gamma_{f(x)} = \frac{\sum_x \left(f(x) - \overline{f(x)}\right)^3}{card(f(x))\sigma^3_{f(x)}} \right] \qquad \text{(Eq. 7)}$$

$$\text{D3} : \bar{x} = \frac{\sum_i x_i f(x_i)}{\sum_i f(x_i)} \qquad \text{(Eq. 8)}$$

$$\text{D4} : \left[\bar{x} = \frac{\sum_i x_i f(x_i)}{\sum_i f(x_i)} \quad \sigma^2_x = \frac{\sum_i f(x_i) x_i^2}{\sum_i f(x_i)} - \bar{x}^2 \quad \gamma_x = \frac{\sum_i \left(x_i - \bar{x}\right)^3 f(x_i)}{\sigma^3_x \sum_i f(x_i)} \right] \qquad \text{(Eq. 9)}$$

La description des signaux par ces informations de nature statistique nous permet d'extraire des caractéristiques représentatives des codages et de descendre à une taille d'espace caractéristique plus appropriée à une classification.

2.5. Evaluation du choix du descripteur en correspondance avec la technique de codage

La pertinence réelle du couple facteur de formes et descripteur ne peut être quantifiée a priori, mais, une évaluation est possible par une analyse en composantes principales au travers du plan des observations.

Le motif le plus grossier est de dimension 6, et un descripteur extrayant une seule valeur par bâton nous permet de rester dans cette taille d'espace de représentation. Cependant, le choix d'une description plus complète est envisagé. Quoi qu'il en soit, l'espace de caractérisation est inabordable pour une représentation visuelle et mentale. Aussi, l'analyse en composantes principales va nous permettre une représentation des données sur un plan, avec une perte minimale d'informations.

Les représentations des données selon les descripteurs choisis permettent d'en évaluer le pouvoir discriminant. Cependant, on ne cherche pas à quantifier la pertinence mais simplement à déterminer un ordre de pertinence des descripteurs proposés pouvant guider le choix de l'un d'entre eux.

L'analyse en composantes principales est réalisée sur les 20 premières signatures de chacun des 20 scripteurs, et chaque scripteur est représenté par une étiquette graphique différente dans le plan des observations. Dans chaque analyse, les signatures d'un scripteur sont représentées par la même étiquette graphique. Pour une visualisation simplifiée, la technique de l'analyse en composantes principales nous permet de retenir les 2 axes factoriels les plus représentatifs de l'information présente. Il faut remarquer que le plan factoriel d'observation n'est pas gradué de la même façon selon les différentes expériences, ceci étant dû au phénomène de contraction de l'ACP. Toutefois, le niveau de recouvrement que l'on cherche à

observer est indépendant des échelles du plan d'observation, nous pouvons ainsi en évaluer le niveau d'une expérience à l'autre.

La première analyse en composantes principales concerne la technique de projection de R. Sabourin couplée avec le descripteur moyenne, l'ensemble formant le couple de référence. Le motif choisi est le plus grossier pour permettre une assez bonne interprétation des transformations réalisées. Cette analyse pose la base permettant la comparaison, sachant que R. Sabourin a déjà réalisé une évaluation a posteriori de sa méthode. Ainsi, nous serons capables d'estimer si un descripteur semble offrir un potentiel intéressant. La Figure 24 correspond au plan des observations pour la technique de référence et montre un recouvrement moyen entre les différentes classes de signatures, donc des performances raisonnables pour ce couple. En effet, les simulations poursuivies sur ce couple nous donnent pour le motif considéré un taux d'erreur moyen de vérification de 2,36% selon le protocole présenté dans le chapitre 1 [Sabourin 92]. Cette valeur sert de calibre pour la confusion des classes des observations et permet ainsi d'estimer le niveau de pertinence des autres propositions de ce couple.

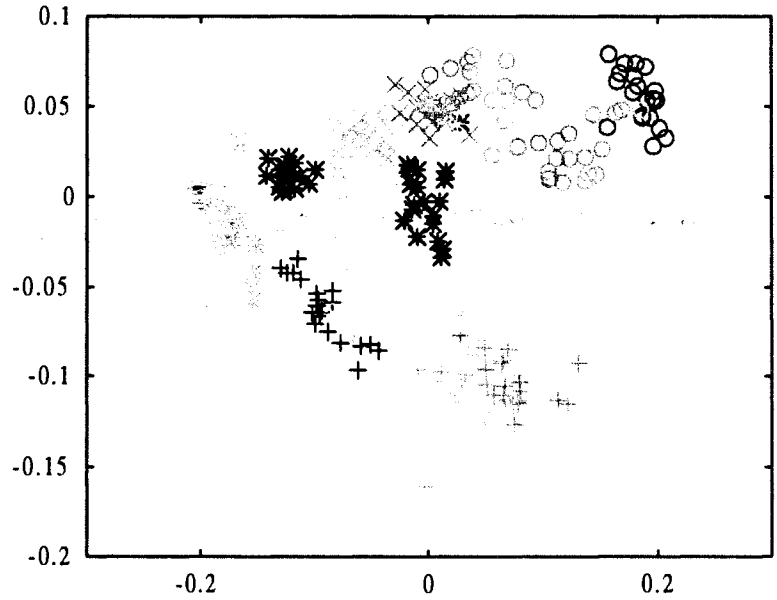


Figure 24 : Plan des observations selon la méthode de référence projection binaire/moyenne du signal (D1)

Le descripteur D2 associé à la méthode de projection de R. Sabourin induit le plan des observations de l'analyse en composantes principales de la Figure 25.

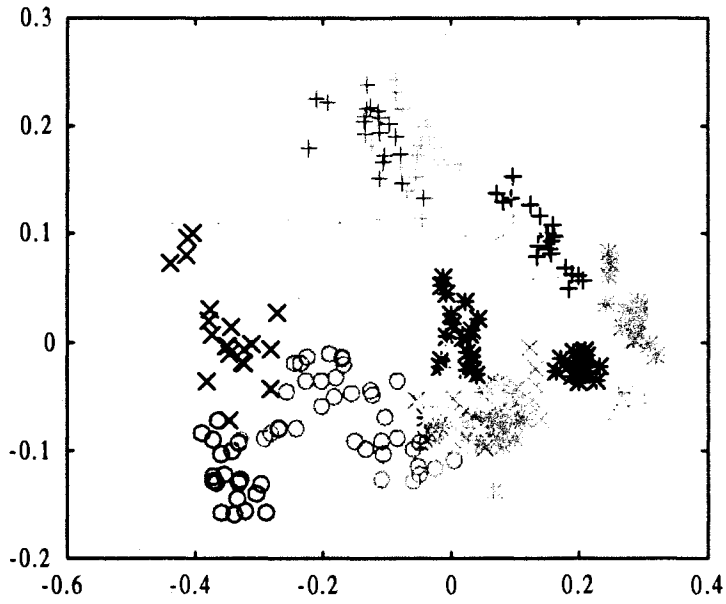


Figure 25 : Plan des observations selon la méthode de référence projection binaire/3 moments du signal (D2)

L'organisation des classes de signatures ne montre pas une différence sensible au niveau des séparations ou du niveau de recouvrement par rapport au descripteur D1. Le phénomène de contraction de l'analyse en composantes principales ne change pas ce niveau de recouvrement. Il est clair que l'utilisation de la moyenne seule fournit un taux de vérification aussi élevé que les trois moments du signal pour des calculs moins importants.

La même analyse est réalisée avec le descripteur D3. Conjointement, cette analyse nous permet de cerner le potentiel du descripteur D3 dans le cadre d'autres méthodes de codage.

Par comparaison aux résultats des analyses précédentes, il apparaît que le descripteur D3 ne montre pas un taux de recouvrement moins important que les descripteurs D1 ou D2. Il est donc tout à fait probable qu'il ne convienne pas aux autres méthodes de projections proposées.

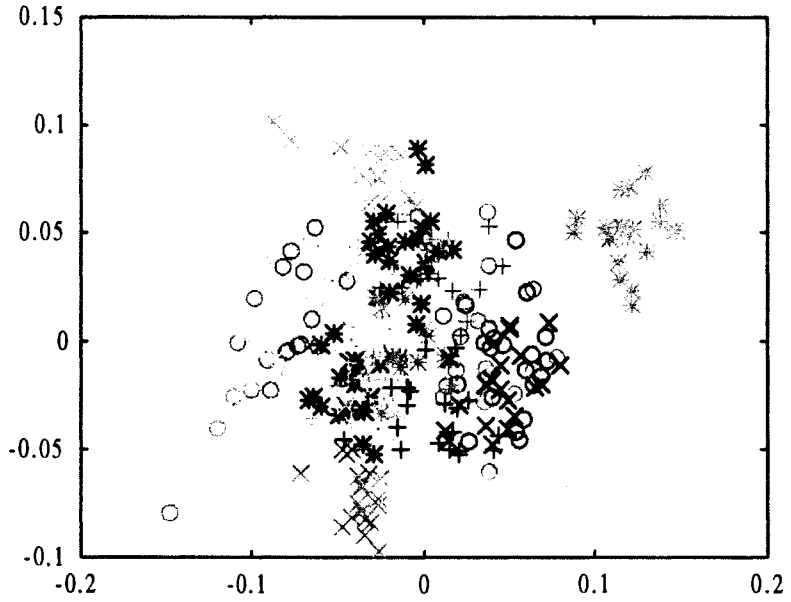


Figure 26 : Plan des observations selon la méthode de référence projection binaire/moyenne de la distribution (D3)

Enfin, l'observation des résultats de l'analyse en composantes principales du descripteur D4 associé à la méthode de R. Sabourin montre les mêmes résultats (Figure 27) et induit les mêmes conclusions que pour le descripteur D3.

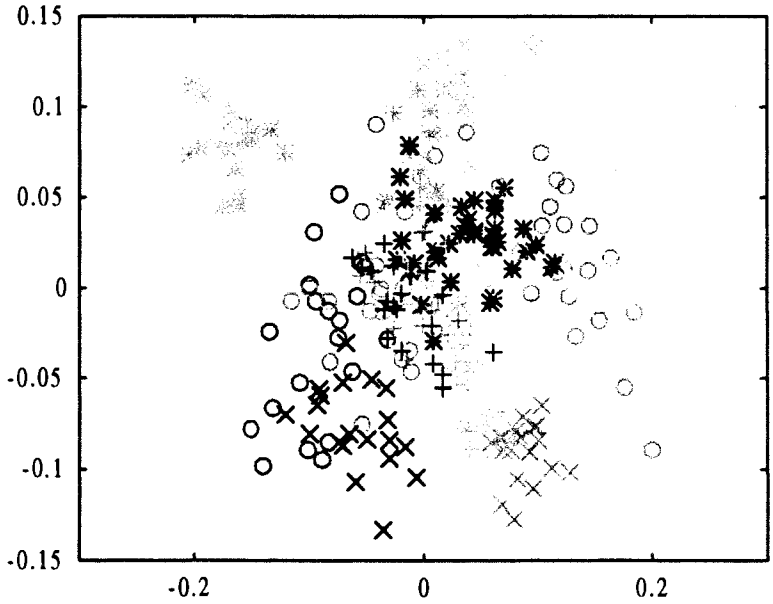


Figure 27 : Plan des observations selon la méthode de référence, projection binaire/3 moments de la distribution (D4)

L'évolution du taux de recouvrement, a fortiori du taux d'erreur de vérification, fournie par l'analyse en composantes principales montre que le descripteur D1 est le plus approprié à la description des signaux issus de la méthode de projection binaire proposée par R. Sabourin. Cependant, il reste à éclaircir la question du potentiel des descripteurs D2 à D4 couplés aux autres méthodes de codage.

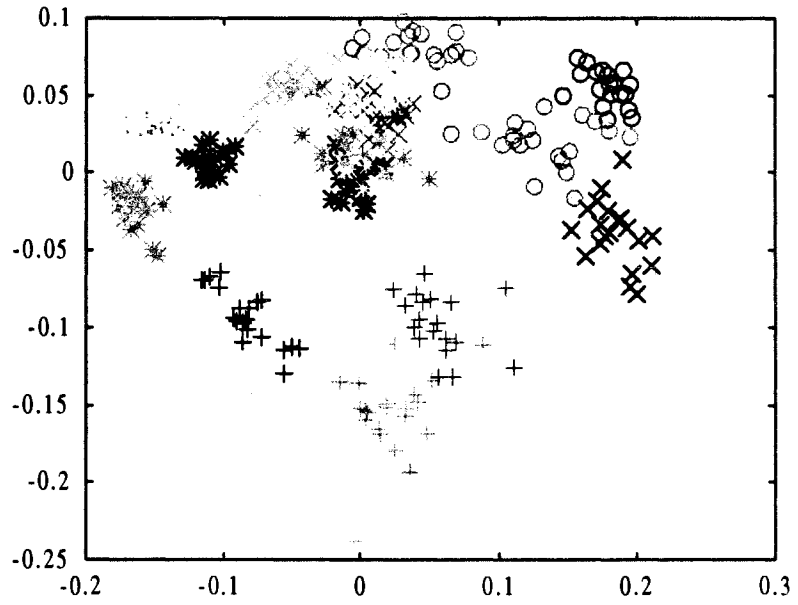


Figure 28 : Plan des observations selon la méthode de codage de la distance, technique de projection de R. Sabourin/D1

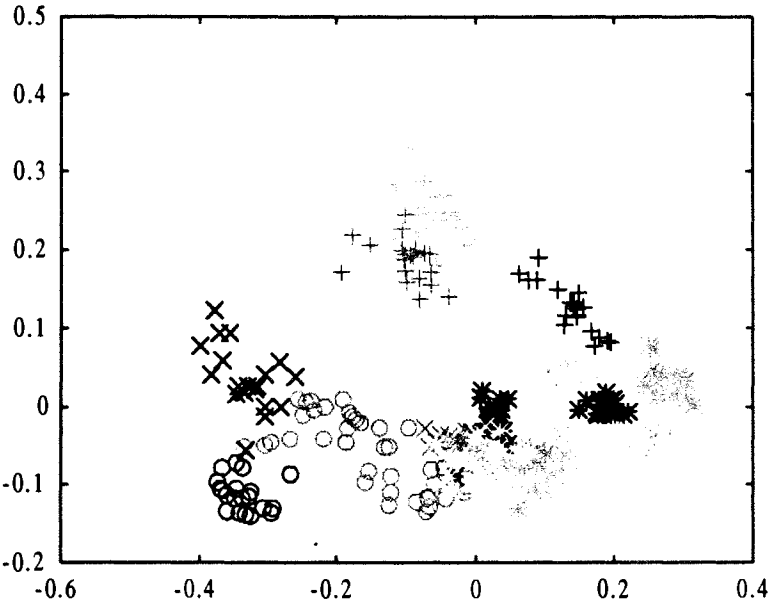


Figure 29 : Plan des observations selon la méthode de codage de la distance, technique de projection de R. Sabourin/D2

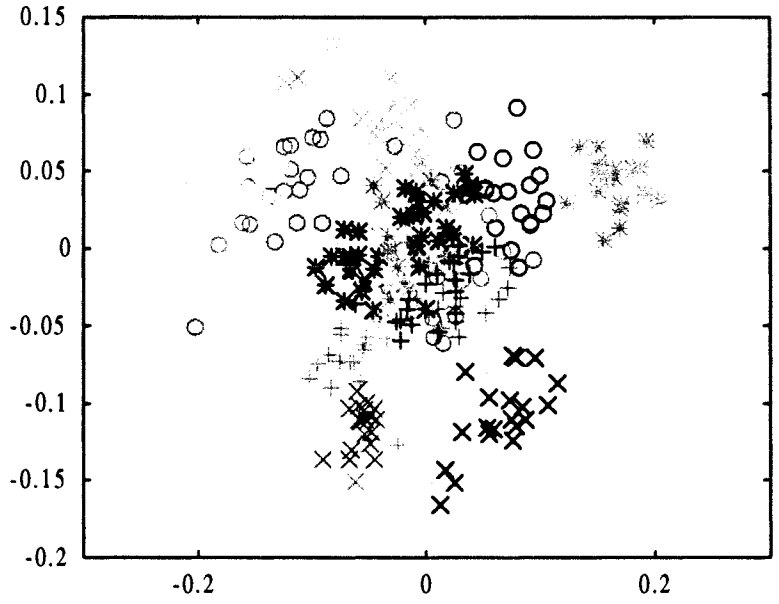


Figure 30 : Plan des observations selon la méthode de codage de la distance, technique de projection de R. Sabourin/D3

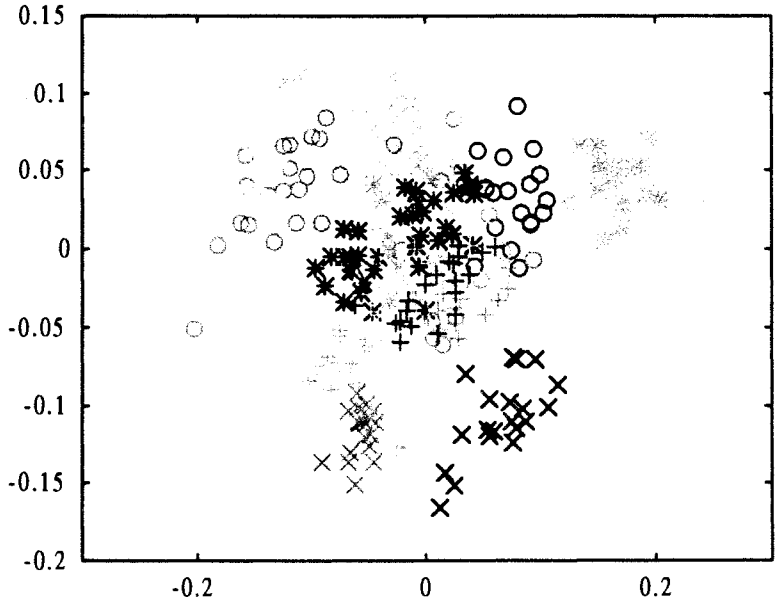


Figure 31 : Plan des observations selon la méthode de codage de la distance, technique de projection de R. Sabourin/D4

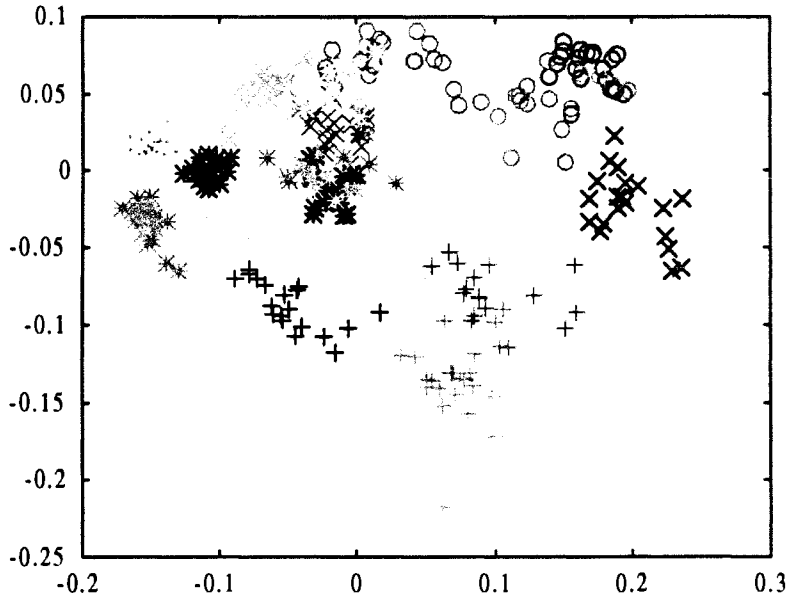


Figure 32 : Plan des observations selon la méthode de codage de la distance, nouveau mécanisme de projection aux plus proches bâtons/D1

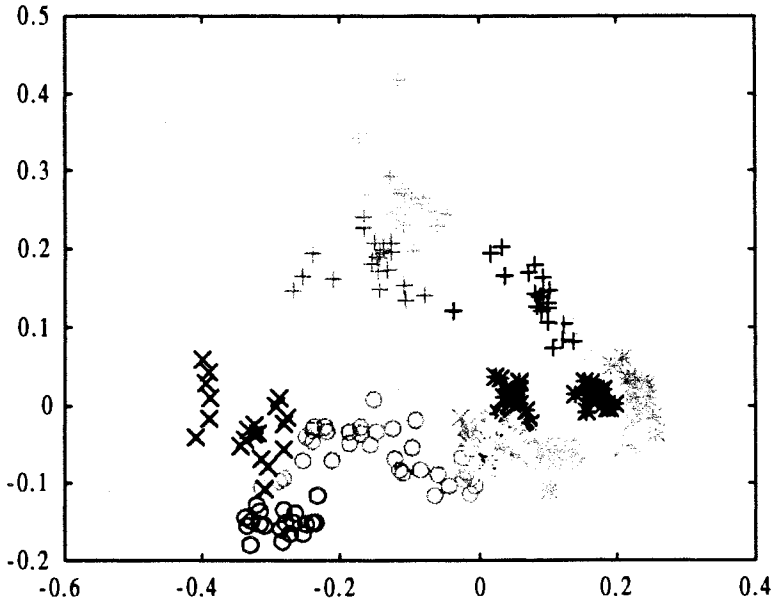


Figure 33 : Plan des observations selon la méthode de codage de la distance, nouveau mécanisme de projection aux plus proches bâtons /D2

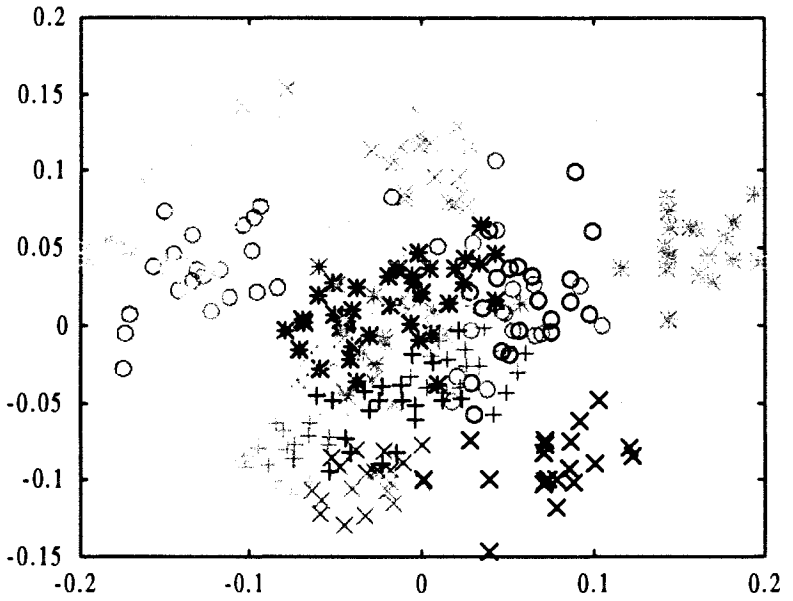


Figure 34 : Plan des observations selon la méthode de codage de la distance, nouveau mécanisme de projection aux plus proches bâtons /D3

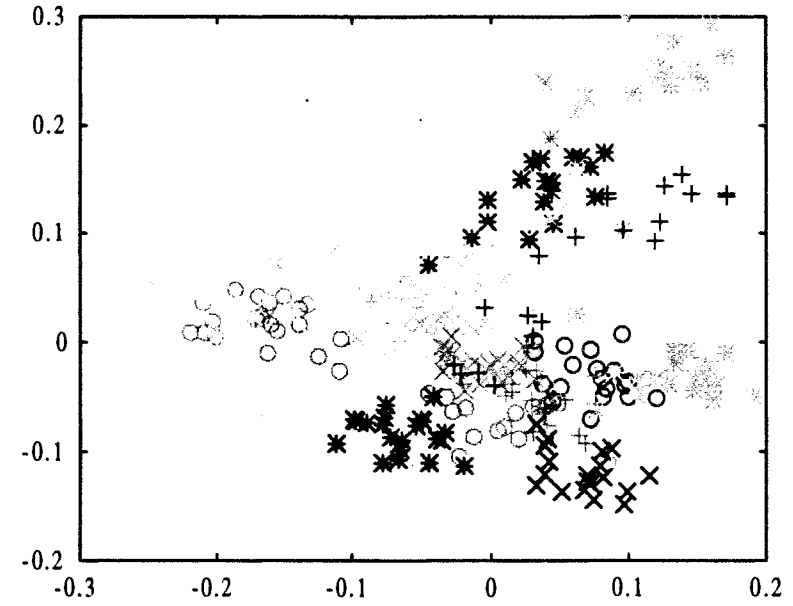


Figure 35 : Plan des observations selon la méthode de codage de la distance, nouveau mécanisme de projection aux plus proches bâtons /D4

A partir des Figures 28 à 35, il apparaît clairement que le choix du descripteur D1 convient mieux à toutes les méthodes que les autres descripteurs statistiques. Cette remarque tend à confirmer celle faite à partir de l'analyse des transformées de signaux où la moyenne du signal paraissait la plus caractéristique. Cependant, l'analyse en composantes principales n'est réalisée que pour le motif *a*, c'est-à-dire le plus grossier. Une analyse complémentaire doit être réalisée sur les autres motifs.

2.5.1. Compléments d'analyse pour la confirmation du choix du descripteur des bâtons

Pour évaluer la pertinence du facteur de formes constitué de la méthode de codage et du descripteur, nous avons repris le protocole de test donné au paragraphe 2.3.3.3 du chapitre 1, de façon à obtenir des résultats de classification directement comparables avec ceux de l'approche proposée par R. Sabourin. Pour poser les performances de référence, nous rappelons que l'évaluation de cette erreur dans le cas du descripteur D1, correspond aux résultats obtenus par R. Sabourin.

Conjointement à l'évaluation du couple codage/description, l'évaluation de la méthode de codage de R. Sabourin associée aux quatre descripteurs (D1, D2, D3, D4) permet de confirmer le choix du descripteur D1 fait à partir de l'analyse en composantes principales. En observant les résultats de l'ACP sur les différentes méthodes de codages des signatures, les résultats de l'étude complémentaire sur la pertinence du facteur de formes de R. Sabourin seront extrapolés aux autres méthodes pour le choix du descripteur.

Les résultats de l'évaluation du couple codage de R. Sabourin/ Descripteur D1 sont repris dans le Tableau 2 et situent le niveau de référence auquel on devra se reporter pour l'évaluation de nos apports :

K	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
1	2.360	1.230	0.595	0.423	0.198	0.125	0.190	0.153	0.098	0.130	1.530	0.480	0.170	0.015	0.023
3	2.190	1.173	0.555	0.523	0.273	0.085	0.235	0.173	0.118	0.073	1.560	0.690	0.215	0.033	0.025
5	2.175	1.095	0.785	0.973	0.360	0.185	0.313	0.220	0.180	0.163	1.590	1.405	0.283	0.063	0.070
7	2.055	1.105	0.995	1.033	0.495	0.308	0.405	0.345	0.240	0.223	1.618	1.850	0.395	0.115	0.083
9	2.208	1.338	1.290	1.480	0.805	0.648	0.685	0.535	0.530	0.585	1.903	2.175	0.603	0.330	0.248

Tableau 2 : Erreur totale en % avec k plus proches voisins pour le couple codage de R. Sabourin/D1

Le descripteur D1 fournit des informations globales à partir d'informations locales de la signature codée sur le bâton. Les vecteurs caractéristiques selon les 15 motifs donnent une transcription des objets dans des espaces de dimension 6 pour le motif le plus grossier et 276 pour la résolution la plus fine.

Les tableaux 3, 4 et 5 reprennent les évaluations des descripteurs D2 à D4 avec la méthode de codage de R. Sabourin.

K	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
1	2.436	2.660	1.219	2.467	4.696	3.448	2.577	2.089	2.103	3.312	8.772	4.054	1.655	0.715	0.256
3	2.249	4.601	1.216	2.059	5.724	5.073	3.758	2.470	2.308	3.577	10.277	4.431	1.819	1.134	0.906
5	2.393	4.559	1.925	2.090	6.989	6.477	4.907	2.827	3.327	4.232	11.088	4.627	2.038	1.580	1.120
7	2.486	4.600	3.076	2.294	7.464	7.618	5.992	3.438	4.208	4.570	11.782	5.196	2.513	2.056	1.562
9	2.532	5.984	3.257	2.519	7.809	8.193	7.464	3.943	5.061	5.315	12.731	5.751	2.941	2.694	2.212

Tableau 3 : Erreur totale en % avec k plus proches voisins pour le couple codage de R. Sabourin/D2

K	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
1	5.525	6.305	2.465	3.42	4.353	3.763	4.063	3.838	3.613	4.533	8.005	4	2.468	1.57	0.558
3	4.853	8.593	4.033	4.423	6.868	6.293	6.303	6.463	4.068	6.413	9.783	5.25	2.38	2.428	0.595
5	6.485	9.678	4.995	5.998	10.34	8.698	7.975	7.438	4.65	9.628	10.46	7.738	3.278	3.213	0.61
7	6.52	10.26	7.3	6.903	12.06	9.138	9.483	8.488	5.338	10.03	12.18	9.93	3.778	3.853	0.793
9	7.153	11.87	8.485	8.545	14.96	12.45	10.86	10.2	7.238	10.95	12.88	10.56	4.153	4.398	1.118

Tableau 4 : Erreur totale en % avec k plus proches voisins pour le couple codage de R. Sabourin/D3

K	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*
1	10.815	7.903	5.466	6.415	1.426	0.623	2.048	3.013	0.349	0.634	7.473	6.053	0.673
3	11.039	8.876	5.586	6.775	1.823	0.720	3.001	3.610	0.624	0.787	9.480	6.697	1.078
5	11.733	10.022	6.099	7.327	2.240	1.435	3.256	4.069	0.898	1.195	10.192	7.329	1.191
7	11.571	10.801	6.708	7.700	2.696	1.855	3.659	4.486	1.273	1.880	11.870	7.912	1.540
9	12.427	12.092	8.139	8.226	4.734	2.399	5.709	5.672	2.081	2.542	11.837	8.974	2.121

Tableau 5 : Erreur totale en % avec k plus proches voisins pour le couple codage de R. Sabourin/D4

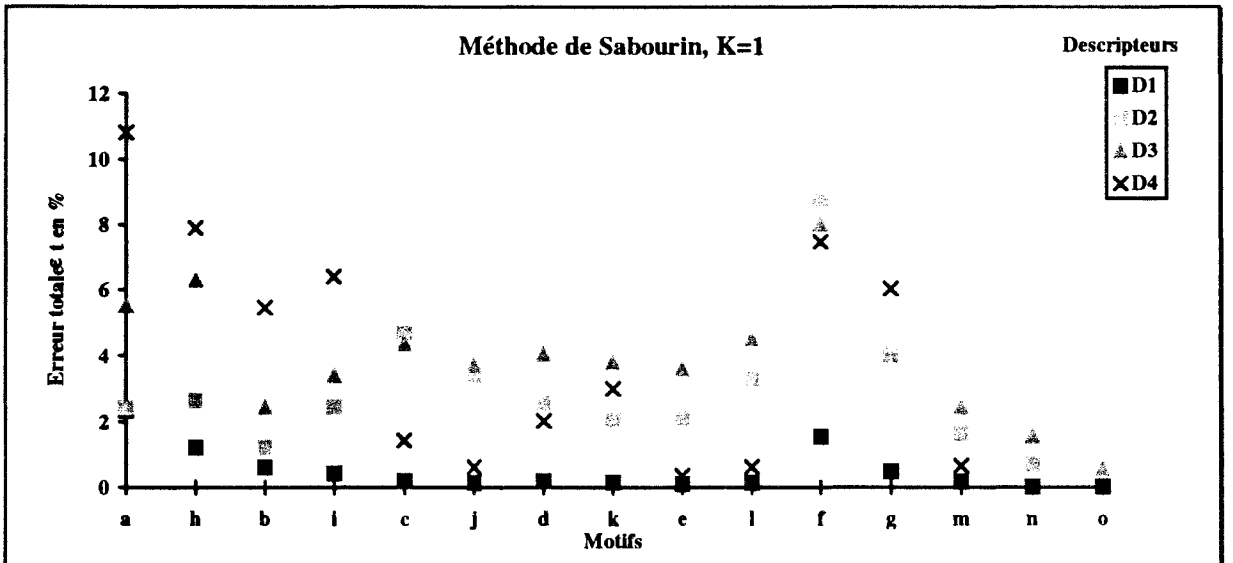


Figure 36 : Comparaison des performances des descripteurs de bâton pour tous les motifs

Par comparaison des performances obtenues pour les différents descripteurs, le descripteur D1 apparaît le mieux adapté. Le choix intuitif du descripteur D1 suscité par les résultats de l'analyse en composantes principales est confirmé par cette comparaison plus réaliste, puisque les essais sont réalisés sur toute la base de données selon le même protocole. Les erreurs plus importantes pour D3 et D4 viennent confirmer les résultats de l'ACP illustrés par un taux de recouvrement important (Figure 26, Figure 31).

Le descripteur D1 est donc globalement plus intéressant. L'utilisation du descripteur D2 n'amène pas une amélioration sensible dans les résolutions grossières, motifs *a*, *h*, *b*, *i*, et au contraire engendre des performances nettement moins intéressantes dans les autres motifs. Les descripteurs D3 et D4 ne sont pas appropriés à la description des signaux de codage des signatures. Les performances obtenues sont clairement en dessous de celles de référence.

Néanmoins, il est nécessaire de relativiser les performances obtenues dans les motifs de grande dimension. En effet, leur fiabilité est sujette à discussion quant au nombre de données utilisées pour l'évaluation par rapport à la dimension de l'espace caractéristique lié au motif. Le protocole de test définit l'évaluation à partir de 115 données et dans le cas de codages décrits par une seule caractéristique, l'espace est de 6 à 276 dimensions. Le problème de « curse of dimensionnality » s'exprime alors clairement dès que l'on dépasse une résolution de l'ordre de 12 [Bellman 58].

Etant données les performances quantifiées, mais également le coût de calcul, le descripteur D1 semble le plus approprié vis-à-vis du codage des projections de signature. Dans ce contexte, les évaluations des modifications de la méthode initiale de R. Sabourin qui seront entreprises seront réalisées avec ce descripteur.

2.6. *Quantification de l'apport du codage de la distance*

Afin d'obtenir un codage plus représentatif de la forme réelle de la signature, nous avons introduit la notion de distance entre les points de la signature et les bâtons du motif. L'évaluation de la pertinence du codage de la distance décrit par le descripteur D1 est réalisée selon le protocole prédéfini. La méthode de scrutation de R. Sabourin est conservée, ce qui permet une bonne évaluation de l'apport du codage de la distance indépendamment de la scrutation. Les performances obtenues sont regroupées dans le Tableau 6 :

K	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
1	1.068	0.838	0.618	0.830	0.473	0.178	0.968	0.708	0.075	0.213	0.770	0.928	0.265	0.098	0.150
3	1.030	1.133	0.905	1.255	0.630	0.318	0.825	0.775	0.185	0.178	1.135	1.425	0.385	0.118	0.390
5	1.103	1.298	1.200	1.950	0.745	0.385	0.970	0.890	0.368	0.228	1.318	2.168	0.503	0.225	0.645
7	1.258	1.185	1.468	2.503	0.820	0.615	1.208	1.115	0.545	0.328	1.368	2.638	0.623	0.338	0.888
9	1.420	1.460	1.718	3.195	1.170	0.815	1.398	1.280	0.853	0.603	1.398	3.050	0.898	0.698	1.113

Tableau 6 : Erreur totale en % avec k plus proches voisins pour le couple codage de la distance/D1

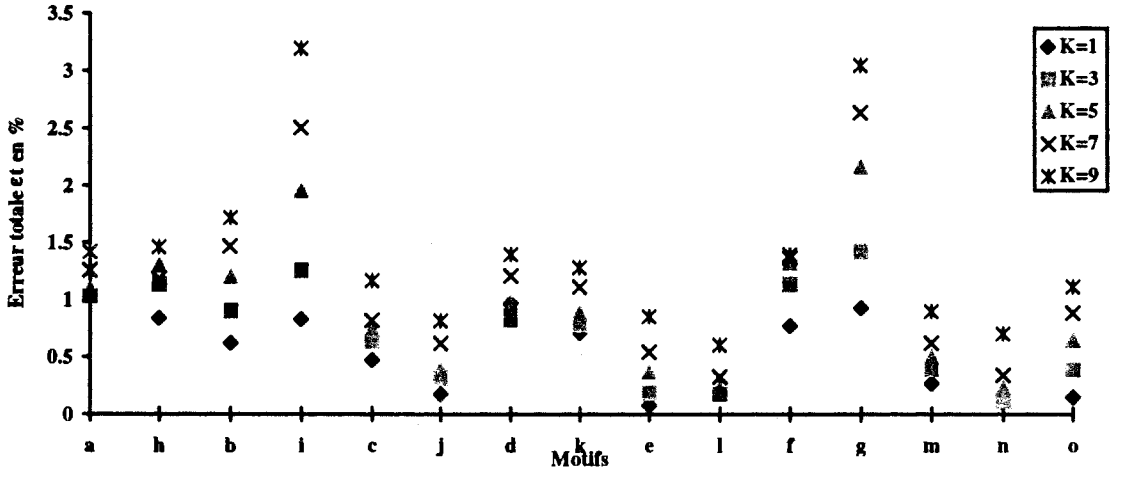


Figure 37 : Performances du codage de la distance selon le mécanisme de projection de R. Sabourin/D1

Les résultats illustrent l'importance de l'introduction du codage de la distance au niveau des performances de vérification. On note une très forte amélioration des performances de ce facteur de formes pour les motifs grossiers *a*, *h*, *f*. En effet, la distance est plus importante quand les motifs sont de faible résolution. A forte résolution, la distance est sensible au pas de quantification et quasiment toute l'information est portée par la résolution et la forme du motif. La résolution devient l'information la plus pertinente. La

Figure 38 présente des résultats comparatifs de la méthode utilisant le codage de la distance avec la méthode de référence de R. Sabourin (ESC) selon la résolution des motifs.

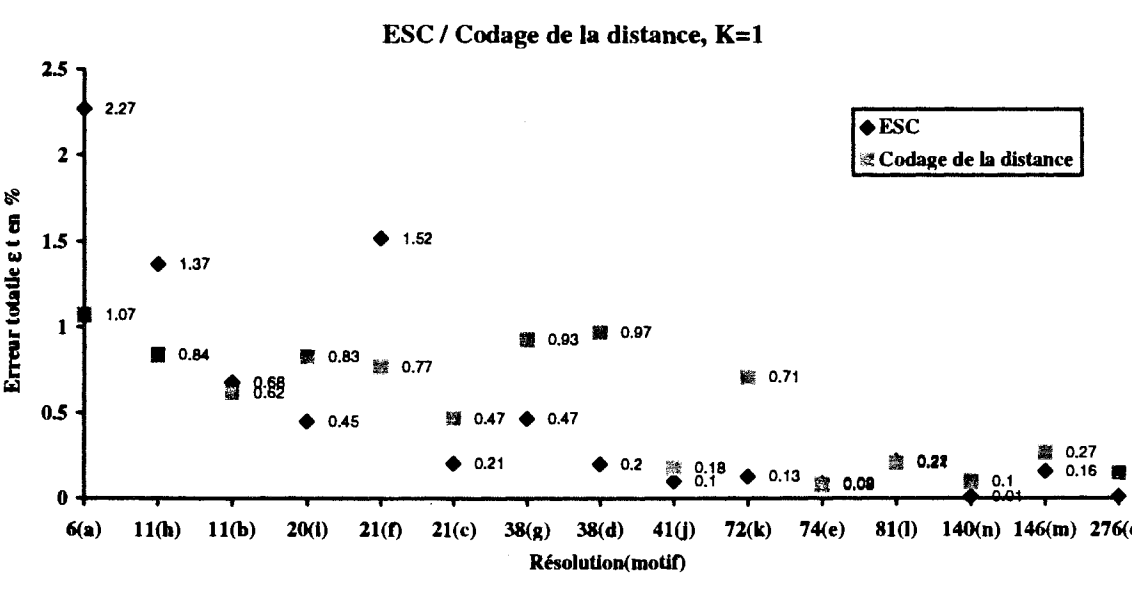


Figure 38 : Performances comparées selon la résolution des motifs

Dans le cas du motif *a*, la hauteur du cadre de la signature est de 128 pixels. Une modification de 5% en position, que l'on peut considérer raisonnable pour une signature, représente environ 6 pixels. Pour le motif *o*, cela représente environ 20% de la hauteur d'un

motif. Si on refait la même analyse pour la longueur, la longueur du cadre est de 512 pixels, 5% donnent 25 pixels, ce qui fait environ 80% de la longueur d'un motif. On comprend donc le problème de représentativité de l'information de distance dans les motifs fins.

La forme des motifs intervient donc beaucoup sur la pertinence du facteur de formes (

Figure 38), puisque les proportions hauteur/largeur conditionnent aussi les diagonales des motifs. On peut se poser la question de l'intérêt de la forme des motifs vis-à-vis des signatures. Les signatures nord-américaines sont généralement sur l'axe horizontal et d'une hauteur assez stable. L'information portée par les bâtons verticaux et diagonaux est redondante lorsque les motifs sont allongés dans l'axe vertical (*e.g.* motif *l*). En revanche, si l'on doit traiter des signatures européennes (paraphes), la structure entière des motifs sera nécessaire du fait de l'inclinaison de ces paraphes.

L'analyse des résultats permet de considérer que le codage de la distance est intéressant puisqu'il améliore les performances pour les motifs de faible résolution. En effet, les diverses analyses des simulations présentent toujours un gain de performance avec le codage de la distance par rapport à l'ESC à descripteur équivalent et pour les motifs grossiers. L'amélioration des performances est moins nette pour les motifs à haute résolution.

On peut également se demander de quel type de bâtons vient l'accroissement des performances (horizontaux, verticaux et diagonaux). Sur ce point, une analyse sur le pouvoir discriminant des types de bâtons peut révéler l'origine directionnelle de l'apport. Sur les signatures nord-américaines, on peut supposer que le plus gros apport vient des bâtons horizontaux car c'est dans l'axe vertical que l'on a le plus d'informations significatives.

On constate de façon générale que le niveau de résolution induit le taux de performances. Plus la résolution est fine, plus les performances sont fortes. Mais à résolution équivalente, un échantillonnage plus fin des bâtons horizontaux tend à donner de meilleurs résultats. Ce dernier constat est valable jusqu'à la résolution 81 (motif *l*) car au-delà, les comparaisons sont difficiles et incertaines, la résolution prend le pas sur la forme.

L'utilisation de trois moments (moyenne, variance, dissymétrie) comme descripteur statistique du signal, n'apporte rien de significatif. On retrouve même une diminution des performances suite au problème de "curse of dimensionality". Avec un descripteur de cette taille, l'espace métrique des paramètres dans lequel on place la discrimination va de $\Re^{18}$ (motif *a*) à $\Re^{828}$ (motif *o*), alors que le nombre de données est de 115. On sait, de plus, que l'augmentation de la taille du descripteur n'entraîne pas obligatoirement celle des performances; c'est d'ailleurs un des problèmes de la reconnaissance de formes.

Il faut également juger les performances selon un autre critère, c'est-à-dire comparer l'apport des méthodes en relation avec la résolution des motifs et leur orientation. Il faut noter que la comparaison des performances sur des motifs fins est peu fiable du fait du faible nombre de données par rapport à la résolution (cf. [Milgram 93], qui préconise un rapport supérieur à 10 entre données et résolution).

2.7. Quantification pour la méthode de codage de la distance avec modification de la scrutation

La modification de la technique de scrutation vise à adapter la technique de l'Extended Shadow Code aux signatures. L'apport est négligeable pour les motifs dont la diagonale possède un coefficient directeur différent de 0.5, mais important pour les motifs carrés (*e.g.* c, e, o). Les performances des simulations concernant cette méthode de codage de distance associée à une modification de la technique de scrutation sont présentées dans le Tableau 7 :

K	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
1	1.133	0.410	0.628	0.635	0.133	0.155	0.275	0.178	0.083	0.200	0.835	0.780	0.260	0.025	0.035
3	1.153	0.615	0.720	0.938	0.170	0.300	0.325	0.315	0.218	0.215	1.163	1.453	0.393	0.068	0.070
5	1.155	0.598	0.920	1.495	0.330	0.365	0.475	0.373	0.378	0.245	1.380	2.158	0.425	0.115	0.113
7	1.313	0.883	1.203	1.873	0.455	0.545	0.673	0.540	0.508	0.370	1.530	2.925	0.540	0.233	0.223
9	1.638	1.160	1.648	2.300	0.748	0.740	1.020	0.940	0.873	0.663	1.810	3.258	0.958	0.595	0.488

Tableau 7 : Erreur totale ε_t, en %, obtenue par modification de la scrutation et codage de la distance pour les différents motifs, par la règle des k plus proches voisins

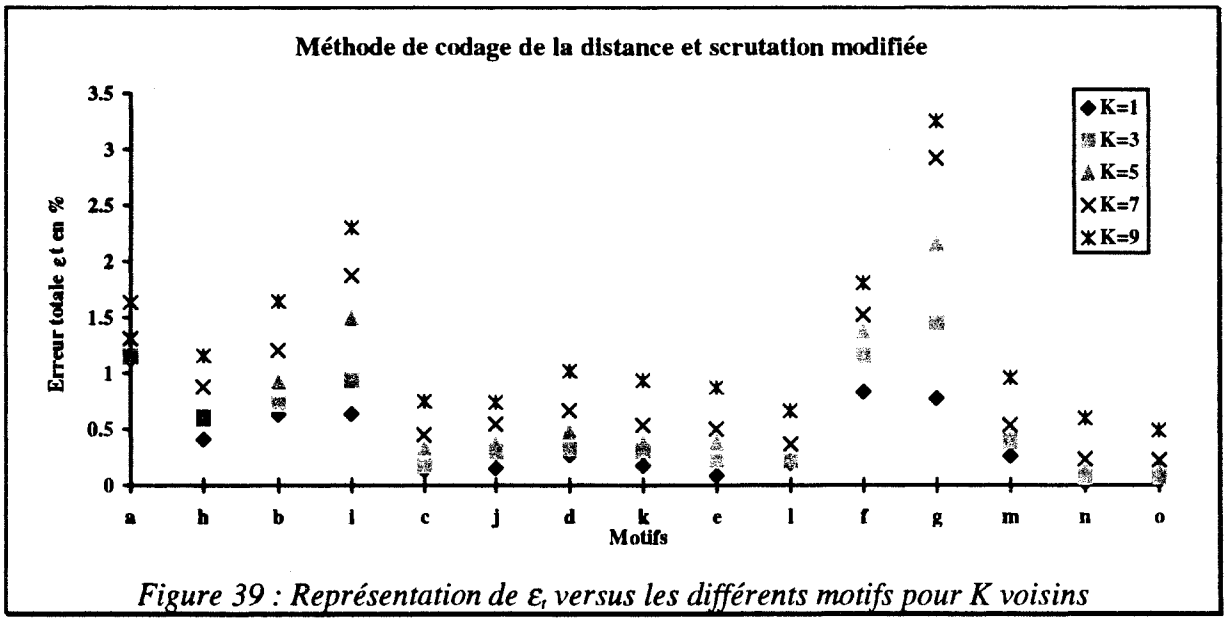


Figure 39 : Représentation de ε_t versus les différents motifs pour K voisins

Comparativement à la méthode proposée par R. Sabourin (Figure 36), l'erreur totale est diminuée pour les motifs grossiers comme *a*, *h*, *c*, *f*, puis relativement égale à celle obtenue avec les motifs *b*, *c*, *j*, *d*, *k*, *e* et *l*. On peut donc dire que la méthode de scrutation associée au codage de la distance semble être prometteuse à basse résolution. De plus, par rapport à la méthode de codage de la distance sans modification de la scrutation (Figure 37), quasiment toutes les performances sont améliorées, à l'exception de celles liées aux motifs à résolution fine. Là encore, les performances sont à relativiser par rapport à la résolution (Figure 40). A ce niveau de finesse, les performances sont dues à la résolution et la réponse que nous avons proposée pour le phénomène d'aliasage ne se ressent plus car le point rajouté porte la même information que ses voisins directs.

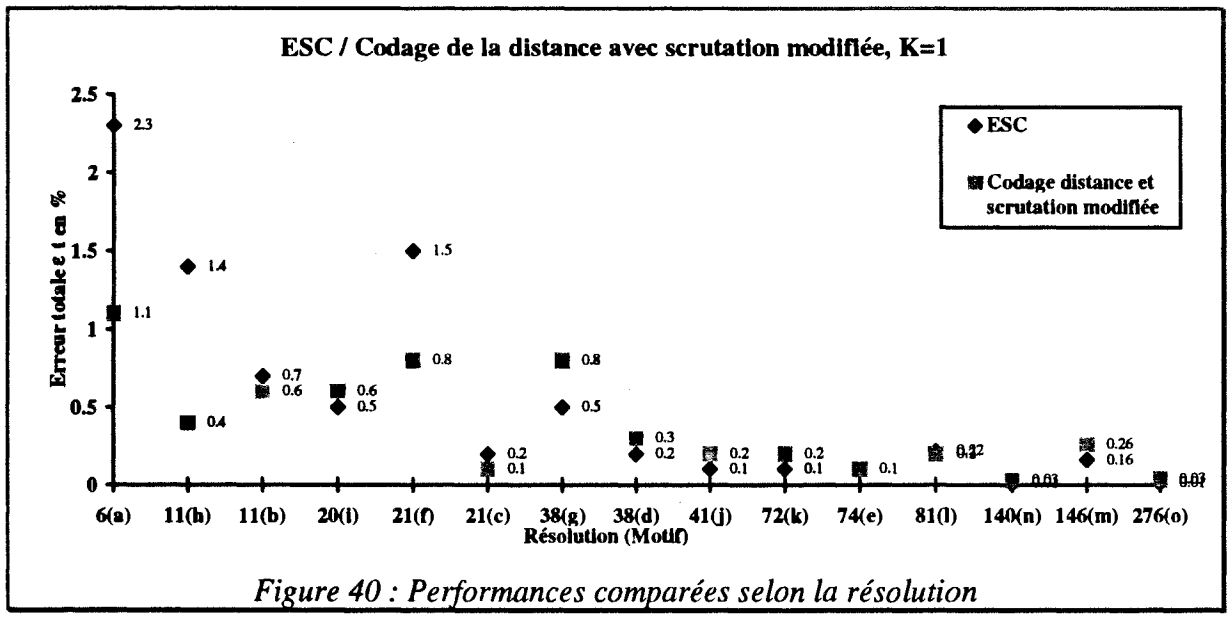


Figure 40 : Performances comparées selon la résolution

2.8. *Conclusion*

En conclusion, on montre que la méthode de scrutation proposée améliore les performances, notamment pour les motifs assez grossiers. A forte résolution, l'apport est moins sensible. On suppose donc, par déduction, que la scrutation est à l'origine de ces améliorations. En comparant les résultats obtenus avec ceux des méthodes ESC et ESC avec codage de la distance, on constate un certain intérêt de la modification du mode de scrutation dans les motifs grossiers. En revanche, le descripteur de bâton semble absorber l'amélioration dans la mesure où moins de motifs sont concernés.

L'apport général de la transformation de la méthode de l'Extended Shadow Code pour l'approche statique du problème est globalement concluante. L'apport est très intéressant dans les motifs grossiers, mais la notion de présence est plus significative dans les motifs fins. Cependant, on peut se fixer un autre critère de pertinence du facteur de formes, celui de la constance de son pouvoir discriminant. En effet, sur notre banque de données, notre apport permet de lisser les performances des différentes résolutions. Les différences sont moins importantes entre les pouvoirs discriminants des motifs, conduisant à une plus grande homogénéité des niveaux de performances des réponses des classificateurs, ce qui semble être très important dans le cadre d'une coopération de décision.

3. Un facteur de formes pour une approche pseudo-dynamique de la vérification

La technique de l'Extended Shadow Code, associée au codage de la distance, a permis de mieux répondre au problème de vérification de signatures dans le cas des faux aléatoires, et par extrapolation dans le cas des faux simples.

Cependant, l'objectif général de la vérification de signatures ne se limite pas au traitement des faux aléatoires, mais demande le traitement de faux plus complexes comme les faux par transfert optique ou les faux par imitation libre. Une "spécialisation" de l'Extended Shadow Code pour le traitement de ces faux, impliquerait une duplication du facteur de formes dans la mesure où la nature du faux n'est pas connue a priori. Aussi, il apparaît plus judicieux d'élaborer un seul facteur de formes pertinent vis-à-vis des différents types de faux. Pour cela, il est nécessaire d'introduire dans la technique de l'Extended Shadow Code des informations discriminantes révélatrices de ces types de faux.

3.1. Introduction du niveau de gris dans l'Extended Shadow Code avec codage de la distance

Locard [Locard 36] a montré que l'aspect dynamique de la signature peut être observé sur une image en niveaux de gris. En particulier, les traits clairs et foncés dénotent un déplacement plus ou moins rapide de la plume sur le papier. Le trait foncé caractérise un déplacement lent, mais surtout à vitesse constante. Le trait clair caractérise un déplacement rapide. Toutes ces caractéristiques, montrant l'aspect dynamique de la signature, sont propres à chaque scripteur et sont très difficiles à imiter. Elles présentent donc un intérêt discriminatoire certain.

La dynamique d'une signature peut donc être accessible, dans un traitement hors-ligne de l'image à partir des niveaux de gris. L'hypothèse de Locard fait abstraction de la pression du scripteur, alors qu'Ammar [Ammar 86] a étudié et utilisé l'influence des variations de pression sur les valeurs de niveaux de gris. Aussi, le niveau de gris traduit en même temps une combinaison de la vitesse et de la pression lors de l'acte d'écriture et cette combinaison est une caractéristique propre au scripteur. Le niveau de gris est une information composée qu'on qualifiera de pseudo-dynamique. La combinaison est relative et suppose certaines variations qui dépendent des conditions de réalisation. Le facteur de formes doit donc être assez souple, pour relativiser les variations de la combinaison entre différentes signatures du même scripteur, sans être trop tolérant.

Le facteur de formes que nous allons définir conserve la structure proposée précédemment, c'est-à-dire l'Extended Shadow Code avec la modification du mécanisme de projection au plus proche bâton, ainsi que la nouvelle technique de scrutation. Ce choix vient du fait que cette amélioration offre la possibilité de coder plusieurs caractéristiques dans l'objectif de discrimination des faux par transfert optique. De plus, elle a montré une certaine efficacité par rapport au problème de la vérification des faux aléatoires.

Notre objectif est de traiter les faux par transfert en même temps que les faux aléatoires. Aussi, les deux informations discriminantes, statiques et pseudo-dynamiques accessibles à partir des mêmes points du tracé, doivent être codées par le facteur de formes.

3.2. Codage simultané de la pseudo-dynamique et de la statique

Nous avons vu dans la première partie de ce chapitre, que le codage de la distance séparant les points de la signature des bâtons pouvait être facilement réalisé. Implicitement, nous codons, par la distance, la présence ou l'absence de points à projeter. Ici, le problème est plus difficile, puisque les deux informations à coder sont d'égale importance.

Le codage des niveaux de gris des points de la signature ne peut être fait sans considération des facteurs qui l'influencent. En effet, puisque le niveau de gris est l'image d'une combinaison de vitesse et de pression de réalisation de la signature, il nous semble judicieux d'étudier cette combinaison. Classiquement, une signature est formée de trois composantes globales : les posés, le corps et les levés [Gupta 79], [Sabourin 90].

Les posés sont les zones de la signature où le scripteur a posé son stylo, soit pour débuter son geste, soit pour une reprise. Ces zones de la signature sont caractérisées par une faible vitesse et une forte pression. Après le posé, la vitesse d'écriture augmente tandis que la pression diminue. Les posés correspondent donc à des portions de la signature où le niveau de gris est le plus foncé.

A l'opposé, les levés sont les zones de la signature où le scripteur s'apprête à stopper son geste, soit pour terminer sa signature, soit pour la reprendre à une autre position. Les levés sont caractérisés par une vitesse d'écriture importante et qui croît, ainsi que par une pression plutôt faible qui diminue. Ces portions de l'image de la signature possèdent des niveaux de gris plus clairs que le reste de la signature.

Notons que les levés sont moins constants, dans leur position relative dans la signature, que les posés. L'amplitude du mouvement du scripteur peut être limitée par une mauvaise position d'apposition ou par une contrainte spatiale, limitant ainsi les levés. En revanche, les posés sont des points d'ancrage, par rapport auxquels la signature se développe.

Le corps de la signature est formé par les parties de la signature où la vitesse et la pression se situent entre les extremums des posés et des levés. De plus, dans le corps, la vitesse et la pression varient localement. Ces variations sont caractéristiques de l'ombrage constitué des pleins et des déliés de la signature. Les déliés sont caractérisés par un trait fin, peu appuyé, de luminance forte, donc proche d'un levé de plume. En revanche, un plein est caractérisé par un trait appuyé qui le rapproche, en terme de luminance, d'un posé [Gupta 79]. Les niveaux de gris du corps de la signature se situent alors entre ceux des posés et des levés.

Par ailleurs, dans l'aspect dynamique du geste, le corps est une succession continue de pleins, de déliés et d'un niveau moyen de vitesse et de pression du corps, et donc une alternance entre des niveaux de gris faibles et forts. Au cours de la réalisation, il y a passage d'un plein vers un délié en passant par le corps de manière continue et graduelle. Il semble donc difficile de distinguer précisement les constituants du corps selon une caractérisation en niveaux de gris, pour former des caractérisations spécifiques des ombrages. Ainsi, on peut

considérer que le corps est formé des points de niveau de gris moyen, et que les passages vers des niveaux de gris extrêmes sont les pleins et les déliés.

L'analyse que nous venons de faire sur les niveaux de gris des signatures est globale et purement qualitative. Nous nous orientons donc vers une caractérisation qualitative de l'échelle des niveaux de gris plutôt que vers une quantification. L'échelle des niveaux de gris peut être caractérisée par trois symboles, Posé, Corps et Levé, relatifs aux différentes composantes de la signature.

Au cours de l'acte d'écriture, les conditions de réalisation du posé influent sur la suite de la gestuelle. Le point d'accroche implique une attention particulière de la part du scripteur, et revêt une grande importance pour l'acte dynamique et géométrique. Le corps reflète un acte continu, plus rapide et moins dirigé que le posé. Le scripteur libère son attention pour accompagner le stylo dans les déplacements spatiaux et dynamiques. Etant moins contrôlés, ils sont moins stables entre plusieurs signatures. Ils ont donc une importance moins grande que le posé. Enfin, le levé de plume, qui termine la signature au niveau duquel on retrouve parfois des phénomènes de dégénérescence, est une partie peu contrôlée par le scripteur. La dynamique et la géométrie de cette dernière phase de l'acte est peu fiable, car moins contrôlée et moins stable entre signatures authentiques.

A partir de cette description, un classement d'importance des phases d'exécution de l'acte peut être donné. Le posé se retrouve moins souvent, mais apparait crucial pour l'acte. Le corps représente la majeure partie de la géométrie, mais semble moins fiable dynamiquement, il est donc légérement moins important que le posé. Enfin, le levé de plume est d'une faible fiabilité géométrique et dynamique, il est donc l'événement le moins important.

En ce qui concerne la distance séparant les points de la signature des bâtons du motif, nous pourrions faire une analyse du même type si nous cherchons à relier la distance aux niveaux de gris. En effet, chaque partie de l'acte a une incidence spatiale et est spécifiquement caractérisée géométriquement. Ainsi, plus le tracé est proche d'un bâton, plus il ressemble au bâton, et plus il l'active, et inversement s'il s'en éloigne. Le principe est le codage de la proximité de façon à caractériser la signature par l'ensemble géométrique formé des bâtons.

Ainsi, une signature authentique est la réalisation d'un geste formé des événements spatio-temporels de posé, de corps et de levé, dont l'importance décroît au cours de sa réalisation, aussi bien géométriquement que dynamiquement. Pour caractériser une propriété spatio-temporelle de groupes de signatures authentiques, il faut donc relier les caractérisations de la géométrie et de la dynamique.

Parallèlement, nous avons vu au cours des paragraphes précédents, que ces caractérisations ne peuvent être que globales non-disjointes. De ce fait, le codage sur les bâtons résulte plus d'une relation logique des caractérisations que d'une combinaison analytique. La définition de cette relation, pour la réalisation du codage, peut se déduire linguistiquement des analyses de l'importance des événements rencontrés au cours de l'acte d'écriture. Ces dernières remarques nous ont conduits à utiliser une formulation d'inférence floue pour définir le codage.

3.3. Définition d'un opérateur flou de codage de signatures

Une caractérisation linguistique générale des informations pertinentes, ainsi que de leur relation, nous amène à proposer l'utilisation d'un outil de la théorie des ensembles flous pour réaliser le codage des signatures. Pour cela, nous retenons le principe le plus simple et le plus connu, ayant généré de nombreux développements en commande : le modèle de Mamdani [Mamdani 75]. La particularité de ce modèle est d'intégrer une formulation linguistique, sur la base des caractérisations floues des informations d'entrée et de sortie. Ce modèle convient donc parfaitement à la traduction de l'analyse globale que nous avons précédement réalisée.

Le principe du modèle flou de Mamdani repose sur trois étapes. Les variables d'entrée et de sortie doivent être caractérisées par des ensembles flous, c'est-à-dire par la définition de termes linguistiques. Une première étape consiste en une conversion des valeurs numériques d'entrée dans l'univers flou décrit par les termes linguistiques, c'est ce que l'on nomme la fuzzification. Une deuxième étape de conjuguaison de ces descriptions floues, l'inférence floue, relie les combinaisons des termes d'entrée avec les termes de sortie, selon des règles d'inférence définies. Enfin, les entrées floues inférées activent les termes de sortie, et les termes activés sont transcrits en une valeur numérique de sortie, c'est la défuzzification.

3.3.1. Fuzzification

La fuzzification, nécessaire au fonctionnement d'un modèle en logique floue, est un opérateur de symbolisation et de caractérisation sous forme d'ensembles flous définissant des états globaux différents, mais non-disjoints, d'une variable.

Cette caractérisation doit, si possible, s'appuyer sur un fondement sémantique. Dans notre cas, l'univers de discours est, pour la pseudo-dynamique, une combinaison de vitesse et de pression. Nous avons vu que ces combinaisons sont associées à des événements particuliers au cours de l'acte, et s'expriment par les niveaux de gris. Rappelons que 3 termes principaux caractérisent la pseudo-dynamique de la signature : « Posé », « Corps » et « Levé ».

La fuzzification des niveaux de gris de la signature, donc de la pseudo-dynamique, est posée à partir de l'histogramme caractéristique des images de signatures manuscrites. Cependant, l'histogramme exprime également les niveaux de gris du fond alors que la fuzzification doit caractériser uniquement les niveaux de gris de la signature. Aussi, pour accéder à la plage de variation des niveaux de gris de la signature et définir correctement les termes de la fuzzification, le seuil d'Otsu est conservé pour réaliser la séparation entre le fond de la scène et la signature.

Le seuil d'Otsu nous donne une borne supérieure des niveaux de gris de la signature, et l'analyse de l'histogramme nous permet de définir la valeur inférieure de cette plage. On pose alors les trois termes de fuzzification caractérisant le niveau de gris (Figure 41).

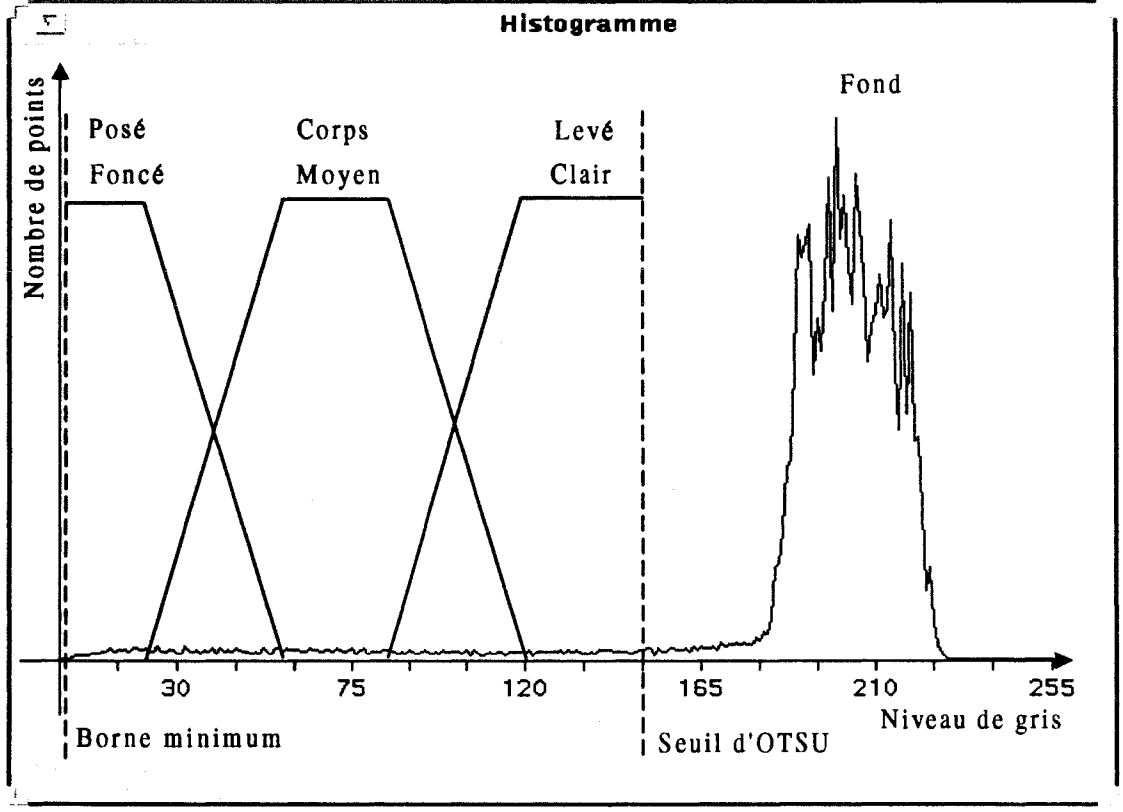


Figure 41 : Fuzzification de la pseudo-dynamique

Les termes linguistiques sont pris de forme trapézoïdale. Le choix de cette forme est lié à la caractérisation imprécise des termes. Les termes de « Posé » et « Levé » sont pris trapézoïdaux pour signifier une caractérisation plus sûre dans les parties extrêmes de la plage des niveaux de gris de la signature. Le terme « Corps » est choisi trapézoidal pour tolérer les variations autour du niveau de gris moyen, correspondant à l'ombrage, et également pour assurer la condition d'orthogonalité des degrés d'appartenance.

Parallèlement à la fuzzification de la pseudo-dynamique, la géométrie doit être fuzzifiée. La notion de proximité aux bâtons de l'Extended Shadow Code est caractérisée par 3 termes triangulaires « Proche », « Moyen » et « Loin ». Ces 3 termes sont choisis de forme triangulaire et équirépartis, afin d'obtenir une définition raisonnable des niveaux de proximité. Les termes extrêmes « Loin » et « Proche » sont nécessaires et qualifient sémantiquement la proximité. Le terme intermédiaire assure la condition d'orthogonalité et permet une bonne description de la distance (Figure 42).

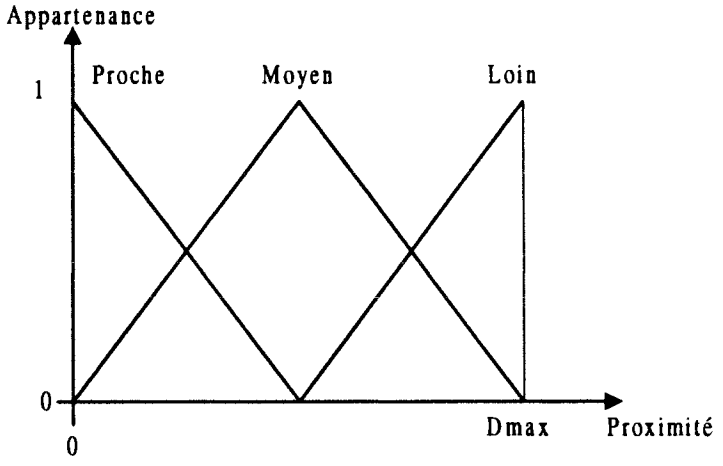


Figure 42 : Fuzzification de la distance

Enfin, la caractérisation floue de l'univers de sortie correspond au niveau de codage sur le bâton. N'ayant pas une sémantique particulière à rattacher à cette caractérisation, nous avons pris 3 termes correspondant à un codage « Petit », « Moyen » et « Grand ». Les termes sont choisis équirépartis et de formes triangulaires de façon similaire à la Figure 42.

3.3.2. Règles d'inférence

L'établissement des règles d'inférence est fait de façon intuitive. Le critère que nous avons utilisé est basé sur l'importance des événements dynamiques au cours de l'acte d'écriture, tout en tenant compte de la distance, qui doit rester caractéristique de la géométrie. Notons que ce critère peut être modifié, suivant l'interprétation souhaitée dans la combinaison de ces deux notions. Ainsi, le « Posé » « Proche » est une information considérée très importante, comme le « Corps » « Proche ». En revanche, le « Levé » « Proche » est pris moins important de par sa faible répétabilité. Parallèlement, le corps est essentiellement caractérisé par la proximité afin de conserver le potentiel de discrimination des faux aléatoires en traduisant la géométrie de la signature. Par ailleurs, le « Posé » « Loin » et le « Posé » « Moyen » sont plus importants qu'un « Corps » « Loin » ou qu'un « Corps » « Moyen ». Finalement, le « Levé » est considéré moins important que les autres notions, à proximité équivalente.

La définition linguistique des règles que nous venons de poser peut alors se traduire sous une forme d'implication « Si... Alors.. ».

- Si la proximité est « Loin » et la pseudo-dynamique est « Levé » alors le codage est « Petit ».

- Si la proximité est « Proche » et la pseudo-dynamique est « Posé » alors le codage est « Grand ».

Les autres combinaisons sont dérivées de ces cas extrêmes, et la dynamique est placée comme modificateur de la distance. Un point « Proche » est globalement important. Cependant, si son niveau de gris devient clair, il est moins certain qu'il soit aussi important pour la signature,

alors son intérêt baisse, ce qui implique que le codage doit baisser. Sur cette base, la table d'inférence proposée est présentée Figure 43 :

Pseudo-dynamique \ Distance relative	Loin	Moyen	Proche
Levé	Petit	Petit	Moyen
Corps	Petit	Moyen	Grand
Posé	Moyen	Moyen	Grand

Codage

Figure 43 : Table d'inférence pour la fusion de la distance et du niveau de gris

Cette table de règle est une proposition, on peut penser que d'autres tables sont possibles selon les caractères que l'on veut renforcer dans le codage. Les opérateurs de combinaison/projection [Foulloy 92] sont le min. et le max., le produit des entrées est effectué par l'opérateur min. Le lecteur trouvera dans [Foulloy 92], [Bouchon-Meunier 94] et [Benoit 93], des compléments d'information sur l'inférence floue, que nous n'avons pas jugé nécessaire de développer ici.

3.3.3. Défuzzification

Dans la mesure où plusieurs règles peuvent agir dans un modèle flou, ce qui permet d'assouplir le comportement réel vis-à-vis de celui que l'on veut donner, plusieurs termes de sortie peuvent être activés. Or, le codage sur le bâton est purement numérique et demande une valeur unique. L'opérateur de défuzzification force alors une agrégation en une seule valeur numérique de sortie. Afin de tenir compte des diverses propositions de la table d'inférence et observer un certain consensus sur la sortie, l'opérateur de défuzzification retenu est celui du centre de gravité (Figure 44).

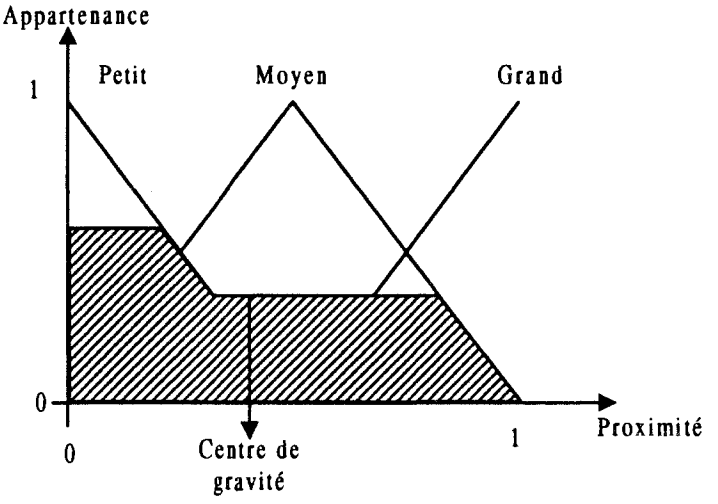


Figure 44 : Défuzzification du codage par le centre de gravité

3.4. Pertinence du FESC pour la discrimination des faux aléatoires

De la même façon que pour l'Extended Shadow Code, la performance est évaluée selon le protocole d'évaluation présenté au paragraphe 2.3.3.3 du chapitre 1. Le Fuzzy Extended Shadow Code est évalué dans le cas des faux aléatoires. Un de nos objectifs est d'approcher les performances de l'Extended Shadow Code avec ce facteur de formes, pour le traitement de ce type de faux, sachant qu'il n'est pas uniquement dédié à ce problème, mais peut traiter d'autres types de faux.

Le descripteur de bâton utilisé au cours de l'évaluation de la performance du facteur de formes est le descripteur D1, c'est-à-dire la "moyenne du signal". Ce choix est justifié par les précédentes études. Les résultats obtenus sont présentés dans le Tableau 8 et sur la Figure 45.

K	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
1	1.290	0.628	0.465	0.430	0.143	0.130	0.334	0.091	0.191	0.216	1.164	0.518	0.247	0.147	0.095
3	1.360	1.099	0.378	0.764	0.193	0.163	0.493	0.249	0.358	0.145	1.615	0.921	0.403	0.299	0.188
5	1.495	1.251	0.665	1.240	0.310	0.409	0.661	0.541	0.527	0.253	1.835	1.790	0.542	0.408	0.244
7	1.728	1.404	1.050	1.874	0.530	0.920	0.774	0.778	0.655	0.345	1.996	2.714	0.681	0.689	0.413
9	2.008	1.773	1.400	2.219	0.681	1.121	1.134	1.080	1.086	0.714	2.475	3.324	1.131	1.144	0.774

Tableau 8 : Erreur totale en % avec k plus proches voisins pour le couple codage Fuzzy Extended Shadow Code/D1

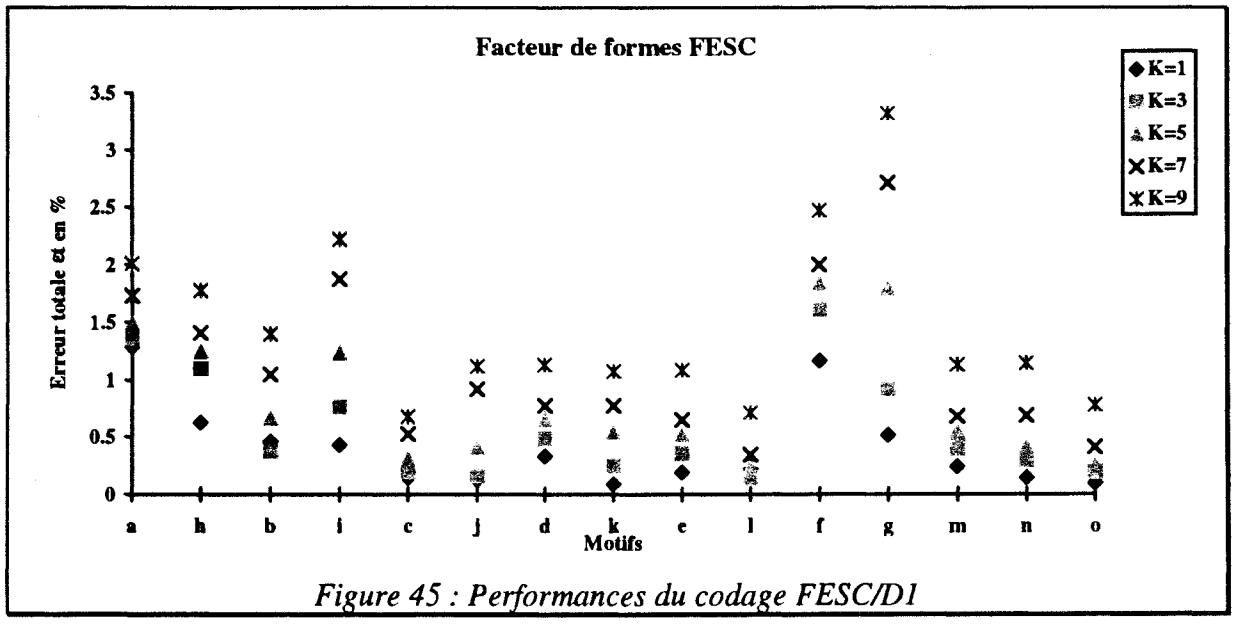


Figure 45 : Performances du codage FESC/D1

Ces résultats montrent que le Fuzzy Extended Shadow Code présente des performances très correctes en réponse au problème des faux aléatoires, pour lequel il n'est pas entièrement dédié. Les performances obtenues pour le plus proche voisin restent toujours dans le même ordre de grandeur que celles obtenues avec les autres méthodes. Elles sont toujours

améliorées pour les motifs grossiers par rapport à la méthode ESC, et présentent un taux d'erreurs légèrement plus élevé pour les motifs fins.

La Figure 46 permet de comparer les performances de ce nouveau facteur de formes/descripteur par rapport à celles des précédentes méthodes. Au vu des performances de la méthode de référence proposée par R. Sabourin, le FESC est intéressant. Pour les motifs *a, h, b et f*, les performances sont meilleures. La description horizontale des signatures par les motifs reste grossière, c'est ce qui semble justifier le gain de performances par l'apport de l'information de niveau de gris. Pour les motifs *c, i, j, k, l* et *g*, les performances sont assez proches de celles de l'ESC, avec toutefois des variances de performances un peu moins bonnes pour la plupart des motifs. Encore une fois, les motifs assez fins présentent des performances moins intéressantes avec un apport d'information, ce qui semble toujours conclure à l'utilisation de la présence avec ces motifs.

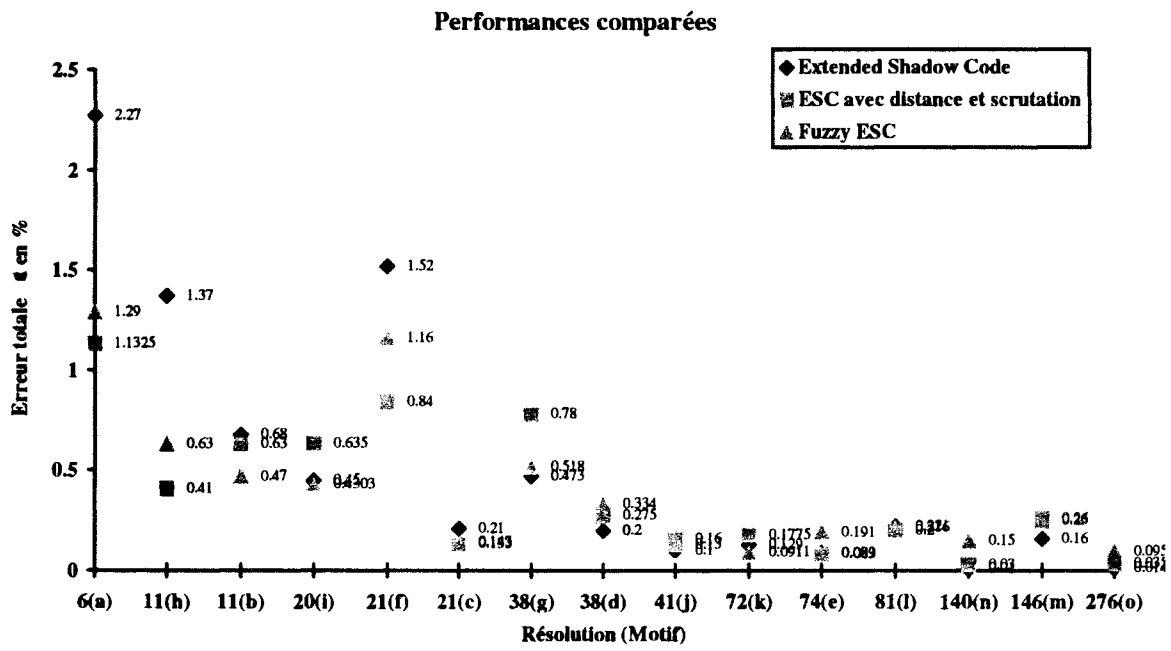


Figure 46 : Performances comparées de l'ESC, de l'ESC avec codage de la distance et nouvelle scrutation et le Fuzzy ESC.

Par comparaison avec l'ESC et codage de la distance, les performances pour les motifs *b, i, k* et *g* sont meilleures avec le FESC. De surcroît, les variances présentées dans le Tableau 9 sont sensiblement égales ou meilleures. Pour les motifs *a, d, e, j et l,* les performances sont proches, avec des variances globalement meilleures. Pour les motifs fins, encore une fois les performances sont moindres.

Variances de et	a	h	b	i	c	j	d	k	e	l	f	g	m	n	o
ESC	0.27	0.12	0.11	0.17	0.03	0.07	0.06	0.03	0.06	0.08	0.24	0.15	0.06	0.02	0.01
FESC	0.21	0.16	0.12	0.15	0.10	0.08	0.02	0.06	0.03	0.10	0.20	0.08	0.05	0.07	0.06

Tableau 9 : Variances des erreurs totales et pour les méthodes ESC et FESC

Après analyse des performances, l'introduction du niveau de gris agrégé à la distance par la table d'inférence floue semble apporter des performances assez intéressantes pour une utilisation dans le traitement des faux aléatoires, alors qu'il n'est pas uniquement dédié à cette opération. On aurait pu craindre que l'apport d'une information supplémentaire agrégée intuitivement à la distance réduise les performances, mais ce n'est globalement pas le cas.

Vis-à-vis de la méthode que nous avons proposée pour le traitement des faux aléatoires, le FESC est tout aussi bon, et dans certains cas meilleur. Cependant, pour les motifs fins, la méthode de référence présente de meilleures performances.

3.5. Amélioration du FESC

La fuzzification en trois termes des niveaux de gris, et par conséquent de la pseudo-dynamique, est basée sur la frontière entre les niveaux de gris de la signature et ceux du fond. Cependant, le seuil d'Otsu donne une valeur pessimiste de cette limite et a tendance à tronquer une plage assez importante de niveaux de gris de la signature (Figure 41) caractérisant les levés de plume et les déliés.

Pour pallier cet inconvénient, une fuzzification en 4 termes est proposée. Aux 3 termes caractérisant la signature, un terme pour le fond est ajouté. L'image de la signature est traitée globalement, et non plus au travers d'une sélection préalable de l'information utile par le seuil (Figure 47) séparant la signature du fond. Nous avons ainsi une approche image de la signature manuscrite.

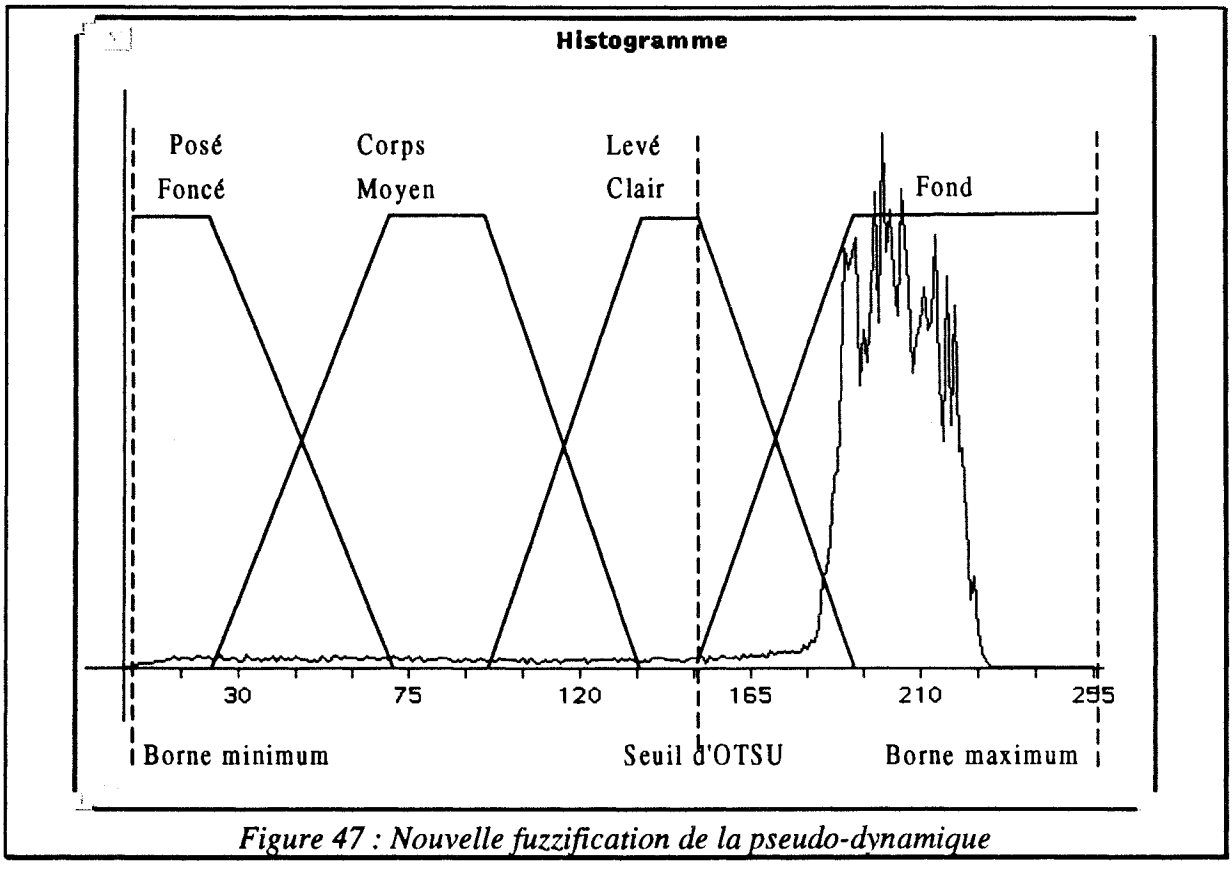


Figure 47 : Nouvelle fuzzification de la pseudo-dynamique

La fuzzification présentée offre alors une plus grande richesse d'information et il n'y a plus d'opérateur "dur" dans la chaîne de traitements. Elle est ainsi plus globale. Cette nouvelle fuzzification permet de prendre en compte l'ambiguïté entre la signature et le fond, par un recouvrement des termes relatifs à ces zones. En conséquence de la modification de la fuzzification, la table d'inférence elle-même est modifiée. Elle reste tout de même construite selon le même principe, mais le terme « Fond » implique le rajout de règles donnant un codage « Petit » en sortie. L'ensemble de règles d'inférence est proposé Figure 48.

Distance relative

Codage

Pseudo-dynamique	Loin	Moyen	Proche
Fond	Petit	Petit	Petit
Levé	Petit	Petit	Moyen
Corps	Petit	Moyen	Grand
Posé	Moyen	Moyen	Grand

Figure 48 : Table d'inférence pour une approche image de signatures

La table d'inférence peut donner l'impression d'une combinaison quasi proportionnelle des entrées. Or, les formes des termes de fuzzification, ainsi que les opérateurs utilisés contribuent à définir la forme de la surface d'agrégation donnant le codage sur le bâton. Nous pouvons constater sur la Figure 49, que cette surface est non linéaire, et correspond bien à la table d'inférence proposée.

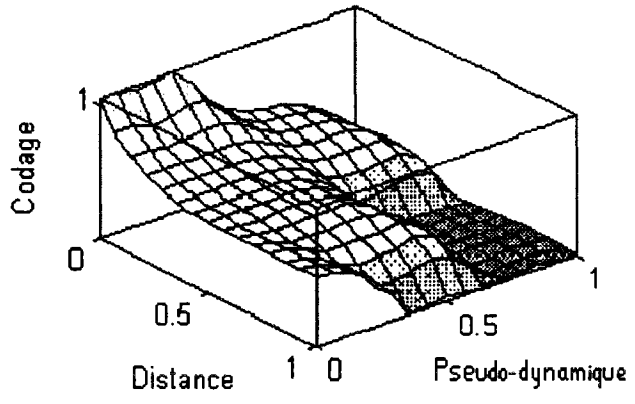


Figure 49 : Surface de codage en fonction des entrées

3.6. Pertinence du nouveau facteur de formes

L'évaluation de la pertinence de ce facteur de formes est réalisée selon les mêmes conditions que précédemment, avec le même protocole. Les résultats sont regroupés dans le Tableau 10 :

K	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
1	0.880	0.446	0.415	0.535	0.160	0.167	0.397	0.257	0.287	0.203	0.998	0.809	0.296	0.045	0.047
3	0.980	0.664	0.665	0.842	0.246	0.175	0.571	0.262	0.382	0.198	1.173	1.333	0.396	0.195	0.089
5	1.373	0.757	0.788	1.414	0.330	0.269	0.617	0.351	0.507	0.290	1.416	2.088	0.496	0.279	0.188
7	1.525	1.040	1.125	1.917	0.433	0.537	0.796	0.733	0.658	0.432	1.672	2.734	0.627	0.428	0.279
9	1.790	1.348	1.510	2.396	0.762	0.818	1.220	1.039	0.966	0.774	1.972	3.121	1.080	0.785	0.505

Tableau 10 : Erreur totale en % avec k plus proches voisins pour le couple codage Fuzzy Extended Shadow Code/D1

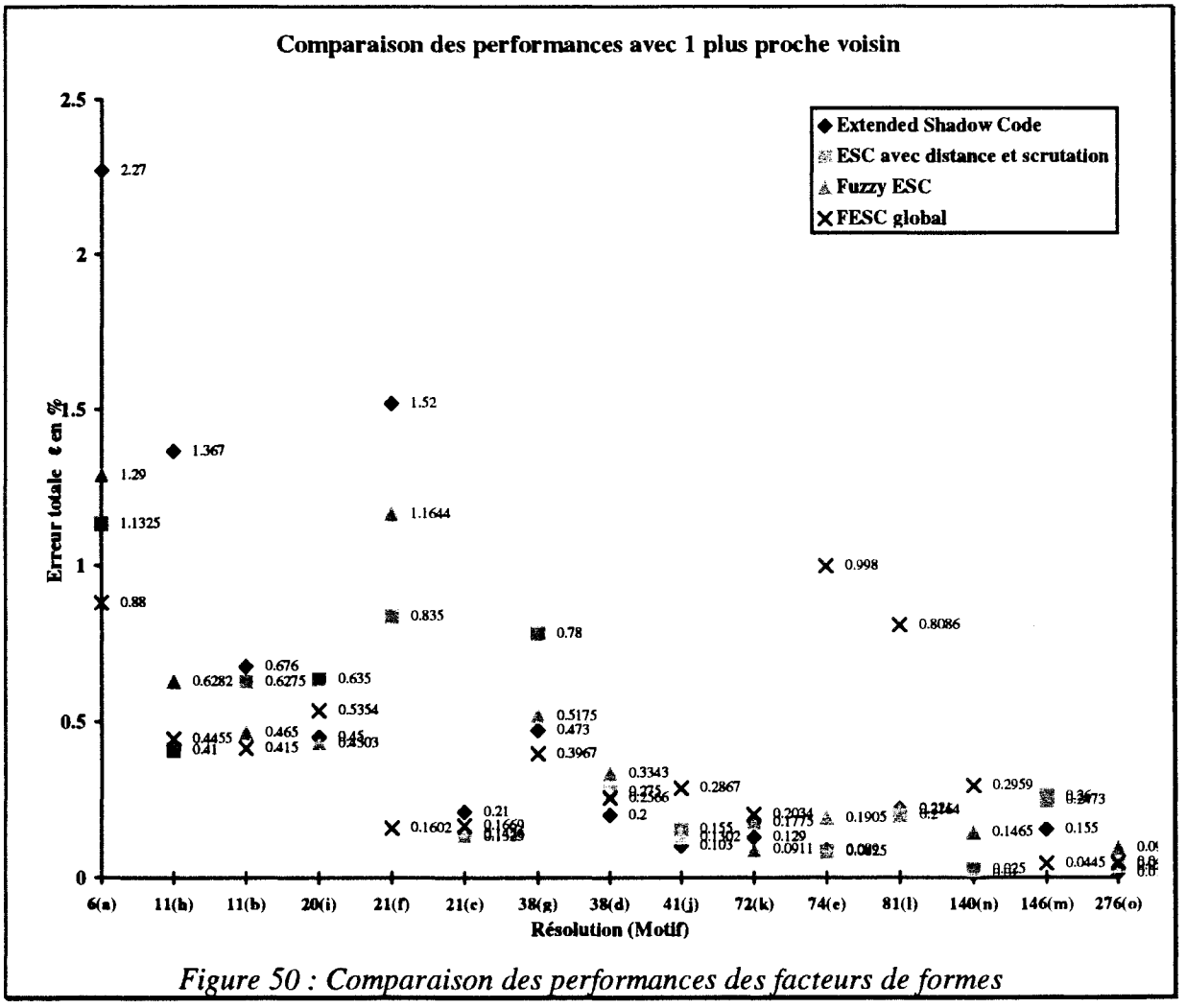


Figure 50 : Comparaison des performances des facteurs de formes

La comparaison des performances du FESC global avec celles de la méthode de R. Sabourin est motivante (Figure 50). Les résultats pour les motifs *a, h, b* et *f* sont vraiment meilleurs avec en plus des variances plus faibles présentées dans le Tableau 11. Pour les motifs *i, k, l, c* et *j*, les performances sont proches de celles de l'ESC, avec des variances sensiblement

proches. Enfin, les résultats pour les motifs *d, e, g* et *m* sont moins bons, mais restent toujours assez proches.

Par rapport à l'ESC avec codage de la distance, les résultats pour les motifs *a, b* et *i* restent meilleurs. Ils sont sensiblement équivalents, avec des variances également équivalentes pour les motifs *h, j, k, l* et *g* (Tableau 11). Enfin, avec les motifs *d, e, f* et *m,* le FESC global est légèrement moins performant. Nous remarquons cependant que, pour le motif *f,* les performances sont assez proches de celles de l'ESC avec codage de la distance et que la variance de l'erreur est réduite. On pourrait donc considérer que pour ce motif particulier, il y a équivalence de l'ESC et du FESC en terme de performances.

Finalement, par comparaison des performances avec celles du FESC, les résultats pour les motifs *a, h* et *f* persistent à être meilleurs. Ceux pour les motifs *b, c, j, l, m, n* et *o* sont assez équivalents avec des variances intéressantes ce qui laisserait supposer une légère supériorité du FESC global. Les performances des motifs *i* et *d* sont légèrement moins bonnes, mais là encore la variance est substantiellement améliorée (Tableau 11). Enfin, les performances avec les motifs *g, e, k* sont très légèrement moins significatives.

Variances de et	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
ESC	0.27	0.12	0.11	0.17	0.03	0.07	0.06	0.03	0.06	0.08	0.24	0.15	0.06	0.02	0.01
ESC avec distance	0.24	0.14	0.12	0.15	0.05	0.07	0.07	0.04	0.07	0.08	0.18	0.14	0.07	0.02	0.027
FESC	0.21	0.16	0.12	0.15	0.10	0.08	0.02	0.06	0.03	0.10	0.20	0.08	0.05	0.07	0.06
FESC global	0.32	0.15	0.15	0.11	0.07	0.06	0.11	0.07	0.08	0.07	0.16	0.15	0.06	0.03	0.021

Tableau 11 : Variances des méthodes ESC, ESC avec distance et scrutation FESC et FESC global.

Ce nouvel opérateur de codage des signatures présente un potentiel certain sur la base des performances mesurées précédement. Par rapport au codage unique de la distance, il offre de bonnes performances pour un assez grand nombre de motifs, même s'il n'est pas typiquement dédié au problème des faux aléatoires. Néanmoins, les apports sont moins appréciables avec la finesse du motif. Comme pour le codage de la distance, les résolutions fines poussent à l'utilisation de la présence. Il reste que le niveau de gris pourrait intervenir sans association avec la distance à partir d'une résolution donnée. De fait, si le niveau de gris est codé alors la présence l'est également, car il n'y a pas de niveau de gris de la signature sans la présence d'un tracé.

3.7. Application à la discrimination des faux par transfert

Le développement du facteur de formes FESC est orienté dans le traitement de multiples types de faux. Nous avons évalué son pouvoir discriminant dans le cadre des faux aléatoires, et nous devons également quantifier sa pertinence dans le traitement des faux par transfert optique.

Ces faux sont des signatures dont la dynamique est différente de celle des vraies signatures mais dont la statique est proche. Dans ce cas, tout facteur de formes ne caractérisant

pas la notion de dynamique, comme l'ESC, sera incapable de détecter une falsification par transfert optique.

3.7.1. Simulation d'un faux par calque

Nous ne possédons pas de bases de données de faux par calque. Aussi, nous utilisons la base de données initiale, et nous transformons les signatures destinées au test, pour simuler une imitation par transfert optique.

Le principe de transformation que nous proposons est une LUT, dont l'objectif est de pratiquer sur une image de signature, une opération de traitement de l'image, similaire à celle d'une photocopieuse. Les photocopieuses actuelles sont de très bonne qualité, ce qui leur permet d'obtenir une photocopie très proche de l'original. Nous avons supposé que la photocopieuse que nous voulons simuler, opère essentiellement par une modification linéaire du contraste de l'image de la signature. La LUT proposée Figure 51, nous permet de simuler une modification de contraste ajustable par l'intermédiaire d'un facteur δ.

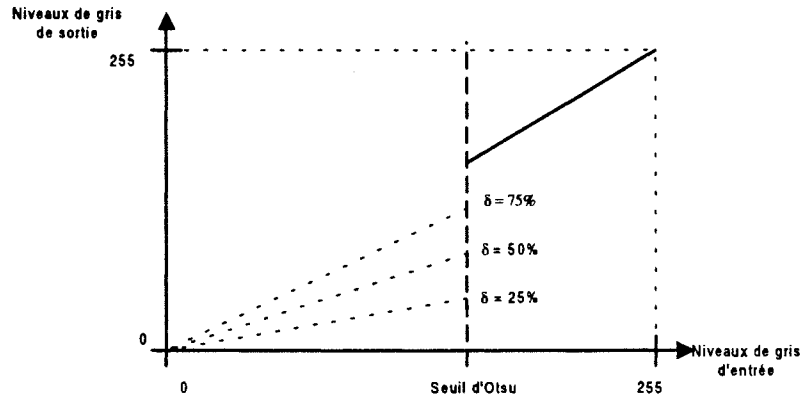


Figure 51 : LUT pour la simulation d'une photocopie

Un facteur de 100% simule une photocopie parfaite, et plus le facteur δ baisse moins la qualité de la photocopieuse sera bonne. Les valeurs du facteur δ les plus proches de 100% représentent donc la plus grande difficulté pour la détermination d'un faux.

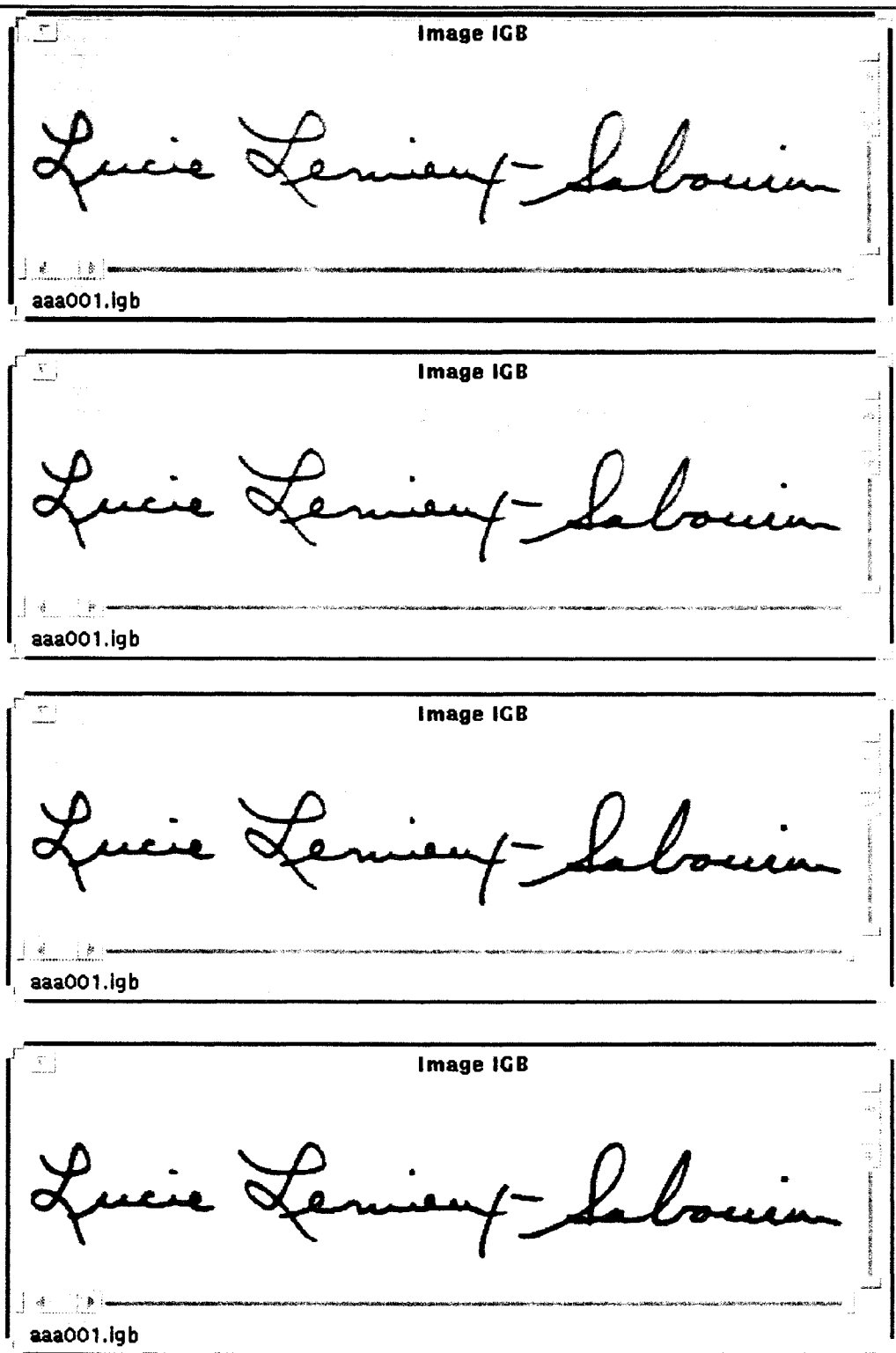


Figure 52 : Image initiale et ses photocopies avec δ = 75%, 50% et 25%

La Figure 52 présente l'image initiale ainsi que ses photocopies après transformation par la LUT que nous avons définie. La seconde image présentée résulte d'une transformation avec un facteur δ de 75%. Une telle transformation représente un faux d'une excellente qualité

puisque simultanément, la statique et la pseudo-dynamique sont très proches de l'authentique. La dynamique s'est très légèrement amoindrie, mais la statique reste identique. Les images suivantes présentent la même signature après transformation avec des facteurs de correction de 50% et 25%. La dynamique de la signature s'exprime d'autant moins que le facteur est petit, ce qui correspond bien à notre objectif de simulation d'une photocopieuse. La statique reste cependant la même, quel que soit le niveau de transformation introduit par la LUT.

Ce problème de faux est donc dramatiquement complexe. Cette transformation est appliquée avant l'entrée dans le système de codage, afin de nous placer dans les conditions réelles d'utilisation. De ce fait, comme la procédure de placement des termes de fuzzification de la dynamique est automatique, sur la base du seuil d'Otsu, ces termes vont se cadrer sur la plage de niveaux de gris de la signature. Par là même, ils caractériseront des zones dynamiques similaires à celles des authentiques. Ainsi, la différence de codage se situe essentiellement dans les zones frontières des termes comme le montre la Figure 53.

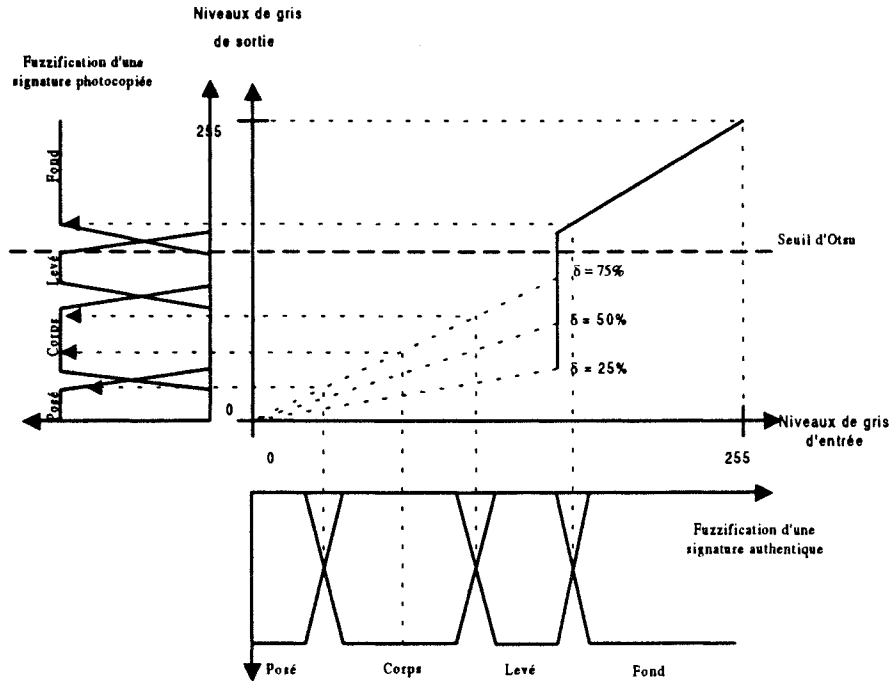


Figure 53 : Correspondance des termes de fuzzification pour une signature authentique et pour sa photocopie avec δ = 75%

Les niveaux de gris de la signature vont donc balayer de façon légèrement différente les trois termes de fuzzification se rapportant à la dynamique, et le niveau de dynamique moyen sera dépendant du niveau de transformation lié à δ, et donc plus ou moins proche du niveau moyen d'une signature vraie. Les pleins de la signature initiale ont tendance à être caractérisés comme des posés, le corps reste principalement codé comme un corps ainsi que les déliés. Par ailleurs, certains levés seront pris comme des éléments du corps. La différence reste cependant très fine, notamment avec un facteur de correction de 75%.

3.7.2. Pertinence pour le traitement des faux par photocopies

Pour évaluer la pertinence du facteur de formes FESC global sur le traitement des faux par photocopie, le protocole précédemment utilisé pour l'évaluation du facteur de formes dans le cas des faux aléatoires doit être modifié. La constitution des ensembles de référence à partir de ω_1 et ω_2 reste identique, puisque nous ne pouvons intégrer la multitude de faux par transfert. La modification du protocole d'évaluation se focalise essentiellement sur la constitution des ensembles de test. Ces ensembles sont constitués de signatures de la base de données, transformées par la LUT que nous avons proposée, c'est-à-dire qu'ils ne contiennent que des faux par transfert optique. Ainsi, nous ne recherchons plus deux affectations distinctes en sortie du système d'évaluation, mais une affectation unique à la classe des faux. De ce fait, l'erreur ε_1, caractérisant le taux de refus de vraies signatures, doit être maximum, puisqu'une signature vraie "photocopiée" ne doit pas être considérée vraie. En revanche, ε_2 caractérisant les erreurs d'affectation de fausses signatures photocopiées prises comme vraies, doit rester minimum. La pertinence est donc donnée par le taux d'erreur de classification d'un k plus proches voisins, à partir du taux d'erreur et calculé selon l'équation (Eq. 10).

$$\varepsilon_t = \frac{(1-\varepsilon_1)+\varepsilon_2}{2} \qquad \textit{(Eq. 10)}$$

Le taux d'erreur ε_1 est en fait un taux de réussite de détection de vraies signatures photocopiées, mais nous continuerons à le considérer comme un taux d'erreur.

Protocole pour l'évaluation des faux par photocopie.

POUR motif = *a* à *o*
  POUR itr = 1 **à** 25
    POUR individu 1 **à** 20
      Choisir les échantillons ω_2 de référence
      Constituer l'ensemble de référence
      Choisir les échantillons de test "photocopiés" avec un facteur δ préfixé
      Calculer les distances et évaluer les erreurs ε_1 et ε_2 pour k=1, 3, 5, 7, 9
    FIN POUR
    Evaluer les erreurs du système (pour chaque itération)
  FIN POUR
FIN POUR
Evaluer l'erreur moyenne ε_t du système ainsi que la dispersion, pour k=1,3,5,7,9

Le descripteur de codage que nous avons pris pour l'évaluation, est toujours le descripteur D1, c'est-à-dire la moyenne du signal de codage.

Les performances du FESC global à traiter les faux par photocopie sont présentées dans le Tableau 12. Ce tableau regroupe les erreurs totales obtenues avec la règle des k plus proches voisins, lorsque le facteur de correction des images de signatures δ est de 75% (Tableau 12a, b) et 50%(Tableau 12c). Il regroupe également les erreurs de type I, ε_1, qui spécifie le taux de signatures photocopiées considérées comme fausses.

et	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
k=1	39.87	40.40	41.93	36.80	43.82	45.87	38.65	43.37	39.56	46.73	39.37	38.67	40.04	43.35	43.80
k=3	47.56	48.14	48.40	46.34	49.80	50.01	49.11	49.40	49.82	50.04	48.66	47.61	49.66	49.95	50.11
k=5	47.75	48.16	49.04	46.39	49.87	50.20	49.29	49.68	49.93	50.17	48.60	47.70	49.77	49.91	50.05
k=7	48.29	48.29	49.27	46.90	49.86	49.93	49.40	49.95	50.04	50.25	48.80	47.44	49.83	49.99	50.29
k=9	48.43	48.46	49.63	47.23	50.02	50.29	49.72	50.14	50.46	50.54	49.09	47.12	50.01	50.24	50.38

a) Erreur totale en %

et	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
k=1	21.33	20.05	16.70	27.53	12.60	8.43	23.13	13.58	21.15	6.75	22.05	23.85	20.05	13.50	12.55
k=3	6.23	4.78	4.10	8.83	0.75	0.25	2.40	1.73	0.85	0.20	3.68	6.45	1.03	0.45	0.00
k=5	6.23	5.08	3.10	9.15	0.80	0.13	2.25	1.40	0.80	0.10	4.08	6.78	0.93	0.68	0.25
k=7	5.53	5.13	3.08	8.58	1.10	0.80	2.38	1.25	0.95	0.15	4.10	7.65	1.10	0.88	0.05
k=9	5.75	5.40	2.90	8.78	1.33	0.70	2.30	1.30	0.78	0.08	4.18	8.78	1.25	1.10	0.40

b) Erreur de type I en %

1 voisin	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
et; d = 50%	31.93	30.60	32.59	27.33	37.73	42.14	26.67	33.89	30.42	45.34	30.64	29.79	20.05	13.50	12.55
e1;d = 50%	37.6	40.1	35.65	46.73	24.83	15.93	47.18	32.53	39.55	9.55	39.7	41.6	37.63	26.68	23.75

c) Erreur totale et erreur de type I en %

Tableau 12 : Erreur totale et erreur de type I en %, avec la règle des k plus proches voisins, sans référence de photocopies dans l'ensemble de référence

Dans le cas d'un facteur de correction de 75%, les taux d'erreur ε_1 se situent entre 7 et 27% de détection de photocopies. Le taux d'erreur global ε_t, montre également une bonne discrimination des photocopies de faux aléatoires. La diminution du facteur de correction δ est directement observée sur les résultats. En effet, les taux de détection de photocopies augmentent systématiquement par rapport au cas où δ=75%. Ces taux se situent alors entre 27 et 45 %, et dépassent donc le meilleur taux de discrimination obtenu précédement.

Ces résultats ne sont pas totalement négatifs, car le protocole utilisé dans cette évaluation est directement extrait de celui destiné à l'évaluation des faux aléatoires. Or, les photocopies sont également des faux, et le facteur de formes utilisant la table d'inférence floue ne peut que transformer légèrement les signaux de codage puisque la géométrie de la signature est conservée. Par ailleurs, les photocopies sont naturellement plus proches des vraies signatures que des faux aléatoires, puisque ces derniers ont des caractéristiques géométriques globalement différentes. Cela implique que dans l'espace des caractéristiques, les photocopies seront à proximité des vraies signatures. De ce fait, comme cette transformation est faible, et

que la règle des k plus proches voisins est sans rejet de distance, les performances ne peuvent être que moyennes. Nous pouvons constater cet état de fait à l'aide d'une visualisation sur le plan factoriel d'une analyse en composantes principales réalisée sur le codage par le motif grossier des signatures. Dans cette analyse, les signatures naturelles du premier scripteur de la base sont présentes, ainsi que ses mêmes signatures photocopiées avec différentes valeurs du facteur de correction δ (Figure 54).

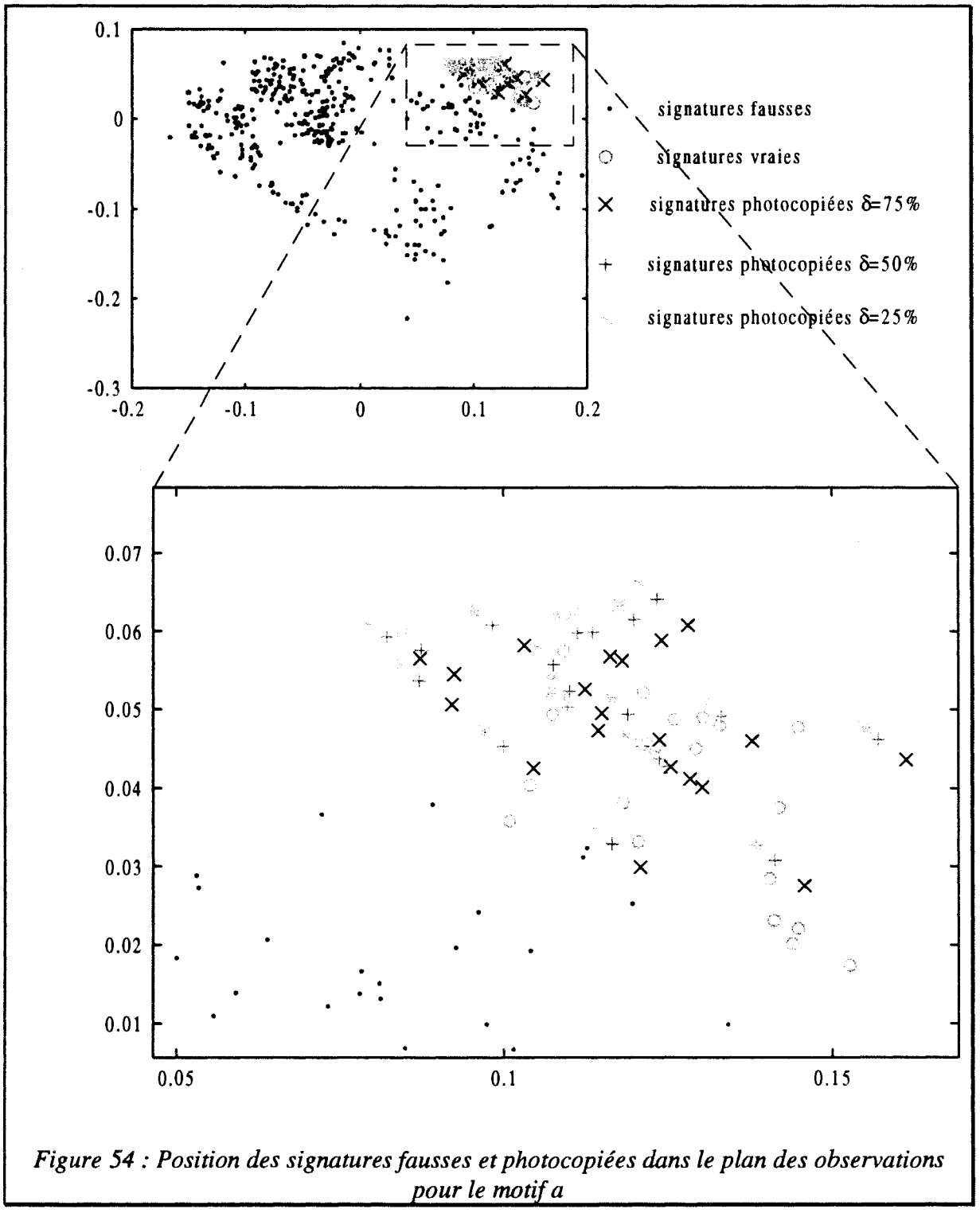


Figure 54 : Position des signatures fausses et photocopiées dans le plan des observations pour le motif a

Nous proposons donc un second protocole destiné à montrer que les signatures photocopiées sont effectivement différentes des signatures authentiques. Pour cela, nous introduisons cinq signatures photocopiées parmi l'ensemble de référence des fausses signatures. Ces cinq signatures vont caractériser une zone de l'espace des caractéristiques, à laquelle les signatures photocopiées pourront se rattacher.

Second protocole d'évaluation :

POUR motif = *a* à *o*

  POUR itr = 1 **à** 25

    POUR individu 1 **à** 20

      Choisir les échantillons ω_2 de référence, et 5 signatures photocopiées du scripteur de référence

      Constituer l'ensemble de référence

      Choisir les échantillons de test "photocopiés" avec un facteur δ préfixé

      Calculer les distances et évaluer les erreurs ε_1 et ε_2 pour k=1, 3, 5, 7, 9

    FIN POUR

    Evaluer les erreurs du système (pour chaque itération)

  FIN POUR

FIN POUR

Evaluer l'erreur moyenne ε_t du système ainsi que la dispersion, pour k=1,3,5,7,9

Les performances obtenues à partir de ce second protocole d'évaluation sont regroupées dans le Tableau 13, où nous avons retenu les performances pour la règle du plus proche voisin, en fonction du facteur de correction δ.

1 voisin	*a*	*h*	*b*	*i*	*c*	*j*	*d*	*k*	*e*	*l*	*f*	*g*	*m*	*n*	*o*
ε_t, δ=75%	23.04	17.50	20.97	16.43	21.67	23.51	16.51	20.40	16.75	25.98	16.88	18.78	17.99	18.99	21.23
ε_1, δ=75%	55.33	65.93	59.18	68.43	57.35	53.45	67.73	59.88	67.03	48.53	67.15	63.93	64.70	62.60	58.15
ε_t, δ=50%	14.94	9.48	13.13	8.72	13.70	15.97	15.97	9.78	9.79	21.46	8.20	9.17	10.69	12.04	13.71
ε_1, δ=50%	71.80	82.35	75.08	84.23	73.23	68.53	87.20	81.03	81.03	57.63	84.63	83.30	79.28	76.5	73.2

Tableau 13 : Erreur totale et erreur de type I en % avec 1 plus proche voisin, avec 5 références de photocopies

A partir des Tableaux 12 et 13, nous constatons que la discrimination des faux par photocopies est possible. Les taux de rejet de photocopies obtenus sans introduction de références de photocopies sont néanmoins élevés. Cependant, le Fuzzy Extended Shadow Code montre une certaine capacité à transformer les signaux de codage, afin de renforcer la

dissemblance d'une photocopie et d'une signature authentique, sur la base d'une modification de la pseudo-dynamique. En outre, l'introduction de références de photocopies permet de montrer que cette différence dans l'espace des caractéristiques reste locale du fait de la conservation d'une géométrie similaire des signatures.

Nous pouvons également constater que les performances sont moins dépendantes de la forme et de la résolution du motif que dans le cas des faux aléatoires. Ceci peut s'expliquer par la nature de la transformation des niveaux de gris. En effet, cette transformation renforce le contraste de la signature, et simule une diminution homogène de vitesse. Ainsi, toutes les parties de la signature sont transformées. Le posé, qui représente un faible volume d'information conserve sa caractéristique. Le corps, qui représente la majeure partie de la signature, est un peu plus transformé. Néanmoins, si nous décomposons le corps en trois parties correspondant aux pleins, valeur moyenne du corps et déliés, ce sont principalement les déliés qui subissent la transformation la plus importante. Les levés, moins présents, subissent également une assez forte transformation. De ce fait, comme la photocopie réduit les niveaux de gris les plus élevés de la signature, la transformation porte essentiellement sur les déliés et les levés. Ces parties de la signature sont cependant assez faiblement représentées, ce qui implique une faible modification au niveau du codage par rapport à une signature authentique, et représente donc un problème difficile.

Par ailleurs, cette transformation étant globale et linéaire, la pertinence du FESC global se ressent dès les motifs grossiers, et nous retrouvons une certaine équivalence des performances avec les motifs plus fins.

La simulation de photocopie est un cas particulier de faux par transfert optique. L'utilisation d'une autre technique de falsification apporterait des modifications particulières de la signature. Ainsi, une falsification par calque pourrait transformer le corps de la signature, ce qui modifierait profondement les signaux de codage et impliquerait leur discrimination. Des modifications plus locales de la signature, notamment sur les posés et les levés, sans grandes modifications du corps, seraient potentiellement discriminées par des motifs plus fins puisque ceux-ci seraient susceptibles de localiser ces parties transformées. Néanmoins, un tel problème restera fort difficile à résoudre.

3.8. Conclusion

Le cas des faux par transfert optique est un problème très complexe, et nous n'avons exploité qu'une partie du potentiel de notre facteur de formes par la simulation de photocopies. Le problème est d'autant plus délicat que la statique des signatures est similaire à celle des authentiques. Nous avons tout de même montré que le FESC global ouvrait la possibilité de discriminer en partie les faux par photocopies, et par extension les faux par transfert optique. Une étude plus poussée de ce facteur de formes devra néanmoins être réalisée pour juger de sa pertinence, pour le cas général des faux par transfert.

Une telle évaluation devrait permettre de situer l'importance des différents paramètres intervenant dans le facteur de formes. En effet, nous pourrons évaluer le choix des fonctions d'appartenance, tant en forme qu'en nombre, la table d'inférence, ainsi que la méthode de défuzzification choisie. Leur définition, lors de l'élaboration du facteur de formes, est restée jusqu'à présent heuristique. Le nombre de termes sur la fuzzification des niveaux de gris peut

cependant être rattaché à la décomposition des éléments constitutifs d'une signature en tenant compte des pleins et des déliés. Néanmoins, la définition de leur position exacte, ainsi que de leur forme restera heuristique. En revanche, la table d'inférence évoluera pour prendre en compte ces nouveaux termes et nous pourrons spécifier un comportement plus précis, mais également plus particulier du codage, augmentant sa sensibilité, aux fausses signatures comme aux vraies.

4. Conclusion

La définition d'un facteur de formes permet de doter un système de reconnaissance d'une fonction de perception en rapport avec le problème à résoudre. Il apparaît clairement que le facteur de formes doit être en adéquation avec le problème, en synthétisant les informations pertinentes, et doit être adapté à la méthode de reconnaissance, en lui fournissant les informations nécessaires.

Quant à l'approche que nous avons suivie, deux caractéristiques singulières peuvent être retirées des expériences réalisées; le typage perceptif des motifs; les informations extraites par les méthodes de codage vis-à-vis du problème posé.

Ce sont essentiellement la forme et la résolution des motifs qui typent leur pertinence. D'un point de vue général, un découpage fin dans le sens vertical, et grossier dans le sens horizontal, va forcer les bâtons horizontaux à décrire très globalement l'information telle que les orientations et la hauteur moyenne de la signature. Les bâtons verticaux vont traduire simultanément les largeurs moyennes par section de motif et des comportements particuliers comme des hampes, certains déliés, des accents dont la variabilité est d'autant plus grande que leur réalisation est faite avec une dynamique importante. Un découpage fin horizontal traduit mieux l'enveloppe de la signature et s'attache plus à des proportions locales, des espacements entre les lettres. Un découpage fin dans les deux sens permet la description de nombreuses caractéristiques, comme les pentes, les hauteurs relatives des lettres, les coupures dans la signature et des alignements.

Le motif le plus grossier (motif *a*) apporte une perception très globale de l'enveloppe de la signature. Le filtrage des signatures est important puisque l'on passe d'un espace image dans $\Re^{512x128x256}$ à un espace des caractéristiques dans $\Re^{6}$. La possibilité de percevoir deux signatures proches dans l'espace image par un point unique dans l'espace des caractéristiques est importante. Cela entraîne une confusion naturelle assez grande.

Le motif *b* permet une meilleure perception de l'enveloppe horizontale des signatures, par la subdivision des bâtons horizontaux. Les bâtons verticaux s'attachent plus à une description globale de la signature dans les sections du motif.

Le motif *h* opère un découpage horizontal de la signature ce qui permet de traduire le niveau moyen du tracé du corps pour le bâton horizontal central, exprimant par la même une caractérisation globale de l'orientation générale de la signature. On peut rapprocher cette caractérisation d'une ligne guide. Les bâtons horizontaux externes caractérisent des portions spécifiques du tracé comme des hampes, des points d'attaque, des déliés amples et des accents. Ces actions sont en général dynamiques et moins stables en position et en vitesse dans le temps.

Les découpages successifs, donnant la multi-échelle, vont permettre d'affiner la perception des différentes caractéristiques géométriques des signatures que nous avons citées. Par exemple, sur la base du motif *h*, un découpage vertical en deux parties définit le motif *i*. Le motif peut également venir d'un découpage horizontal du motif *b*. On constate que le motif *i* permet de mieux cerner le niveau moyen de la signature sur les bâtons horizontaux centraux, et simultanément une meilleure caractérisation des proportions globales verticales hautes et

basses des signatures. L'échelle finale, donnée par le motif *o,* approche très finement la géométrie de la signature tant sur le plan externe, qu'interne. Une telle finesse de perception décrit toute la géométrie, les particularités du scripteur comme les hampes, les déliés, ou les points d'attaque, les dégénérescences de la signature, ... Dans le corps de la signature, le motif *o* donne presque une caractérisation lettre par lettre, c'est-à-dire proche d'une segmentation en lettre sans pourtant l'être puisque cette approche est insensible au texte.

Ce typage perceptif impose que, pour chaque scripteur, de par la typologie de sa signature, un certain nombre de motifs soient plus sensibles et plus appropriés. Dans le cas de notre base, seules des signatures nord-américaines sont utilisées, ce qui entraîne généralement une moins bonne réponse des motifs ayant une prépondérance dans le découpage vertical comme les motifs *h*, *f*, *g*. Les signatures européennes définies comme des parafes sont beaucoup plus stylées et d'orientations générales différentes. De ce fait, les motifs à forme horizontale prépondérante seront plus appropriés, et d'autres conviendront moins. L'approche multi-échelle nous permet donc de mieux aborder les différents types de signatures et ainsi de pouvoir répondre au problème de vérification de signatures manuscrites, sans contraintes de régionalisation.

Cependant, nous n'avons pas cherché à sélectionner des motifs pour un scripteur particulier, car nous pensons que la base de données disponible n'est pas assez représentative pour un scripteur, ni pour tous les scripteurs.

La seconde caractéristique essentielle de l'approche de perception que nous avons proposée, est l'introduction de la théorie des ensembles flous pour construire un facteur de formes plus à même de répondre au problème général de la vérification de signatures manuscrites. La possibilité de construction intuitive de la table d'inférence en logique floue permettant l'agrégation de notions hétérogènes, nous a permis de généraliser un peu plus le procédé et d'ouvrir de nouvelles perspectives de traitement comme celui des faux par photocopie.

La souplesse de la logique floue nous a également permis de réduire l'influence du seuil d'Otsu, qui maintenant n'est plus pris comme seuil de binarisation, mais comme une information pour la construction de la fuzzification de la pseudo-dynamique. L'intérêt primordial est de réduire les opérations "dures" dans les premières étapes du processus de reconnaissance de formes, celles-ci étant souvent une source importante d'erreur.

On peut également préciser que la logique floue nous a permis de construire un facteur de formes n'ayant plus une action purement "réflexe", mais qui possède une partie plus "réfléchie", plus "humaine" pour décrire les signatures. Ainsi, l'apport de "raisonnement" dès cette étape va favoriser le potentiel général de la chaîne de reconnaissance, et l'étape de classification/décision doit, sur cette base, offrir de meilleurs résultats.

Il faut préciser que le problème posé par les photocopies ne peut être traiter en classification que par des méthodes qui ne tiennent pas compte directement de la distance, ou qui autorisent un rejet de distance. L'évaluation par la règle de discrimination du plus proche voisin en a montré l'exemple. La méthode à choisir devra donc tenir compte de cette caractéristique, et offrir la possibilité d'un rejet de distance, puisque nous ne pouvons pas introduire de photocopies dans les ensembles de référence, pour la vérification proprement dite.

Chapitre 3
Raisonnement : Discrimination intégrée

1. Introduction

La vérification hors-ligne de signatures manuscrites est une problématique sérieuse de reconnaissance de formes. Dans ce cadre, nous avons choisi de privilégier une approche globale, pour laquelle il faut dissocier les deux fonctions essentielles que sont la perception, basée sur le facteur de formes, et le raisonnement, porté par l'étape de classification/décision. Le chapitre 2 se dédiait à la perception, nous avons proposé une approche d'extraction de caractéristiques statiques et pseudo-dynamiques des images de signatures, nous ouvrant la possibilité de discriminer simultanément des authentiques, les faux aléatoires et les faux par transfert optique.

Ce chapitre est dédié à la présentation de notre proposition d'un outil de discrimination intégrée, exploitant au mieux les potentialités du nouveau principe de codage et de l'approche de perception multi-échelle.

La première partie de ce chapitre est consacrée au rappel des caractéristiques fondamentales du problème de vérification automatique de l'identité, puis nous détaillons l'ensemble de discrimination proposé par R. Sabourin, afin de cerner son adéquation avec le problème à résoudre. Nous présentons ensuite les différentes sources d'améliorations possibles de cette approche, afin de fixer les critères de choix d'une méthode susceptible de répondre à notre problème.

Nous introduisons ensuite la méthode retenue, les k plus proches voisins flous de Béreau, que nous détaillons dans l'objectif de montrer comment elle peut être adaptée à notre application. Nous explicitons également comment l'introduction de la théorie des ensembles flous peut conduire à une augmentation de la sécurité dans la décision. Après avoir présenté les différentes adaptations de la fonction de discrimination de Béreau, nous proposons quelques expérimentations. La dernière partie de ce chapitre est consacrée à la définition de l'outil de discrimination intégrée, grâce à deux types de coopération de discrimination, dont nous présentons et discutons les résultats d'application à notre base de données.

La conclusion présentera la synthèse de nos contributions à l'étape de raisonnement.

2. Le Raisonnement pour la vérification de signatures manuscrites

Le problème de discrimination que nous devons résoudre pour la vérification automatique de l'identité par l'analyse de la signature manuscrite est un problème à deux classes, les vraies et les fausses signatures associées à un scripteur donné. Le contexte d'application formel est important, puisqu'il impose une décision privilégiant la sécurité par rapport au confort.

La classe des vraies signatures est formée de sous-classes de signatures, telles que les signatures formelles, informelles et fugitives, perçues au travers des caractérisations horizontale, verticale et diagonale du facteur de formes. Dans l'espace de représentation des signatures, ces sous-classes sont de formes géométriques complexes. Cette complexité est provoquée par les différents modificateurs de la signature énoncés par Saferstein [Saferstein 82]. Par conséquent, il ne peut pas y avoir de réplique exacte entre authentiques.

La classe des fausses signatures est encore plus complexe. Elle regroupe toutes les catégories de faux présentées dans le chapitre 1, à l'intérieur desquelles les signatures subissent les mêmes modifications que celles énoncées par Saferstein pour les signatures vraies.

L'établissement d'une règle de décision sur une base aussi complexe est extrêmement difficile. La règle à déterminer ne peut s'attacher à des cas particuliers, puisqu'il n'y a pas de réplique exacte d'authentiques, ni uniquement à des cas généraux, à cause de la complexité suggérée par Saferstein. La diversité des signatures entraîne un problème de complétude des données d'apprentissage, afin d'obtenir une connaissance suffisante sur les classes de vraies et de fausses signatures.

Pour déterminer une règle de discrimination, nous disposons d'un ensemble d'apprentissage de signatures vraies émises par le scripteur lors de son abonnement. Cet ensemble d'apprentissage contient un nombre limité d'éléments certifiés, puisqu'il est impossible de contraindre l'abonné à fournir beaucoup de signatures en un court laps de temps, ce qui rendrait les signatures moins naturelles. De plus, il est de toute façon impossible d'avoir une représentation complète des différents types de signatures. Les prototypes ne sont donc qu'une représentation, à un instant donné et dans des conditions particulières, de l'ensemble des signatures vraies.

Cependant, malgré cette faible représentativité des prototypes de la classe des vraies signatures, on peut considérer qu'ils nous permettent de mieux connaître cette classe que celle des fausses signatures. En effet, dans l'apprentissage, les seuls faux que l'on puisse utiliser comme prototypes sont des faux aléatoires certifiés, c'est-à-dire des signatures d'autres scripteurs de la base. Dans la phase d'apprentissage, il faudra surtout chercher à exploiter les connaissances sur la classe des vraies, et utiliser les fausses signatures pour la limiter. La définition de la classe des vraies signatures doit être assez insensible à celle des fausses. De plus, comme l'ont montré les analyses en composantes principales présentées dans le chapitre 2, les classes de vraies et de fausses signatures se recouvrent et parfois se superposent.

Etant donné le peu de connaissances disponibles sur les classes en présence, on ne peut pas poser un modèle simple et sûr de leur statistique, il n'est donc pas réaliste d'aborder ce problème de discrimination par une approche statistique paramétrique. Cependant, nous posons l'hypothèse que la classe des vraies signatures est caractérisée par une loi statistique composée d'un mélange de lois correspondant aux sous-classes de signatures formelles, informelles et fugitives. Aussi, dans la suite de nos travaux, nous nous orienterons vers les méthodes de discrimination statistiques non paramétriques.

3. *La base de nos travaux*

La réponse apportée par R. Sabourin à cette problématique de vérification est, comme nous l'avons vu dans les chapitres 1 et 2, basée sur une coopération parallèle de 15 classificateurs à seuil, à distance minimum et 15 k plus proches voisins [Sabourin 95b]. L'apprentissage des classificateurs est supervisé non permanent et la coopération est de type vote majoritaire. Les classificateurs opèrent dans l'espace de représentation des formes et relèvent d'une approche géométrique de la discrimination. Cette structure forme un classificateur intégré [Sabourin 94b] [Sabourin 95a].

Le seuil de discrimination de chacun des classificateurs est déterminé à partir d'un ensemble d'apprentissage formé de signatures vraies et fausses (faux aléatoires) selon le protocole présenté dans le chapitre 1. N_{comp} signatures servent de références pour déterminer le seuil de distance, minimisant l'erreur de vérification sur l'ensemble d'apprentissage. La distance choisie est une distance euclidienne, la faible quantité de formes d'apprentissage ne permet pas d'utiliser les caractéristiques statistiques des classes dans la distance. La Figure 55 présente un exemple de cette procédure, où après la recherche d'une zone « vraie signature » autour de chacune des N_{comp} références de vraies sélectionnées, les limites de la classe des vraies signatures sont définies par le seuil. Les fausses signatures servent à délimiter la classe des vraies dans l'apprentissage, la détermination du seuil de décision est donc sensible au choix des faux. Aussi, R. Sabourin a-t-il pris la précaution de valider son approche en choisissant différentes configurations des vraies et fausses signatures d'apprentissage et en calculant le pourcentage d'erreur moyenne totale de vérification selon le protocole présenté dans le chapitre 1.

La discrimination des formes est réalisée en utilisant le seuil et l'ensemble des N_{comp} formes de référence des vraies signatures conservées après l'apprentissage. Pour discriminer une forme, il suffit de calculer les distances euclidiennes la séparant de chacune des N_{comp} références, puis de déterminer la plus faible, et enfin de comparer cette distance avec le seuil défini dans l'apprentissage. Si cette distance est inférieure au seuil, la forme est affectée à la classe des vraies, sinon elle est déclarée fausse.

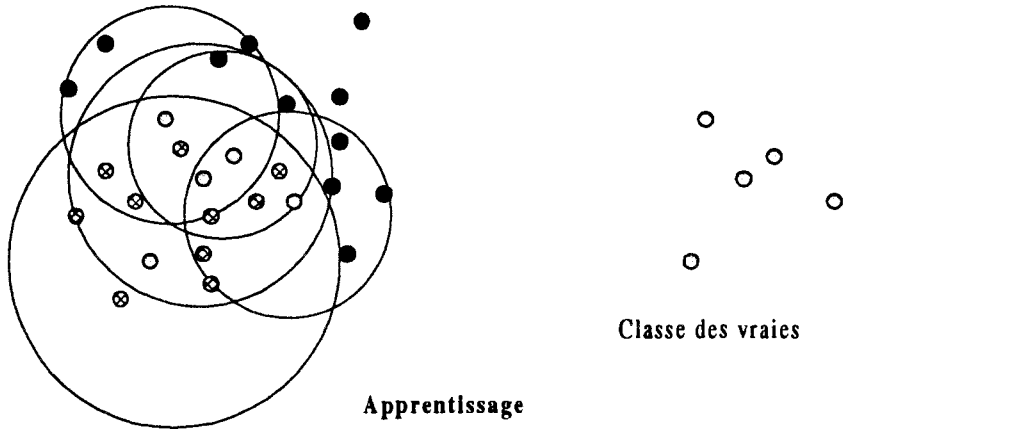


Figure 55 : Exemple de définition et d'application du seuil de discrimination

Les classificateurs associés à chacun des motifs de perception fournissent, après la discrimination, une réponse binaire vraie/fausse signature. La coopération considère chacune des réponses comme celle d'un expert répondant à une même question. Une loi de type vote majoritaire synthétise les réponses et fournit la réponse majoritaire parmi les différents experts. Un rejet de la forme est possible, s'il n'y a pas de consensus parmi les experts.

Cette coopération a bien sûr pour objectif de sécuriser la décision, en multipliant les points de vue, mais elle vise également à augmenter les performances de vérification de chacun des classificateurs. R. Sabourin a montré dans [Sabourin 94b] que les performances de ce

classificateur intégré sont très satisfaisantes. Le taux d'erreur moyenne total de vérification, évalué sur la base de données selon le protocole présenté dans le chapitre 1, pour différentes associations de classificateurs (trois plus mauvais, trois meilleurs, ensemble total) permet de conclure à l'amélioration systématique des performances par rapport à l'utilisation d'un seul classificateur. Rappelons que la vérification portait sur les faux aléatoires.

Cette procédure de classification/discrimination, particulièrement simple, est très peu gourmande en temps de calcul et en espace mémoire. En effet, seules les N_{comp} signatures vraies de référence sont nécessaires à la discrimination. Elle ne nécessite qu'une seule phase d'apprentissage, la règle de discrimination étant ensuite appliquée pour toutes les formes à discriminer.

Ces résultats encourageants nous ont conduits à exploiter le principe de coopération et à chercher les améliorations possibles de cette approche dans un contexte général de vérification. Comme nous l'avons précisé dans l'introduction de ce mémoire, il ne s'agit pas ici de rivaliser de performances avec les travaux de R. Sabourin. Nos objectifs sont l'amélioration globale des performances de l'étape de raisonnement et l'augmentation de la sécurité de la décision, quel que soit le type de fausses signatures rencontrées. Notre intérêt se porte sur la prise en compte de l'évolution de la forme de la signature authentique, la particularisation de chaque classificateur selon les qualités discriminantes du facteur de formes associé, et enfin sur une coopération informée des classificateurs.

Aussi, nous passons en revue les différentes composantes de ce classificateur intégré, afin de cerner ses points faibles par rapport au problème de discrimination en présence de faux par calque ou transfert optique.

Le premier point concerne les capacités de la classification/discrimination à absorber l'évolution des signatures. En effet, comme nous l'avons vu au chapitre 1, de nombreux facteurs influencent l'expression de l'identité par la signature, et engendrent une variabilité importante dans le temps de la forme des signatures authentiques.

Le seuil de décision, calculé lors de l'apprentissage, deviendra inefficace dès que les vraies ou les fausses signatures ne seront plus situées dans les zones délimitées par le seuil. La règle de discrimination de chacun des classificateurs n'exploite pas la connaissance disponible. Aussi, si la signature se déplace dans le temps et qu'elle sort de la zone des vraies (Figure 56 : point n°5), elle sera automatiquement qualifiée de fausse. Ce phénomène s'opérera même si, à l'apprentissage, aucune fausse signature ne se plaçait dans cette partie de l'espace de représentation. Une signature vraie peut être déclarée fausse (Figure 56 : points n°5 et 6) et le superviseur du système de vérification demandera à l'abonné de confirmer cette décision. Une signature fausse peut être déclarée vraie (Figure 56 : point n° 4) et le client avertira le superviseur de cette erreur. Si ce phénomène se répète, le système devrait normalement tenir compte de ces erreurs et intégrer les nouvelles connaissances. Une procédure d'apprentissage sera alors remise en route, cette opération se répétera aussi souvent que des erreurs importantes seront commises. On voit donc qu'il n'y a aucune possibilité dans cette approche, de prendre automatiquement en compte l'évolution probable des habitudes du scripteur.

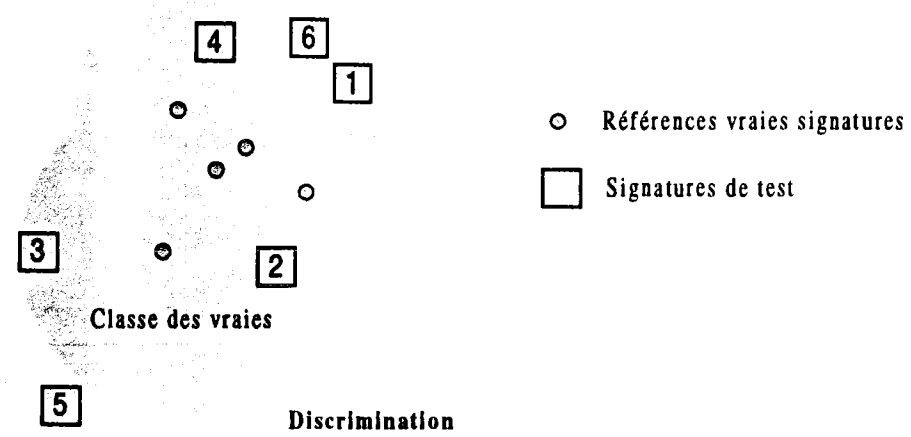


Figure 56 : Exemples d'erreurs possibles de vérification

Le second point d'intérêt de cette approche concerne la coopération des classificateurs.

Dans [Sabourin 94], l'auteur précise que la coopération s'effectue sur la base d'informations abstraites, les étiquettes fournies par les classificateurs. Cette démarche nous semble tout à fait pertinente. Cependant, les réponses des discriminateurs sont des décisions binaires et orthogonales (vraie ou fausse signature), dans lesquelles aucune information concernant la difficulté locale de discrimination ne peut être contenue. Aussi, la coopération ne peut réaliser la synthèse du jury que sur une base d'affirmations décontextualisées.

4. Les améliorations possibles

4.1. L'apprentissage

Dans un contexte réel d'application, un apprentissage permanent du classificateur nous paraît indispensable. Il permettrait de s'adapter, en ligne, aux variations possibles des habitudes du scripteur et aux modifications circonstancielles de la signature, par l'acquisition de nouvelles connaissances. L'apprentissage doit être de type semi-supervisé, puisqu'un opérateur peut certifier l'authenticité d'une signature, bien que cela ne puisse être le mode unique d'apprentissage. Pour l'apprentissage sans certification, le processus de classification ne doit pas prendre de risque par rapport à sa connaissance.

Bien sûr, le choix d'un apprentissage permanent alourdit notablement le système, mais il peut lui procurer des performances plus stables et plus robustes dans le temps. Cependant, l'évaluation de ces performances sur la base de données disponible ne serait pas démonstrative, puisque nous ne disposons pas d'assez de formes, par scripteur, pour simuler un apprentissage permanent.

Une conséquence directe de notre volonté de prendre en compte l'évolution possible des habitudes du scripteur, est de conserver des références de fausses signatures afin de limiter l'espace de représentation. La différence entre cerner la classe des vraies par les fausses signatures, et limiter ou borner l'espace de représentation autour des vraies, est très importante et source de généralité. Le faux ne doit plus être déclaré faux par absence de vrai, mais par

ressemblance au faux et dissemblance au vrai. Nous choisirons donc une méthode de discrimination dans laquelle une classe est définie à l'aide de ses éléments, ceux de l'autre classe intervenant indirectement comme contraintes sur son enveloppe, l'empêchant de se placer au même endroit dans l'espace de représentation.

4.2. La coopération des classificateurs

Dans un principe de coopération parallèle de classificateurs, tel qu'il est proposé par R. Sabourin, les décisions binaires prises par les différents classificateurs ne renseignent pas le module de coopération sur leurs difficultés locales de discrimination. Aussi, ce module n'est pas en mesure de faire une synthèse objective des avis des classificateurs, puisqu'il en a une vue tronquée. En d'autres termes, et pour citer une phrase chère à A. Kaufmann, « la décision binaire de discrimination locale fait chuter l'entropie trop tôt ». Il serait plus pertinent que la coopération s'opère sur des informations, certes abstraites, mais portant sur des propositions non binaires et non orthogonales de discrimination locale. Seule, la discrimination globale (coopération) ferait chuter l'entropie sur la base d'informations beaucoup plus riches au niveau sémantique. La décision finale serait alors encore plus sécurisée, elle pourrait être vraie signature, fausse signature et rejet de signature (rejet de distance ou d'ambiguïté) lorsque les propositions des classificateurs ne permettent pas de prendre une décision. Cette coopération peut être facilement envisagée en utilisant des classificateurs flous, fournissant des degrés d'appartenance non binaires aux classes.

Une autre coopération peut être envisagée dans les procédures de discrimination locales. Nous savons que chaque classificateur voit les mêmes formes au travers de différents filtres de l'information (le facteur de formes). Lorsqu'ils réalisent la discrimination propre à leurs points de vue, ils fonctionnent en aveugle par rapport à leurs voisins. Cette situation n'est pas problématique, cependant les classificateurs pourraient partager une information commune, dans le processus de décision. Sur la base de cette information commune, chacun des classificateurs apporterait ensuite sa réponse suivant sa spécificité. Cette proposition de coopération n'est pas applicable à tous les classificateurs existants, il faut bien sûr que l'information partagée soit abstraite. La méthode des k plus proches voisins est un exemple de discrimination qui permet cette coopération intra-classificateurs.

Nous présenterons dans la suite de notre mémoire une discrimination intégrée exploitant ces différentes possibilités de coopération.

4.3. Prise en compte de l'approche multi-échelles

La perception des signatures est réalisée au travers du facteur de formes ESC ou FESC, et selon plusieurs motifs de formes et de résolution différente. Le facteur de formes engendre, naturellement, une perte d'information, et oriente l'information perçue vers les invariants que l'on a heuristiquement privilégiés dans sa conception. Au travers de la perception et des divers facteurs de formes, un point de vue différent est pris sur les données selon la globalité du motif. Une signature physique est vue selon 15 points de vue différents, mais reste unique.

Le motif et la méthode de codage des signatures forment un filtre sur l'information réelle où l'imprécision est d'autant plus grande que le motif est grossier. Cette imprécision introduit naturellement une possibilité de recouvrement des sous-classes et des classes, ainsi qu'une possibilité de confusion entre vraies et fausses signatures. Les Figures 24 à 35, présentées dans le chapitre 2, permettent de visualiser ce recouvrement. La dimension de l'espace de représentation des formes a une influence directe sur la décision prise par chaque classificateur. La possibilité de confondre une signature fausse avec une vraie est plus grande dans l'espace $\Re^6$ que dans $\Re^{276}$, alors que la possibilité de ne pas reconnaître une vraie signature est plus importante dans $\Re^{276}$ que dans $\Re^6$. La Figure 57 illustre cette remarque, s'appuyant sur les propos de Qi [Qi 95] concernant les images de signatures : *« D'un point de vue global, toutes les signatures sont semblables. D'un point de vue local (pixels), toutes les signatures sont différentes ».*

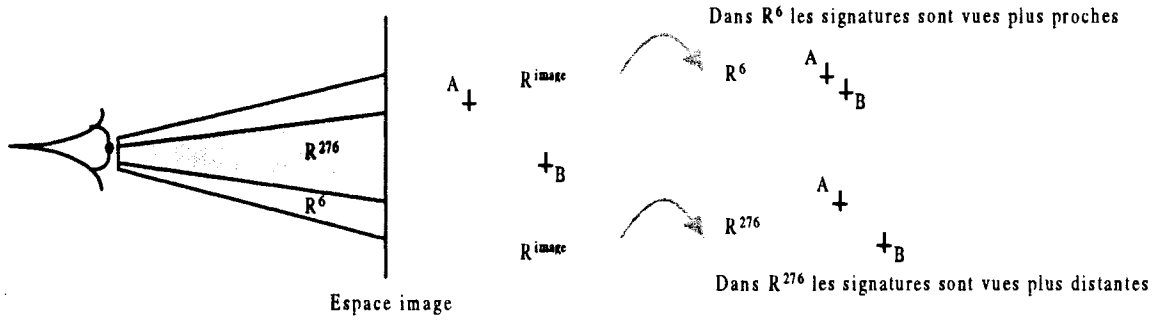


Figure 57 : Evolution de la notion de ressemblance selon le motif de perception

Les connaissances mémorisées au cours de l'apprentissage des classificateurs sont basées sur les mêmes formes perçues depuis l'espace image. Le point de vue qu'apporte chaque discrimination dépend de son niveau de perception, c'est-à-dire de la forme et de la résolution du motif. Le comportement de chaque classificateur, associé à chaque niveau de perception, doit donc être différent, pour pouvoir maintenir la même connaissance.

Dans les résolutions grossières, la possibilité de confusion est plus grande que dans les résolutions plus fines, aussi la formation de classes convexes est facilitée. Le phénomène d'interpolation des connaissances est intrinsèque à une résolution grossière, et la méthode de discrimination choisie n'a pas besoin de le provoquer. Simultanément, le phénomène d'extrapolation des connaissances ne doit pas être facilité, car la confusion est déjà largement présente.

En revanche, dans une résolution fine, l'ambiguïté naturelle de la perception est réduite, rendant la formation de classes convexes plus difficile. Pour compenser ce problème, on essayera de faciliter l'interpolation entre points de la même classe pour mieux synthétiser la connaissance intra-classe. L'ambiguïté étant moins présente à ce niveau de perception, on pourra laisser un peu plus de souplesse au niveau de l'extrapolation pour permettre l'acquisition de nouvelles connaissances, intervenant dans le renforcement des classes formées.

Ces remarques nous conduisent à rechercher des classificateurs dans lesquels un rejet de distance est possible et ajustable selon la résolution sur laquelle ils s'appuient. Pour les

motifs grossiers, on limitera l'interpolation et l'extrapolation, tandis que dans les motifs fins, on les favorisera.

La dimension de l'espace de représentation a une autre influence sur la décision des classificateurs, lorsqu'on la relie à la taille de l'ensemble d'apprentissage. Comme l'ont confirmé les différents résultats présentés dans le chapitre précédent, lors de la validation de la pertinence du facteur de formes, l'erreur de vérification diminue lorsque la résolution du motif de perception augmente. Ce phénomène n'est pas dû uniquement à la meilleure pertinence du motif *o* par rapport aux autres motifs, mais également à la sensibilité du classificateur à la taille de l'ensemble d'apprentissage. En effet, comme nous l'avons précisé dans les chapitres 1 et 2, plus la taille de l'espace de représentation est grande, plus le nombre de prototypes d'apprentissage doit être conséquent, si l'on ne veut pas courir le risque d'obtenir des résultats peu significatifs. Dans notre application, le nombre initial de formes de référence est faible au regard de la taille maximale de l'espace de représentation des formes (276). Sous l'hypothèse d'un apprentissage permanent, la taille de cet ensemble doit augmenter dans le temps, ce qui réduira ce risque. Nous chercherons donc à intégrer le rapport entre nombre de références et dimension de l'espace de représentation de la classe, comme information relativisant la réponse de chacun des classificateurs.

Dans le contexte d'une évolution de la connaissance grâce à l'apprentissage permanent, les classes et leurs caractéristiques sont liées à la résolution, mais aussi au temps et aux données qui permettent d'extraire cette connaissance. Toutes les informations relatives à la connaissance doivent être entretenues continuellement au cours du processus d'apprentissage, puisque à un instant donné, elles ne sont pas représentatives de la réalité définitive, mais elles convergent vers cette réalité [Harnad 87].

5. *L'algorithme de discrimination locale*

Notre choix d'une méthode de discrimination a été conduit pour satisfaire l'ensemble des critères et des propositions que nous venons de présenter, et que nous pouvons résumer succinctement :

1. principe de reconnaissance basé sur une représentation vectorielle des formes,
2. apprentissage permanent semi-supervisé,
3. utilisation d'un lot de prototypes constitué de vraies et de fausses signatures,
4. approche statistique non paramétrique,
5. réponse du classificateur sous la forme de propositions d'affectation non binaires et non orthogonales,
6. possibilité de prise en compte de la pertinence de la décision en rapport avec la dimension de l'espace de représentation,
7. possibilité d'adaptation de la réponse du classificateur selon le motif de perception,
8. possibilité de prise en compte du recouvrement des classes,
9. possibilité de rejet de distance et d'ambiguïté,
10. partage possible d'informations abstraites entre les différents classificateurs.

Les caractéristiques 5 et 8 ne peuvent être vérifiées correctement que si l'on choisit une discrimination basée sur les concepts de la théorie des ensembles flous autorisant la notion d'appartenance graduelle. Aussi, parmi les méthodes existantes de discrimination ordinaires et floues, dont une partie a été présentée dans le chapitre 1, nous avons retenu l'algorithme des k plus proches voisins flous proposé par Béreau [Béreau 86], [Dubuisson 90]. Cet algorithme possède un fonctionnement qui semble convenir à notre problème sur l'ensemble des critères retenus.

Ce choix ne constitue pas une originalité en soi, et on pourrait nous objecter que la règle de discrimination des k plus proches voisins n'est pas la plus adaptée à notre problème. En effet, elle force la mémorisation d'un ensemble infini de prototypes si l'on veut être sûr de minimiser la probabilité d'erreur au sens de Bayes. De plus, bien que des procédures de recherche optimisée des k plus proches voisins [Kittler 78], [Richetin 80], [Yunck 76], [Fukunaga 75], ou des procédures de réduction du nombre de prototypes par l'édition [Devijver 80], [Wilson 72], aient été proposées, elle reste gourmande en temps de calcul, lors de la recherche des plus proches voisins.

Cependant, il ne faut pas négliger la difficulté de notre problème. Les faux par photocopie, que nous avons simulés, sont particulièrement difficiles à discriminer. Leur position dans l'espace de représentation est très proche des signatures authentiques, relativement à la classe des fausses. Nous avons volontairement proposé une situation très critique, mais pas impossible, de vérification qui nécessite d'engager des moyens à la hauteur de nos objectifs de sécurisation de la décision.

Les méthodes de discrimination, vérifiant la condition d'orthogonalité d'appartenance aux classes, ne seront pas en mesure de distinguer ces fausses signatures des vraies. La méthode de discrimination géométrique utilisée par R. Sabourin en est un exemple. La règle des k plus proches voisins classique, qui constitue pourtant une référence dans le domaine de la reconnaissance de formes, ne permet pas plus la discrimination.

Cette incapacité est également due à l'absence du concept de rejet de distance pour traiter le recouvrement, voire la superposition de classes, mais aussi à la faible représentativité des prototypes de vraies et de fausses signatures, au départ du fonctionnement du système de vérification.

Dans ce contexte, la méthode développée par Béreau peut apporter une solution, puisqu'on sait qu'elle permet d'obtenir une erreur de reconnaissance moins grande que la règle des k plus proches voisins classique ou avec rejet d'ambiguïté [Devijver 77], [Devijver 82], dans les situations où il y peu de prototypes [Béreau 86], [Bezdek 86]. De plus, selon le cadre réel d'application, nous pourrons par la suite envisager des solutions, telles que la recherche optimisée des voisins, afin de minimiser ses inconvénients majeurs (temps de calcul et espace mémoire). Notons enfin que, compte tenu des capacités des calculateurs actuels, ces inconvénients risquent de ne pas être très pénalisants en regard des possibilités de l'algorithme.

5.1. L'algorithme des k plus proches voisins flous de Béreau

La procédure générale proposée par Béreau s'articule en deux phases reposant sur une extension, par le concept flou, de la règle de discrimination des k plus proches voisins.

La première étape consiste à réaliser une partition floue d'un lot de données d'apprentissage non étiquetées, par une classification automatique floue. Cette étape structure les connaissances portées par les données. Chaque classe est caractérisée par des paramètres, tels que le nombre de prototypes, l'entropie floue de la classe, ou les seuils d'affectation, utilisés dans la suite de l'algorithme.

La seconde phase consiste en une discrimination, dont la structure est similaire à l'algorithme de classification. Cette étape permet d'étiqueter les nouvelles données, et d'apprendre de façon permanente, en rejetant certaines formes et en créant à partir d'elles de nouvelles classes. On peut également renforcer la connaissance acquise lorsque les données sont réintégrées dans les classes existantes.

5.1.1. La procédure de classification automatique

Une analyse globale d'un lot d'apprentissage initial E_L permet de former une première classe floue. On détermine ensuite séquentiellement le degré d'appartenance des autres points d'apprentissage à cette classe. Si celui-ci est en dessous d'un seuil prédéfini, les points sont rejetés, sinon ils sont étiquetés. Après un nombre fixé a priori de rejets, on crée une nouvelle classe, dans la limite du nombre maximal de classes prédéfini. Les degrés d'appartenance des formes à chacune des classes, sont calculés selon l'équation (Eq. 11) :

$$\mu_i = \max_{x^j \in [1,k]} \mu_i(x^j) \exp\left[-\lambda\left(\frac{d_j}{d^{m_i}}\right)^2\right], 1 \le i \le c \qquad \textit{(Eq. 11)}$$

où

- λ est un facteur de modulation de la distance relative,
- x^j sont les plus proches voisins au sens de la distance choisie,
- $\mu_i(x^j)$ est le degré d'appartenance du voisin x^j à la classe C_i,
- d_j distance entre le point analysé et le voisin x^j,
- d^{m_i} distance moyenne intra-classe,
- k nombre de voisins les plus proches,
- c le nombre de classes.

La procédure est itérée jusqu'à épuisement du lot d'apprentissage. Le degré d'appartenance est fourni par le voisin de la forme dont la contribution est maximale. Cette

contribution s'exprime par la valeur du degré d'appartenance du voisin sélectionné, modulée par une fonction d'attraction que nous expliciterons en détail dans le paragraphe suivant.

5.1.2. La procédure de discrimination

Cette phase débute par la discrimination des points de l'ensemble d'apprentissage à l'aide de la fonction présentée dans l'équation (Eq. 12). L'algorithme réaffecte les points de l'ensemble d'apprentissage aux classes existantes. Cette nouvelle partition sert à la discrimination des formes inconnues.

La phase suivante de l'algorithme consiste en la discrimination des points de l'ensemble de test E_t par l'équation (Eq. 12). Chaque nouveau point est analysé séquentiellement. Il y a rejet des formes selon le même mode que celui décrit dans l'étape de classification, et on aboutira, de manière analogue, à la création et fusion de classes.

Cette étape est supportée par une seule fonction de discrimination exprimant une transition graduelle entre l'appartenance et la non-appartenance. Cette fonction a été conçue par Béreau par analogie avec la fonction de Fermi-Dirac [Béreau 86], [Dubuisson 90]. La fonction d'appartenance a la forme suivante (Eq. 12):

$$\mu_i = \frac{1}{1+\exp\left(\frac{\frac{k}{2}-t_i}{b_i}\right)} \quad \text{avec} \quad t_i = \sum_{x^j \in [1,k]} \mu_i(x^j)\exp\left[-\lambda\left(\frac{d_j}{d^{m_i}}\right)^2\right] \qquad \textit{(Eq. 12)}$$

où

- k est le nombre de voisins,
- b_i est défini comme l'entropie floue H_i (Eq. 13) de la classe C_i ,

et les autres paramètres ont la même forme que dans l'équation (Eq. 1).

Tous les k plus proches voisins de la forme à discriminer contribuent au calcul de son degré d'appartenance. On évite ainsi une transition trop brutale telle que celle réalisée par la fonction présentée dans l'équation (Eq. 11) à l'étape de classification.

L'entropie floue proposée par De Luca et Termini [De Luca 72] mesure le degré de flou de la classe C_i dans l'ensemble d'apprentissage. Elle est définie par :

$$H_i = -\sum_{x \in E_L} \mu_i(x)\log_2(\mu_i(x)) + (1-\mu_i(x))\log_2(1-\mu_i(x)) \qquad \textit{(Eq. 13)}$$

La Figure 58 présente le comportement de la fonction d'appartenance en fonction de H_i.

Si H_i s'annule, la classe n'est pas floue (tous les degrés d'appartenance sont binaires) et la règle de discrimination devient celle des k plus proches voisins ordinaire. L'autre cas limite est

représenté par H_i =1 signifiant que la classe est complètement floue (tous les degrés d'appartenance sont à 0.5). Dès lors, μ_i ne peut atteindre au mieux que des valeurs extrêmes, dépendantes de b_i. Les limites haute et basse de valeurs d'appartenance, précisent que l'affectation est sujette à une très forte incertitude.

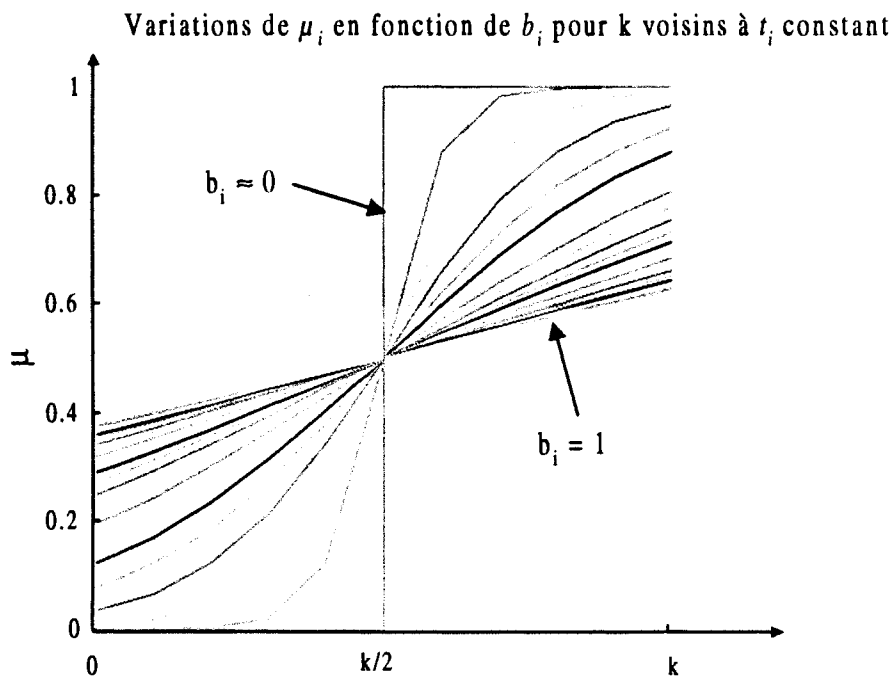


Figure 58 : Comportement de $\mu_i(y)$ en fonction du nombre de voisins et selon b_i

La quantité t_i permet de prendre en compte les degrés d'appartenance des plus proches voisins du point analysé, modulés par une pondération. Cette pondération établit une courbe d'attraction autour de chacun des prototypes x^j. Si le point considéré a comme voisin le plus proche le prototype x^j, alors son degré d'appartenance ne tiendra pas compte de celui de x^j s'il n'est pas dans la zone d'attraction de x^j. Cette zone d'attraction a pour effet de relativiser l'éloignement du point avec ses voisins, par la distance moyenne intra-classe. Ainsi, une grande distance relativement à une classe dense représente un grand éloignement, alors qu'une grande distance pour une classe « lâche » diminue l'éloignement du point. Le paramètre λ règle l'élargissement de la zone d'attraction, sa valeur est définie par Béreau dans l'intervalle [0,1]. La Figure 59 présente l'allure de la pondération en fonction de la distance relative $\left(\frac{d_j}{d^{m_i}}\right)$ entre la forme et un prototype, élevée au carré, pour différentes valeurs de λ. Si λ est proche de 0, l'ouverture de la courbe d'attraction est grande et on favorise la contribution des voisins, même éloignés, à la définition du degré d'appartenance du point considéré. A contrario, si λ est proche de 1, l'ouverture de la courbe d'attraction diminue et seule la contribution des voisins très proches est prise en compte.

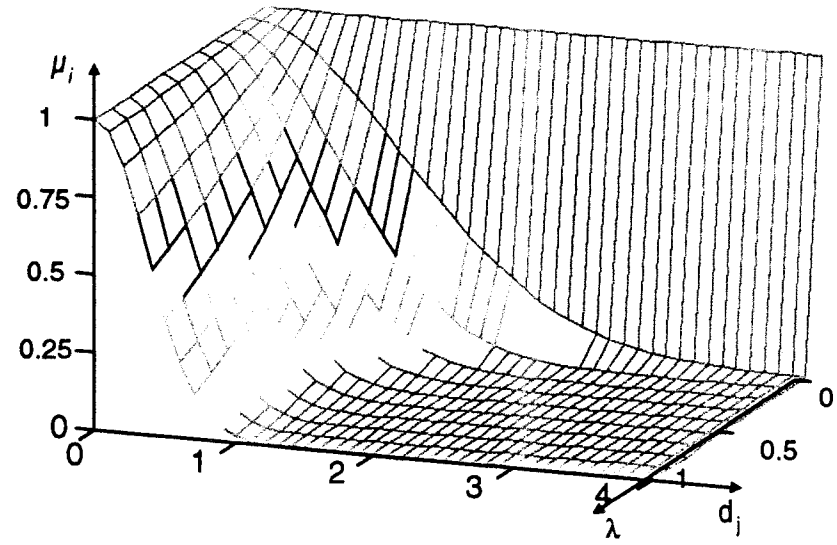


Figure 59 : Evolution de la zone d'attraction d'un prototype selon la valeur de λ

Cette fonction de discrimination permet d'introduire un rejet de distance des formes, lorsque celles-ci ne sont pas situées dans des zones d'attraction des prototypes de vraies ou fausses signatures.

La procédure de discrimination est résumée par l'algorithme ci-dessous, dans lequel la structuration initiale de l'ensemble d'apprentissage n'est pas présentée.

Algorithme de discrimination de BEREAU

1) Initialisation

- *Définition des paramètres : seuil d'affectation, λ, c, k, b_i*
- *Lecture des éléments de l'ensemble d'apprentissage*

2) Discrimination floue de l'ensemble d'apprentissage

- *Calcul de l'entropie logarithmique de la classe considérée à partir du résultat de la classification automatique sur l'ensemble d'apprentissage E_L*
- *Détermination pour chaque forme de ses k plus proches voisins dans E_L et calcul de μ aux classes floues selon Eq.12*
- *Calcul de l'entropie résultante de chaque classe H_i*
- ***3) Analyse et discrimination du point courant y de l'ensemble de test***
- *Détermination des k plus proches voisins de y dans E_L et dans les éléments de E_t affectés auparavant aux classes*

- *Calcul des μ selon l'équation (Eq.12)*
- *Affectation à la classe pour laquelle μ >seuil d'affectation*
- *Si pour une classe C_i, $\mu_i(y) \geq \mu_{i\min}$ alors :*

  intégrer y dans l'ensemble étiqueté, mise à jour des d^{m_i}, aller en 5)

- *Sinon rejeter le point*

4) Analyse des points rejetés

- *Si le nombre de points rejetés atteint le nombre maximal préfixé alors :*
  - *Si le nombre maximum de classes floues n'est pas atteint alors générer une nouvelle classe en utilisant Eq.1, remettre à jour les degrés d'appartenance des formes déjà classées, en tenant compte de cette nouvelle classe*

    Retour à 3).

  - *Sinon, intégrer ces points dans l'ensemble étiqueté, sans créer une nouvelle classe. Si ce n'est pas possible de fusionner deux classes.*
- *Analyser les points suivants, en calculant seulement leur degré d'appartenance aux classes actuelles.*

5) Fusionner deux classes

- *Si un certain nombre de points sont analysés alors :*
- *Examiner les classes deux par deux et fusionner les classes floues C_i et C_j si pour chaque forme étiquetée y, la différence entre $\mu_i(y)$ et $\mu_j(y)$ n'excède pas un seuil ε_{reg}.*
- *Retour à 3).*

5.1.3. Discussion

L'ensemble de cette procédure présente un certain nombre d'intérêts. La structure en deux étapes présente un cycle de raisonnement cohérent. La première étape définit et structure les connaissances, la seconde étape décide et apprend sur la base de la connaissance acquise. La règle générale de discrimination utilisée est celle des k plus proches voisins sous une forme modifiée par la théorie des ensembles flous.

L'algorithme ne demande pas de connaissances a priori pour fonctionner. Selon l'initialisation des nombreux paramètres, l'algorithme conduit à différentes partitions. Il n'y a pas de critère explicite pour optimiser la partition floue au cours de l'apprentissage permanent, ce qui permet l'apprentissage séquentiel. La condition d'orthogonalité n'est pas de mise dans cet algorithme, ce qui permet d'intégrer la notion de rejet, et offre la possibilité de détecter

l'évolution de la connaissance. Enfin, la fonction de discrimination floue nuance l'interprétation de l'affectation par un étiquetage graduel, permettant de prendre en compte le phénomène d'évolution durant l'apprentissage permanent, ainsi que le recouvrement possible des classes.

Cette méthode a été développée dans un objectif de diagnostic par reconnaissance de formes. Ces formes ne sont pas étiquetées et le nombre de classes n'est pas connu a priori. L'apprentissage débute donc par la structuration de l'ensemble d'apprentissage initial. L'algorithme de discrimination s'appuie ensuite sur ce partitionnement flou, ce qui suppose que les données soient suffisamment informantes pour spécifier les classes formées. En effet, les caractéristiques attachées aux classes sont déterminées en fin d'apprentissage initial et sont conservées pour la discrimination et l'apprentissage permanent.

Un grand nombre de paramètres sont à régler, pour un bon fonctionnement de l'algorithme, et conduisent à des partitions floues différentes. Ainsi, il est possible de typer le comportement de la procédure, bien que le contrôle explicite de ce comportement ne soit pas aisé. En effet, la représentativité des formes et leur ordre d'étiquetage ont également une influence sur ce comportement.

Nous notons également que parmi ces paramètres, un certain nombre de seuils permettent le rattachement des données à une des classes, et bien que le degré d'appartenance nuance ce rattachement, l'entropie chute brutalement. En effet, dans la procédure de Béreau, un choix inadéquat des valeurs de rattachement peut conduire à un échec de l'algorithme et à la définition d'une partition floue mal appropriée.

5.1.4. Le flou : assouplissement ou rigidification de la contrainte

L'action du flou, introduit dans la fonction de discrimination, peut être analysée selon deux points de vue. D'une part, comme dans beaucoup d'applications, le flou permet d'assouplir, de relaxer, une règle ordinaire. Cette relaxation intervient lorsque le seuil de décision d'affectation fixé sur les degrés d'appartenance est inférieur à 0.5. Ce cas est cependant trop tolérant, et on préfère généralement des niveaux d'appartenance plus significatifs pour considérer une affectation à une classe.

Lorsque le seuil d'affectation fixé est supérieur à 0.5, la règle devient plus stricte que celle des k plus proches voisins. En effet, plus on exige une appartenance forte pour réaliser l'affectation d'une forme, plus ses voisins doivent contribuer fortement, en valeur et en quantité, au calcul de son degré d'appartenance. Cette remarque est illustrée par la Figure 60, où pour avoir un degré d'appartenance égal à 1 sur la première courbe floue, il faut environ que les 4k/5 plus proches voisins de la forme à discriminer appartiennent totalement à la classe.

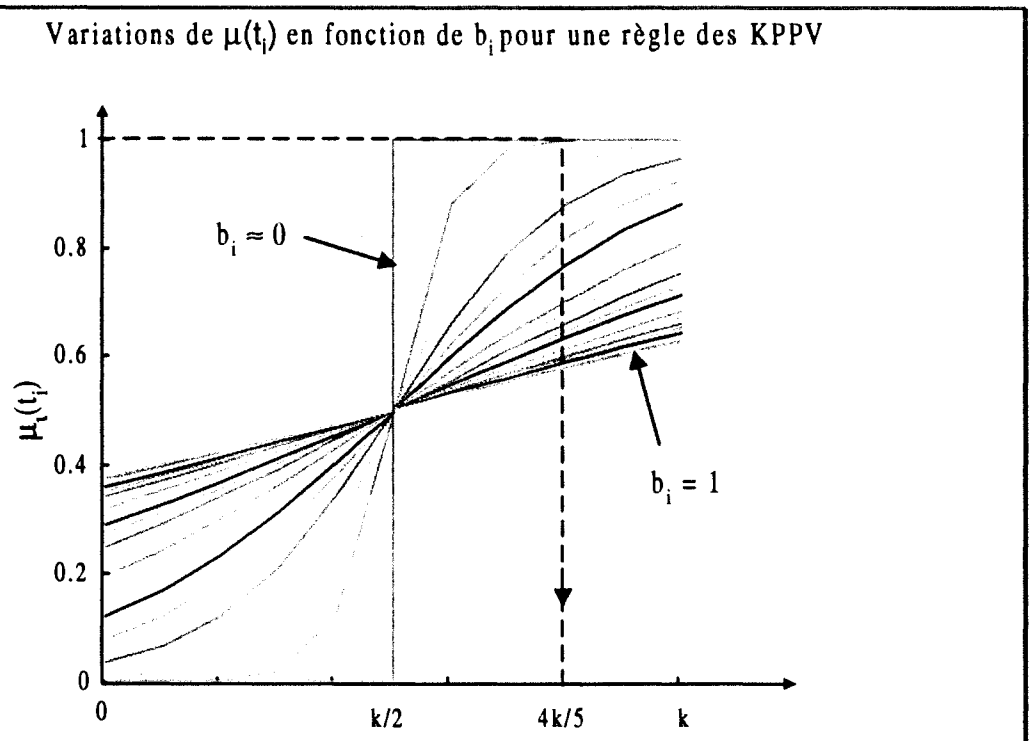


Figure 60 : Modification de la contrainte selon b_i

L'angle sous lequel on analyse l'influence du flou, dans cette règle de discrimination, nous conduit à voir une relaxation, un assouplissement de la règle des k plus proches voisins ordinaires (seuil de décision < 0.5) ou une rigidification de la règle selon la valeur de b_i (seuil de décision > 0.5). Cette dernière possibilité présente, pour notre application, beaucoup d'intérêt, puisque l'on peut moduler la force de la règle de décision suivant des facteurs tels que la représentativité du lot d'apprentissage, par rapport à la complexité du problème de discrimination.

6. Adaptation de la discrimination pour la vérification automatique de l'identité

Notre application est particulière et demande que l'on adapte cette méthode pour la rendre opérationnelle.

La partie essentielle que nous conservons est la phase de discrimination, et plus particulièrement la fonction de discrimination (Eq. 12) et la procédure d'acquisition permanente de connaissances.

En revanche, la phase de classification des données de référence n'est pas conservée puisqu'elle n'a pas lieu d'exister dans notre application. En effet, les quelques références que nous possédons à la mise en place du système de vérification sont entièrement supervisées. Si une phase préalable de structuration était réalisée sur l'ensemble d'apprentissage, il se pourrait que des signatures vraies soient classées fausses et vice-versa. Ceci entraînerait une mauvaise définition initiale de la partition. Le problème est abordé directement en discrimination de nouvelles formes.

La création/fusion de classes ne sera pas exploitée puisque nous ne disposons, avec certitude, que de deux classes.

Enfin, nous ne conserverons, comme nouveaux prototypes, que les données supervisées vraies et fausses signatures fournies au cours du fonctionnement du système de vérification, ainsi que les données non supervisées, ayant une appartenance à la classe des vraies signatures supérieure à un seuil prédéfini, et non réfutées par le superviseur du système de vérification. Cette restriction est guidée par le contexte d'application du système qui demande à sécuriser au mieux la décision.

6.1. Particularisation de la fonction de discrimination

Une fonction de discrimination particulière est associée à chaque classe (vraies et fausses signatures) et à chaque échelle de perception. Les paramètres b_i et λ permettent d'en adapter le comportement, selon l'incertitude d'affectation et l'imprécision de la perception.

6.1.1. Incertitude d'affectation

Comme nous l'avons présenté dans le §5.1.2, b_i quantifie le degré de flou de la partition initiale, il est représentatif de la quantité d'incertitude existant lors de la décision d'affectation à une classe [Dubuisson 90].

Dans notre application, l'étiquetage supervisé force les degrés d'appartenance aux valeurs limites 0 ou 1 et entraîne l'annulation de l'entropie H_i. Cette situation traduit une incertitude d'affectation nulle et implique une décision binaire.

Cependant, nous considérons que le nombre initial de prototypes est insuffisant pour affirmer que l'incertitude d'affectation est nulle. Dans ce cadre, la décision binaire est trop souple par rapport à la notion de sécurité imposée à la décision.

En effet, si l'on considère l'hypothèse de base de la règle des k plus proches voisins, pour converger vers une bonne estimation de la densité de probabilité d'une classe, il faut que le nombre d'échantillons soit infini. Cela signifie qu'autour d'une forme connue, il existe un volume de petite taille où la probabilité d'appartenance de la forme à une classe est la même que celle de ses voisins les plus proches. Si le nombre de prototypes est faible, alors le volume autour d'une forme peut devenir grand, et la probabilité d'erreur sur la décision est augmentée [Dubuisson 90], [Milgram 93], [Postaire 87]. La fonction de discrimination de Béreau possède l'avantage de limiter ce volume grâce à la pondération de distance de la contribution des voisins, mais n'annule pas l'incertitude d'affectation.

En outre, se pose le problème de la représentativité des prototypes par rapport à la dimension de l'espace de représentation. Ainsi nous savons que, plus la dimension de l'espace de représentation est grande, plus il est nécessaire de disposer d'un grand nombre de prototypes pour fournir une décision d'affectation certaine. Or, au départ du système, nous

disposons de peu de formes d'apprentissage pour fournir une décision d'affectation dans des espaces de représentation de dimensions 6 à 276.

Aussi, nous proposons de modifier la formulation de b_i, tout en conservant ses valeurs aux limites, pour intégrer au mieux ces deux notions. Nous proposons également d'évaluer la valeur de b_i en continu durant l'apprentissage permanent, afin de maintenir une information représentative de la connaissance apportée à chaque instant, par les prototypes réintégrés supervisés ou non.

Ainsi, la quantité b_i doit représenter l'entropie lorsqu'il y a suffisamment de prototypes représentatifs de chaque classe, et dans le cas contraire, une information quantifiant le niveau de représentativité.

6.1.2. Formulation de b_i

Dans [Cover 65] et [Milgram 93], les auteurs proposent une fonction f_i à valeurs dans [0,1] définie par l'équation (Eq. 14), informant sur la représentativité des prototypes de chaque classe C_i, par rapport à la dimension de l'espace de représentation :

$$\text{Si} \quad m > n+1 \quad f_i(m,n) = \frac{1}{2^{m-1}} \sum_{i=0}^{n} C_i^{m-1} \qquad \text{sinon} \quad f_i(m,n) = 1 \qquad \textit{(Eq. 14)}$$

où :

C_i^{m-1} *est l'opérateur de combinaison*

m est le nombre de prototypes représentatifs de la classe Ci,

n est la dimension de l'espace de représentation.

Lorsque le nombre de données disponibles, m, est inférieur à la dimension de l'espace de représentation, f_i prend la valeur 1. Puis lorsque m augmente, f_i tend vers zéro selon une courbe monotone décroissante (Figure 61). Plus f_i est grand, plus le système d'inéquations permettant de séparer les classes est sous-déterminé et peut être plus facilement satisfait. Mais, la séparabilité trouvée n'est pas forcément significative.

Bien que cette fonction soit destinée initialement à l'estimation de la séparabilité linéaire de classes, elle nous fournit une estimation optimiste de la représentativité des prototypes par rapport à l'espace de représentation [Milgram 93]. En effet, une estimation plus pessimiste pourrait avoir comme conséquence de transformer l'augmentation de la sécurité de la décision en absence totale de décision (rejet systématique des formes).

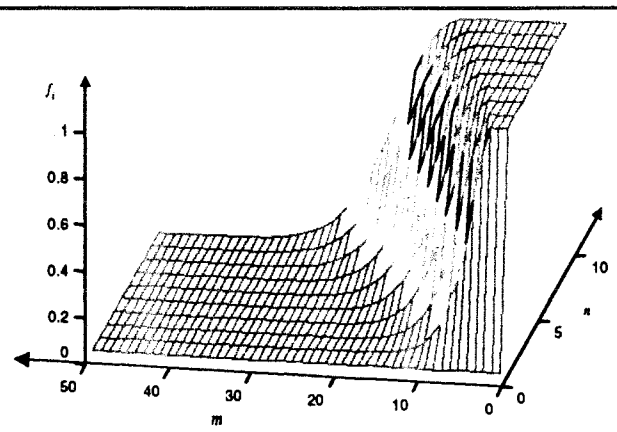
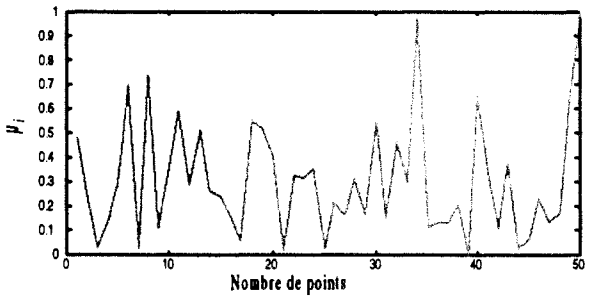


Figure 61 : Comportement de f_i selon n et m

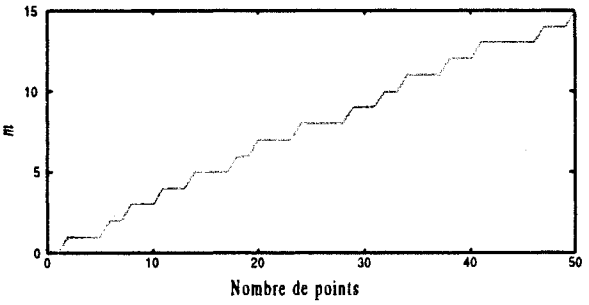
Une légère adaptation de f_i est cependant nécessaire, puisque sa formulation tient compte d'une participation complète des données à leur classe d'affectation. Or, dans notre cas, nous avons une attribution partielle définie par le degré d'appartenance. Ainsi, plus le degré d'appartenance à la classe est faible, moins le point contribue à la définition de la classe par rapport au problème de discrimination. Nous proposons alors d'introduire les degrés d'appartenance dans le calcul de $f_i(m, n)$ pour tenir compte de l'attribution partielle, m est défini par l'équation (Eq. 15) :

$$m = \textit{partie entière}\left(\sum \mu_i\right) \qquad \textit{(Eq. 15)}$$

La Figure 62 présente le comportement de f_i lorsque m est pris comme la contribution partielle des points à la définition de la classe C_i, exprimée par le degré d'appartenance à la classe.



Degré d'appartenance de formes à la classe i



Comportement de m selon l'équation (Eq.15)

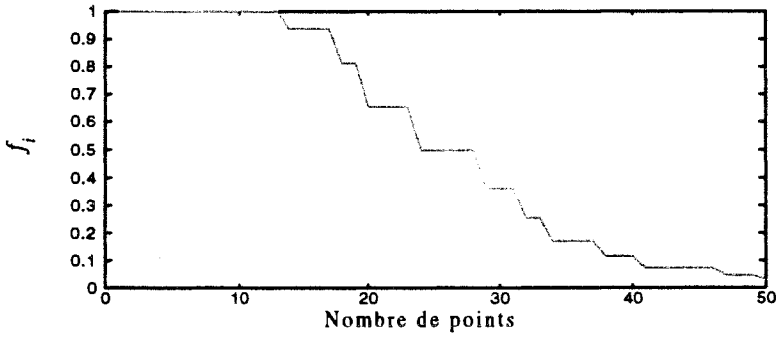


Figure 62 : Comportement de f_i en fonction de m pour un espace à 6 dimensions

Dans l'expression de la fonction de discrimination, on désire que b_i soit peu différent de 1 si le nombre de données par rapport au nombre de dimensions est faible, quelle que soit la valeur de l'entropie H_i . b_i doit être proche de l'entropie si le nombre de prototypes par rapport au nombre de dimensions est élevé. On désire également que b_i soit supérieur à l'entropie lorsque le problème n'est que moyennement déterminé par les prototypes. Le comportement désiré de b_i selon f_i et H_i peut être obtenu simplement par un *OU logique flou* entre f_i et H_i. Toutes les T-conormes différentes de max. peuvent être choisies pour cela. En effet, nous souhaitons avoir une contribution simultanée de f_i et H_i. Si l'on choisit la somme probabiliste nous obtenons pour b_i l'équation (Eq. 16) :

$$b_i = f_i + H_i - f_i H_i \qquad \textit{(Eq. 16)}$$

La Figure 63 présente le comportement de b_i, pour cet opérateur, selon f_i et H_i.

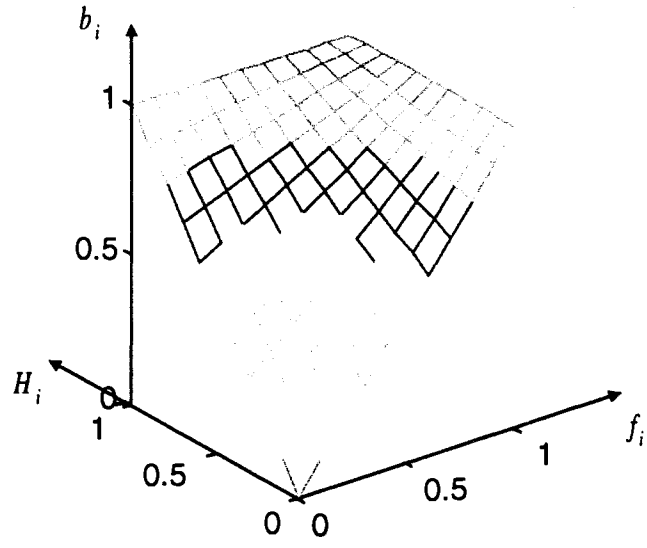


Figure 63 : Comportement de b_i

L'expression de b_i permet d'adapter la fonction de discrimination au fur et à mesure de l'apprentissage. A chaque nouveau point analysé, la fonction de discrimination et le niveau de sécurité évolueront avec la connaissance apportée. En effet, un point possédera un fort degré d'appartenance à la classe C_i, si la contribution de ses voisins est d'autant plus forte que b_i est faible. Cette considération convient parfaitement à la notion de sécurité pour le problème de vérification de signatures où la décision sera plus stricte initialement. Ensuite, lorsqu'une meilleure définition des classes sera obtenue, la décision sera plus souple.

6.1.3. Intégration de la pertinence de la perception

Le second point de particularisation de la fonction de discrimination concerne la prise en compte de l'imprécision engendrée par le niveau de perception associé. Comme nous l'avons précisé dans le §4.3, le niveau de résolution du motif influe sur l'appréciation de la ressemblance entre les formes.

Aussi, il convient d'adapter le calcul de la quantité t_i selon le niveau de perception, afin d'homogénéiser la réponse des différents classificateurs pour leur coopération ultérieure. Pour

cela, nous pouvons moduler l'étendue de la zone d'attraction centrée sur les prototypes, selon la valeur du paramètre λ.

Dans une résolution grossière, la distance moyenne entre les points d'une même classe (vraies signatures et fausses signatures aléatoires) est plus faible que dans une résolution plus fine. Ainsi, les classes sont vues plus compactes, elles se recouvrent plus facilement. Les faux par calque ou transfert optique sont, quant à eux, plus proches des signatures authentiques dans les résolutions fines que dans les résolutions grossières. Aussi, si l'on souhaite discriminer ou rejeter aussi efficacement ces fausses signatures dans les différentes résolutions, il convient d'adapter les paramètres favorisant la discrimination et le rejet pour les classes de vraies et fausses signatures.

6.2. Applications de la fonction de discrimination

Ce paragraphe a pour objectif de présenter les performances de la fonction de discrimination retenue, pour une vérification mono-motif avec décision binaire, en présence de faux aléatoires et faux par transfert optique. Dans notre analyse, nous nous intéressons à l'influence des paramètres de la fonction de discrimination, k, λ, et b_i, sur les performances de vérification.

Le facteur de formes utilisé est le Fuzzy Extended Shadow Code global, présenté dans le §3.5 du chapitre 2. Les fausses signatures par transfert optique sont obtenues par la procédure de modification de la pseudo-dynamique de la signature, présentée au §3.7.1 du chapitre 2. Pour nous placer dans les conditions les plus difficiles de vérification, nous ne présenterons au système que des faux correspondant à un coefficient de modification de δ=75%.

Enfin, le protocole de test est composé des protocoles d'évaluation des faux aléatoires, présentés au §2.3.3.3 du chapitre 1, et le premier protocole pour l'évaluation des performances du système pour des faux par transfert optique présenté au §3.7.2 du chapitre 2. Dans les deux cas, nous présentons le taux d'erreur moyen de vérification pour les erreurs de type I (vraies signatures prises pour des fausses) et de type II (fausses signatures prises pour des vraies). La possibilité de rejet de formes présente dans la fonction de discrimination locale, impose de présenter également les pourcentages de rejet moyen pour les deux types d'erreurs et pour les deux types de fausses signatures.

Les différents tableaux de synthèse proposés présentent donc pour les faux aléatoires les erreurs moyennes de vérification en pourcentage E_f1 et E_f2 ainsi que les taux de rejet associés R_f1 et R_f2, et pour les faux par transfert optique E_p1, E_p2, R_p1 et R_p2. Ces erreurs moyennes correspondent à :

- E_f1: Pourcentage de vraies signatures déclarées fausses,
- R_f1: Pourcentage de vraies signatures rejetées (demande de confirmation),

- E_f2: Pourcentage de faux aléatoires déclarés vrais,
- R_f2: Pourcentage de faux aléatoires rejetés (demande de confirmation),

- E_p1: Pourcentage de photocopie de vraies signatures déclarées fausses,
- R_p1: Pourcentage de photocopie de vraies signatures rejetées (demande de confirmation),

- E_p2: Pourcentage de photocopies de fausses signatures déclarées vraies,
- R_p2: Pourcentage de photocopies de fausses signatures rejetées (demande de confirmation).

Comme nous l'avions précisé dans le chapitre 2 §3.7.2, le taux E_p1 ne représente pas une erreur de vérification mais une vérification correcte.

Si l'on relie ces taux d'erreurs aux notions de sécurité et de confort, nous obtenons :

E_fT : Pourcentage d'erreur moyenne totale : $E_fT = \frac{E_f1 + E_f2}{2}$

Sécurité : **E_f2, R_f2, E_p1, R_p1, E_p2, R_p2.**

Confort : E_f1, R_f1.

Nos résultats ne pourront être analysés qu'en les comparant à ceux obtenus avec la règle de discrimination des k plus proches voisins classiques présentés dans les Tableaux 10, 12 et 13 du chapitre 2.

6.2.1. Influence de b_i

Comme nous l'avons précisé dans le §5.1.2 de ce chapitre, la valeur de b_i modifie la réponse floue de la fonction de discrimination, en limitant les valeurs maximale et minimale des degrés d'appartenance.

Lorsque la configuration permet, sans aucun doute possible, l'affectation d'une forme à l'une ou l'autre des deux classes de vraies ou de fausses signatures, la valeur de b_i fixe la valeur maximale du degré d'appartenance de la forme. Pour la classe d'affectation de la forme, une valeur de b_i proche de 0 pousse le degré d'appartenance vers 1; en revanche, une valeur de b_i proche de 1 l'attire vers la valeur 0.55. Pour l'autre classe (non affectation), si b_i est proche de 0, le degré d'appartenance de la forme est poussé vers 0, alors que si b_i tend vers 1, il est attiré vers 0.45.

L'augmentation de la valeur de b_i conduit donc à l'écrasement de l'échelle des degrés d'appartenance. Une valeur fixe du seuil d'affectation (0.65 par exemple) conduira, suivant les valeurs de b_i, pour une même configuration de formes, à affecter ou à rejeter une forme. En revanche, b_i ne peut pas faire passer une forme d'une classe à une autre.

Cette influence est donc bien à relier à la notion de certitude d'affectation. En effet, si un motif n'offre pas toutes les caractéristiques de confiance que l'on souhaite, b_i permet de

relativiser sa réponse. Cette possibilité est à exploiter dans l'utilisation conjointe ou coopérante de plusieurs échelles de perception. Les manipulations présentées dans les paragraphes suivants ont toutes été conduites avec une valeur de b_i égale à 0.01 et un seuil d'affectation de 0.65.

6.2.2. Influence du nombre de voisins

Afin de visualiser l'influence du nombre de voisins, sur les performances de vérification, nous avons réalisé des manipulations de vérification associées à 4 motifs de perception, 3 grossiers et 1 fin, de dimensions 6, 11, et 74, pour différentes valeurs de k (1, 3, 5 et 7). Nous n'avons pas poussé plus loin le nombre de voisins, afin de rester dans un contexte réaliste d'application. Les valeurs des paramètres de la fonction de discrimination sont fixes pour les différentes expériences, λ_{vraies} vaut 0.01, $\lambda_{fausses}$ prend la valeur 0.5. Ces réglages correspondent à une zone d'attraction très grande pour les vraies signatures et moyenne pour les fausses. Le Tableau 14 regroupe les résultats de ces différentes expériences.

		$\lambda_{vraies} = 0.01$, $\lambda_{fausses} = 0.5$, $b_i = 0.01$								
		Photocopies (δ 75%)				**Faux aléatoires**				
Motifs	*k*	E_p1	R_p1	E_p2	R_p2	E_f1	R_f1	E_f2	R_f2	E_fT
Motif *a* $\Re^6$	1	5.96	0	1.68	0	1.32	0	0.65	0	0.985
	3	6.82	0	2.33	0	1.57	1.35	0.40	0.79	0.985
	5	7.26	0	2.59	0	1.57	0	1.20	0	1.385
	7	7.78	0	2.96	0	1.5	0	1.63	0	1.565
Motif *h* $\Re^{11}$	1	4.18	0	1.49	0	0.57	0	0.30	0	0.43
	3	7.45	0	1.88	0	0.85	1.14	0.09	0.62	0.47
	5	6.14	0.16	2.01	0.05	0.85	0	0.70	0.01	0.775
	7	6.88	0.33	2.41	0.15	1.11	0	1.06	0.07	1.085
Motif *b* $\Re^{11}$	1	3.10	0	0.98	0	0.46	0	0.36	0	0.41
	3	4.04	0	1.31	0	0.85	3.03	0.16	0.61	0.505
	5	4.86	0	1.77	0.1	0.85	0	0.71	0.06	0.78
	7	4.68	0.14	2.26	0.25	0.71	0.39	1.24	0.20	0.975
Motif *e* $\Re^{74}$	1	1.25	0	0.60	0	0.43	0	0.18	0	0.305
	3	1.18	0	0.77	0	0.5	0	0.25	0	0.375
	5	1.18	0	1.06	0	0.5	0	0.46	0.08	0.08
	7	1.32	0	1.52	0	0.43	0.07	0.8	0.39	0.39

Tableau 14 : Influence du nombre de voisins k sur le taux d'erreur de vérification en %

Nous pouvons, dans un premier temps, nous intéresser aux performances de vérification selon la nature des fausses signatures.

Pour les faux aléatoires, on retrouve les remarques faites dans le chapitre 2, §3.6, c'est-à-dire que E_f2 augmente avec le nombre de voisins. Le pourcentage d'erreur moyenne totale $\frac{E_f1+E_f2}{2}$ est globalement identique à celui obtenu avec la règle des k plus proches voisins ordinaire, si l'on excepte les rejets R_f2 et R_f1. En effet, on constate que R_f1 et R_f2, bien que très faibles, ne sont pas toujours nuls.

Pour les faux par transfert optique, les plus mauvais résultats sont obtenus systématiquement pour un seul voisin. Si l'on compare ces résultats avec ceux du Tableau 12 du chapitre 2, on constate que l'augmentation de k a un effet inverse dans les deux expérimentations. En effet, avec la règle des k plus proches voisins classique, la meilleure performance est obtenue avec un seul voisin, elle décroît avec l'augmentation de k. De plus, cette performance pour 1 voisin peut être considérée comme tout-à-fait correcte en regard de la difficulté du problème.

En revanche dans cette expérience, la performance croît avec k pour l'indicateur E_p2 alors que pour E_p1, aucune tendance ne se dégage sur 4 motifs.

Cependant, si les valeurs de E_p1 se situent dans le même ordre de grandeur que les résultats correspondant dans le Tableau 12 du chapitre 2, on constate que pour 1 voisin, les performances ont notablement chuté.

Nous ne sommes pas en mesure d'expliquer ce phénomène qui va à l'encontre de la logique de fonctionnement de la règle du 1 plus proche voisin. En effet, si dans le cas binaire, une forme a comme plus proche voisin un faux, dans le cas flou, ce faux reste le plus proche voisin. Le rejet de distance peut annuler la contribution de ce voisin, et la différence de performances entre les deux règles devrait normalement être transférée dans le taux de rejet R_p1. Or, ce n'est pas le cas dans ces 4 expériences. Aussi, nous continuons nos investigations pour comprendre les raisons de ce phénomène.

La stabilisation du taux d'erreur E_p1 pour les valeurs k=3, 5, 7, peut toujours s'expliquer par l'influence du rejet de distance dans le calcul des degrés d'appartenance. Si l'on augmente le nombre de voisins dans la fonction de discrimination, on étend la zone de l'espace de représentation où l'on recherche ces voisins. Aussi, si l'on dépasse la limite du rejet de distance, ces voisins n'interviennent pas ou peu dans le calcul du degré d'appartenance. Dans ces expérimentations, nous avions pourtant pris la précaution de choisir une valeur λ_{vraies} qui nous place dans une configuration optimiste, puisque le rejet de distance est repoussé à ses limites.

Le taux E_p2 croît avec la valeur de k, ce qui traduit le fait que des prototypes de fausses signatures sont assez proches des authentiques pour que, lorsqu'on étend la zone de recherche des plus proches voisins, les authentiques soient majoritaires. Ce phénomène est local, il concerne des formes situées à la limite des deux classes.

En ce qui concerne le pouvoir discriminant des motifs, nous retrouvons le même type de résultats que dans le Tableau 12 du chapitre 2, c'est-à-dire que les motifs grossiers permettent de mieux discriminer les faux par transfert optique que les motifs fins. Ceci est dû au fait que pour les motifs fins, les faux par transfert optique sont beaucoup plus près des vraies signatures que les faux aléatoires. Le seul moyen de les discriminer sera de les rejeter.

Dans la suite des expérimentations que nous conduiront, nous fixons le nombre de voisins à 5. Ce choix est dicté par le fait que nous souhaitons nous affranchir de variations trop brutales de l'estimation locale de la densité de probabilité des classes de vraies et de fausses signatures, sans pour autant augmenter les taux de rejet. En effet, si la taille du domaine de recherche des plus proches voisins est trop importante, les contributions des prototypes seront

d'autant plus faibles qu'ils sont éloignés. Ainsi, k=5 est un compromis acceptable dans ces conditions.

6.2.3. Influence du facteur de formes λ_{vraies}

Dans ces expérimentations, nous n'avons conservé que 3 motifs, deux grossiers les motifs *a*, *h* et un fin, le motif *e*. Le facteur d'étalement de la zone d'attraction des fausses signatures est placé au centre de sa plage de variation. Le nombre de voisin est pris égal à 5. Nous faisons varier la valeur du facteur d'étalement de la zone d'attraction des vraies signatures entre ses bornes minimale et maximale, selon 3 modalités, *0.01*, *0.4* et *0.7* (Tableau 15).

		$\lambda_{fausses} = 0.5$, $k = 5$, $b_i = 0.01$								
		Photocopies (δ 75%)				**Faux aléatoires**				
Motifs	λ_{vraies}	E_p1	R_p1	E_p2	R_p2	E_f1	R_f1	E_f2	R_f2	E_fT
Motif *a* $\Re^6$	0.01	7.26	0	2.59	0	1.57	0	1.20	0	1.385
	0.4	7.21	12.89	0.82	1.75	0.67	1.96	0.38	1.17	0.525
	0.7	7.21	29.17	0.37	2.20	1.57	6.21	0.22	0.98	0.895
Motif *h* $\Re^{11}$	0.01	6.14	0.16	2.02	0.05	0.85	0	0.70	0.01	0.775
	0.4	6.14	10.41	0.34	1.72	0.82	0.82	0.09	0.83	0.455
	0.7	6.64	34.6	0.14	2.10	0.85	6.11	0.02	0.69	0.435
Motif *e* $\Re^{74}$	0.01	1.17	0	0.77	0	0.5	0	0.46	0.08	0.48
	0.4	1.17	0	1.06	0	0.5	0.25	0.05	0.5	0.275
	0.7	1.32	0	1.52	0	0.5	4.60	0	0.54	0.25

Tableau 15 : Influence du poids λ_{vraies}, pour 5 voisins , sur le taux d'erreur de vérification en %

Pour les faux aléatoires, l'augmentation de λ_{vraies} fait diminuer le pourcentage d'erreur moyenne totale E_f1+E_f2, et le place en deç‡ des valeurs obtenues dans les mêmes conditions avec la règle des k plus proches voisins ordinaires (motif *a* : $\varepsilon_t=1.373$, motif *h* : $\varepsilon_t=0.757$ et motif *e* : $\varepsilon_t=0.507$). Les meilleures performances sont obtenues pour λ_{vraies} proche de 1, mais on tend à transformer l'erreur en rejet. En effet, le rejet (R_f1 et R_f2) augmente avec la valeur de λ_{vraies} et prend des valeurs assez importantes pour λ_{vraies} proche de 1. Le rejet le plus important porte sur les vraies signatures, ce qui ne va pas dans le sens du confort, mais est satisfaisant pour la sécurité de la décision.

Pour les faux par transfert optique, nous constatons que la valeur de l'indicateur E_p1 n'évolue quasiment pas avec les valeurs de λ_{vraies}, alors que le rejet R_p1 augmente notablement avec λ_{vraies} pour les motifs grossiers. Pour le motif fin, ce rejet est inexistant, le motif *e* est en cela caractéristique de l'ensemble des motifs fins.

En observant l'indicateur E_p2, on constate que l'augmentation de λ_{vraies} fait diminuer le nombre d'erreurs de discrimination des photocopies de fausses signatures pour les motifs grossiers alors qu'il l'augmente pour le motif fin, sans que l'on puisse donner une raison particulière. Le rejet R_p2 évolue logiquement avec l'augmentation de λ_{vraies} pour les motifs grossiers, alors qu'il est nul pour le motif fin.

Cette expérimentation est extrêmement informante, puisque nous voyons que pour une valeur de λ_{vraies} de 0.7, la somme de E_p1 et R_p1 pour les motifs *a* et *h*, vaut respectivement 36.38 % et 41.24 % de discrimination de fausses signatures. Ce résultat est particulièrement intéressant, cependant, on peut regretter que la discrimination ne soit pas uniquement portée par E_p1, ce qui irait dans le sens du confort.

Parallèlement, pour cette même valeur de λ_{vraies}, la discrimination des faux aléatoires est également tout à fait satisfaisante puisque l'on obtient, toujours pour les motifs *a* et *h* une erreur moyenne totale de 0.895% et 0.435 %. De plus, la faible capacité de discrimination des faux par transfert optique pour des motifs fins est confirmée.

6.2.4. Influence du facteur $\lambda_{fausses}$

Dans cette expérimentation, nous souhaitons évaluer l'influence de $\lambda_{fausses}$ de la classe des fausses signatures sur les performances de vérification. Etant donnée la faible capacité des motifs fins à discriminer les faux par transfert optique, nous focalisons ici notre attention sur les motifs grossiers *a*, *b* et *h*. Nous choisissons la valeur 0.7 pour la classe des vraies signatures, ce qui correspond à une zone d'attraction relativement faible (Tableau 16).

		$\lambda_{vraies} = 0.7$, $k = 5$, $b_i = 0.01$								
		Photocopies (δ 75%)				**Faux aléatoires**				
Motifs	$\lambda_{fausses}$	E_p1	R_p1	E_p2	R_p2	E_f1	R_f1	E_f2	R_f2	E_fT
Motif *a*	0.01	7.21	29.17	0.37	2.20	1.57	6.21	0.22	0.98	0.895
	0.5	7.21	29.17	0.37	2.20	1.57	6.21	0.22	0.98	0.895
	0.9	6.89	29.5	0.37	2.25	1.57	6.21	0.22	0.99	0.895
Motif *h*	0.01	6.31	34.50	0.15	1.86	0.86	6.1	0.02	0.68	0.44
	0.5	6.64	34.6	0.14	2.10	0.86	6,11	0.02	0.69	0.44
	0.9	5.90	34.91	0.15	2.31	0.86	6.11	0.02	0.86	0.44
Motif *b*	0.01	4.85	27.92	0.12	1.64	0.85	7.07	0.03	0.67	0.44
	0.5	4.85	27.92	0.12	1.74	0.85	7.07	0.03	0.73	0.44
	0.9	4.14	28.64	0.12	2.12	0.71	7.21	0.03	1.03	0.37

Tableau 16 : Influence du poids $\lambda_{fausses}$, pour 5 voisins, sur le taux d'erreur de vérification en %

La principale influence de $\lambda_{fausses}$, pour les fausses signatures, s'observe dans le cas des faux aléatoires sur le taux de rejet R_f2 qui augmente avec $\lambda_{fausses}$. En effet, si l'on augmente la taille de la zone d'attraction des fausses signatures, ce qui correspond à une diminution de la valeur de $\lambda_{fausses}$, on augmente la contribution des fausses signatures dans le calcul des degrés d'appartenance, il est donc plus facile d'affirmer que la signature est fausse. On constate que le rejet ne concerne pas des signatures fausses, qui étaient déclarées vraies, mais celles qui étaient déclarées fausses, puisque le taux E_f2 n'évolue pas.

En ce qui concerne E_f1, on observe une variation uniquement sur le motif *b* pour $\lambda_{fausses} = 0.9$, les autres situations n'étant pas influencées. Cette variation n'est pas illogique, puisque si l'on diminue la zone d'attraction des fausses signatures, celles-ci interviennent moins dans le calcul du degré d'appartenance si elles sont sélectionnées comme voisins, et les formes

peuvent être déclarées vraies, ou rejetées si la contribution des vraies signatures n'est pas suffisante. On observe ce phénomène sur R_f1 qui n'est modifié que dans le motif *b* pour $\lambda_{fausses} = 0.9$.

Pour les faux par transfert optique, l'influence de $\lambda_{fausses}$ est plus marquée sur tous les indicateurs sauf E_p2. L'indicateur E_p1 diminue lorsque $\lambda_{fausses}$ augmente, c'est-à-dire que la contribution des plus proches voisins faux diminue. Il est alors plus difficile de déclarer fausse une photocopie de vraies signatures, elle peut être alors rejetée ou déclarée vraie. En l'occurence, si l'on observe R_p1, on constate qu'une majorité de ces signatures est rejetée. Ce comportement va dans le sens de la sécurité, il n'est donc pas trop problématique. Les indicateurs E_p2 et R_p2 indiquent, quant à eux, que les faux photocopiés ne sont pas non plus considérés comme vrais lorsque $\lambda_{fausses}$ augmente, mais ils sont rejetés.

Pour la suite de nos expérimentations, nous choisirons $\lambda_{fausses}$ de la classe des fausses signatures dans l'intervalle [0.5, 1] afin de maximiser le rejet R_p1.

6.2.5. Synthèse des pouvoirs discriminants

Nous avons regroupé dans le Tableau 17, les performances individuelles de l'ensemble des motifs pour les valeurs suivantes des paramètres $\lambda_{vraies} = 0.7$, $\lambda_{fausses} = 0.5$ et 5 voisins.

	$\lambda_{vraies}=0.7$, $\lambda_{fausses}=0.5$, $k = 5$, $b_i=0.01$								
	Photocopies (δ 75%)				**Faux aléatoires**				
Motifs	E_p1	R_p1	E_p2	R_p2	E_f1	R_f1	E_f2	R_f2	E_fT
a	7.21	29.17	0.37	2.20	1.57	6.21	0.22	0.98	0.895
h	6.64	34.6	0.14	2.10	0.85	6.11	0.02	0.70	0.435
b	4.85	27.92	0.12	1.74	0.85	7.07	0.03	0.73	0.44
i	9.64	29.07	0.33	2.91	1.64	6.14	0.06	1.18	0.85
c	0.89	26.57	0.03	0.85	0.25	8.28	0.01	0.46	0.13
j	05	15.28	0	0.66	0.07	7.53	0	0.50	0.035
d	2	43.14	0	1.61	0.89	5.53	0	0.37	0.445
k	1.63	27	0	1.22	0.25	7.64	0	0.46	0.125
e	1.17	0	1.06	0	0.5	4.61	0	0.54	0.25
l	0.28	0	0.52	0	0.14	7.35	0	0.50	0.07
f	4.36	32.25	0.15	2.01	2.03	6.75	0.01	0.95	1.02
g	7.53	26.93	0.49	2.79	1.71	7.5	0.19	1.93	0.95
m	0.96	0	0.71	0	0.60	6.0	0.01	0.60	0.30
n	1.08	0	0.68	0	0.12	7.28	0	0.49	0.06
o	0.39	0	0.60	0	0.07	8.32	0	0.47	0.03

Tableau 17 : Récapitulatif des performances par motif

Pour les valeurs choisies des paramètres, on constate tout d'abord que les performances de discrimination des faux aléatoires E_f2 sont relativement homogènes et très faibles, pour l'ensemble des motifs. Les valeurs maximales sont obtenues pour les motifs grossiers *a*, *g* et *i*. Plusieurs motifs se partagent la valeur minimale. De plus, le rejet R_f2 reste tout à fait limité pour l'ensemble des motifs.

En ce qui concerne E_f1, les plus mauvaises performances (supérieures à 1%) sont obtenues pour les motifs *a*, *i*, *f* et *g*. Le rejet de vraies signatures est dans l'ensemble assez important allant de 5.53% (motif *d*) à 8.32% (motif *o*), on ne peut pas tirer ici de conclusion très marquée. Le confort se trouve pénalisé pour l'ensemble des motifs.

Le taux d'erreur moyenne totale de discrimination est inférieur ou égal à 1% pour l'ensemble des motifs. Les valeurs maximales se trouvent pour les motifs grossiers où les trois plus mauvais sont les motifs *a*, *f* et *g*, les valeurs minimales pour les motifs fins *m*, *n*, *o* ou pour les motifs *c*, *e*, *j*, *k*, *l*. On retrouve l'ordre de pertinence des motifs, établi au chapitre 2.

La discrimination des faux par transfert optique est nettement moins à l'avantage de ces derniers motifs, puisque l'ordre de pertinence tend à s'inverser pour E_p1. En effet, les motifs grossiers ont une meilleure capacité à discriminer les faux par transfert optique, et présentent des taux de rejets de ces faux assez importants. Ces caractéristiques vont dans le sens de l'augmentation de la sécurité de la décision. L'erreur moyenne E_f2 est faible pour l'ensemble des motifs, et le rejet R_p2 est sensiblement plus fort dans les motifs grossiers que dans les fins.

6.2.6. Conclusion

En conclusion, il se confirme donc que les motifs fins sont plus performants pour la discrimination des faux aléatoires, bien que les motifs grossiers aient des performances tout à fait honorables, puisque supérieures à toutes celles obtenues auparavant si l'on excepte le rejet.

Les motifs grossiers ont en plus une capacité certaine à discriminer les faux par transfert, en les déclarant faux ou en les rejetant. Bien que ce comportement puisse être jugé satisfaisant selon notre critère de sécurité, il serait intéressant que le rejet puisse se transformer en une affectation à la classe des fausses signatures afin d'améliorer le confort du système.

7. Une discrimination intégrée pour la vérification

7.1. Introduction

L'approche multi-échelle nous amène à percevoir les images de signatures selon des points de vue différents, caractérisés par la résolution et la forme des motifs. Plutôt que de retenir le motif de perception le plus significatif pour un scripteur donné, il est plus judicieux de profiter du point de vue de chacun des motifs. La coopération de classificateurs vise à améliorer la décision, tant au niveau des performances, qu'au niveau de la sécurité exigée par la vérification de signatures.

Comme nous l'avons précisé dans le chapitre 1, la coopération parallèle exploite les capacités de discrimination de différents classificateurs pour fournir une décision plus certaine, par le vote majoritaire démocratique ou pondéré des classificateurs. Dans le vote majoritaire démocratique, chaque méthode de classification utilisée contribue avec le même poids à la décision finale prise sur la forme à discriminer. Le vote majoritaire pondéré tient compte d'une pertinence plus affinée de certains classificateurs par rapport au problème. Comme ils conduisent plus naturellement à la bonne discrimination, ils sont favorisés lors de la décision. Ainsi, quelle que soit la coopération parallèle, chaque classificateur prend une décision binaire, qui lui est propre, selon sa connaissance et à partir de la perception qu'il a des formes à discriminer. Cette décision aveugle ne tient pas compte de la façon dont les autres classificateurs appréhendent le problème de reconnaissance.

Dans notre application, les classificateurs ressortent tous de la même approche, et opèrent dans des espaces de représentation différents, pour discriminer les mêmes formes à partir des mêmes prototypes. L'approche de coopération proposée par R. Sabourin a un fonctionnement qui s'apparente à celui d'un jury d'experts graphologues, munis chacun d'une loupe avec un grossissement spécifique, et d'un ensemble d'exemples de vraies et fausses signatures. Ces experts seraient chargés de fournir une réponse sur la validité de la signature par rapport à leurs connaissances et leurs points de vue. La décision finale est prise à la majorité.

Par analogie avec le comportement d'un expert graphologue opérant une vérification, nous proposons une approche différente de coopération :

Un seul expert disposant de 15 loupes différentes et d'un ensemble d'exemples de vraies et fausses signatures, examine chaque nouvelle signature selon les différentes focales des loupes. Chaque niveau d'analyse lui fournit une tendance qu'il corrobore avec les autres tendances, pour établir une proposition informée selon chacun des niveaux d'analyse (grossier à fin), qu'il agrège pour fournir sa réponse. Si toutes les propositions ne concourent pas vers une décision certaine, il exprime son incapacité à décider.

Dans cette structure de discrimination intégrée, on fait apparaître deux coopérations à deux niveaux de la méthode.

La première coopération apparaît dans "*Chaque niveau d'analyse lui fournit une tendance qu'il corrobore avec les autres tendances*". Il s'agit ici de faire coopérer les différentes fonctions de discrimination, selon le principe introduit dans le §3.2, avant qu'elles ne proposent localement un degré d'appartenance à chacune des classes. Nous appelons cette coopération : coopération intra-discrimination ou intra-classificateur.

La seconde coopération apparaît dans : "*une proposition informée selon chacun des niveaux d'analyse (grossier à fin), qu'il agrège pour fournir sa réponse*". Cette coopération dite "coopération extra-classificateur" reprend le principe d'agrégation du vote majoritaire, mais est basée sur des propositions plutôt que sur des décisions.

Le paragraphe suivant présente en détail ces deux propositions de coopération.

7.2. Intégration des discriminations

7.2.1. La coopération extra-classificateur

Avec l'utilisation de l'approche floue des k plus proches voisins, chaque classificateur établit une décision sur l'incriminé, à partir de mêmes voisins, quel que soit le niveau de perception. Selon la définition des paramètres de l'algorithme, chaque classificateur propose une affectation sous la forme de degrés d'appartenance aux classes des vraies et fausses signatures. L'étiquette de l'incriminé, exprimée au travers du degré d'appartenance, est fournie par la fonction de discrimination.

Une prise de décision locale attachée à chaque niveau de perception donne de l'importance aux erreurs locales de décision et l'appartenance graduelle donnée par la règle des k plus proches voisins flous est perdue. Puisque chaque classificateur attribue à la forme à discriminer un degré d'appartenance à chaque classe, on peut substituer la décision de chaque classificateur au profit d'une proposition locale d'appartenance, et concevoir la coopération des propositions comme un k plus proches classificateurs. Ainsi, l'importance des décisions locales erronées est réduite.

En concevant la coopération des classificateurs pour la décision finale comme une règle des k plus proches classificateurs, nous pouvons directement adapter la règle de discrimination proposée dans l'algorithme des k plus proches voisins flous (Eq. 12). Cependant, les quantités b_i, t_i et k ne peuvent avoir la même signification. Le paramètre k déterminant le nombre de voisins considérés dans la règle, est transformé en un nombre K de classificateurs impliqués.

La quantité t_i détermine initialement la contribution des voisins à la définition du degré d'appartenance de la forme à discriminer. Elle conserve donc la même fonctionnalité, mais traduit maintenant la proposition des classificateurs sous forme de degré d'appartenance aux classes vraies et fausses signatures.

Enfin, la quantité b_i traduisait l'incertitude d'affectation, par l'entropie floue. Elle conserve la même fonction mais se rapporte au taux d'erreurs normalisé de vérification. Ainsi, plus on commet d'erreurs, plus il faudra de consensus dans la décision (considérer 10 propositions sur 15, voire plus, plutôt que 8 sur 15), et inversement si peu d'erreurs sont commises.

On obtient par cette procédure une adaptation de la décision selon le taux d'erreur de vérification, visant à augmenter la sécurité de la décision finale. La fonction de discrimination globale est la suivante (Eq. 17) :

$$\mu_i(y) = \frac{1}{1+\exp\left(\frac{\frac{K}{2}-t_i}{b_i}\right)} \quad \text{avec} \quad t_i = \sum_{j=1}^{K} \mu_{i,Dj}(y) \qquad \textit{(Eq. 17)}$$

où :

$i \in \{vraies, fausses\}$

$\mu_i(y)$ est le degré d'appartenance de la forme y à la classe i,

K est le nombre de classificateurs choisis,

$\mu_{i,D_j}(y)$ est la proposition d'appartenance de y à la classe i, fournie par la fonction de discrimination D_j du classificateur j

b_i est le taux d'erreurs de vérification,

$$b_i = n_i/n_{total} \qquad (Eq.\ 18)$$

avec n_i le nombre d'erreurs commises pour la classe i et n_{total} le nombre total de formes vérifiées.

7.2.2. Expérimentations

Nous présentons dans ce paragraphe, deux expériences permettant de montrer l'intérêt de la coopération de classificateurs associés à différents niveaux de perception de la signature. Pour cela, nous avons sélectionné 2 ensembles de 3 motifs, fondamentalement différents dans leurs caractéristiques et leurs performances évaluées dans le paragraphe précédent. Le premier ensemble regroupe les motifs *a*, *f* et *g*, le second est constitué des motifs *m*, *n* et *o*. Pour les deux expériences, nous avons conservé des valeurs identiques de paramètres, c'est-à-dire $k=5$, $\lambda_{vraies}=0.7$ et $\lambda_{fausses}=0.7$.

Nous n'avons pas relativisé les réponses des différentes fonctions de discrimination car nous n'avons pas pris le risque d'une mauvaise analyse des résultats en multipliant les sources de modification. De plus, les motifs en présence possèdent à notre sens le même niveau de pertinence. Aussi, nous fixons la valeur des b_i des discriminations locales.

Les paramètres b_i de la fonction de discrimination globale sont calculés à chaque nouvelle forme discriminée, selon l'équation (Eq. 18). Ils correspondent au taux d'erreur de vérification pour les classes des vraies et des fausses signatures. Le seuil d'affectation global est fixé à *0.65*.

Le Tableau 18 présente les résultats obtenus par la coopération des motifs *a*, *f* et *g*.

	$\lambda_{fausses} = 0.7$, $\lambda_{vraies} = 0.7$, $k = 5$							
	Photocopies (δ 75%)				**Faux aléatoires**			
Coopération	E_p1	R_p1	E_p2	R_p2	E_f1	R_f1	E_f2	R_f2
Motifs *a, f et g*	34.87	0	0.11	1.09	3.59	0	0.04	0.6

Tableau 18 : Coopération extra-classificateurs, motifs a, f et g

La première remarque à faire concerne E_p1 et R_p1, puisque l'on constate immédiatement que R_p1 est nul alors que E_p1 a pour valeur 34.87%. Ainsi, la coopération de

ces trois motifs, qui n'avaient pas un taux d'erreur moyenne E_p1 supérieur à 7.53% (Tableau 17) mais de grosses capacités de rejet, permet la transformation du rejet en discrimination réelle. Les résultats sont également concluants pour E_p2 et R_p2. Les performances sont améliorées puisque l'on diminue le pourcentage d'erreur moyenne d'affectation de fausses signatures photocopiées à la classe des vraies, tout en diminuant le rejet.

Ce résultat est particulièrement intéressant car il signifie que la coopération a exploité les propositions et a levé les difficultés des discriminations locales. L'affectation est plus prononcée, la sécurité est améliorée, mais au détriment du confort.

Pour les faux aléatoires, les résultats sont plus partagés. En effet, le pourcentage d'erreur moyenne E_f2 ainsi que le rejet R_f2 sont inférieurs aux valeurs obtenues motif par motif, alors que l'erreur E_f1 a notablement augmenté (cf. Tableau 17). L'erreur moyenne totale s'en trouve donc pénalisée. L'objectif d'amélioration de la sécurité est donc atteint pour les deux types de fausses signatures, mais on note une perte de confort pour les faux aléatoires.

Le Tableau 19 présente la même expérience conduite sur les motifs *m*, *n* et *o*.

	$\lambda_{fausses} = 0.7$, $\lambda_{vraies} = 0.7$, $k = 5$							
	Photocopies (δ 75%)				**Faux aléatoires**			
Coopération	E_p1	R_p1	E_p2	R_p2	E_f1	R_f1	E_f2	R_f2
Motif *m, n et o*	19.38	0	0	0.88	6.4	0	0	0.52

Tableau 19 : Coopération extra-classificateur, motifs m, n et o

On peut faire ici les mêmes remarques que précédemment, puisque la coopération des motifs permet de discriminer 19,38 % de faux par transfert alors que les motifs pris un à un en sont strictement incapables (cf. Tableau 17). On ne diminue pas la performance de discrimination des fausses signatures représentée par E_f2; en revanche, on déclare fausses des signatures vraies dans des proportions plus importantes. Là aussi, on gagne en sécurité, mais on perd en confort et un compromis devra être trouvé selon le contexte d'application.

7.2.3. La coopération intra-classificateur

L'apprentissage d'un classificateur se fait en aveugle par rapport aux autres sur son niveau de perception. Or, pour notre problème, chaque classificateur possède les mêmes références, définies de façon supervisée, à des niveaux de perception différents et doit exprimer une décision sur le même objet. Afin d'éliminer cette coopération externe et aveugle, nous proposons de réaliser une coopération interne des classificateurs.

Tous les classificateurs sont basés sur le même type de règle de décision avec leurs propres paramètres, et possèdent la même structure de fonctionnement. Dans le principe des k plus proches voisins, chaque classificateur va rechercher les k plus proches voisins selon sa propre perception à son niveau de résolution. Les voisins déterminés ne sont pas forcément les

mêmes suivant les différents classificateurs. Or, au niveau physique, les plus proches voisins de la signature sont bien définis et uniques. Aussi, pour décider sur les mêmes sources d'information, nous proposons d'utiliser les mêmes voisins au sens multi-échelle pour les différents classificateurs. On peut voir ce principe comme une mise en commun d'informations, ressemblant au principe de l'ingénierie concourante. Les voisins multi-échelles se rapprochent du concept de formes fortes qui sont clairement caractéristiques du problème et sont potentiellement plus voisines de l'image de la signature. A cette fin, nous proposons une procédure de coopération que l'on appellera coopération interne, réalisée en 2 étapes :

- Détermination des voisins, au sens de la distance, dans chaque résolution,
- Détermination des voisins forts, au sens de tous les niveaux de perception.

Pour déterminer les voisins multi-échelles, il est nécessaire que chaque classificateur fournisse la liste des voisins, selon son point de vue et la forme à discriminer. Puis, sur ces propositions de voisins, une liste de voisins multi-échelles est établie en recherchant les voisins dont l'indice de sélection, en terme de proximité, est minimum dans chaque résolution. Il en résulte automatiquement qu'un proche voisin multi-échelle a une somme d'indices de proximité minimale. La Figure 64 présente ce principe de détermination de voisins multi-échelles.

Points par odre d'apparition dans l'ensemble de test	Distance dans R^6, indice de classement au plus proche		Distance dans R^{276}, indice de classement au plus proche	Sélection par Σ indice et σ indice
1	1	-------	3	4()
2	3		4	7()
3	5		1	6()
4	6		2	8()
.	.		.	.
N	N		N	X()

Figure 64 : Détermination des voisins au sens multi-échelles

La somme des indices de proximité peut conduire à une égalité pour certains voisins; aussi, un critère supplémentaire de sélection peut être donné par la variance minimale des indices de proximité.

La coopération intra-classificateur a pour effet de modifier le choix des voisins pour l'estimation de la densité de probabilité locale. Pour comprendre le principe de cette modification, il suffit de raisonner dans un premier temps sur une règle du type un plus proche voisin.

Dans un espace de représentation donné, si le voisin sélectionné initialement comme plus proche voisin d'une forme, n'est pas retenu, le domaine d'estimation n'est plus symétrique, centré sur la forme, mais déformé vers le prototype sélectionné. Si le prototype est éloigné, le domaine d'estimation se déforme fortement et le rejet de distance peut être sollicité pour rejeter la forme. Ainsi, la déformation du domaine de recherche du plus proche voisin dégrade la règle de discrimination locale.

Si l'on augmente notablement le nombre de voisins recherchés, (11, 13, ...), on minimise la possibilité de déformation du domaine de recherche des voisins, mais on minimise également le bénéfice de la coopération.

Par ailleurs, afin de minimiser la déformation du domaine et maximiser l'intérêt du remplacement d'un voisin local erroné pour un voisin multi-échelle pertinent, nous avons choisi de fixer le nombre de voisins à 5.

7.3. La structure de la discrimination intégrée

La Figure 65 présente la synthèse graphique de nos différentes contributions à la problématique de la vérification automatique de l'identité, sous la forme d'un système composé d'une perception multi-échelle et d'une discrimination intégrée floues. L'ensemble formant un système de perception finalisée.

Lorsque l'on parcourt ce schéma depuis l'arrivée séquentielle des données, on trouve, en premier lieu, les différents motifs de perception basés sur le Fuzzy Extended Shadow Code fournissant les vecteurs caractéristiques des signatures à chacune des fonctions de discrimination associées. Ces fonctions coopèrent dans la détermination des k plus proches voisins de la forme multi-échelles, puis elles proposent une affectation floue de la signature à chacune des classes selon leur niveau de perception.

Les différentes propositions sont agrégées dans le module de discrimination globale par une coopération extra-classificateur, pour fournir un degré d'appartenance à chacune des classes.

Le seuil de décision global permet ensuite d'affecter à la signature une étiquette vraie, fausse ou rejet. Lorsque la signature authentique est rejetée, le client la certifiera, elle sera alors intégrée dans l'ensemble des prototypes de vraies signatures. Si la signature est déclarée vraie, elle sera intégrée à l'ensemble des vraies signatures si son degré d'appartenance dans les différentes discriminations le justifie. En outre, si une signature fausse est déclarée vraie, le client signalera cette erreur et la signature sera intégrée dans l'ensemble des fausses signatures. Aucune autre fausse signature ne devra être intégrée dans l'ensemble des signatures fausses. En cas d'erreur de vérification, les paramètres de la fonction de discrimination globale seront ajustés en conséquence.

Chaque nouvelle signature intégrée dans les classes de vraies ou de fausses signatures contribuera à ajuster les paramètres des fonctions de discrimination locales. Notons que pour la classe des fausses signatures, l'entropie floue reste nulle puisque l'on intègre que des fausses signatures certifiées. Ainsi, seuls les paramètres f_i évolueront avec l'apport de connaissances.

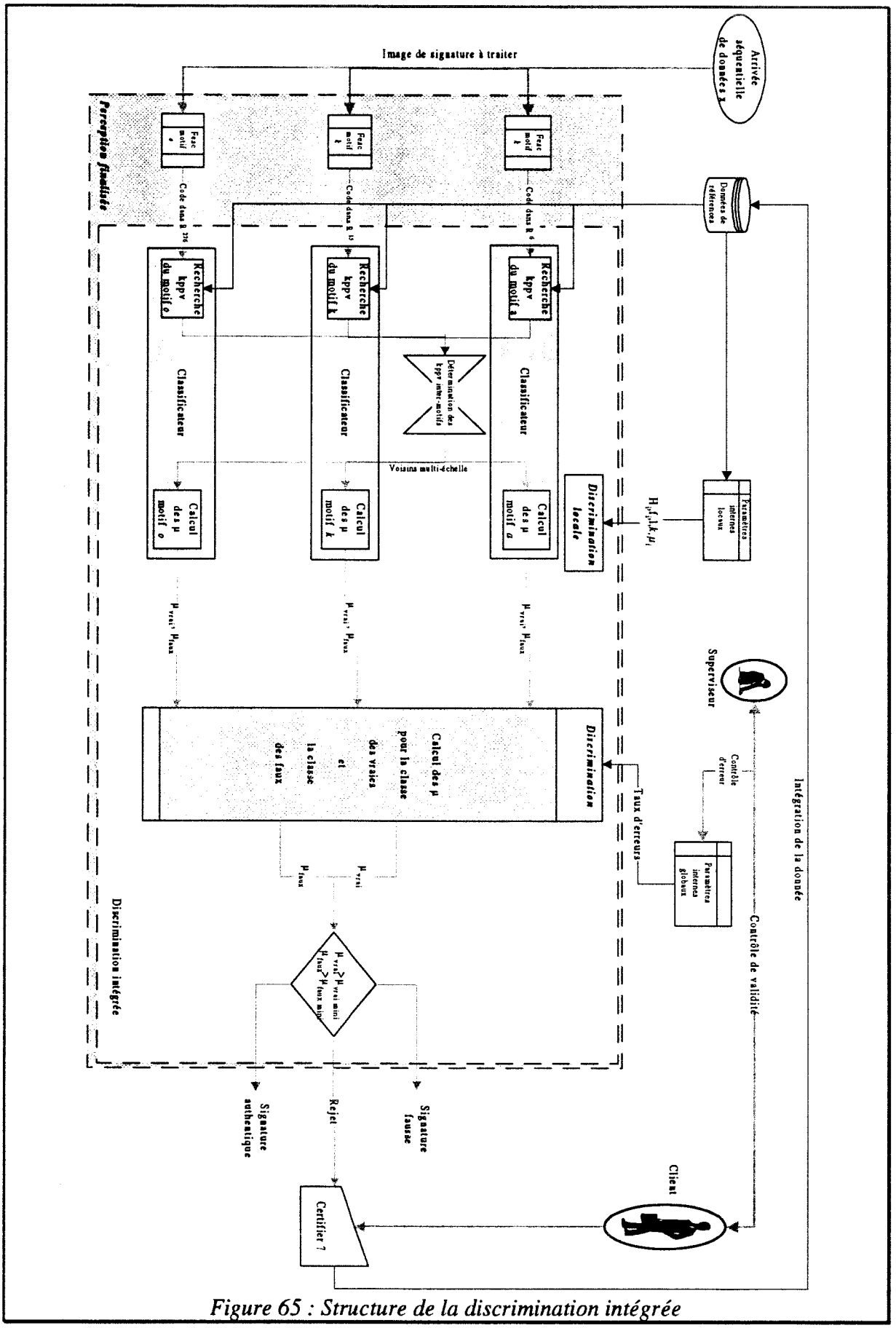


Figure 65 : Structure de la discrimination intégrée

Ce système répond à l'ensemble des critères que nous nous étions fixés pour résoudre le problème de la vérification automatique de l'identité, en présence de faux aléatoires et de faux par transfert optique. Cependant, l'évaluation des performances de ce système dans un contexte d'apprentissage permanent n'est pas réalisable sur notre base de données. En effet, les différents paramètres de ce système ne peuvent être ajustés sur moins de 40 signatures authentiques par scripteur. De plus, certaines questions telles que le choix de la configuration initiale de prototypes de vraies et de fausses signatures suivant le nombre de prototypes ou la nature téléologique de la signature ne peuvent trouver de réponse sans une étude approfondie que nous n'avons ni le temps ni les moyens matériels de conduire.

Aussi, la seule évaluation de performance que nous puissions proposer porte sur les capacités de discrimination du système, à partir d'un ensemble fixe de références. Cet ensemble est choisi selon le même protocole que les expériences précédentes, ce qui nous permettra de comparer les performances du système proposé dans une version dégradée.

7.4. Expérimentations

La dernière série d'expérimentations que nous présentons dans ce chapitre, concerne l'évaluation des performances de vérification du système dégradé. Nous proposons deux expériences conduites sur les deux ensembles de motifs *a*, *f*, *g*, et *m*, *n*, *o*. Les valeurs des paramètres λ_{vraies}, $\lambda_{fausses}$, k, b_i local sont identiques à l'expérience précédente ($k=5$, $\lambda_{vraies}=0.7$, $\lambda_{fausses}=0.7$, et $\forall\, i \in \{vraies, fausses\}$, $\forall\, motif \in \{a, f, g\}$, $b_i=0.01$, b_i global $\in \{b_{vraies\_global}, b_{fausses\_global}\}$ est calculé en ligne et le seuil d'affectation est à *0.65*). Le Tableau 20 présente les résultats obtenus pour le premier groupe de motifs.

	$\lambda_{fausses} = 0.7$, $\lambda_{vraies} = 0.7$, $k = 5$							
	Photocopies (δ 75%)				**Faux aléatoires**			
Coopération	E_p1	R_p1	E_p2	R_p2	E_f1	R_f1	E_f2	R_f2
Motif *a*, *f* et *g*	43.42	0	0.16	1.78	0.67	0	0.83	0.06

Tableau 20 : Performances de vérification du système dégradé, motifs a, f et g

Dans cette expérimentation, nous pouvons également quantifier l'apport de la coopération intra-classificateur, puisque nous nous sommes placés dans les mêmes conditions que les deux expériences précédentes.

La première remarque qui vient lorsque l'on analyse les résultats du Tableau 20, porte sur l'augmentation de la performance de discrimination des photocopies de vraies signatures. En effet, on observe une augmentation du taux E_p1 de 8.5%, sans que les autres taux E_p2, R_p1 et R_p2 en soient particulièrement affectés. Ainsi, la mise en commun des k plus proches voisins multi-échelle permet une meilleure caractérisation locale des photocopies de vraies signatures.

En ce qui concerne les faux aléatoires, l'erreur moyenne totale se situe à 0.75% et les taux de rejet R_f1 et R_f2 sont insignifiants. On améliore donc par la coopération intra-classificateur les capacités de discrimination de l'ensemble des motifs grossiers *a*, *f* et *g*.

Le Tableau 21 présente les mêmes caractéristiques pour les faux par transfert optique.

	$\lambda_{fausses} = 0.7$, $\lambda_{vraies} = 0.7$, $k = 5$							
	Photocopies (δ 75%)				**Faux aléatoires**			
Coopération	E_p1	R_p1	E_p2	R_p2	E_f1	R_f1	E_f2	R_f2
Motif *m*, *n* et *o*	29.53	0	0	1.66	10.32	0	0	1.01

Tableau 21 : Performances de vérification, motifs m, n et o

Le taux de détection de signatures vraies photocopiées E_p1 progresse de 10%, sans une dégradation significative des autres indicateurs.

En revanche, l'amélioration des performances de discrimination des faux aléatoires, constatée pour les motifs grossiers *a*, *f*, et *g*, n'est pas aussi positive. En effet, le taux E_f1 de vraies signatures déclarées fausses augmente de 4% lorsqu'on on fait coopérer ces motifs localement.

On peut noter cependant, que dans les configurations *a*, *f*, *g* et *m*, *n*, *o*, les performances sont améliorées dans le sens de la sécurité, et pour les motifs m, n, o au détriment du confort.

Les résultats peuvent s'expliquer, si l'on se focalise sur la perception différente des signatures selon deux ensembles de motifs.

Dans les motifs fins, nous savons que les photocopies de vraies signatures sont plus proches des authentiques que les faux aléatoires. Aussi, si l'on améliore le taux de discrimination de photocopies, on détériore simultanément le taux de reconnaissance des authentiques et l'indicateur E_f1 croît.

En revanche, dans les motifs grossiers, les photocopies sont aussi éloignées des authentiques que les faux aléatoires et parfois plus proches des faux aléatoires que des authentiques.

Aussi, la coopération intra-classificateur, qui a pour objectif d'améliorer la pertinence du choix des voisins, permet d'augmenter le taux E_p1 sans dégrader E_f1.

7.5. Conclusion

Ces dernières expérimentations ont montré, pour deux associations de trois motifs, que la coopération intra-classificateurs augmente le taux E_p1, caractéristique de la discrimination des photocopies de vraies signatures. Cependant, si pour l'association de motifs grossiers, cette augmentation de performances s'applique aussi à la discrimination des faux aléatoires,

pour les motifs fins, la tendance est inverse. Bien que le taux d'erreur E_f2 n'évolue pas, le taux E_f1, caractérisant l'erreur d'étiquetage des signatures authentiques, augmente notablement.

Ainsi, l'accroissement de la sécurité dans la vérification, quel que soit le type de fausses signatures, force à sacrifier le confort pour les motifs fins, tandis que pour les motifs grossiers, le confort est amélioré.

Ces résultats, bien qu'encourageants, demandent à être confirmés lors d'une série d'expérimentations plus complète, portant sur différentes associations de motifs.

8. Conclusion

Dans ce chapitre, nous avons explicité les contraintes de la vérification automatique de signatures manuscrites, puis détaillé le classificateur intégré proposé par R. Sabourin, afin de dégager les caractéristiques essentielles que doit posséder un système automatique pour remplir la fonction de vérification.

Nous avons présenté ensuite les améliorations possibles et souhaitables de cette approche, pour traiter les faux par calque ou transfert optique. Ces améliorations portent sur l'apprentissage, la coopération des discriminations locales et sur la prise en compte de la perception multi-échelles dans la discrimination. Ces propositions nous ont conduits à poser les critères de sélection d'une méthode de discrimination locale. Notre choix s'est porté sur la méthode des k plus proches voisins flous proposée par Béreau. Nous avons ensuite proposé des adaptations de la fonction de discrimination de cette méthode, afin de la rendre opérationnelle pour notre application. Ces adaptations portent sur la formulation de l'incertitude d'affectation et sur l'intégration de la pertinence de la perception.

Une série d'expérimentations a permis de présenter les capacités de discrimination locale de cette fonction, selon différentes valeurs de paramètres, pour l'ensemble des motifs de perception. Les résultats indiquent que les motifs fins sont peu appropriés à la discrimination des faux par calque lorsqu'ils sont pris isolément, tandis que les motifs grossiers présentent des capacités tout à fait intéressantes de discrimination de ces fausses signatures, si on somme les erreurs et les rejets de formes. De plus, leur capacité de discrimination des faux aléatoires est tout à fait acceptable.

Nous avons ensuite introduit la structure de discrimination intégrée que nous proposons. Elle est basée sur le facteur de formes Fuzzy Extended Shadow Code, sur une coopération parallèle globale des discriminations locales et sur une coopération interne des discriminations, fournissant les k plus proches voisins d'une forme au sens multi-échelles. Nous n'avons pas pu proposer une évaluation des performances de vérification de ce système à apprentissage permanent, puisque la base de données à notre disposition ne contient pas assez d'authentiques. Cette évaluation devrait se faire sur une base de données beaucoup plus conséquente, et présentant différentes natures d'authentiques (formelles, informelles, fugitives).

Aussi, nous avons proposé d'évaluer les performances de ce système de vérification en le dégradant, c'est-à-dire en travaillant sur un ensemble de prototypes figé, formé de 115 signatures vraies et fausses. Pour cela, nous avons sélectionné deux associations de trois motifs

fins et grossiers. L'association des deux coopérations des discriminations locales permet d'obtenir des résultats de discrimination de faux par calque tout à fait acceptables, étant donné le peu de différences entre les faux par calque et les authentiques. Là encore, l'ensemble de motifs grossiers est globalement plus performant pour la discrimination simultanée des faux aléatoires et des faux par calque, que les motifs fins. Pour les motifs grossiers, la sécurisation de la décision ne se fait pas aux dépens du confort, ce qui n'est pas le cas pour les motifs fins.

D'autres expérimentations devraient permettre de valider ces résultats et d'évaluer les pouvoirs discriminants d'autres associations de motifs.

Conclusion et perspectives

Conclusion

La perception ne relève pas du domaine de la reconnaissance de formes proprement dit, mais elle constitue, pour nous, une partie essentielle de l'élaboration d'un scénario de reconnaissance. Elle est forcément à définir en fonction du problème, en relation avec l'approche de reconnaissance envisagée et doit lui fournir les caractéristiques pertinentes des formes. Dès lors, elle a réalisé une forme de synthèse de l'information privilégiant la qualité à la quantité. La perception offre une potentialité de discrimination à exploiter par la méthode de reconnaissance.

Au cours de cette étude, nous avons vu et spécifié la nature de la vérification de l'identité par l'image de la signature manuscrite. La concentration de nos efforts sur l'image de la signature, nous a conduits à envisager la construction de l'étape de perception en fonction des informations caractéristiques des différents types de fausses signatures que l'on peut rencontrer. Ainsi, la restriction du problème à la discrimination des faux aléatoires oriente la perception vers des informations statiques. Les nombreux développements de R. Sabourin dans le domaine de la vérification de signatures sont dans ce sens.

Le libellé est une caractéristique évidente permettant la discrimination d'un faux aléatoire. Cependant, il est sensible au texte, et l'accession à la sémantique de la signature n'est plus pertinente dans le cadre d'autres types de faux. En développant la technique de l'Extended Shadow Code, R. Sabourin a pris une approche globale de la reconnaissance de formes, se détachant ainsi de toute sémantique. Une approche multi-échelle décrit tous les aspects de la signature, allant d'un point de vue grossier à un point de vue fin. Cette approche exploite l'information spécifique en relation avec la graphie particulière du scripteur. Selon l'échelle adaptée, on décrit alors la panoplie des caractéristiques statiques, ouvrant potentiellement le système de vérification au traitement des faux simples.

Sans remettre en cause la perception développée pour le traitement des faux aléatoires, nous avons constaté que l'information de présence du tracé dans l'Extended Shadow Code pouvait être enrichie pour caractériser la géométrie de la signature. En introduisant la distance séparant les supports de codage, du tracé de la signature, nous avons enrichi l'information codée pour obtenir des performances sensiblement améliorées dans le cas de motifs grossiers. En revanche, dans les motifs fins, cette information perd de son intérêt car le maillage de l'Extended Shadow Code est très proche du tracé. Cela signifie que l'ESC caractérise finement la géométrie de la signature et l'apport de la distance devient alors négligeable. Cette amélioration présente néanmoins un intérêt certain puisque nous avons atteint un taux global d'amélioration de 25% en considérant tous les motifs.

Afin d'étendre la vérification à d'autres types de faux, nous avons introduit simultanément la distance et le niveau de gris de la signature dans le codage. Le niveau de gris est une information pseudo-dynamique qui permet de traiter les faux par transfert optique, nous permettant d'envisager à court terme le traitement des faux serviles. En effet, le niveau de gris de la signature est caractéristique de la vitesse et de la pression lors de l'acte d'écriture. Or, d'après les caractéristiques discriminantes pour les types de faux, la pseudo-dynamique est essentiellement associée à la discrimination des faux par transfert optique. Ainsi, en combinant

les deux informations, nous offrons la possibilité de traiter simultanément les faux aléatoires et les faux par photocopie. Par ailleurs, la réalisation de la fusion de ces deux notions hétérogènes a été possible grâce à la théorie des ensembles flous. Bien que d'autres solutions soient possibles, la définition d'une table d'inférence pour réaliser cette combinaison présente l'avantage de supporter directement une analyse qualitative. La fuzzification des niveaux de gris nous permet de caractériser symboliquement et globalement les notions de pression et de vitesse, associées à un point dans l'image. Le découplage de ces deux notions est impossible à partir du niveau de gris, et leur association exprimée est imprécise. Par ailleurs, nous ne pouvons nous attacher à la valeur exacte du niveau de gris. Aussi la caractérisation floue permet de remonter à un niveau plus global, symbolique. Cette caractérisation floue facilite l'exploitation des informations en passant à une sémantique plus élevée puisque nous avons rattaché les niveaux de gris à des événements caractéristiques au cours de l'acte d'écriture comme le posé, le corps, ou le levé de plume et c'est certainement l'intérêt de cette approche.

Le codage résultant de cette combinaison nous permet donc d'étendre la vérification des faux aléatoires aux faux par transfert. En effet, les simulations du FESC dans le cadre des faux aléatoires montrent une amélioration de 23% par rapport à l'ESC, tous motifs confondus, donc des performances proches de celles de l'ESC avec codage de la distance dont le champ d'application est réduit au traitement des faux aléatoires.

L'étape de perception est performante, l'étape de raisonnement ne l'est pas moins si elle réalise l'adaptation de la discrimination à la complexité du problème et à la forme dans laquelle la perception exprime les données. L'hypothèse d'un modèle statistique pouvant synthétiser l'ensemble des données est d'autant plus délicate que leur nombre est faible et rend alors impossible une estimation paramétrique. Cela conduit directement au choix d'une approche non paramétrique afin d'exploiter le maximum d'informations car une décision doit nécessairement être prise. Cette contrainte sur la quantité de données nous conduit également à l'utilisation d'une méthode autorisant un apprentissage permanent susceptible d'acquérir la connaissance nécessaire à une meilleure décision. Une telle procédure vise à nous affranchir des évolutions des habitudes du scripteur. Nous avons évoqué cette possibilité dans le cadre d'une application réelle. Cependant, la faible quantité de données par scripteur ne nous a pas permis d'en montrer l'intérêt.

La complexité de la discrimination des faux par photocopie, ainsi que la faible connaissance apportée par le nombre restreint de données, obligent à ne pas trop extrapoler au-delà des formes connues pour lesquelles le risque de commettre une erreur devient grand. Un rejet de distance est donc nécessaire. Il faut également noter qu'en reconnaissance de formes, les classes sont le plus souvent non-disjointes et qu'une méthode non adaptée à leur recouvrement conduit à l'erreur. Pour ces différentes remarques, nous avons choisi une extension floue de la règle des k plus proches voisins dont les propriétés sont bien adaptées au problème de vérification de signatures.

Cet algorithme présente également d'autres avantages. Comme chaque échelle de perception est orientée selon le motif, le comportement de chacun des classificateurs associés à ces échelles ne peut pas être similaire. Différents paramètres de l'algorithme peuvent alors être réglés en fonction de l'échelle. Par ailleurs, on peut remarquer que les classes "vrai" et "faux" sont indépendantes l'une de l'autre, on peut ainsi particulariser chaque classificateur.

Dans cet algorithme, le concept des ensembles flous est essentiel. Il nous permet de caractériser l'ambiguïté et le rejet de distance au travers du degré d'appartenance. Par ce degré, une meilleure caractérisation des formes est réalisée et, dans le domaine flou, les erreurs sont moins sensibles. Grâce à l'approche multi-échelle et à une fonction globale de discrimination, la caractérisation floue des formes n'est pas défuzzifiée, et on évite ainsi une erreur locale de décision. Les degrés d'appartenance fournis par chaque classificateur associé à chaque échelle restent concrètement rattachés à la forme analysée. Ils sont de nature homogène quelle que soit l'échelle de perception. La défuzzification qui se traduit par la décision est le moment critique d'un système flou. En élargissant le domaine flou jusqu'à la discrimination, nous limitons les erreurs locales de décision, pour ne faire chuter l'entropie qu'au dernier instant. Par ailleurs, l'incertitude introduite dans le raisonnement, par l'imprécision de la perception, est réduite par la coopération externe des différentes propositions d'appartenance des classificateurs. En outre, la fonction finale de discrimination s'adapte afin de sécuriser la décision et exprime son incapacité à décider si l'incertitude est trop grande. Par ailleurs, une coopération interne définit des prototypes comme formes fortes au sens multi-échelle. En outre, c'est une procédure innovante car on peut rapprocher ce principe de celui d'une opération humaine sur des informations multi-sources où les objets sont caractérisés par plusieurs sens simultanés comme la vue, l'odorat et le toucher. Ensuite, la coopération externe se charge de prendre une décision sur les différentes propositions émises selon les différents points de vue.

Perspectives

Par ces travaux, nous avons atteint un certain objectif dans le développement de la vérification de signatures manuscrites. Un certain nombre d'activités peuvent suivre le développement que nous avons proposé, tant sur le plan de la perception que sur le plan du raisonnement.

Au niveau de la perception, le FESC a ouvert de nouvelles perspectives, et demande de nouveaux développements et analyses. En effet, le choix du nombre de termes linguistiques pour la caractérisation floue de la pseudo-dynamique a été fait en rattachant ces termes à des événements précis au cours de l'acte d'écriture. Une caractérisation plus fine de l'échelle des gris offrirait peut-être une plus grande pertinence du facteur de formes et conduirait à des performances de vérification plus élevées.

Sur le plan du raisonnement, le principe de discrimination intégrée proposé semble cohérent, il est nécessaire cependant de le valider dans le temps avec un plus grand nombre de données. Cette remarque est d'autant plus justifiée que c'est dans ces conditions que notre méthode devrait dévoiler ses meilleurs atouts.

Par ailleurs, la méthode globale mise en place n'est pas seulement dédiée à la vérification de signatures, on peut et doit envisager son extension à d'autres problèmes, comme par exemple l'identification de produits marqués ou la discrimination d'objets perçus par des capteurs multi-sources.

Références bibliographiques

[Ahmed 75] Ahmed N., Rao K.R., *"Orthogonal Transforms for Digital Signal Processing"*, Springer-Verlag, 1975.

[Al-Yousefi 92] Al-Yousefi H., Upda S., "Recognition of Arabic caracters", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8), pp. 853-857, 1992.

[Ammar 86] Ammar M., Yoshida Y. and Fukumura T., "A new effective approach for off-line verification of signatures by using pressure features", 8th International Conference on Pattern Recognition, Paris, pp. 566-569, 1986.

[Ammar 90] Ammar M., Yoshida Y. and Fukumura T., "Structural description and classification of signatures images", *Pattern Recognition*, 23(7), pp. 697-710, 1990.

[Anigbogu 92] Anigbogu J., Belaïd A., "Reconnaissance de caractères multifontes par vote à la majorité", CNED 92, pp. 91-99, Nancy, France, 1992.

[Basseville 88] Basseville M., "Detecting changes in signal and systems - A survey", *Automatica*, 24(3), pp. 309-326, 1988.

[Belaïd 92] Belaïd A., Belaïd Y., *"Reconnaissance de formes : Méthodes et applications"*, InterEditions, série Informatique Intelligence Artificielle, Paris, 1992.

[Bellman 58] Bellman R., "Dynamic Programming and stochastic control processes", *Information and control*, 1(3), pp 228-239, 1958.

[Benoit 93] Benoit E., "*Capteurs symboliques et capteurs flous : un nouveau pas vers l'intelligence*", Thèse de Doctorat, Université Joseph Fourier, Grenoble, France, 1993.

[Béreau 86] Béreau M., "*Contribution de la théorie des sous-ensembles flous à la règle de discrimination des k plus proches voisins en mode partiellement supervisé*", Thèse de Doctorat de l'Université de Technologie de Compiègne, URA CNRS Heudiasyc, 1986.

[Bezdek 81] Bezdek J.C., "*Pattern Recognition with Fuzzy Objective Function Algorithms*", Plenum Press, New York and London, 1981.

[Bezdek 86] Bezdek J.C., "Generalized k-nearest neighbour rules", *Fuzzy Sets and Systems*, 18(3), pp. 237-256, 1986.

[Bobrowski 91] Bobrowski L., Bezdek, J.C., "C-means clustering with the l(1) and l(∞) norms", *IEEE Transactions on Systems, Man, and Cybernetics*, 21(3), pp. 545-554, 1991.

[Bouchon-Meunier 94] Bouchon-Meunier B., *"La logique floue"*, Edition "Que sais-je ?", Presses universitaires de France, Paris, 1994.

[Brockelhurst 85] Brockelhurst E., "Computer Methods of Signature Verification", *Journal of Forensic Science Society*, 25, pp. 445-457, 1985.

[Burr 87] Burr D.J., "Experiments with a connectionnist text reader", IEEE International Conference on Neural Networks, San Diego, 1987.

[Burr 88] Burr D.J., "Experiments on Neural Net Recognition of Spoken and Written Text", *IEEE Transactions on Acoustics, Speech and Signal Processing*, 36(7), pp. 1162-1168, 1988.

[Celeux 88] Celeux G., "Le traitement des données manquantes dans le logiciel SICLA", Rapport Technique INRIA, n°102, 1988.

[Chuang 77] Chuang P., "Machine verification of handwritten signature images", International Conference on Crime Contermesures Sciences and Eng., pp. 105-109, 1977.

[Cohen 91] Cohen E., Hull J, Srihari S., "Understanding handwritten text in a structured environment : Determining ZIP codes from adresses", Character and Handwritten Recognition : Expanding Frontiers, World Scientific Series in *Computer Science*, P.S.P. Wang Editor, 30, pp. 221-264, 1991.

[Cover 65] Cover T., "Geometrical and statistical properties of linear inequalities with application to pattern recognition", *IEEE Transactions on Electronic Computer*, EC-14(3), pp. 326-334, 1965.

[Cover 67] Cover T.M., Hart P.E., "Nearest neighbor pattern classification", *IEEE Trans. Info. Theory*, IT-13, pp. 21-27, 1967.

[Craine 94] Craine W., "Characterizations of fuzzy interval graphs", *Fuzzy Sets and Systems*, 68(2), pp. 181-194, 1994.

[Crettez 89] Crettez J.P., "Lecture automatique de documents imprimés", Texte du GRCE, 1989.

[Dave 92a] Dave R., "Generalized fuzzy C-shells clustering and detection of circular and elliptical boundaries", *Pattern recognition*, 25(7), pp. 713-721, 1992.

[Dave 92b] Dave R., Bhaswa K., "Adaptive fuzzy c-shells clustering and detection of ellipses" *IEEE Transactions on Neural Networks*, 3(5), pp. 643-662, 1992.

[De Luca 72] De Luca A., Termini S., "A definition of a non probabilistic entropy in the setting of fuzzy sets theory", *Information Control*, n°20, pp. 301-318, 1972.

[Denoeux 95] Denoeux T., "A k-nearest neighbor classification rule based on Dempster-Shafer Theory", *IEEE Transaction on Systems, Man. and Cybernetics*, 25(5), 1995.

[Devijver 77] Devijver P.A., *"Reconnaissance des formes par la méthode des plus proches voisins"*, Thèse de 3ème cycle, Paris VI, Juin 1977.

[Devijver 80] Devijver P.A., Kittler J., "On the edited nearest neighbor rule", 5th Int. Conf. on Pattern Recognition, pp. 72-80, Miami Beach (USA), 1980.

[Devijver 82] Devijver P.A., Kittler J., *"Pattern recognition : a statistical approach"*, Prentice Hall, 1982.

[Dubois 92] Dubois D., Prade H., "When upper probabilities are possibility measures", *Fuzzy Sets and Systems*, 49(1), pp. 65-74, 1992.

[Dubuisson 90] Dubuisson B., *"Diagnostic et reconnaissance de formes"*, Traité des nouvelles technologies, Eds Hermès, Paris, 1990.

[Duda 73] Duda R., Hart P., "*Pattern classification and Scene analysis*", John Wiley and Sons, New York, 1973.

[Durand 94] Durand D., "*La systémique*", Edition "Que sais-je ?", Presses universitaires de France, Paris, 1994.

[Epstein 94] Epstein R., Yuille A., "Training a general purpose deformable Template", IEEE Int. Conf. on Image Processing, ICIP 94, 1, pp 203-207, AUSTIN, U.S.A, 1994.

[Evett 85] Evett I.W., Totty R.N., "A study of the variation in dimensions of guenine signatures", *Journal of the Forensic Science Society*, 25, pp. 205-215, 1985.

[Forre 86] Forre R. and Debruyne P., "Signature verification with elastic image matching", International CARNAHAN Conference on Security Technic, Golthenburg, Sweden, pp. 113-118, 1986.

[Foulloy 92] Foulloy L., Gallichet S., "Représentation des contrôleurs flous", 2^{ièmes} Journées Nationales, Les applications des ensembles flous, pp. 129-136, Nîmes, 2-3 novembre 1992.

[Fu 68] Fu K.S., "*Sequential Methods in Pattern Recognition and Machine Learning*", Academic Press, New York and London, 1968.

[Fukunaga 75] Fukunaga K, Narendra P.M., "A branch and bound algorithm for computing k nearest neighbors", *IEEE Trans. on Computers*, 24, pp. 750-753, 1975.

[Fukunaga 90] Fukunaga K., *"Introduction to statistical pattern recognition"*, Academic Press, New-York, 1990.

[Gaillat 79] Gaillat G., Berthod M., "Panorama des méthodes d'extraction de traits caractéristiques en lecture optique de caractères", *Revue Technique THOMSON-CSF*, 11(4), pp. 943-959, 1979.

[Gayet 61] Gayet J., *"Manuel de police scientifique"*, Editions Payot, Paris, 1961.

[Geva 89] Geva A., Gath I., "Unsupervised optimal Fuzzy clustering", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7), pp. 773-781, 1989.

[Girod 93] Girod J.P., "Les réseaux de neurones : application à la reconnaissance de caractères", Rapport 93-2, G.R. Automatique, Pôle Automatisation Intégrée, Projet Commande Symbolique et Neuromimétique, Avril 1993.

[Grabisch 95] Grabisch M., "Pattern classification and feature extraction by fuzzy integral", EUFIT 95, 3, pp. 1465-1469, Aachen, Germany, 1995.

[Guan 93] Guan J.W., Bell D.A., "Bayesian decision theory, certainty factors, and evidence theory", EUFIT 93, 2, pp. 961-967, Aachen, Germany, 1993.

[Gupta 79] Gupta S.K., "Protecting signatures against forgery", *Journal of the Forensic Science Society*, 19, pp. 19-23, 1979.

[Gustafson 79] Gustafson D., Kessel E., "Fuzzy clustering with fuzzy covariance matrix", Int. Conf. Control and Decisions, 2, pp. 761-766, 1979.

[Hall 80] Hall P., Dowling G., "Approximate string matching", *Computing Surveys*, 12(4), pp. 381-402, 1980.

[Harmon 72] Harmon L.D., "Automatic recognition of print and script" *Proceeding of IEEE*, 60(10), pp. 1165-1176, 1972.

[Harnard 87] Harnard S., *"Categorical Perception : The groundwork of Cognition"*, Cambridge University Press, New York, 1987.

[Harrisson 81] Harrisson WR., *"Suspect Documents, Their Scientific Examination"*, Nelson-Hall, Chicago, 1981.

[Heutte 94] Heutte L., "*Reconnaissance de caractères manuscrits : Application à la lecture automatique des chèques et enveloppes postales*", Thèse de Doctorat en physique option informatique industrielle de l'Université de Rouen, 1994.

[Hilton 71] Hilton O., "Do we really have adequate signature standards ?", *Journal of the Forensic Science Society*, 11, pp. 122-133, 1971.

[Huang 92] Huang K.Y., Yen H.T., "Robust recognition of printed chinese caracters using multilayer perceptron and Walsh function", *SPIE*, 1709, pp. 315-324, 1992.

[Hubert 94] Hubert Y., "*Contribution à la modélisation qualitative du confort de sièges automobiles*", Mémoire d'ingénieur CNAM, 18 décembre 1994.

[Jajuga 91] Jajuga K. "L1-norm based fuzzy clustering", *Fuzzy Sets and Systems*, 39(1), pp. 43-50, 1991.

[Kapoor 85] Kapoor T.S., Kapoor M., Sharma G.P., "Study of the form and extent of natural variation in genuine writtings with age", *Journal of the Forensic Science Society*, 25, pp. 371-375, 1985.

[Kissiov 92] Kissiov V., Hadjitodorov S. "A fuzzy version of the KNN method", *Fuzzy Sets and Systems*, 49, pp. 323-329, 1992.

[Kittler 78] Kittler J., "A method for determining k nearest neighbors", *kibernetes*, 7, pp. 313-315, 1978.

[Krattenthaler 94] Krattenthaler W., Mayer K.J., Zeiller M., "Point correlation : A reduced cost template matching technique", IEEE Int. Conf. on Image Processing, ICIP 94, 1, pp 208-212, AUSTIN, U.S.A, 1994.

[Krishnapura 92] Krishnapura R., Nasraoui O., Frigui H., "The fuzzy C spherical shells algorithm : a new approach", *IEEE Trans. on Neural Networks*, 3(5), pp. 673-671, 1992.

[Kuncheva 95] Kuncheva L., Kounchev R., Zlatev R., "Aggregation of multiple classification decisions by fuzzy templates", EUFIT 95, 3, pp. 1470-1474, Aachen, Germany, 1995.

[Lachaux 94] Lachaux H., "Lecture automatique de caractères sur pièces de fonderies par réseaux de neurones", Rapport de fin de contrat, CRAN PRAISSIH, Septembre 1994.

[Lan 95] Lan S., Zhiwen M., "Closure of finite state automaton languages", *Fuzzy Sets and Systems*, 75(3),pp. 393-398, 1995.

[Lebourgeois 91] Lebourgeois F., Emptoz H., "Organisation d'un système de lecture automatique adapté aux documents imprimés multi-polices", Proc. 8ème Congrès AFCET-RFIA, Lyon, France, pp 713-718, 1991.

[Levrat 96] Levrat E., Ruelas R., Bremont J. Andre I., "Comfort evaluation using the fuzzy sets techniques" à paraître dans EUFIT 96.

[Locard 36] Locard E., *"L'expertise des documents ecrits"*, Traité de Criminalistique, Joannès Desvignes et Cie, 1936.

[Locard 59] Locard E., *"Les faux en écriture et leur expertise"*, Editions Payot, Paris, 1959.

[Lorrette 92] Lorette G. and Lecourtier Y., "Reconnaissance et interprétation de textes manuscrits hors-ligne : un problème d'analyse de scène?", CNED 92, Nancy, France, pp. 109-135, 1992.

[Mamdani 75] Mamdani E.H., Assilian S., "An experiment in linguistic sysnthesis with a fuzzy logic controller", *International Journal of Man Machines Studies*, 7, pp. 1-13, 1975.

[Mantas 86] Mantas J., "An overview of character recognition methodologies", *Pattern recognition*, 19(6), pp. 425-430, 1986.

[Masson 85] Masson J. F., "Felt Tip Pen Writing : Problems of Identification", *Journal of Forensic Sciences*, 30(1), pp. 172-177, 1985.

[Mathyer 61] Mathyer J., "The expert examination of signatures", *Journal of Criminal Law, Criminology and Police Science*, 5(3),1961.

[Miclet 84] Miclet L., *"Analyse structurelle pour la reconnaissance de formes"*, Eyrolles, Paris, 1984.

[Milgram 93] Milgram M., *"Reconnaissance des formes: Méthodes numériques et connexionnistes"*, Société des Nouvelles Editions Liégeoises, 1993.

[Monjallon 63] Monjallon A., *"Eléments de statistiques mathématiques"*, Eds Vuibert, 1963.

[Moon 77] Moon H.W., "A survey of handwritting styles by geographic location", *Journal of the Forensic Sciences*, 4, 1977.

[Nagel 77] Nagel R. and Rosenfeld A., "Computer detection of freehand forgeries", *IEEE Transaction on Computers*, 26(9), pp. 895-905, 1977.

[Nouboud 88] Nouboud F., Crozzo R., Callot M., et al., "Authentification de signatures manuscrites par programmation dynamique", PIXIM, Paris, pp. 345-360, 1988.

[Otsu 79] Otsu N., "A threshold selection method from gray level histogram", *IEEE Transactions on Systems, Man and cybernetics*, 9(1), pp. 62-66, 1979.

[Parzen 62] Parzen E., "On estimation of a probability density function and mode", *Ann. Math. Stat.*, 33, pp. 1065-1076, 1962.

[Pascual 92] Pascual E., Virbel J., "Connaissances linguistiques et morphodispositionnelles pour le contrôle de la décomposition structurelle de documents", CNED 92, pp. 217-224, Nancy, France, 1992.

[Pedrycz 90] Pedrycz W. "Fuzzy sets in pattern recognition: methodology and methods", *Pattern Recognition Letter*, 23(1), pp. 121-146, 1990.

[Phelps 82] Phelps R.J., "A holistic approach to signature verification", International Conference on Pattern Recognition, 2, pp. 1187, 1982.

[Plamondon 89] Plamondon R. and Lorette G., "Automatic signature verification and writter identification. The state of art", *Pattern Recognition*, 20, pp. 107-131, 1989.

[Plamondon 90] Plamondon R., Lorette G. and Sabourin R., "Automatic processing of signatures images: static techniques and methods", *Computer processing of Handwriting*, World Scientific Publishing, pp. 49-63, 1990.

[Postaire 87] Postaire J.G., "*Analyse des images numériques et théorie de la décision*", Dunod informatique, 1987.

[Qi 95] Qi Y. and Hunt R.H., "A multi resolution approach by computer verification of handwritten signatures", *IEEE Transactions on image processing*, 4(6), pp. 870-874, 1995.

[Raudys 91] Raudys S.J., Jaine A.K., "Small sample size effects in statistical pattern recognition : Recommendations for practitioners", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(3), pp. 252-264, 1991.

[Richetin 80] Richetin M., Rives G., Naranjo M., "Algorithme rapide pour la détermination des k plus proches voisins", R.A.I.R.O. Informatique, 14(4), pp. 369-378, 1980.

[Rocha 94] Rocha J., Pavlidis T., "A shape analysis model with applications to a character recognition system", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(4), pp. 393-404, 1994.

[Sabourin 90] Sabourin R., "*Une approche de type compréhension de scène appliquée au problème de la vérification automatique de l'identité par l'image de la signature manuscrite*", Thèse de Philosophiae Doctorat, 1990.

[Sabourin 92] Sabourin R., Plamondon R., Lorette G., "Off-line identification with handwritten signature images : Survey and Perspectives", *Structured Document Image Analysis*, Springer Verlag, pp. 219-234, 1992.

[Sabourin 93] Sabourin R., Cheriet M. and Genest G., "An extended-shadow-code based approach for off-line signature verification", 2nd IAPR Conference on Document Analysis and Recognition, Ibraki, Japan, pp. 1-5, October 20-22,1993.

[Sabourin 94a] Sabourin R. and Genest G., "An Extended-Shadow-Code based approach for off line signature verification : Part-I- Evaluation of the bar mask definition", 12th ICPR International Conference on Pattern Recognition, Jerusalem, Israël, October 9-13, 1994.

[Sabourin 94b] Sabourin R. and Genest G., "Coopération de classificateurs pour la vérification de signatures manuscrites", Colloque National sur l'Ecrit et le Document, Rouen, France, 6-8 Juillet, 1994.

[Sabourin 95a] Sabourin R., Genest G., "Coopération de classifieurs pour la vérification automatique de signatures", *Traitement du Signal*, 12(6), 1995.

[Sabourin 95b] Sabourin R., Genest G., "An Extended-Shadow-Code based approach for off line signature verification : Part-II- Evaluation of several Multi-classifier combination strategies", Proceeding of the third IAPR Conf. on Document Analysis and Recognition, Montréal, Canada, pp. 197-201, August 14-16, 1995

[Saferstein 82] Saferstein R., "*Forensic Science Handbook*", Prentice-Hall, pp. 687-694, 1982.

[Schneider 92] Schneider M., Lim H., Shoaff W., "The utilization of fuzzy sets in recognition of imperfect strings", *Fuzzy Sets and Systems*, 49(3), pp. 331-338, 1992.

[Shaw 80] Shaw D., "Proof of identity - A review", International Conference Security, Technical University, Berlin, Germany, September, 1980.

[Simon 93] Simon C., Levrat E., Sabourin R., et al., "Vérification de signatures manuscrites par classification floue", Les applications des ensembles flous, Nîmes, France, pp. 127-133, 1993.

[Smolarz 87] Smolarz A., "Un algorithme de discrimination adaptatif avec rejet", *R.A.I.R.O APII*, 21(5), pp. 449-474, 1987.

[Stanislaw 93] Stanislaw H., "Bayesian updating distribution of uncertainty", EUFIT 93, 3, pp. 1606-1609, Aachen, Germany, 1993.

[Strackeljan 95] Strackeljan J., Behr D., "A feature selection method for acoustic analysis problems", EUFIT 95, 2, pp. 1175-1180, Aachen, Germany, 1995.

[Sudkamp 92] Sudkamp T., "On probability-possibility transformations", *Fuzzy Sets and Systems*, 51(1), pp. 73-81, 1992.

[Suen 93] Suen CY., Legault R., Nadal C., Cheriet M., Lam L., "Building a new generation of handwritting recognition systems", *Pattern recognition letters*, 14(4), pp. 303-315, 1993.

[Tappert 90] Tappert C.C., Suen C.Y., Wakahara T., "The state of the art in on-line Handwritten recognition", *IEEE Trans on Pattern Analysis And Machine Intelligence*, pp. 787-807, 1990

[Totty 83] Totty R.N., "The dependence of slope of handwritting upon the sex and handedness of the writer", *Journal of the Forensic Science Society*, 23, pp. 237-240, 1983.

[Viard-Gaudin 92] Viard-Gaudin C. and Barba D., "Extraction robuste et structuration des informations par une approche multirésolution pour la localisation du bloc adresse sur les objets postaux plats", CNED 92, Nancy, France, pp. 48-56, 1992.

[Wang 95] Wang X., De Baets B., Kerre E., " A comparative study of similarity measures", *Fuzzy Sets and Systems*, 73(2),pp. 259-268, 1995

[Wilson 72] Wilson D.L., "Asymptotic properties of nearest neighbor rules using edited data", *IEEE Trans. on Systems, Man and Cybernetics*, SMC-2(3), pp. 408-421, 1972.

[Xu 92] Xu L., Krzyzak A., Suen C.Y., "Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition", IEEE Trans. on Systems, Man and Cybernetics, 22(3), pp. 418-435, May/June 1992.

[Yunck 76] Yunck T.P., "A technique to identify nearest neighbors", *IEEE trans. on Systems, Man and Cybernetics*, SMC-6(10), pp. 678-683, 1965.

Université HENRI POINCARÉ NANCY 1

Nom : **SIMON**

Prénom : **Christophe**

DOCTORAT de l'UNIVERSITE HENRI POINCARE, NANCY-I

en AUTOMATIQUE

VU, APPROUVÉ ET PERMIS D'IMPRIMER

Nancy, le 23 JUIL. 1996 n° : 57

Le Président de l'Université

UNIVERSITÉ HENRI POINCARÉ, NANCY I — LE PRESIDENT

J.P. FINANCE

3C-F9

UNIVERSITÉ HENRI POINCARÉ - NANCY 1
24-30, rue Lionnois B.P.3069 54013 Nancy cedex
Tél : 83 32 81 81 Fax : 83 32 95 90
Adresse électronique : _____ @presn1.u-nancy.fr

Résumé : La vérification automatique de l'identité par l'image de la signature manuscrite est une problématique particulière de la reconnaissance de formes. Au cours du premier chapitre, nous situons la vérification de signatures dans les problèmes de l'écrit, et nous posons le scénario de reconnaissance de formes selon une approche globale contribuant à la solution du problème.

Ainsi, dans le deuxième chapitre, une méthode de perception multi-échelle des signatures est dans un premier temps améliorée afin de mieux répondre au cas restreint des faux aléatoires. Dans un deuxième temps, la méthode de perception est transformée à l'aide d'une table d'inférence floue exploitant les informations nécessaires au traitement de faux plus complexes, et notamment les faux par photocopie.

Enfin, au cours du troisième chapitre, un classificateur intégré est élaboré sur la base des différents points de vue donnés par les échelles de perception. Ce classificateur se base sur des classificateurs de type k plus proches voisins flous associés à chaque échelle de perception. Leurs propriétés autorisent un apprentissage permanent et une coopération interne par la sélection de prototypes définis au sens multi-échelle. La décision finale est prise par une règle floue du type k plus proches classificateurs. Cette règle est ajustée en tenant compte des erreurs de vérifications de manière à rigidifier la décision afin de la sécuriser car les domaines d'application se situent majoritairement dans le domaine bancaire.

Mots-clés : Reconnaissance de formes, Vision, Vérification de signatures manuscrites, Logique floue, Classification floue.

Summary : Automatic identity verification by Handwritten signature image is a particular problem of pattern recognition. Firsteval, we place the signature verification in written problems. Then, we define the pattern recognition scenari contributing to solve the verification problem.

Secondly, we assess a multi-scale method of perception in oder to give a best respons to the problem of random forgeries. To treat more difficult problems of forgeries, we transform the method of perception by fuzzy logic that work with the pertinent informations taking on signature images. By developping this new method, we are able to decide on forgery obtained by transfer technics like photocopy.

Thirdly, an integrated classifier is developped on different point of view given by percepting scale. This classifier is based on fuzzy k nearest neighnors classifiers attached to each scale. Their properties have the advantage of allowing continuous learning and an internal cooperation by selecting multi-scale prototypes. Then, the discrimination rule is given by a fuzzy k nearest classifiers rule that is automatically adjusted by taking account the verification errors in order to become strong. This adjustment secures the discrimination and is adapted in bank application.

Keyword : Pattern recognition, Vision, Handwritten signatures verification, Fuzzy logic, Fuzzy classification.