



RAPPORT DE MASTER RECHERCHE – Juin 2005

Extraction de similarités de formes dans les images de traits

**Contribution à la compression sans perte de documents
manuscrits du patrimoine**

Djamel GACEB
LIRIS UMR 5205, INSA de Lyon

Gaceb.djamel@yahoo.fr

Encadrement :

Véronique EGLIN & Stéphane BRES
LIRIS, INSA de Lyon

Extraction de similarités de formes dans les images de traits

Contribution à la compression sans perte de documents manuscrits du patrimoine

Résumé

Les formats de compression existants ont montré leurs limites sur les images de documents très abîmés présentant de fortes irrégularités tant sur les zones de textes que sur l'arrière plan. Les documents manuscrits anciens, en particulier les brouillons d'auteurs, les documents en langues arabes, mais également les schémas et les plans illustrent très bien ce genre de difficultés. Les formats de compression usuels utilisés sur les images de documents fonctionnent de manière très imparfaite sur les documents très irréguliers. Cet échec est en partie lié à la grande complexité des formes (dans le cas du manuscrit, il s'agit des tracés présents dans les textes) rendant leur localisation souvent bien difficile et pour laquelle aucune méthode de segmentation n'est actuellement efficace. Pour remédier à ces difficultés, nous nous sommes orientés vers une nouvelle approche permettant de localiser précisément les zones d'informations, de repérer les formes redondantes et de les indexer sur une page entière. Nous proposons dans ce mémoire une approche originale de recherche de similarités de formes dans les tracés manuscrits issus de documents dégradés du patrimoine ancien. Pour cela, nous avons développé un outil permettant de signer les motifs présents dans les traits des écritures à partir de l'extraction de segments orientés et de points de changements de directions présents sur les contours. Cette recherche de similarités a été conçue pour être robuste aux changements d'échelle, aux rotations et bien sûr aux translations. Les premiers résultats obtenus sur les images de manuscrits anciens sont très encourageants. Leur application à la compression sans perte en dépend tout naturellement, c'est pour cette raison que cette contribution est très importante dans le domaine car il n'existe pas à ce jour de travaux antérieurs comparables.

Mots clés : *Numérisation du patrimoine écrit, compression, bancs de filtres directionnels de Gabor, détecteur de Deriche, décomposition des traits, similarités partielles, signature hiérarchique du tracé.*

Abstract

The existing compression formats have shown their limits on very damaged documents images presenting strong irregularities in textual regions as well as in the background. Ancient manuscript documents, and author drafts in particular, documents in Arabic language, as well as drawings and diagrams illustrate very well this kind of difficulties. Usual compression formats work in a very unsatisfying manner on irregular documents. This failure is therefore partially related to the great complexity of shapes making their localisation most of the time much difficult and for which no segmentation method is currently efficient. To tackle these difficulties, we propose a new approach allowing to localize precisely the information zones, to detect all redundant shapes and to index them all over an entire page. We propose in this work an original approach for similarities retrieval in degraded ancient written collections. For that purpose, we have developed a tool that signs all oriented partial shapes in the contours of handwritten lines. This similarity approach is robust to scale changes, rotations and of course to translations. The first results that have been obtained on handwritten ancient documents are very promising. The application to lossy compression is a challenge that has never been realized before. That is the reason why this contribution is very important in this research field.

Keywords: *Digitalization of written patrimony, compression, GaborBank filtering, Deriche detector, partial similarities, multi-scale signature of drawing.*

Sommaire

I. INTRODUCTION

- I.1 Contexte de l'étude autour de projets de valorisation du patrimoine écrit
- I.2 Transmettre et diffuser la patrimoine ancien : un défi au temps et aux techniques

II. SYNTHESE DE L'EXISTANT POUR UNE MISE EN LIGNE DES OUVRAGES : Compression de documents et recherche de similarités

- II.1 La compression générique avec perte : une solution peu adaptée aux documents
- II.2 Les approches binaires pour une compression sans perte d'informations
- II.3 La compression par recherche d'auto-similarités - méthode fractale
- II.4 La compression par recherche des redondances de formes
- II.5 Généralisation autour de la recherche de similarités

III. NOTRE PROPOSITION DE RECHERCHE DE SIMILARITES

- III.1 Principe général de l'approche
- III.2 Analyse fréquentielle des manuscrits
- III.3 Segmentation coopérative pour une décomposition du tracé en formes connexes
- III.4 Signature des motifs redondants et estimation des similarités

IV. EXPERIMENTATION ET ANALYSE DES RESULTATS

- IV.1 Analyse comparative des trois approches de recherche de similarités
- IV.2 Application à la recherche de similarités sur une image complète de document
- IV.3 Intégration au processus de compression adaptée aux tracés manuscrits
- IV.4 Perspectives directes : développement d'un outil générique

V. CONCLUSION GENERALE

VI. REFERENCES BIBLIOGRAPHIQUES

I. INTRODUCTION

I.1 Contexte de l'étude autour de projets de valorisation du patrimoine écrit

Nos bibliothèques regorgent d'ouvrages riches en informations (images de documents anciens ou historiques, des catalogues de vente par correspondance, des formulaires, des archives gouvernementales ou commerciales, des publications scientifiques ou techniques,...). La numérisation de documents du patrimoine scientifique et culturel facilite leur accès à un public plus large et offre de nouveaux services comme la consultation en ligne de documents rares, la duplication rapide et économique des ouvrages, la navigation simplifiée, une recherche d'information plus performante et enfin la possibilité de partager des connaissances avec d'autres lecteurs, [DE00, LE04]. A l'heure actuelle, il n'existe pas d'outils réellement adaptés à une exploitation à distance des collections numérisées. Jusqu'à maintenant, le besoin majeur était de stocker et d'archiver les données provenant généralement de collections et de corpus volumineux, sans se soucier réellement de la qualité de la transmission puisque les documents étaient généralement consultés à partir de CD Rom ou sur des réseaux locaux sans risque de pertes de paquets de données et garantissant des temps d'accès aux données très convenables, [DE00].

Du point de vue de l'utilisateur de la bibliothèque numérique, on peut constater un besoin de plus en plus grand de fonctionnalités qui ne peuvent pas encore être offertes en temps réel. Les usagers cherchent d'abord une plus grande accessibilité aux ouvrages, aux images et à leur contenu. L'accessibilité traduit une possibilité de consultation en ligne facile et rapide, ainsi qu'une possibilité d'interrogation avancée des contenus pour trouver l'information désirée. Les projets actuels de numérisation d'ouvrages anciens du patrimoine se réalisent le plus souvent entre partenaires distants nécessitant de mettre en place des protocoles très étudiés de consultation à distance et nécessitant en tout premier lieu une transmission fiable et rapide des données, <http://www.culture.gouv.fr/culture/mrt/numerisation/>, [EM03], [WE03]. L'ensemble de ces besoins nécessite la mise en place de méthodes de compression des images spécifiques plus efficaces ainsi que la définition de formats de données qui permettent une interrogation à distance de leurs contenus. Les formats doivent aussi permettre de gérer tous les contenus, par définition très hétérogènes, d'une bibliothèque numérique (textes, images, structures, annotations par des objets multimédia, hyperliens...) et doivent pouvoir être robuste aux perturbations et aux erreurs de transmission.

De nombreux projets de numérisation se sont ainsi vu concerner par de tels besoins et ont ainsi vu le jour autour de ces thématiques émergentes. Le travail de ce MASTER s'inscrit précisément dans le cadre du projet d'ACI MADONNE autour de la sauvegarde et de la valorisation de données patrimoniales, [MA05]. Cette ACI traite des « masse de données », représentée ici par les collections d'ouvrages numérisés, sous forme d'images scannées, pour lesquelles les techniques traditionnelles des bases de données sont inopérantes. L'exploitation et la valorisation à venir de ces collections d'images n'ont pas encore trouvé de réponses satisfaisantes, du fait même de leur caractère faiblement structuré. La génération de ces entrepôts de données, présentés sous forme de collections de documents hétérogènes faiblement structurés soulève le problème de la recherche d'information et de la navigation au sein de ces corpus.

Dans le cadre de ce vaste projet, il a été proposé une démarche commune d'exploitation et de valorisation des corpus s'adaptant aux différentes représentations de l'information que l'on peut rencontrer dans ces documents patrimoniaux souvent très hétérogènes (documents textuels, imprimés ou manuscrits, images, ou encore illustrations graphiques, plans et cartes territoriales). La consultation en mode image des documents patrimoniaux qui est visé par la plupart des projets de numérisation et de valorisation de corpus suppose dans un premier temps un archivage qui exige d'examiner de manière approfondie les possibilités spécifiques de compression de ces masses de documents.

Il n'existe pas de travaux antérieurs relatifs à la compression des données du patrimoine écrit. Les seuls que l'on peut recenser et qui ont ouverts la voie à ces nouvelles recherches concernent le patrimoine ancien imprimé, notamment autour des premiers ouvrages de la fin du Moyen Age et de la Renaissance [DE00].

I.2 Transmettre et diffuser la patrimoine ancien : un défi au temps et aux techniques

Le volume des images de documents numérisés pose des problèmes importants en termes de stockage et de transmission sur des réseaux limités en débit. La compression de ces données devient incontournable. Les formats de compression existants ont montré leurs limites sur les images de documents très abîmés présentant de fortes irrégularités tant sur les zones de textes que sur l'arrière plan (voir section II). Les documents manuscrits anciens, en particulier les brouillons d'auteurs, les documents en langues arabes, mais également les schémas et les plans illustrent très bien ce genre de difficultés, [EM03, WE03, MA05]. Les formats comme DjVu et DEBORA¹ sur lesquels nous reviendrons dans la suite du mémoire fonctionnent de manière très satisfaisante sur des documents imprimés pour lesquels il est possible de comprimer l'image en utilisant la redondance des formes de caractères. Sur les documents très irréguliers, comme les documents manuscrits anciens que nous avons privilégiés pour cette étude, ces approches ne sont pas satisfaisantes. L'échec de ces approches est donc en partie lié à la grande complexité des formes (dans le cas du manuscrit, il s'agit des tracés présents dans les textes) rendant leur localisation souvent bien difficile et pour laquelle aucune méthode de segmentation n'est actuellement efficace. La difficulté réside ainsi dans la localisation de ces redondances qui ne se limite pas aux caractères mais à toute partie de l'image qui présente des similarités distribués sur toute la surface de la page, voir figure 1. La notion de similarité devient ainsi centrale.

Les documents manuscrits sur lesquels nous travaillons présentent typiquement une vraie difficulté en terme de segmentation en lignes et en mots liés en partie à la dégradation naturelle (ou d'usage) de ces documents au fil du temps et principalement à l'imprécision des tracés et à la variabilité des formes présentes dans les textes écrits (plus spécialement des textes et brouillons d'auteurs qui ont servi d'images de tests). Nous avons choisi de centrer notre contribution sur cette catégorie d'images. Nous avons pu faire le premier constat que l'étape de la localisation des lignes des textes à l'aide des méthodes existantes était inopérante compte tenu de la non linéarité des mots et des lignes rencontrés, [AB05]. Qu'il s'agisse de la mise en page de cette classe de documents non modélisable a priori ou de la forme du tracé liée à l'allure générale et à la courbure des lignes de texte, nous avons pu souligner l'inefficacité des approches usuelles de segmentation des formes présentes dans les textes imprimés ou les manuscrits très réguliers (approches par recherche de connexités binaires, filtrages directionnels strictement horizontaux, projections en lignes et en colonnes...). Il faut ainsi pouvoir tenir compte des irrégularités de structures (organisation sur la page) mais également des irrégularités des formes (irrégularité des espaces entre les lignes qui entraînent de nombreux contacts entre les caractères de lignes voisines, irrégularité des espaces entre les caractères entraînant la fusion involontaire des caractères, voir figure 1).

Pour remédier à ces difficultés, nous nous sommes orientés vers une nouvelle méthodologie permettant de localiser précisément les zones de textes, de repérer les formes redondantes sans avoir recours à une recherche de formes connexes binaires difficiles sur ce type de support et conduisant à un partitionnement de l'image centré autour des zones à fortes variabilités (les traits). Rappelons ici que le partitionnement de l'image est à la base de nombreuses méthodes de compression et qu'il peut être réalisé de façon idéale par une recherche guidée par les régions d'intérêt. A ce jour, peu de techniques permettent un tel partitionnement : on peut toutefois citer l'approche de la compression fractale basée sur la recherche d'auto-similarités entre blocs (à partir d'un partitionnement sous forme géométrique, rectangulaire ou triangulaire, [ZH95] et les techniques de compression DjVu et DEBORA qui procèdent en une séparation par couches d'informations à l'aide d'un partitionnement de zones binaires exclusivement textuelles, [DE00]. Dans ce domaine où la recherche est encore à ses débuts, notre contribution s'inscrit dans la mise en place d'un outil de partitionnement des images contrôlant les

¹ DEBORA: Digital accEss to BOoks of the RenAissance. Projet européen dont l'objectif est de concevoir un ensemble d'outils permettant l'accès distant et collaborative à des livres numérisés du XVI^{ème} siècle.

régions d'intérêt rassemblées autour des traits (de plume) et des formes redondantes. Cet outil sera également au service de la recherche et de la caractérisation des similarités fortement corrélées au contenu des images. Nous avons cependant cherché à le rendre le plus générique possible en faisant des choix adaptés à la plus grande variabilité de documents, voir figure 1.

Ce mémoire s'articule autour de trois parties : une partie de synthèse de l'existant présentant les différents domaines de recherche impliqués dans ce travail. Nous aborderons d'une part, les approches de compression d'images et leurs performances sur des images de documents anciens et d'autre part les méthodes de recherche de similarités sur des images naturelles et les images de traits binaires. Ces deux domaines de recherche nous permettent de poser les bases de notre contribution pour laquelle aucun travail antérieur ne peut directement être cité en exemple. Nous présentons ensuite au chapitre III les différentes étapes de notre approche de recherche et de codage des similarités partielles présentes dans les formes des écritures et proposons finalement les premières étapes d'intégration de ce travail à la compression sans perte d'images de documents du patrimoine.



Figure 1 : Illustration de la grande diversité des manuscrits du patrimoine (du Moyen Age au 18^{ème} siècle), par la forte irrégularité des tracés et des styles, [EC05], [IR05].

II. SYNTHÈSE DE L'EXISTANT POUR UNE MISE EN LIGNE DES OUVRAGES: Compression et recherche de similarités

Les usagers des bibliothèques numériques désirent de plus en plus de fonctionnalités qui ne peuvent pas encore être offertes en temps réel. L'accessibilité aux ouvrages, aux images et à leurs contenus nécessite une consultation en ligne facile et rapide. La compression est devenu un outil indissociable d'un système de stockage et de mise en ligne d'images. Les formats traditionnels utilisés pour la compression des images (JPEG, DjVu, DEBORA, méthode fractale) ne tiennent pas compte de la spécificité des ouvrages très dégradés rendus peu lisibles et contenant des tracés manuscrits irréguliers, une disposition des formes souvent imprévisible et des caractéristiques des tracés très variables. Dans cette partie, nous présentons les spécificités de ces méthodes en soulignant la nécessité de reconsidérer les approches de la compression dans le cas des documents anciens.

II.1 La compression générique avec perte : solution peu adaptée aux documents

Les méthodes de compression d'images avec perte utilisent les imperfections de l'œil humain pour détruire l'information non perceptible par l'homme. Cette méthode consiste à enlever toutes les hautes et moyennes fréquences d'une image sans modifier la perception de celle-ci. Dans une image de documents, ces hautes et moyennes fréquences sont des informations qui concernent les contours des traits et des caractères. Le format le plus répandu de compression avec perte est JPEG (Joint Photographic Expert Group), [DE00]. Il offre un gain moyen de compression autour de 1:10 avec des dégradations très importantes qui nuisent aux traitements automatiques par ordinateur (voir figure 2).

D'autre part, il n'existe pas de filtre inverse qui corrige les effets de la compression de JPEG, [KI97]. Ainsi les fonds documentaires déjà numérisés et stockés sous ce format resteront difficilement interprétables par ordinateur. Le nouveau standard JPEG2000, basé sur la compression en ondelettes, améliore très peu la compression des images de textes, [HO96]. L'échec de cette méthode s'explique par la forte présence de traits dans ces images dont la complexité peut être difficilement reproduite par un codage générique. JPEG n'est pas adapté à la complexité des traits des caractères et à leurs localisations aléatoires dans une image de texte.



Figure 2 : 2.1 Image originale. 2.2 Image Comprimée JPEG à 90% qualité (extrait d'un ouvrage du CESR de Tours)

L'étude des méthodes génériques de compression montre enfin qu'aucune technique n'est actuellement disponible et efficace en terme de compression et de rendu sur les images du patrimoine ancien. La solution passe ainsi par des moyens de compressions adaptés aux images des documents et à la conservation des informations essentielles pour la lecture et le traitement informatique. Des approches plus récentes, telles que DjVu, Tiff-FX développé par Xérox et procédant par une séparation en couches compressées de l'information ont fourni des résultats prometteurs allant jusqu'à des taux avoisinant les 1 :100.

II.2 Les approches binaires pour une compression sans perte d'informations

La plupart des bibliothèques numériques fournissent des accès aux documents par leurs versions binaires faciles à exploiter et pour lesquels il existe des formats de compression très intéressants sans perte d'information. Le CCITT (comité consultatif international télégraphique et téléphonique) propose des formats basés sur le codage des séquences binaires les plus fréquentes, [DE00]. Les formats TIFF et JBIG par exemple sont fondés sur une telle approche de codage. JBIG produit un encodage prédictif des séquences mais n'est pas réellement adapté à une transmission progressive sur un réseau haut débit, [KI97], [HO98].

Par ailleurs, les différentes tentatives de l'adapter à ce type de transmission n'ont pas réellement abouti à une compression satisfaisante du fait de la perte de précision par une prise en compte trop faible des structures fines. Finalement, les améliorations successives du schéma JBIG par le codage des images en niveaux de gris respectant les standards internationaux ISO ont permis à des formats sans perte très compétitifs de voir le jour, [HO96] : des formats étendus de TIFF, TIFF-FX par exemple proposé par Scan Soft, ou encore le format XIFF (eXtended TIFF) proposé par Xerox. Cette méthodologie se retrouve également dans le format DjVu d'AT&T, [DJ05, BO98].

II.3 La compression par recherche d'auto-similarités - méthode fractale

La compression fractale a pour principe qu'une image n'est qu'un ensemble de motifs identiques en nombre limité, sur lesquels on applique des transformations géométriques affines de types translations, rotations, symétries, agrandissements, ou encore réductions. Evidemment, plus l'image possède de motifs vérifiant ces propriétés, meilleur en est le résultat, [ZH95].

Comme pour le format JPEG, l'image est découpée en blocs de pixels, mais ils sont ici de tailles variables. Il faut ensuite détecter les redondances entre ces blocs à diverses résolutions. On parle alors de transformations fractales basées sur un opérateur contractant. Ces transformations décrivent l'image de plus en plus finement. A la fin de ce processus, on ne stocke pas le contenu d'un bloc autant de fois qu'il a été "vu" dans l'image mais seulement les équations mathématiques permettant de représenter le contenu de ces carrés, [ZH95]. Au final, on obtient une structure présentant des caractéristiques similaires à des échelles différentes. Pour retrouver l'image il suffit alors de décrire les transformations qui ont été appliquées aux blocs initiaux. Ce processus rend la compression indépendante de la taille de l'image. De plus, l'image produite est vectorisée et ne subit pas les effets de la pixellisation, contrairement au JPEG. Ce phénomène est surtout visible lors d'un zoom par exemple, l'image fractale peut devenir floue mais ne « pixellise » pas. Ceci est dû au fait que lors de l'agrandissement, ce ne sont pas les pixels qui sont élargis, mais toute l'image qui est recalculée mathématiquement.

Pour la compression fractale, on trouve plusieurs types de partitionnement, les plus connus sont en Quadtree et en rectangle et Delaunay (triangles). La plus efficace de ces méthodes dans le rapport qualité/compression est la méthode par partitionnement Delaunay.

La méthode de compression fractale appliquée aux images de textes (textes imprimés principalement) permet de trouver dans les imageries (ou domaines qui sont des blocs de pixels qui partitionnent l'image) les éléments qui peuvent être retrouvés à des échelles différentes dans l'image. Avec l'ensemble des Domaines extraits et utilisés dans la phase initiale de compression, on peut reconstruire une base de référence des traits pour l'alphabet utilisé sur ces images de textes. Cette base de référence représente les similarités internes contenues dans l'écriture, [SE04]. Il faut noter que la forme géométrique des imageries des domaines dépend du type de partitionnement utilisé. Cependant l'absence de similarité interne à grande échelle dans les écritures montre l'échec de cette méthode et le mauvais taux de compression (1:16) obtenu sur les différentes images que nous avons testées.

redondance des formes sur tout un livre et effectue la compression page par page de façon indépendante ce qui limite considérablement le nombre de formes redondantes caractérisables.

La méthode DEBORA

Cette méthode est adaptée aux images de documents imprimés et repose sur la très forte redondance des caractères par rapport aux éléments graphiques de formes aléatoires. Comme pour les méthodes précédentes, on sépare l'arrière plan représentant l'image de support pour le comprimer avec pertes. Le principe consiste à diviser l'image de document en trois plans dont chacun est compressé différemment avec un modèle de compression adapté.

- le plan textuel (image binaire) : il comporte les images de caractères qui ont la propriété d'être fortement redondants sur l'ensemble du livre. Ce plan peut être compressé avec un taux extrêmement élevé de compression en utilisant l'appariement des formes.
- le plan graphique (image binaire) : il contient les éléments graphique qui ne sont pas des caractères et qui se caractérisent par une très faible redondance. Il n'est pas utile de compresser ces éléments vue la faible redondance de formes dans le livre.
- l'arrière plan (image couleur) : il représente l'image du support papier privé des caractères et des éléments graphiques. Il supporte des compressions avec pertes comme JPEG.

Pour la lisibilité et l'interprétation des images, le remplacement des formes réelles par leur formes prototypées (contenu dans le plan textuel) ne modifie généralement par le confort de lecture ni les résultats d'un traitement. En revanche, les petites variations contenues à la frontières de formes ou au cœur des formes elles-mêmes sont bien souvent fondamentales pour le chercheurs et l'historiens. Il est donc important pour transmettre une information fidèle à l'origine d'estimer une carte de compensation qui contient les données résiduelles obtenues en faisant la différence entre les formes prototypées et les formes réelles. Ce rendu fidèle à l'origine redonne à la page numérisée sa signification d'origine. La figure 4 présente un exemple de plan de compensation calculé par l'approche DEBORA en effectuant une différence logique entre les deux images binaires : l'image d'origine et l'image reconstruite à partir du remplacement des caractères par leur prototype.

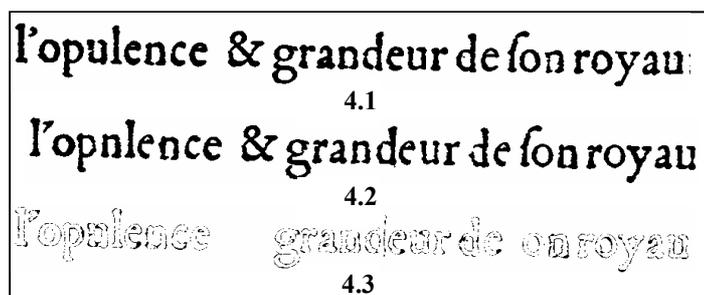


Figure 4 : La compensation. 4.1 Image binaire originale. 4.2 Image binaire reconstruite à partir des prototypes de caractères. 4.3 Plan de compensation (contenant les « résidus ») issu de la différence entre les images 4.1 et 4.2, [DE00].

Enfin, la figure 5 montre les différences visuelles résultant de l'utilisation de différents algorithmes avec leurs taux de compression respectifs. Les résultats visuels de la compression de DEBORA sont comparables avec ceux de la version sans perte de la compression DjVu. Cependant, la compression avec perte de DjVu introduit de nombreuses erreurs qui peuvent ne pas satisfaire l'utilisateur ou l'historien qui cherche à faire un travail plus minutieux sur l'image.

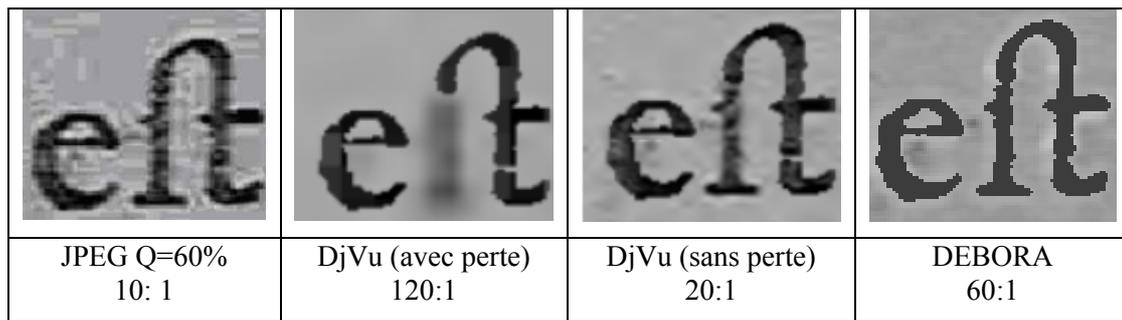


Figure 5 : Comparaison entre méthode de compression sur un extrait imprimé de la Renaissance.

On peut finalement conclure que pour les documents manuscrits (mais également l'imprimé dégradé, les cartes et les plans contenant une information partagée entre le texte et le dessin), il est difficile de séparer les différentes couches d'informations et d'appliquer une approche de compression par redondance des formes (exploitation de l'appariement et de la substitution des formes). Pour ces raisons, les méthodes de compression par prototypage et substitution de formes binaires ne sont pas adaptées à la compression des images de documents manuscrits : il faut envisager une approche plus robuste à la grande variabilité des formes présentes dans le tracé et pouvoir les segmenter correctement.

Il faut donc développer de nouvelles approches exploitant des similarités de formes à des niveaux plus fins. Pour cela, nous avons dégagé dans la partie suivante les principaux outils de recherche de similarités existants en traitement d'images non limités aux images binaires et aux images de documents afin de proposer une contribution en rapport avec les contraintes imposées par les images des documents anciens.

II.5 Généralisation autour de la recherche de similarités

Similarités, silhouettes et points d'intérêt.

La notion de similarité de formes est globalement difficile à définir sur des tracés présentant en tous points des différences. Plutôt que de rechercher à trouver les ressemblances globales, on peut définir les formes à partir d'un ensemble de points caractéristiques, les *points d'intérêt* révélateurs de configurations locales jugées intéressantes, voir figure 6.

Les méthodes généralement développées pour réaliser un appariement de formes se base soit sur des techniques de correspondances de points de contours révélateurs de la silhouette des formes, soit sur des techniques basées sur des descripteurs de formes nécessitant la description d'un grand nombre de prototypes invariants pour établir des correspondance avec les éléments à retrouver, [LA93], [CO95], [SC97]. Ces dernières techniques ont été largement employées pour la localisation et la reconnaissance de chiffres selon un vocabulaire restreint, [LE98] et la recherche de visage dans des environnements naturels [MO00]. Elles demeurent néanmoins inefficaces sur des images de texte contenant un vocabulaire plus étendu et à plus forte raison si le texte est manuscrit.

D'autres techniques plus souples nécessitant moins de connaissance a priori sur les objets à trouver (et à reconnaître) sont basées sur des informations globales de structures, de contours et de silhouettes. On peut notamment citer les travaux basés sur des descripteurs de Fourier dans [ZA72], de recherche de squelettes dans [SH98] utilisant des comparaisons flexibles entre contours à l'aide de distances dynamiques. Les points de contours ont été introduits par Carlsson dans [CA99] qui utilise la notion *d'ordre* de structure. Ces travaux ont trouvé de nombreux échos notamment à travers les propositions de Johnson [JO97], Chui [CH00], [PE91], [SC00]. La catégorisation des

points et leur stabilité à travers les changements d'échelle et de perspective ont été étudiées par Schmidt dans [SC00].

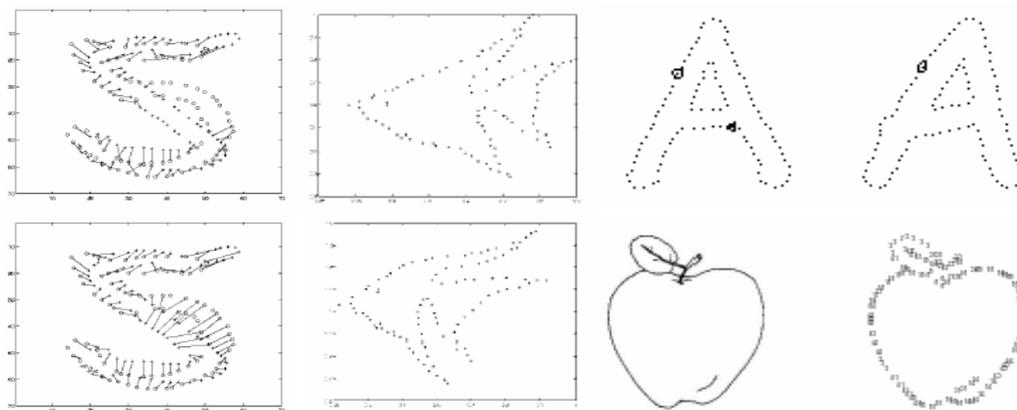


Figure 6 : Codage des frontières des objets en points discrets sur des images en niveaux de gris pour la recherche de similarités et la comparaison de formes, [BE02].

Un bilan des approches pour déterminer les configurations locales aux frontières est présenté dans [BE98, MI04, TE00]. L'échantillonnage des points de contours donne une description des formes par liste de points à partir de laquelle on peut réaliser des estimations de distances. Pentland propose dans [PE91] une approche originale basée sur la correspondance de points de contours robuste aux déformations locales des objets, et invariante au changement d'échelle, voir figure 7. La limitation de ce type d'approches est liée une fois encore au fait qu'un seul objet est à analyser à la fois (une image contient un seul objet), ce qui limite considérablement la complexité globale du problème.

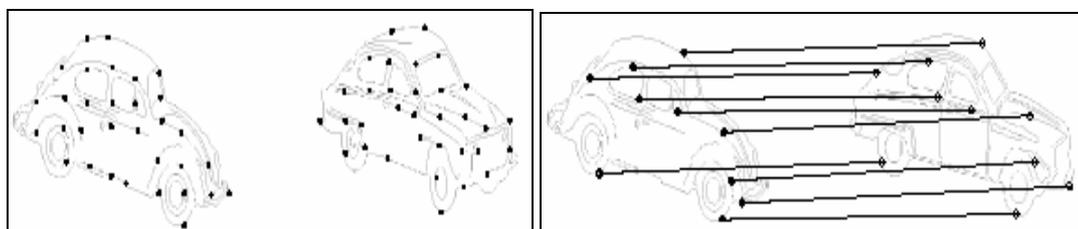


Figure 7 : Comparaison de formes à partir de l'appariement de points de contours, [PE91].

Le détecteur de points d'intérêt de Harris est très souvent utilisé dans ce type d'approche. Il permet de localiser avec une grande précision les maxima locaux et minima locaux, les pics et les vallées présents sur les hautes fréquences des tracés, [BE98]. La recherche de configurations similaires entre deux formes semblables s'effectue ensuite par une analyse de points saillants de contours et des combinaisons stables qu'on peut retrouver entre deux formes, [MI04, SC00]. Cette approche présente l'intérêt de pouvoir traiter les images en couleurs, et en niveaux de gris mais demeurent difficiles à utiliser sur des textes très imparfaits présentant de trop nombreuses variations aux frontières, [AB05]

Ce sont finalement vers des approches hybrides alliant la forme au contour qui ont vu les jour très récemment, [LO04, SC00]. Ces méthodes sont basées sur des estimateurs de gradients locaux traduisant les changements d'intensité aux frontières et définissant des invariants locaux nécessaires à une recherche d'objets insensible à la rotation.

C'est dans cette dernière catégorie d'approches que nous avons choisi d'inscrire notre contribution basée sur une coopération entre la forme et son contour. La méthodologie complète est présentée dans les sections suivantes.

III. NOTRE PROPOSITION DE RECHERCHE DE SIMILARITES ADAPTEE AUX TRACES

III.1 Principe général de l'approche

Les formes redondantes que nous allons chercher ne sont pas nécessairement des mots ni des lettres : elles peuvent être simplement constituées des petites boucles des lettres (les graphèmes) qui se retrouvent localisés à d'autres endroits dans le tracé manuscrit et ceci à une transformation géométrique près (un changement d'échelle, une rotation ou bien une translation). Le graphème est considéré ici comme une unité graphique stable pour un même scripteur associé à un mouvement de plume continu sans reprise ni levé de plume. Au delà de cette unité, les similarités perceptibles deviennent plus incertaines, moins régulières et dans tous les cas très difficiles à évaluer. La figure 8 montre quelques exemples de tracés similaires, sans qu'il s'agisse de lettres similaires. C'est donc bien au niveau du trait et non pas au niveau de ce qui est écrit que nous cherchons la ressemblance.

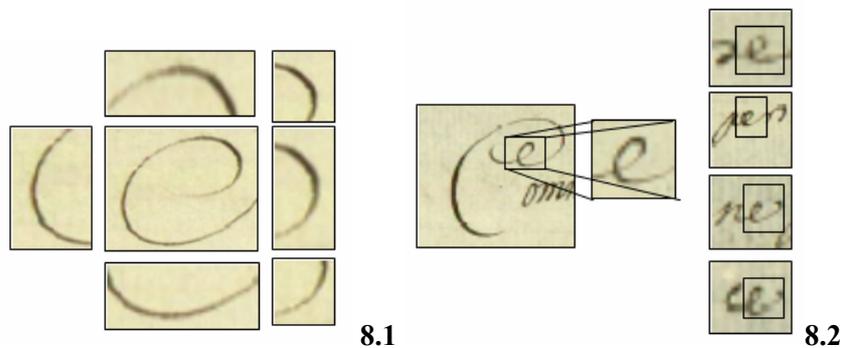


Figure 8 : Granularité des ressemblances. **8.1** Similarité de traits sur la base des graphèmes. **8.2** Tracés similaires sans rapport avec le sens dégagé.

Notre approche consiste donc à caractériser, dans un premier temps dans les images en niveaux de gris, les tracés en localisant une version simplifiée des formes sous la forme de *motifs orientés (élémentaires ou composés)* à partir desquelles nous définissons des indices de similarité caractéristiques de la redondance visuelle des tracés d'origine. Nous avons ainsi développé une approche de localisation des zones de traits et de repérage des formes redondantes en exploitant conjointement des techniques basées sur les filtres directionnels de Gabor et des descripteurs spécifiques associés notamment au détecteur de contours de Deriche. Ces outils nous permettent de définir de manière stable les lignes d'orientations identiques et les zones de changements d'orientations présentes sur les contours des traits. Cette recherche de similarités a été conçue pour être robuste aux changements d'échelle, aux rotations et bien sûr aux translations. Le principe général de la recherche de similarités est présenté ci-dessous (figure 9). Les étapes que nous allons développer y ont été numérotées.

En premier lieu (étape 1), nous allons extraire du tracé de l'image originale quatre *cartes* complémentaires contenant les portions aux frontières des formes dans les directions respectives 0° , 45° , 90° , 135° . Une analyse fréquentielle des images de documents est mise en place à cet effet (section III.2). Le marquage de ces directions est une pré-segmentation des traits d'écriture en graphèmes orientés. Parallèlement à cette étape, nous conduisons une segmentation des formes pleines cette fois (étape 2) à l'aide d'un outil original de croissance de régions, sous contrainte des frontières imposées par le filtre de Deriche (section III.3). L'association des formes pleines et de leurs contours marqués selon les quatre directions conduit à une caractérisation et une signature du tracé (étape 3) nécessaire à la définition complète et à la recherche hiérarchique des similarités partielles (étape 4). Ces deux étapes sont présentées à la section III.4.

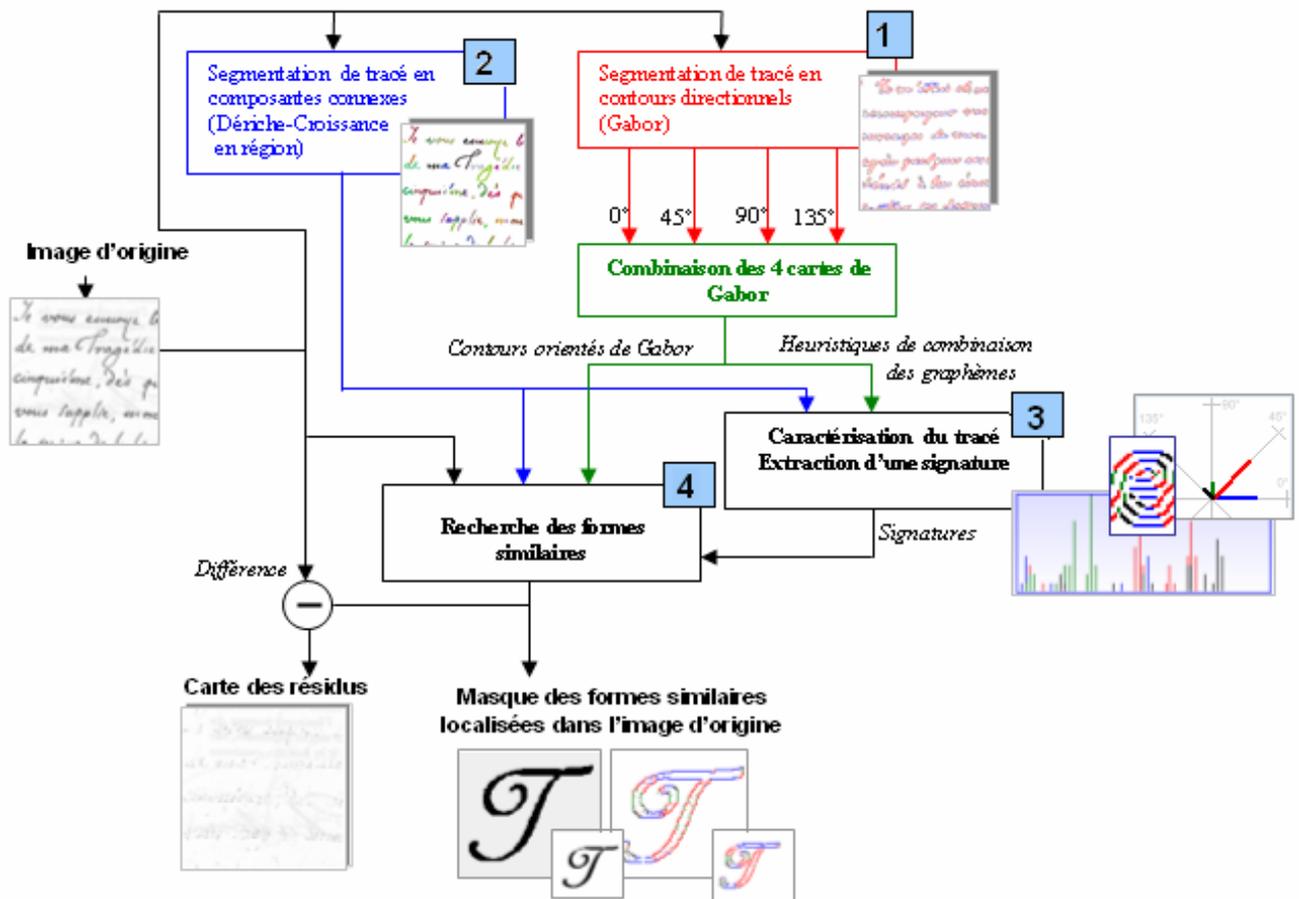


Figure 9 : Synoptique de notre proposition de recherche de similarités partielles pour la compression.

III.2 Analyse fréquentielle des manuscrits

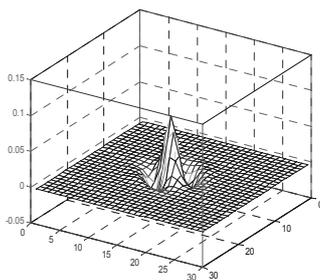
Dans cette partie nous présentons une approche originale de segmentation des images de traits en niveaux de gris (formes cursives et écritures de toutes origines) en régions centrées sur les contours orientés. L'application d'un banc de filtres de Gabor paramétré reposant sur une décomposition du domaine fréquentiel de l'image des documents est réalisée pour localiser la présence de segments orientés dans les écritures.

Le filtre de Gabor : pour un accès au contenu fréquentiel des images de manuscrit.

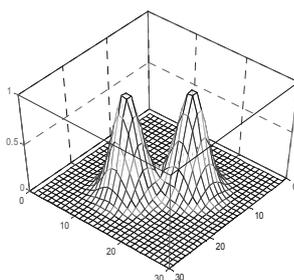
Le filtrage de Gabor constitue un excellent compromis espace - fréquence pour la représentation du contenu des images, [WE98, YA03]. C'est en règle générale cette étape de paramétrage qui est laborieuse et rarement automatisable. Une fonction de Gabor se définit comme une fonction gaussienne modulée par une onde sinusoïdale :

$$G(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}} e^{j2\pi(xf_{x_0} + yf_{y_0})}$$

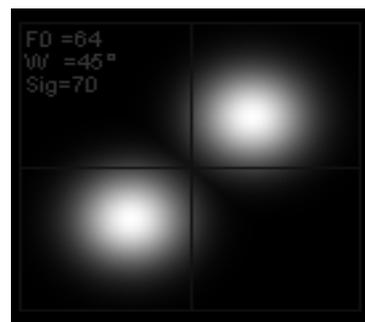
avec σ_x et σ_y la largeur respectivement selon x et y de la fonction gaussienne et f_{x_0} et f_{y_0} les fréquences spatiales de la modulation. Ce filtre, de type passe-bande orienté, a une réponse impulsionnelle complexe, voir figure 10.2.



10.1



10.2



10.3

Figure 10 : 10.1. Filtre de Gabor dans le domaine spatial 10.2. Réponse impulsionnelle de filtre de Gabor avec deux fonctions gaussiennes dans la direction 45°. 10.3 Vue 2D de la réponse impulsionnelle de filtre de Gabor dans la direction 45°.

La fréquence centrale f_0 est donnée par $\sqrt{f_{x0}^2 + f_{y0}^2}$, l'orientation θ par $\arctan\left(\frac{f_{x0}}{f_{y0}}\right)$ et la largeur de bande par σ_x et σ_y . La réponse fréquentielle du filtre de Gabor est donnée à la figure 10.1 et 10.3 selon l'expression suivante :

$$TF(G)(f_x, f_y) = e^{-2\pi^2(\sigma_x^2(f_x - f_{x0})^2 + \sigma_y^2(f_y - f_{y0})^2)}$$

Exploitation du banc de filtres de Gabor et construction des 4 cartes directionnelles

Le filtre de Gabor permet de filtrer les fréquences orientées dans une direction et à une fréquence donnée, [WE98]. Ainsi, en privilégiant les hautes fréquences présentes sur les contours des formes et en filtrant plus massivement les données de fréquences globalement plus faibles représentatives du fond des pages, nous parvenons à un découpage de l'image en plans directionnels sur lesquels repose ensuite la recherche des similarités. Cette analyse directionnelle se base sur un banc de 4 filtres (N=4 avec $\theta=0^\circ, 45^\circ, 90^\circ, 135^\circ$) partitionnant le domaine fréquentiel en différents canaux, chacun étant caractérisé par des secteurs angulaires correspondant aux 4 directions retenues. Le réglage des paramètres f_0, σ est déterminant pour localiser de façon précise les hautes fréquences, Ces deux paramètres influent en particulier les résultats de la segmentation en contours orientés.

L'image est alors décomposée en quatre cartes directionnelles contenant chacune un ensemble de zones de l'écriture orientées dans la même direction. Le regroupement des plans directionnels nous permet de reconstruire le contour de l'écriture de notre image et de confirmer ainsi le recouvrement complet des formes du tracé, autorisant de ce fait à ne pas considérer de directions supplémentaires. Le recouvrement est le résultat de *OU logique* entre les quatre cartes, voir figure 11.

Le paramétrage des filtres de Gabor selon les 4 directions est naturellement complexe. Nous avons choisi d'automatiser les réglages en nous inspirant de la méthode d'optimisation de Pratt, [PR78], [LI05]. Les deux paramètres de filtres (fréquence centrale f_0 , largeur σ) sont initialement choisis de façon expérimentale. La méthode d'optimisation du choix de ces paramètres repose sur une évaluation de résultats de la segmentation en contours pour chaque valeur de f_0 et de σ . Pour cela, on estime à partir d'une image de document bien contrastée et peu dégradée une distance entre sa carte de contours optimale (présentant le meilleur compromis entre bonne localisation et bonne détection des zones de contours) et les différentes cartes de contours choisies avec des valeurs variables de paramètres. Ces valeurs sont ensuite appliquées à l'ensemble des extraits de la base.

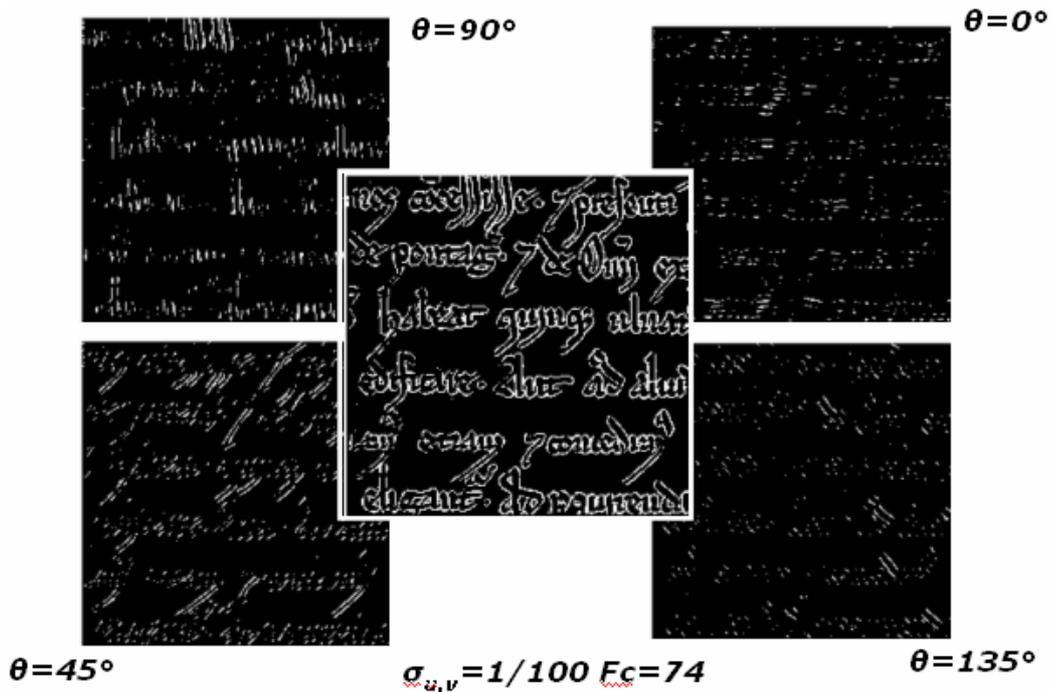


Figure 11 : Reconstruction « brute » de la carte des contours à partir des 4 cartes directionnelles, (l'image initiale a été extraite de la figure 1).

Suppression des non-maxima locaux et reconstruction des contours par hystérésis

Comme on peut le constater à la figure 11, les contours issus du filtrage de Gabor ne sont ni unitaires ni très précis en localisation. Il a donc fallu remédier à cet inconvénient en réduisant la largeur des contours à un seul pixel afin de faciliter leur analyse et leur codage. Pour cela, nous procédons à la suppression des réponses de Gabor non maximales. Le principe développé dans [BR03] et [KR99] est le suivant : dans la direction orthogonale à la direction de contours, seul le point contours (notée G) dont l'amplitude est maximale est conservé. On ne conserve donc que les maxima locaux.

Dans notre application, après la construction « brute » des 4 cartes de directions et la suppression des non maxima locaux, les points de la carte finale reconstruite de Gabor sont filtrés grâce à deux seuils distincts, S_h et S_b ($S_h > S_b$). En dessous du seuil bas S_b , on considère qu'il n'y a pas de contours. Au dessus du seuil haut, on décide qu'il y a contour : ces contours sont ensuite complétés par les pixels compris entre S_b et S_h si et seulement s'ils sont connexes entre eux. En résumé, les étapes de l'algorithme sont les suivantes :

- Si $G(x,y) < S_b$ on considère $P(x,y)$ comme n'étant *pas* un point contour.
- Si $G(x,y) > S_h$ on considère $P(x,y)$ comme étant un point contour.
- Si $S_b \leq G(x,y) \leq S_h$ on ne garde le point $P(x,y)$ que s'il a au moins un pixel voisin de contour supérieur à S_h ($G(x_v, y_v) > S_h$).

Le seuil S_b permet donc de filtrer les réponses indésirables liées au bruit. Le seuillage haut ne garde que les composantes importantes mais incomplètes. La troisième étape permet de compléter ces contours. La figure 12 illustre le résultat des traitements de la carte des directions reconstruite.

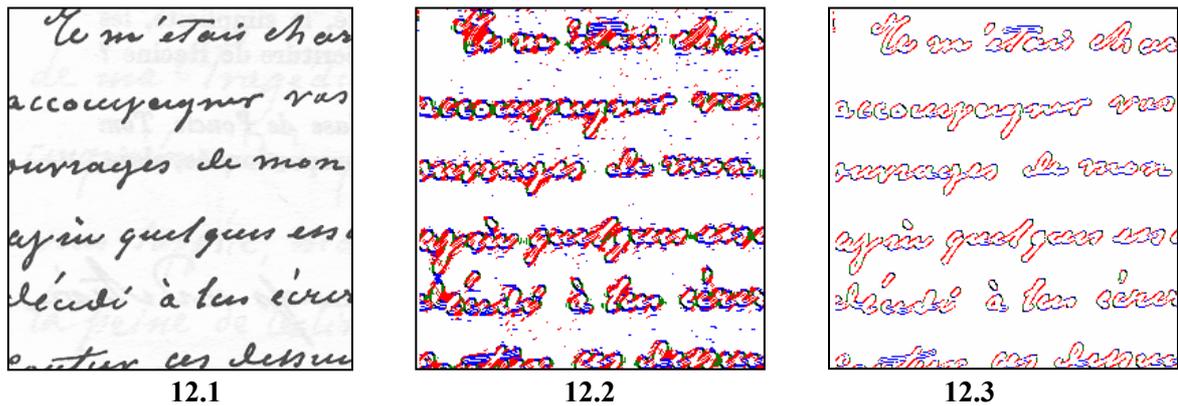


Figure 12 : 12.1. Image d'origine. 12.2. Carte de contours sans suppression des non maxima locaux avec un seuillage simple. 12.3. Carte des contours reconstruite après traitement avec les valeurs $S_b=10$ et $S_h=20$.

III.3 Segmentation coopérative pour une décomposition en formes connexes

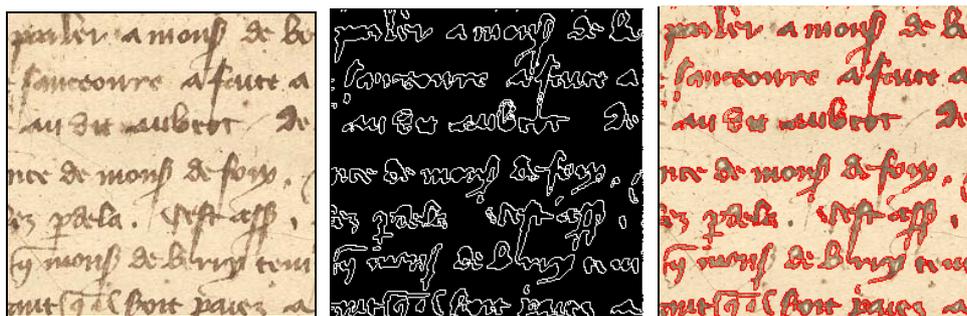
A ce stade, nous disposons de contours orientés. Nous voulons maintenant associer ces contours ou segments orientés à leur forme d'origine, c'est-à-dire retrouver très précisément l'intérieur correspondant à chaque contour. L'approche que nous proposons permet de nous dégager d'une binarisation directe des formes qui n'est pas jugée pertinente sur des images de documents du patrimoine. Nous utiliserons pour cela la méthode de détection de contours de Deriche qui garantit une bonne précision et une grande rapidité de segmentation en contours associée à une méthode de croissance de régions partant d'un germe intérieur aux éléments obtenus par la *carte binaire* et guidé par la *carte des contours de Deriche*. La construction des deux cartes est présentée ci-dessous. Les contours de Deriche représentent ainsi une *contrainte* pour l'évolution du processus de croissance de régions : on parle alors de processus de segmentation par contrainte. On obtient ainsi une représentation très fidèle des formes présentes dans le tracé qu'une simple binarisation n'aurait pas su faire émerger, du fait même de la présence de bruit et du faible contraste existant parfois aux frontières des traits.

L'intérêt de cette coopération est lié à la nécessité de considérer à la fois les contours pour signer les formes et leur contenu pour conserver l'information d'épaisseur de la plume du scripteur. L'association des deux informations est indispensable dans ce cas. La dualité des deux concepts « contours et régions » permet par ailleurs de corriger des erreurs relevant de l'usage exclusif de l'une ou l'autre des approches.

Localisation des contours par l'opérateur de Deriche

La méthode de détection de contours de Deriche (que l'on ne présente plus) est très largement exploitée en traitement d'images pour la localisation des objets et la séparation en régions, [BR03], [SP97]. Les contours de Deriche possèdent des propriétés remarquables liées à la grande précision de localisation des frontières. Cette approche ne permet cependant pas un marquage directionnel des contours. C'est pour cette raison que nous ne pouvons l'utiliser seule, et que nous lui avons associé les filtres de Gabor : Deriche pour localiser précisément les contours, Gabor pour avoir leur orientation.

Cet opérateur ne dépend que d'un seul paramètre, connu sous le nom de α de Deriche. Pour un α grand, le filtre détecte les moindres variations. Pour α petit, il y a peu de réponse, on augmente ainsi la robustesse vis-à-vis du bruit mais cela induit obligatoirement une perte en localisation. Dans la pratique, il est d'usage d'utiliser $\alpha = 1$. La figure 13 montre des résultats de l'application du filtre de Deriche sur des extraits de documents très bruités. Ces résultats permettent de juger la grande précision dans la localisation des frontières des traits.



13.1

13.2

13.3

Figure 13 : Résultats de l'application du filtre de Deriche sur des images de manuscrits médiévaux, [EC05]. **13.1** Images originales. **13.2.** Contours de Deriche. **13.3.** Superposition Deriche / originale.

Binarisation adaptative par K-means.

L'objectif de cette étape est de calculer la *carte binaire* de l'image à l'aide d'un classifieur de type K-means. Cette étape de binarisation sera exploitée pour construire les formes connexes finales en orientant la « germination » au centre des composantes binaires limitées par les contours de Deriche. Durant l'étape de croissance de régions à partir de ces germes, la carte binaire empêche de créer des nouveaux germes sur l'arrière plan de l'image (le fond). Pour réaliser la binarisation adaptative, nous décomposons les images de document en N blocs ($N=2^{21}$ dans la pratique $N=256$), l'algorithme de K-means classe chaque pixels selon leur niveau de gris en K classes ($k=2$ dans notre cas) dans chaque bloc. La première classe représente le tracé alors que la deuxième représente l'arrière plan, [BR03].

Formation de segments connexes par croissance de régions et agrégation de points.

Cette étape consiste à regrouper les points de tracés en formes connexes. Nous avons développé pour cela une méthode de segmentation en régions par agrégation de pixels, imposant à chaque région une limitation aux contours de Deriche et une validation d'un critère d'homogénéité, [TH97].

Le principe de cette technique consiste à regrouper les pixels du tracé en se basant sur un double critère d'homogénéité et d'adjacence, [NA03, MI01]. Chaque pixel est regroupé avec son voisin si leurs niveaux de gris sont suffisamment proches. Finalement par cette approche, tous les points de contours sont associés à la région qu'ils entourent. Le regroupement de l'ensemble de ces régions constitue la carte des formes connexes, voir figure 14. Cette carte sera utilisée dans l'étape de recherche de similarités.

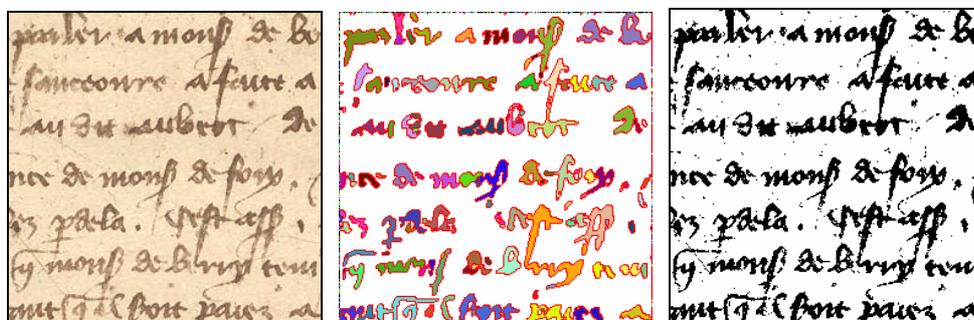


Figure 14 : Résultats de la croissance de régions. Superposition des contours de Deriche et des régions connexes. Visualisation des meilleures versions binaires (mais non satisfaisantes) obtenues ici par seuillage adaptatif, [EC05].

III.4 Signature des motifs redondants et estimation des similarités

Approche hiérarchique de la construction de la signature des motifs

On appelle *motif*, toute forme constituée d'un ensemble de graphèmes, c'est-à-dire un ensemble de segments orientés. Le nombre de segments n'est pas fixé a priori, il varie en fonction de la taille des similarités recherchées. Nous verrons notamment que la taille des similarités optimales visant la compression d'une page ne doit pas dépasser 6 segments orientés. Au-delà de cette valeur, on se projette davantage dans des applications de type *word-spotting* basées sur la recherche de mots similaires pouvant présenter de grandes variations structurelles et en deçà, les similarités limitées à un ou deux segments directionnels ne constituent pas un compromis satisfaisant pour coder une page sur toute sa surface.

La carte des contours directionnels de Gabor contient beaucoup d'informations qui caractérisent une forme ou un ensemble de graphèmes, l'objectif de cette partie est de définir une signature représentative, discriminante et invariante à des transformations simples telles que le changement d'échelle (deux formes similaires de taille différente doivent pouvoir se ressembler), la rotation et la translation. Cette signature sera utilisée pour trouver l'ensemble des formes redondantes dans une image de document. Dans la pratique, nous avons procédé par étapes successives en définissant plusieurs signatures des plus grossières au plus fines structurellement. Les performances de chaque type de signatures sont liées à leur capacité à minimiser les résidus en augmentant la possibilité d'avoir une similarité entre les formes perceptiblement similaires. Rappelons ici que les résidus constituent l'ensemble des différences entre les formes choisies comme références pour le codage de l'image et les formes réelles substituées au moment du codage. Ces différences constituent le plan de compensation qui rétablit les variations parfois infimes entre l'image reconstruite et l'image d'origine (voir section II). Pour réaliser cela, nous proposons une approche *hiérarchique* du calcul des signatures partant des propriétés les plus globales des motifs liés à la distribution et la quantification des segments orientés aux propriétés de structures les plus fines basées sur la localisation et la succession des points de jonction entre segments orientés. Cette démarche hiérarchique de la recherche de similarités dans les écritures permet de traiter des informations fortement bruitées, peu stables et nécessitant de considérer tout d'abord les formes dans leur plus grande globalité. La forme va ainsi être analysée à partir d'une *fenêtre d'analyse* contenant un motif dont la taille n'est pas fixée a priori.

Similarités basées sur les roses de directions, invariance à l'échelle

Dans cette approche nous présentons, une nouvelle signature basée sur le nombre $Nap(\theta_i)$ d'occurrences normalisées des points marqués de contour selon les directions θ_i issues des filtres de Gabor (ces directions peuvent être 0° , 45° , 90° ou 135°). Ce nombre est donné par la formule suivante:

$$Nap(\theta_d) = \frac{1}{N} \cdot Card\{C_{\theta_d}(x, y) | P(x, y) \in contours\}$$

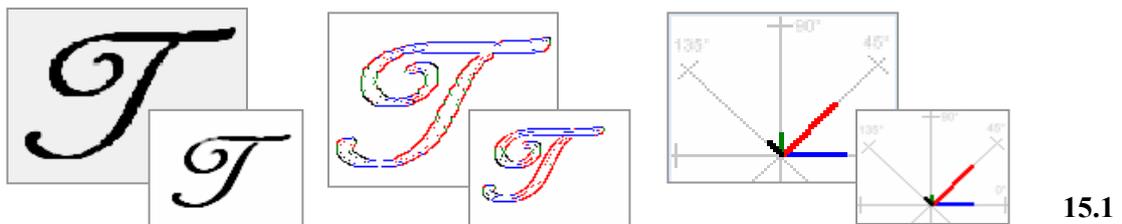
θ_i est la direction d'un point contour $P(x, y)$, $C(x, y)$ est le nombre de points de contour correspondant à cette direction, N le nombre total des points de contour dans la fenêtre d'analyse englobant le motif à étudier. L'ensemble des nombres d'occurrences des points marqués pour une forme S_f est donné par :

$$Nap(S_f) = \{ Nap(0^\circ), Nap(45^\circ), Nap(90^\circ), Nap(135^\circ) \}$$

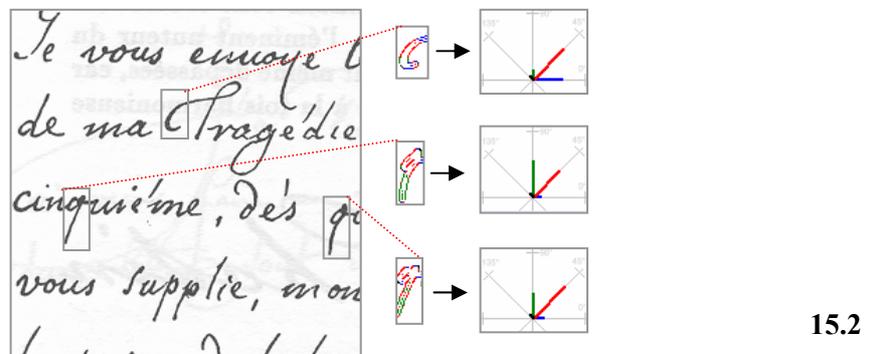
La ressemblance entre les deux formes S_f et S_{ref} est estimée à partir de la distance euclidienne entre les deux ensembles $Nap(S_f)$ et $Nap(S_{ref})$ elle est donné par :

$$D(Nap_{S_f}, Nap_{S_{ref}}) = \sqrt{\sum [(Nap_{S_f}(\theta_i) - Nap_{S_{ref}}(\theta_i))^2]} \quad avec \quad i = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$$

Ces estimations de densités locales des différents points de contours selon les quatre directions sont rendues explicites par la construction d'une rose des directions associée à chaque motif. L'analyse de la rose (selon la méthode présentée) permet d'apporter plusieurs informations importantes sur la nature des motifs comparés. Elle permet d'une part de retrouver deux formes semblables à des échelles différentes, puisque le rapport des différents segments orientés des motifs est invariant par échelle. Cette propriété confère à la mesure une robustesse aux changements de taille fréquemment rencontrés dans les textes manuscrits, voir figure 15.1. Elle permet d'autre part d'écarter aisément des motifs ne contenant pas les mêmes rapports de directions. Dans ce cas, on peut considérer que les motifs ne présentent aucun degré de similarité, voir figure 15.2.



15.1



15.2

Figure 15. 15.1 Roses de directions entre deux formes similaires. Mise en valeur de l'invariance par échelle. 15.2. Rose de directions entre formes similaires et présence d'un intrus.

Cette approche donne des résultats satisfaisants mais toutefois imprécis en raison de la distribution non contrôlable des directions : deux motifs structurellement très différents peuvent avoir des rose de directions semblables sans se ressembler d'un point de vue structurel. Pour cela nous avons exploité une méthode plus fine basée sur l'analyse des histogrammes de directions respectant cette fois la localisation spatiale des directions dans les motifs traités.

Similarités basées sur les histogrammes de projection des segments orientés

Cette seconde approche est basée sur les histogrammes de projections des segments de contour marqués d'un motif. Soit M un motif inclus dans une fenêtre glissante, soit M_{ref} le motif de référence (qui va servir d'élément de comparaison). Soient $H_0, H_{45}, H_{90}, H_{135} / H_{ref0}, H_{ref45}, H_{ref90}, H_{ref135}$ les histogrammes de projections des point de contours directionnels respectivement des motifs M et M_{ref} sur les repères orthogonaux à chacune des directions ($0^\circ, 45^\circ, 90^\circ, 135^\circ$), voir figure 16.

Pour réduire la contribution des points aberrant et renforcer celle des points moyens, ces histogrammes sont convolués avec une fonction gaussienne, de telle sorte à ce que les longueurs des vecteurs histogrammes soient comparables. Ces longueurs sont représentées selon la formulation suivante :

$$T(H_0)=w, T(H_{45})=w+h-1, T(H_{90})=h, T(H_{135})=h+w-1$$

avec w et h , les nombres de colonnes et nombres de lignes nécessaires pour représenter la forme M et la forme M_{Ref} . Le principe de comparaison entre ces deux formes est basé sur la comparaison de leurs histogrammes des projections deux à deux. On dira que deux formes sont similaires si leurs histogrammes de segments orientés projetés se ressemblent suffisamment. La quantification de cette ressemblance est donnée par le calcul d'un coefficient de corrélation entre chaque deux histogrammes d'une même direction.

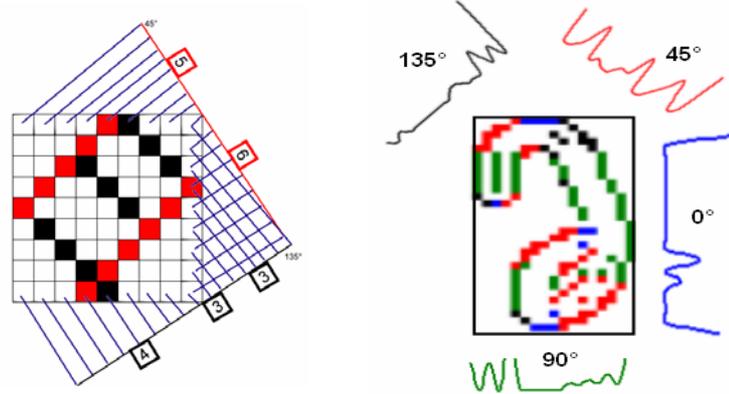


Figure 16. Construction des histogrammes par projection orthogonale des segments orientés.

Le coefficient de corrélation entre deux vecteurs histogrammes $H_i, H_{iref}(i=0^\circ, 45^\circ, 90^\circ, 135^\circ)$ est donné par le produit scalaire normalisé :

$$CoefCor(H_i, H_{iref}) = \frac{H_i \bullet H_{iref}}{\|H_i\| \bullet \|H_{iref}\|}$$

$$\|H\| = \sqrt{\sum_{K=0}^{N-1} h^2(k)} \quad \text{avec } N = \text{la taille de vecteur histogramme}$$

$CoefCor(H_i, H_{iref})$ est compris entre 0 et 1, il est maximum pour $H_i = H_{iref}$. Plus l'écart entre H_i et H_{iref} est important, plus $CoefCor(H_i, H_{iref})$ est faible. Finalement, on estime le degré de similarité Sim défini par : $Sim = \sum P(i) \cdot CoefCor(H_i, H_{iref})$ ($i = 0^\circ, 45^\circ, 90^\circ, 135^\circ$)

Plus Sim est grand (supérieur à un seuil S), plus on considère que les formes M et M_{ref} sont similaires. S fixe le degré de ressemblance entre deux formes, il peut être choisi automatiquement pour l'ensemble des motifs d'une page, voir figure 17.

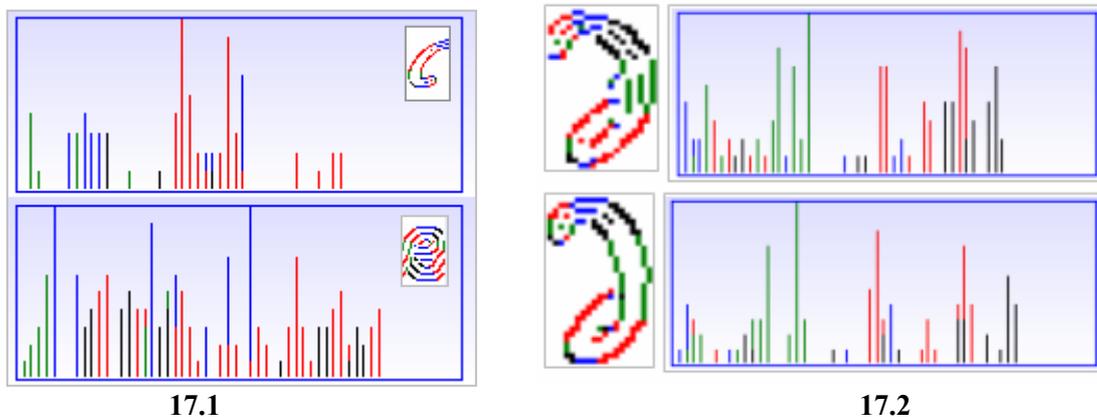


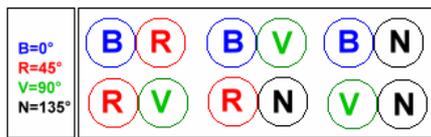
Figure 17. 1. Exemple de deux histogrammes projetés pour deux motifs non similaires. 17.2. Deux formes similaires estimées à partir d'un coefficient de corrélation = 0,55.

Les résultats de l'application de cette approche sur des pages manuscrites complètes sont présentés dans la section d'expérimentation (section 4). La similarité plus précise qui se dégage de cette seconde approche est complétée par le dernier niveau hiérarchique suivant.

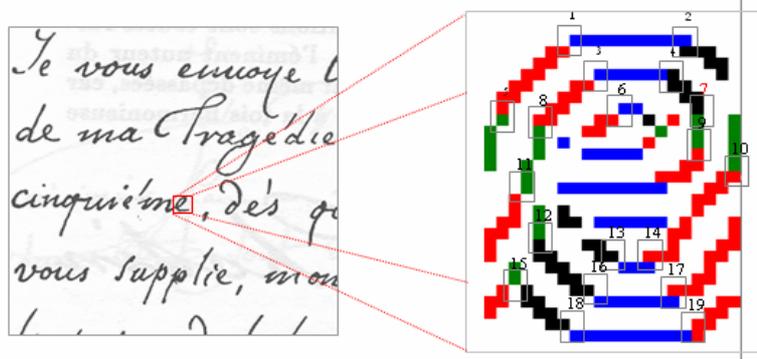
Similarités basées sur la structure des points de jonction

Cette dernière étape se base sur la définition d'une signature composée de la liste des points de jonction aux limites des segments orientés. Le calcul de la signature est basé sur l'agglomération des informations de changement de direction au niveau des jonctions. Il est basé sur une hypothèse de priorités privilégiant une progression allant de haut en bas et de gauche à droite. Avec un seul balayage de la carte des points de contours étiquetés et pour chaque fenêtre d'analyse considérée, on marque les zones de changement de la direction en fonction de la direction courante et de la direction suivante. La figure 18.1 présente les différentes configurations possibles de codage des points de jonctions aux frontières des directions.

$$C(x,y) = \begin{cases} 1 & \text{si } \theta_d(x,y) = 0^\circ \\ 2 & \text{si } \theta_d(x,y) = 45^\circ \\ 3 & \text{si } \theta_d(x,y) = 90^\circ \\ 4 & \text{si } \theta_d(x,y) = 135^\circ \end{cases}$$



18.1



18.2

| N°Jonction | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|
| Courant | 1 | 1 | 1 | 1 | 2 | 1 | 4 | 2 | 3 | 3 | 3 | 3 | 4 | 2 | 3 | 4 | 2 | 4 | 2 |
| Suivant | 2 | 4 | 2 | 4 | 3 | 2 | 3 | 3 | 2 | 2 | 2 | 4 | 1 | 1 | 4 | 1 | 1 | 1 | 1 |

18.3

Figure 18. 18.1. Combinaisons possibles des directions pour le codage des points de jonctions. 18.2. Exemples de points de jonctions sur un motif isolé. 18.3. Signature du motif zoomé.

Pour simplifier les calculs, nous n'avons pas considéré dans cette version les possibilités de points de jonction triples à la frontière de trois (ou plus) directions différentes. Les figures 18.1 et 18.3 illustrent le codage d'un motif en liste de points de jonctions. Tous les points de jonction $J(x,y)$ sont codés par trois valeurs : leur position dans la liste, la direction courante et la direction suivante selon la définition des points contours répondant à la direction $\theta_d(x,y)$.

Soit $S_f(i)$ la signature du motif localisé par la fenêtre d'analyse glissante et $S_{ref}(i)$ la signature de la forme de référence, avec $i=0 \dots N$. N est le nombre de point de jonctions. Nous avons choisi de comparer les signatures des motifs à l'aide de la distance de Hamming qui permet de comparer deux chaînes de même longueur. Le principe de cette distance consiste à compter le nombre de points différents estimés entre les deux signatures S_f, S_{ref} . Par définition, la distance de Hamming est donnée par la formulation suivante :

$$D(S_f, S_{ref}) = \left\{ \text{Card} \left\{ i \in [1, N] \mid S_f(i) \neq S_{ref}(i) \right\} \right\}$$

Le pourcentage de ressemblance entre S_f et S_{ref} s'exprime par :

$$P_s(S_f, S_{ref}) = 100 \cdot \frac{N - D(S_f, S_{ref})}{N}$$

On considère alors que deux motifs sont suffisamment similaires si le rapport $P_s(S_f, S_{ref}) > Seuil$ est vérifié. On pourra fixer le seuil de façon automatique pour un ensemble de pages d'un même ouvrage. La figure 19 présente un résultat d'estimation de similarités basées exclusivement sur les points de jonction avec un seuil fixé à 50%. Dans cet exemple, 9 motifs (contenant des lettres entières, la lettre « d » extraites d'un manuscrit de Montesquieu de 1745 - *Histoire Véritable*) sur les 11 motifs de tests ont été appariés avec succès. Les 2 motifs restants présentaient de fortes irrégularités de structure.

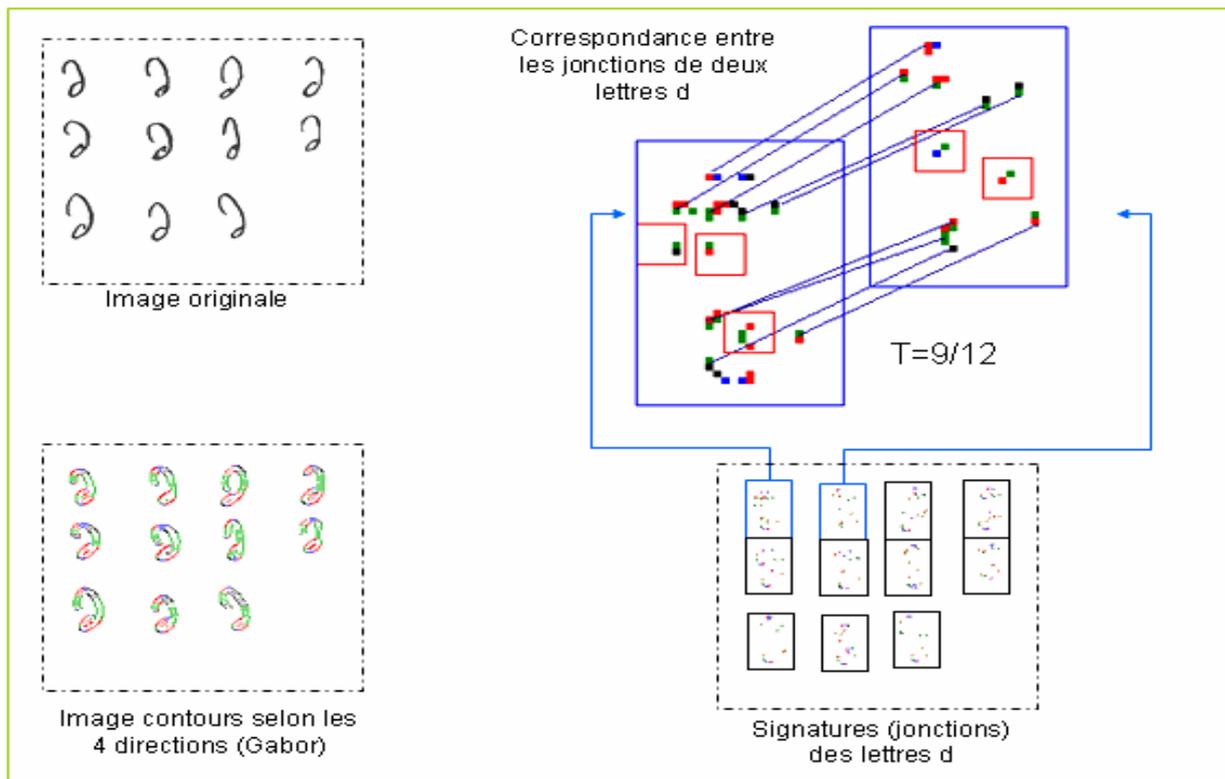


Figure 19. Modélisation basée sur les points de jonction. Correspondance entre motifs.

Pour améliorer les performances de cette approche, nous avons envisagé d'utiliser la distance de Morris et Pratt [LI05] qui permet d'extraire des parties d'une forme ressemblantes à une autre par une recherche d'inclusion de sous-motifs. Cette approche est actuellement en cours de développement. Parallèlement à ces développements, nous testons la coopération entre les trois approches de sorte à faire ressortir les similarités qui minimisent les différences.

IV. EXPERIMENTATION ET ANALYSE DES RESULTATS

IV.1 Analyse comparative des trois approches de recherche de similarités

Quantification des similarités par le tri des ressemblances.

Pour valider notre approche, nous avons sélectionné une base de 130 formes à partir d'une même image, voir figure 20.

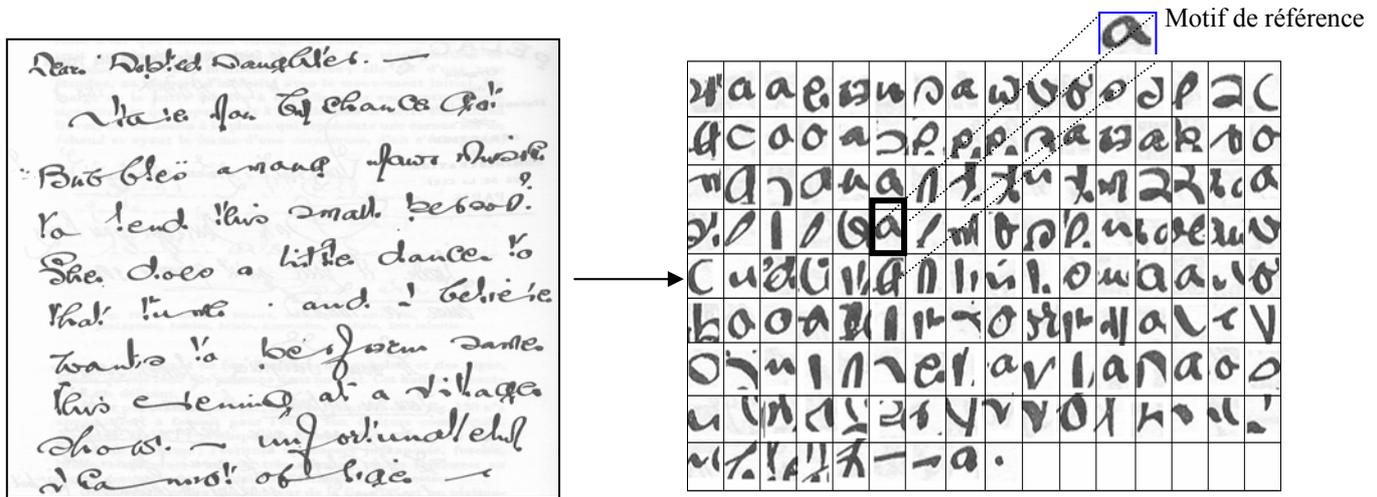


Figure 20 : Image d'origine et base de 130 motifs rangés aléatoirement

La sélection de motifs référents s'est faite de telle sorte à créer la plus grande diversité de formes possibles. A chaque test, le choix du motif de référence a été arbitraire. L'objectif ici est de vérifier que nos approches sont capables de trier les motifs selon un ordre décroissant de similarités en accord avec une réalité terrain estimée visuellement. Le protocole adapté de [CA04] et [FO04] consiste à comparer chaque motif au motif de référence par les trois méthodes de recherches de similarités expliquées ci-dessus (*similarités* par roses de direction « *SDR* », *similarités* par corrélation d'histogrammes de projections des directions « *SHPD* », et enfin *similarités* par codage des points de jonctions « *SPJ* »). La variation de chaque mesure de similarité est représentée par une courbe représentant les résultats estimés pour chacun des 130 motifs de la base selon chacune des approches (figure 21, figure 22 et figure 23).

A partir de ces résultats nous avons reconstruit sur trois images la succession des motifs de la base triés selon leur similarité décroissante au motif référent. Les figures 28, 29, 30 montrent des résultats très satisfaisants de corrélation entre les motifs de la base et le motif de référence. Les classements proposés sur les trois figures sont, à quelques variantes près, très révélateurs des similarités entre formes. La figure 24 justifie le classement proposé par la méthode SHPD par le calcul des résidus (différence entre le motif de référence et les 130 motifs testés). Les surfaces croissantes des résidus estimées à partir des motifs triés permettent de valider le classement.

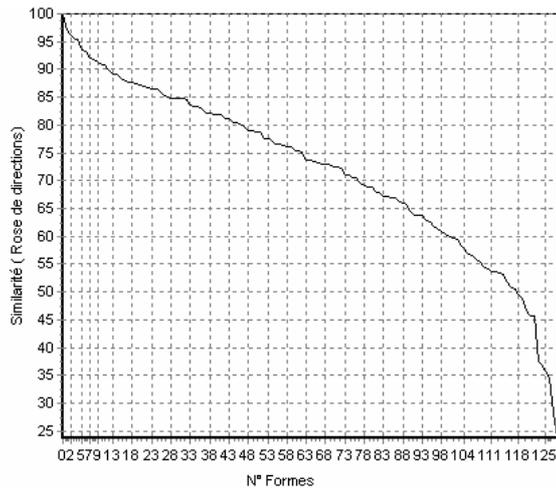


Figure 21 : Variations des ressemblances (selon la méthode basée sur les roses de directions « SDR »)

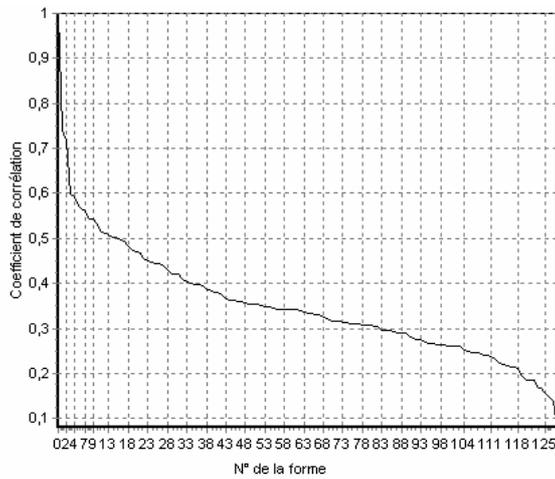


Figure 22 : Variation des ressemblances (selon la méthode basée sur la corrélation entre les histogrammes de projections des directions « SHPD »)

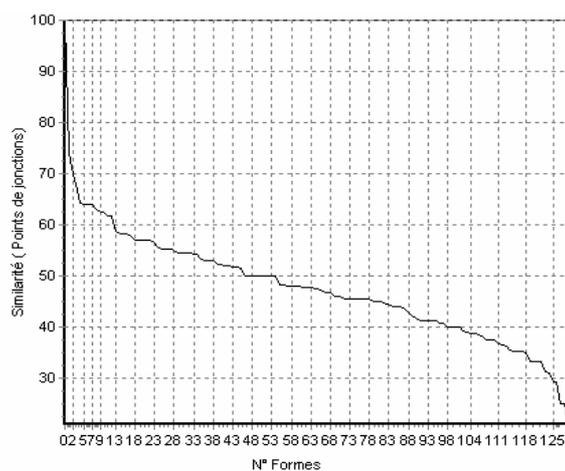


Figure 23 : Variation des ressemblances (selon la méthode basée sur la correspondance entre les points de jonctions « SPJ »)

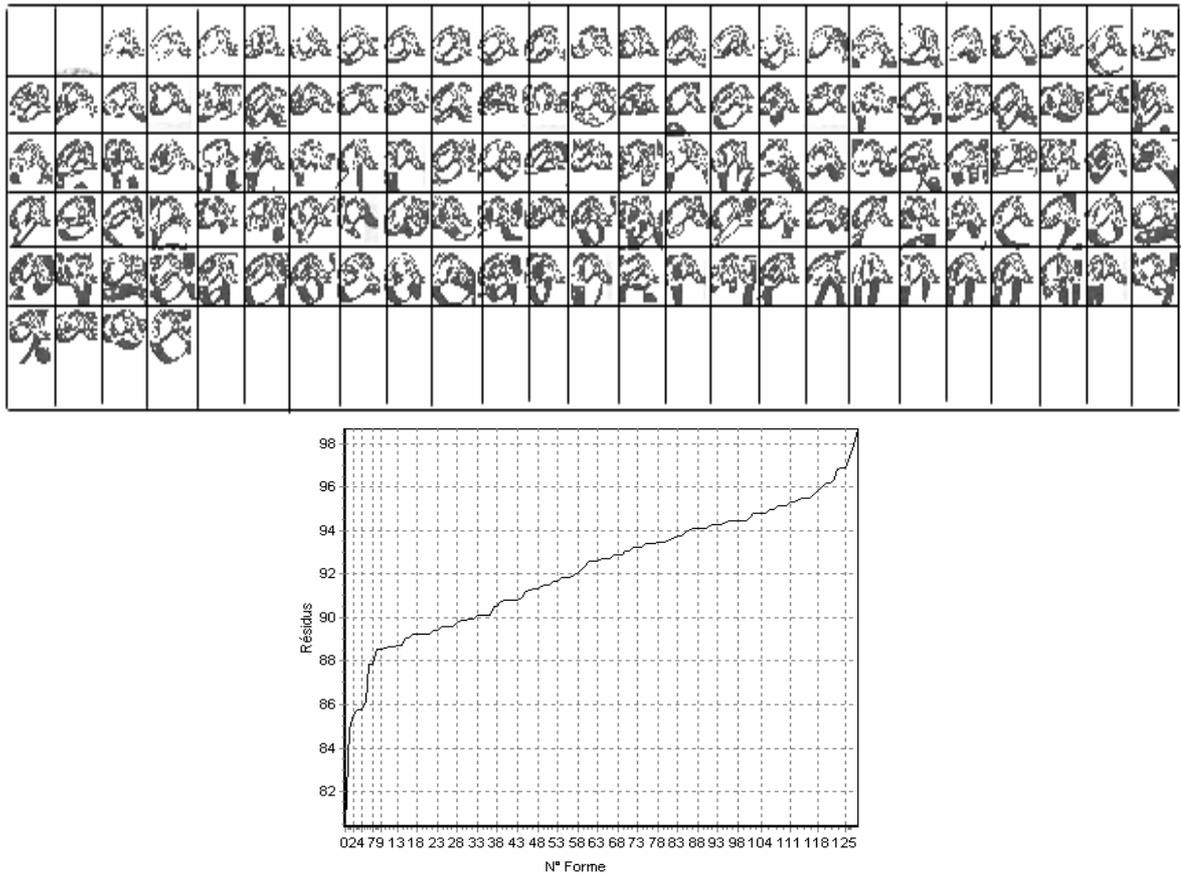


Figure 24. Variation des résidus pour la méthode SHPD. Estimation de la surface des résidus

Appréciation et discussion des résultats.

Sur ces résultats, on peut constater que la similarité est d'autant mieux relevée que l'approche traite la forme avec précision : la pente des courbes est plus importante pour les méthodes SPJ et SHPD qui relèvent d'une analyse plus fine (de structure) des formes. On peut expliquer ce phénomène par le fait qu'une grande précision dans la description de la structure des motifs peut compromettre un appariement car la similarité s'en trouve réduite. L'avantage de disposer de ces trois approches est de pouvoir les combiner selon une analyse hiérarchique. De l'approche la plus globale (SDR) - à utiliser en premier - à la plus fine (SPJ), on passe ainsi d'une extraction grossière des formes similaires à un appariement flexible basée sur les points de structures. Sur les exemples traités ici, la recherche de similarités par analyse des roses de directions permet à plus de 96% de regrouper les formes *réellement* similaires. Pour assurer ensuite un appariement précis, il est nécessaire de se tourner vers un couplage hiérarchique des trois approches. Ce couplage est actuellement à l'essai.

Les méthodes d'appariement existantes procèdent pour la plupart selon un protocole fixe et linéaire ne permettant pas un « bouclage retour » pour affiner ou invalider un résultat, [MI01], [BE02]. Le couplage de nos trois méthodes doit au contraire permettre de remettre en cause un classement des similarités en autorisant d'interclasser des motifs plus éloignés ou de supprimer des motifs invalidés par une analyse plus fine (SHPD ou SPJ). Voyons à présent comment la recherche de similarités s'exprime sur un document entier à partir de la recherche de motifs quelconques (en taille et en nombres de segments orientés).

IV.2 Application à la recherche de similarités sur une image complète de document

Notons pour commencer que notre approche ne repose pas sur un classement de connexité estimés sur l'ensemble du tracé de la page mais sur des motifs dont les tailles peuvent être quelconques et qui sont inclus ou non à l'intérieur de formes connexes (à l'intérieur de mot par exemple). Cette propriété permet de relever des ressemblances cachées à l'intérieur de connexités plus grandes et de ne pas se limiter à une analyse dirigée par la taille des patterns.

Une efficacité prouvée sur du texte imprimé.

Les images de textes imprimés possèdent des formes très redondantes, comme nous l'avons présenté à la section 2. Dans ce type de document et quelle que soit l'approche choisie (parmi SDR, SHPD et SPJ), les résultats sont comparables aux techniques très fiables reposant exclusivement sur des appariements de formes binaires. Ce constat donne à notre proposition une généralité que les méthodes binaires dédiées aux images de textes imprimés ne possèdent pas. La figure 25 présente le résultat de la recherche de similarités basée sur l'approche SPJ appliquée sur des portions de texte imprimé utilisant un seuil de ressemblance de 50%. Le marquage des formes détectées est basé sur la carte des composantes de Deriche.

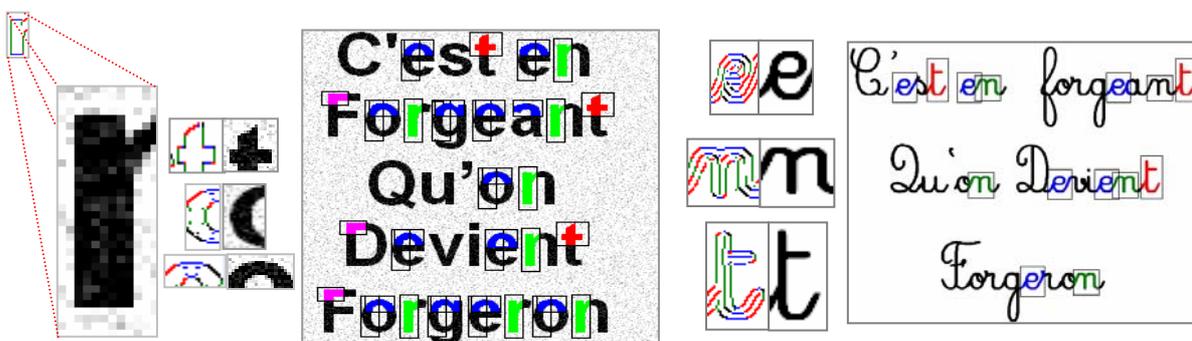


Figure 25. Recherche de trois motifs choisis au hasard sur deux extraits imprimés.

Des résultats très prometteurs sur du manuscrit ancien.

Sur la figure 26, nous avons choisi d'illustrer sur la base d'un motif sélectionné manuellement (figure 26.1) l'efficacité de l'approche. Les éléments jugés similaires à partir de la méthode SPJ (méthode des points de jonction aux limites des segments orientés, figure 26.2) sont marqués de bleu à la figure 26.3. Finalement, la carte des résidus correspondant à la différence entre l'image d'origine et l'image des similarités est représentée à sa droite, figure 26.3.

La grande variabilité des formes présentes sur le texte manuscrit nous a contraint à choisir des seuils de similarité plus faible garantissant ainsi une plus grande tolérance des résultats. L'évaluation des résultats de recherche est plus difficile à réaliser sur des images de manuscrit ancien en raison de la présence de bruit, de l'irrégularité des traits et des motifs de même signification.

L'analyse à grande échelle (sur un corpus complet) nous a conduit à faire le constat que le choix de la granularité du motif (s'exprimant en nombre de segments orientés et en surface) était déterminante : en effet, un motif trop « grand » (englobant une portion entière d'un mot par exemple) peut ne pas être représenté plus d'une ou deux fois sur une même page. Le taux de redondance assorti à cette recherche s'en trouve ainsi considérablement réduit. L'efficacité du codage pour la compression s'en trouve ainsi fortement diminuée. Parallèlement, des motifs trop petits se répèteront très fréquemment sur la page mais la surface occupée par le cumul de ces petits

motifs ne conduira pas à un partitionnement exploitable en terme de compression. Il faudra envisager de coder un très grand nombre de segments et de grossir le dictionnaire des formes similaires de référence. Le choix des motifs qui actuellement n'est pas encore automatisé est fondamental pour la compression.

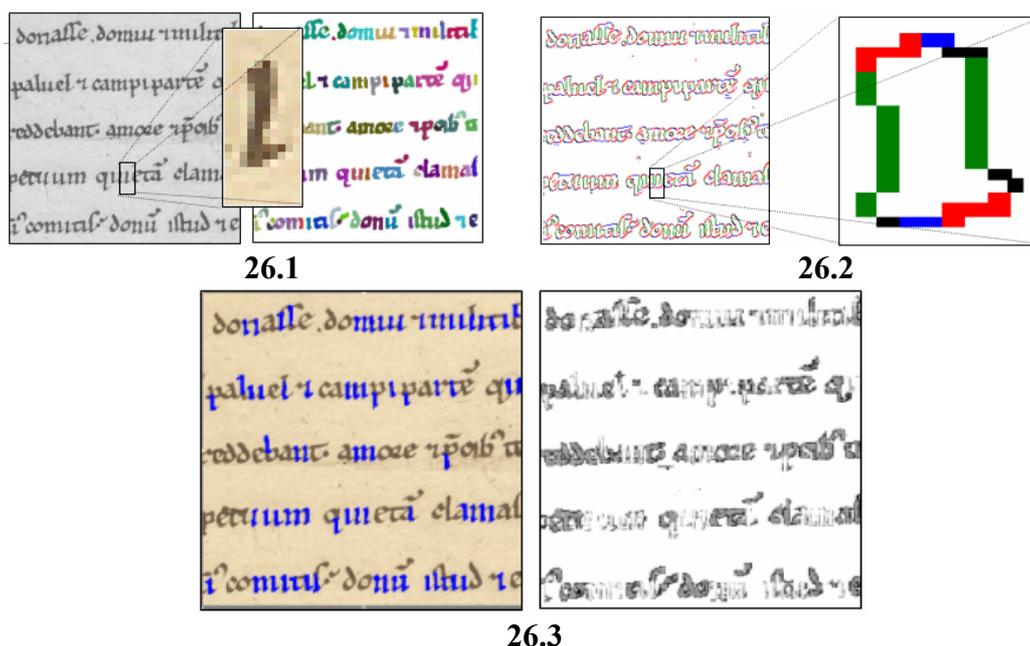


Figure 26. Résultat de la recherche de similarités sur une image de manuscrit à partir d'un motif de granularité fine.

IV.3 Intégration au processus de compression adaptée aux tracés manuscrits

L'exploitation des similarités entre les motifs manuscrits se traduit par un partitionnement de l'image centré tout autour des zones de traits. Nous avons montré précédemment que l'échec des méthodes de compressions usuelles, et même celles dédiées à la compression des images de documents était dû en partie au mauvais partitionnement choisi pour localiser les similarités dans l'image.

L'étape de recherche de similarités va donc pouvoir s'intégrer à un processus complet de compression d'une page d'écriture où les motifs considérés comme similaires pourront être remplacés dans l'image par leur motif de référence codé une seule fois dans un dictionnaire de motifs. Le principe retenu et non développé dans le cadre de ce MASTER consiste à construire ce dictionnaire complet de motifs permettant de mener l'ensemble des substitutions. A ce stade, nous pouvons relever une différence essentielle entre cette approche et les techniques de compression adaptée aux tracés imprimés (DjVu ou DEBORA) : l'existence de motifs de granularité (nombre de segments orientés) et de taille variables et le choix de représentants non binaires.

La seconde étape du processus concerne le codage des résidus : le rendu définitif d'une image décompressée nécessite une estimation d'un plan de compensation faisant une différence entre les données réelles d'origine et les données transmises reconstruites. Ce plan de compensation permet de restituer une image parfaite. Pour reconstruire l'image exacte originale à partir du plan décompressé, le principe retenu (mais non développé) consiste à coder les différences entre le motif de substitution et le motif d'origine. Ces différences seront globalement de faibles amplitudes (sauf au niveau des frontières où les plus grandes différences apparaîtront du fait de l'irrégularité des frontières des traits, voir figure 26) et pourront ainsi être codés sur un volume beaucoup plus faible (codage des pixels). Le plan composite compensé ainsi obtenu restitue sans erreur l'image d'origine.

Il reste alors à s'assurer du codage de l'arrière plan qui possède une propriété remarquable importante : l'existence d'un contraste globalement faible (par rapport à une image naturelle). Ce constat nous permet d'exploiter des techniques de codage prédictif des pixels du fond associé à un codage entropique du résultat (codage de Huffman) assurant ainsi une compression sans aucune perte d'information et sur un faible volume. Cette compression sans perte de l'arrière plan est importante pour plusieurs raisons : d'une part, le fond peut contenir des informations de tracés peu contrastés non détectés par les bancs de filtres directionnels et d'autre part l'arrière plan contient des informations essentielles à une expertise paléographique intéressant également les historiens du livre, les littéraires et les archéologues. Il demeure donc essentiel de lui préserver toutes ses qualités en refusant de lui appliquer des techniques de compression avec pertes.

IV.4 Perspectives directes : développement d'un outil générique

La première amélioration à apporter à ces travaux concerne la localisation automatique des motifs selon un critère de taille lié à la surface moyenne des composantes de Deriche les plus nombreuses sur la page et de nombre de segments orientés contenus dans ces motifs (ne devant pas dépasser une valeur seuil au-delà de laquelle la recherche de similarités n'est plus pertinente). Le balayage du tracé se réalisera ainsi à partir de fenêtres glissantes de taille variables dépendant de celle des fragments connexes parcourus.

La deuxième étape centrale à ce travail consiste à créer une véritable collaboration entre les trois méthodes de recherche de similarités et à choisir, selon le cas, la méthode la plus adaptée au contenu. Enfin, nous envisageons également d'améliorer l'étape de construction des cartes de connexités et de segments orientés en associant les deux approches fréquentielles et différentielles Gabor - Deriche et en généralisant l'invariance à la rotation actuellement opérationnelle pour la méthode SPRD uniquement (le changement d'échelle étant quant à lui déjà validé pour les trois approches). Nous espérons que ces améliorations réalisables à court terme permettront à ces résultats déjà très encourageants d'être exploitables sur l'ensemble des corpus de nos projets de numérisation actuels.

V CONCLUSION GENERALE

Le tracé manuscrit est par définition très complexe à modéliser, ceci est dû à la non régularité des formes, et leur importante variabilité que l'on peut constater d'une écriture à l'autre. Chez un même scripteur d'ailleurs on retrouve une hétérogénéité dans la forme des tracés, parfois même sur une même page d'un seul auteur. Tenant compte de cette forte variabilité propre aux tracés manuscrits, nous avons choisi de nous orienter vers une nouvelle méthodologie permettant de trouver localement des similarités partielles sans avoir recours à une segmentation binaire complète des formes. Nous avons choisi de séparer les traits selon des motifs de taille et de formes variables. Ces motifs redondants ne sont ni nécessairement des mots ni des lettres : ils peuvent être simplement constitués des petites boucles des lettres (les graphèmes) qui se retrouvent localisés à d'autres endroits dans le tracé manuscrit et ceci à une transformation géométrique près (un changement d'échelle, une rotation raisonnable ou bien une translation). La localisation des formes similaires est un étape primordiale à la mise en place d'une méthode de compression efficace et adaptée aux contenus des images de manuscrit.

D'un point de vue technique, nous avons développé une approche de localisation des zones de traits et de repérage des formes redondantes en exploitant conjointement des techniques basées sur les filtrages directionnels de Gabor et des descripteurs spécifiques associés notamment au détecteur de contours de Deriche. Ces outils nous permettent de définir de manière stable les lignes d'orientations identiques et les zones de changements d'orientations présentes sur les contours des traits. Nous proposons ainsi de produire un partitionnement des zones de traits en portions de tailles variables divisibles que nous caractérisons à l'aide d'une signature servant de

base à la mesure de similarité, elle-même basée sur les caractéristiques directionnelles des contours et les points de changement de direction. Cette recherche de similarités a été conçue pour être robuste aux changements d'échelle, aux rotations et bien sûr aux translations. Le processus de compression complet est actuellement en développement, il nécessite d'intégrer la méthode de recherche de similarités proposée associée à un codage optimal des motifs (en terme de bits). Ce travail d'intégration est très long et s'explique par la complexité des méthodes mises en œuvre et par l'absence de travaux antérieurs dans ce domaine. Les premiers résultats obtenus sur les images de manuscrits anciens sont cependant très encourageants. Cette contribution s'inscrit dans une démarche de recherche prospective de long terme.

VI. REFERENCES BIBLIOGRAPHIQUES

- [AB05] EL ABED, L., EGLIN, V., LEBOURGEOIS, F., EMPTOZ, H. *Frequencies decomposition and partial similarities retrieval for patrimonial handwriting documents compression*, ICDAR, Séoul, Aout 2005.
- [BE02] S. BELONGIE, J. MALIK, AND J. PUZICHA, "Shape matching and object recognition using shape contexts," *IEEE Trans. PAMI*, vol. 24, no. 4, pp. 509–522, April 2002.
- [BO98] L. BOTTOU, P. HAFFNER, P.G. HOWARD, P. SIMARD, High quality document image compression with DjVu, *journal of electronics imaging*, 7(3):410-428, 1998.
- [BR03] BRES, S., JOLION, J.M., LEBOURGEOIS, F. *Traitement et analyse des images numériques*, Hermes, 2003, 412 p.
- [CA99] CARLSSON, S. *Order structure, correspondence and shape based categories*, in Shape Contour and Grouping in Computer Vision. Springer LNCS 1681, 1999, pp. 58–71.
- [CH00] H. CHUI AND A. RANGARAJAN, "A new algorithm for non-rigid point matching," in *CVPR*, vol. 2, June 2000, pp. 44–51.
- [CO95] T. COOTES, D. COOPER, C. TAYLOR, AND J. GRAHAM, *Active shape models: training and application*, Computer Vision and Image Understand., 61(1) pp.38–59, 1995.
- [DE00] Projet européen « DEBORA », livre en ligne : <http://debor.enssib.fr>, juin 2000.
- [DJ05] Centre de développement DjVu, <http://djvu.research.att.com/>
- [EC05] Ecole des Chartes, Sorbonne, <http://enc.sorbonne.fr>.
- [EM03] EMPTOZ, H., DALBERA, J.P., COUASNON, B. « Numérisation et patrimoine », *Document numérique*, vol. 7, n°3-4, Hermes, 2003, 188p.
- [FO04] SYLVIE PHILIPP-FOLIGUET, MOHAMED LEKKAT. *Recherche d'images à partir d'une requête partielle utilisant la disposition des régions*, pp.35-43, 2004.
- [HO96] P. HOWARD, lossless and lossy compression of text images by soft pattern matching, *Proc. of the IEEE Data compression Conference* pp. 210-219, 1996.
- [HO98] P. HOWARD, F. KOSENTINI, B. MARTINS, S. FORCHHAMMER, W. RUCKLIDE, F. ONO, The emerging JBIG2 standard, *IEEE transactions on Circuit and Systems for video technology*, vol 8, n°5, September 1998.
- [IR05] IRHT, Institut de Recherche de l'Histoire des Textes, <http://www.irht.cnrs.fr/>
- [JO97] JOHNSON, A. E., HEBERT, M. *Recognizing objects by matching oriented points* in *CVPR*, 1997, pp. 684–689.
- [KI97] O. E. KIA, Document Image Compression and Analysis, *PhD of the university of Maryland*, 1997, 191 p.
- [KR99] KRUIZINGA, P., PETKOV, N., *Nonlinear Operator for Oriented Texture*, IEEE Pattern Aanalysis and Machine Intelligence, vol.3, n°2, pp. 76-88, 1999.
- [LA93] M. LADES, C. VORBRUGGEN, J. BUHMANN, J. LANGE, C. VON DER MALSBURG, R. WURTZ, AND W. KONEN, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. Computers*, vol. 42, no. 3, pp. 300–311, 1993.
- [LE00] F. LEBOURGEOIS, H. EMPTOZ, Document Analysis in Gray level and typography Extraction using Character Pattern redundancies, *proc. of the 5th ICDAR*, sept 20-22, Bangalore India, pp. 177-180, 2000.

- [LE04] F. LE BOURGEOIS, E. TRINH, B. ALLIER, V. EGLIN, H. EMPTOZ, Document Images Analysis Solutions for Digital libraries, 1st International Workshop on Document Image Analysis for Libraries, 2004, Palo Alto, CA, USA. IEEE Computer Society 2004, pp. 2-24
- [LE98] Y. LECUN, L. BOTTOU, Y. BENGIO, AND P. HAFFNER, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, November 1998.
- [LI05] L. LIKFORMAN-SULEM, M. SIGELLE, *Reconnaissance des Formes*, Edition TSI-ENST Paris, 45p. 2005.
- [LO04] D. G. LOWE, "Distinctive image features from scale-invariant keypoints," *Int. Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [MA05] *Projet ACI Masse de données. 2003-2006. Site web <http://l3iexp.univ-lr.fr/madonne/>*
- [MI01] MIKOLAJCZYK, K., SCHMID, C. Indexing based on scale invariant interest points, Proceedings of the 8th International Conference on Computer Vision, Vancouver, Canada pages 525 – 531, 2001.
- [MI04] MIKOLAJCZYK, K. Comparison of Affine Region Detectors, In IJCV, pp. 45-58, 2004.
- [MO00] B. MOGHADDAM, T. JEBARA, AND A. PENTLAND, "Bayesian face recognition," *Pattern Recognition*, vol. 33, no. 11, pp. 1771–1782, November 2000.
- [NA03] NASCIMENTO, M., SRIDHAR, V., LI, X. *Effective and efficient region-based image retrieval*, Journal Title: Journal of Visual Languages and Computing, 2003,14(2), p. 151 – 179.
- [PE91] PENTLAND, A, SCLAROFF, S. *Closed-form solutions for physically-based shape modelling and recognition*. IEEE Trans. Pattern Anal. and Mach. Intel., 13(7):715-729, July, 1991.
- [PR78] PRATT, FAUGERAS, GAGALOWICZ, « *Visual discrimination of stochastic texture field* », IEEE Trans. on systems, Man and Cybernetics, 1978, vol. SMC-8, N°11, pp.796-804.
- [SE04] SEROPIAN, A., GRIMALDI, M., VINCENT, N. *Différenciation entre alphabets dans des textes manuscrits*, Colloque International Francophone sur l'Écrit et le Document, La Rochelle, pp.134_144, 2004.
- [SH98] D. SHARVIT, J. CHAN, H. TEK, AND B. KIMIA, "Symmetry-based indexing of image databases," *J. Visual Communication and Image Representation*, 1998.
- [SC97] SCHMID, C., MOHR, R., *Local grayvalue invariants for image retrieval*. In Pattern Analysis Machine Intelligence, 19(5):530-534, 1997.
- [SC00] SCHMID, C., MOHR, R., BAUCKHAGE, C. *Evaluation of Interest Point Detectors*, International Journal of Computer Vision, v.37 n.2, p.151-172, June 2000.
- [SP97] SPINU, C., *Une approches multi agents pour la segmentation d'images associant estimation et évaluation, thèse de doctorat, 230 p.*, 1997.
- [TE00] TELL D., CARLSSON, S., *Wide baseline point matching using affine invariants computed from intensity profiles*. In ECCV, pp.814-828, 2000.
- [TH97] THIRAN, J.P. *A queue-based region growing algorithm for accurate segmentation of multi-dimensional digital images*, *Signal Processing*, Vol. 60, n° 1, July 1997, voir fiche 279, pp. 1-10, 1997 .
- [WE03] WESTEEL, I., AUBRY, M., *La numérisation des textes et des images : Technique et Réalisations*, Presses de l'université Charles de Gaulle - Lille, 2003, 190 p.
- [WE98] WELDON, TP, HIGGINS, WE. *Algorithm for designing multiple Gabor filters for segmenting multi-textured images*, IEEE Inter. Conf. on Image Proc., 4-7, 1998.
- [WI94] WITTEN, I., BELL, T., EMBERSON, H., INGLIS, S., MOFFAT, A. *Textual image compression : two stage lossy/lossless encoding of textual images*, proc. of the IEEE, 82:878-888, 1994.
- [YA03] YANG, F., LISHMAN, R. *Land Cover Change Detection Using Gabor Filter Texture*, Proceedings of the 3rd Inter. workshop on texture analysis and synthesis, pp. 78-83, 2003.
- [ZA72] C. ZAHN AND R. ROSKIES, "Fourier descriptors for plane closed curves," *IEEE Trans. Computers*, vol. 21, no. 3, pp. 269–281, March 1972.
- [ZH95] YING ZHANG, LAI-MAN PO, *Fractal Color Image Compression using vector distortion measure*, *Proceeding of IEEE Inter. Conf. on Image Processing*, III – 276 to III – 279, D.C, U.S.A, October 1995.

Remerciements

Je tiens tout particulièrement à remercier Monsieur Hubert EMPTOZ (Professeur et responsable de l'axe « Images et Vidéo » du LIRIS de m'avoir accueilli au sein de son équipe.

Je remercie également Véronique EGLIN et Stéphane BRES (Maîtres de conférence au laboratoire LIRIS à l'INSA de LYON) d'avoir accepté de diriger ce stage, d'avoir mis le doigt sur les points difficiles, de m'avoir donné des précieux conseils, et de m'avoir permis de mener ce travail à terme. Sans eux, ce travail n'aurait pas pu aboutir....

Je voudrais remercier également les membres de l'équipe du LIRIS, pour leur aide tout au long de mon stage, ainsi que tous les autres étudiants avec qui j'ai pu échanger.

Finalement, je tiens à exprimer toute ma reconnaissance à mon père qui m'a permis d'aller au bout de ma volonté, qui m'a donné la force d'étudier et de me consacrer à un domaine de recherche qui me passionne réellement.

Je n'oublierais pas non plus le soutien de Dr. Cherfa Yazid, Mr Mohamed TALBIOUI de Nadia, de Riadh, de Fares, de Merouane, de Joh, de Djanine, de Marie, de Céline et de Soumia, et ceux que j'oublie encore de nommer. Merci à vous tous.

Djamel...