
Étude comparative de trois ensembles de descripteurs de texture pour la segmentation de documents anciens

Maroua Mehri^{*,} — Mohamed Mhiri^{***} — Petra Gomez-Krämer^{*}
— Pierre Héroux^{**} — Mohamed Ali Mahjoub^{***} — Rémy Mullot^{*}**

^{*} Laboratoire Informatique, Image et Interaction - L3i, Université de Rochelle, France
{maroua.mehri, petra.gomez, remy.mullot}@univ-lr.fr

^{**} Laboratoire d'Informatique, du Traitement de l'Information et des Systèmes - LI-TIS, Université de Rouen, France
pierre.heroux@univ-rouen.fr

^{***} Systèmes Avancés en Génie Électrique - SAGE, Université de Sousse, Tunisie
mhirimohamed@hotmail.com, medali.mahjoub@ipeim.rnu.tn

RÉSUMÉ. Récemment, des approches basées sur l'analyse des descripteurs de texture ont été largement explorées pour la segmentation d'images de documents anciens numérisés. Il a été prouvé que ces méthodes fonctionnent efficacement en n'ayant pas de connaissances préalables. En outre, il a été démontré qu'elles sont robustes lorsqu'elles sont appliquées sur des documents dégradés ou bruités. Dans cet article, une approche d'évaluation de trois différents ensembles de descripteurs texturaux est présentée pour la segmentation de documents anciens. Sur un vaste corpus expérimental de documents anciens, nous visons tout d'abord à déterminer quels descripteurs de texture pourraient être les mieux appropriés pour séparer les illustrations des régions textuelles d'une part et d'autre part pour discriminer les blocs du texte de différentes tailles et polices de caractères. C'est dans ce but que nous étudions différents descripteurs de texture, extraits de trois outils du traitement d'image, les plus répandus et largement utilisés (la fonction d'auto-corrélation, les matrices de co-occurrence des niveaux de gris et les filtres de Gabor). Un aperçu sur le temps de calcul et la complexité de chaque ensemble de descripteurs de texture est également présenté dans ce travail. Sur une base d'images de plus de 300 documents, une étude expérimentale avec des observations qualitatives et numériques, est proposée pour montrer les performances de chaque ensemble de descripteurs de texture.

ABSTRACT. Recently, texture-based features have been used for digitized historical document image segmentation. It has been proven that these methods work effectively with no a priori knowledge. Moreover, it has been shown that they are robust when they are applied on degraded documents under different noise levels and kinds. In this paper an approach of evaluating

texture-based feature sets for segmenting historical documents is presented in order to compare several texture-based feature sets. We aim to determine which texture features could be better adequate for segmenting on the one hand graphical regions from textual ones and on the other hand for discriminating text in a variety of situations of different fonts and scales. For this purpose, three well-known and widely used texture-based feature sets (autocorrelation function, Grey Level Co-occurrence Matrix (GLCM) and Gabor filters) are evaluated and compared. An additional insight into the computation time and complexity of each texture-based feature set is given. On more than 300 historical documents, qualitative and numerical experiments are given to demonstrate each texture-based approach performance.

MOTS-CLÉS : Documents historiques numérisés, segmentation, texture, multi-résolution.

KEYWORDS: Historical digitized document images, Segmentation, Texture, Multi-scale approach.

1. Introduction

Le développement d'Internet et l'usage fréquent des formats électroniques ont augmenté le besoin des systèmes efficaces et robustes d'analyse et d'interprétation de documents. En outre, la nécessité de garantir une conservation durable et de fournir un accès plus large aux documents, ont généré un intérêt à l'analyse d'image de document. L'analyse d'image de document est devenu un important sujet d'intérêt de nombreux chercheurs et l'un des domaines les plus étudiés, émergents et prospères dans l'analyse et traitement d'images (Nagy, 2000). Dans cet article, nous nous intéressons aux outils non-paramétriques et techniques automatiques de segmentation et caractérisation d'images de documents anciens numérisés.

La littérature montre que divers problèmes peuvent survenir, liés aux particularités des documents anciens, comme une grande variabilité de la mise en page : bruit et dégradations (causés par la copie, la numérisation et le vieillissement), le défaut d'orientation, la complexité de la mise en page, des alignements non-conventionnels, les polices de caractères spécifiques, la présence d'ornement, les variations de l'espacement entre les caractères, mots, lignes, paragraphes et marges, la superposition de plusieurs couches d'information (tampons, notes manuscrites, bruits, interférences recto-verso, . . .) (Agam *et al.*, 2007). Ainsi, le traitement de ce type de documents n'est pas une tâche facile en raison du nombre et des niveaux de dégradations, et ce, sans connaissances *a priori* sur la structure physique ou logique des documents. Crasson et Fekete (Crasson et Fekete, 2004) soulignent la nécessité réelle d'outils et techniques de traitement automatique et non-paramétrique de documents anciens (par exemple des outils pour l'analyse de la mise en page et la séparation texte/illustration) afin de faciliter l'analyse, l'indexation et la navigation dans les corpus de manuscrits anciens.

Il a été prouvé dans la littérature que l'analyse de descripteurs de texture est efficace pour de nombreuses applications, y compris la vision artificielle, la segmentation d'images médicales, la cartographie urbaine, l'analyse d'images satellitaires, la télédétection. . . En outre, les techniques d'analyse de texture ont été largement utilisées au cours des deux dernières décennies, de sorte qu'elles sont devenues les choix les plus appropriés pour l'analyse de nombreux types d'images, en particulier les images de synthèse, médicales et naturelles (Ro, 1998 ; Tuceryan et Jain, 1998). Plus récemment, les approches basées sur l'analyse de texture ont été étudiées pour le traitement d'images de documents, par exemple pour l'identification du script/langue (Busch *et al.*, 2005).

L'objectif de notre travail est d'identifier, sur des images de documents anciens, les régions homogènes, c'est-à-dire des zones contenant des pixels disposant de caractéristiques visuelles communes. Ainsi, nous explorons les différents aspects des descripteurs de texture pour la segmentation des images de documents anciens. Dans notre étude, nous supposons que les régions textuelles et non-textuelles disposent des caractéristiques de texture différentes. En outre, les blocs de texte disposant de polices de caractères différentes sont supposés avoir des apparences visuelles et propriétés de texture distinctes (Julesz, 1962). Dans une approche ascendante de segmentation,

l'analyse des propriétés de texture dans le contenu des documents permet de s'abstraire du manque d'information à disposition (la mise en page, la police de caractère du texte...) en utilisant une approche multi-résolution pour segmenter et caractériser les images de documents anciens (Journet *et al.*, 2008 ; Mehri *et al.*, 2013a).

Par ailleurs, pour la caractérisation d'images de documents anciens, les méthodes basées sur des approches d'analyse de texture ont prouvé leur efficacité pour la segmentation d'images affectées de divers types de dégradations, d'intensité différente, et ce, sans connaissances *a priori* (Eglin *et al.*, 2007 ; Journet *et al.*, 2008 ; Grana *et al.*, 2011 ; Ouji *et al.*, 2011 ; Mehri *et al.*, 2013b ; Mehri *et al.*, 2013c). Dans le cadre de nos travaux présentés dans (Mehri *et al.*, 2013a), nous avons étudié trois ensembles de descripteurs de texture respectivement extraits à partir de la fonction d'auto-corrélation (Petrou et Sevilla, 2006), des matrices de co-occurrence des niveaux de gris (Haralick *et al.*, 1973) et des filtres de Gabor (Gabor, 1946). Ces trois ensembles de descripteurs de texture sont détaillés et évalués sur 25 images de documents anciens simplifiés. Dans ce papier, nous étendons notre travail pour étudier ces trois descripteurs de texture sur plus de 300 documents anciens. Nous analysons en outre la contribution d'une étape de post-traitement dans l'approche à base de texture en intégrant les informations relatives au voisinage des pixels analysés.

Nous proposons dans cette étude d'évaluer et de comparer les différents ensembles de descripteurs de texture, extraits de la fonction d'auto-corrélation, des matrices de co-occurrence des niveaux de gris et des filtres de Gabor. Nos tests sont élaborés pour segmenter 314 documents anciens sans connaissances *a priori* ni sur la mise en page de document scanné, ni son contenu. Cette étude est structurée de la façon suivante. Tout d'abord, la section 2 présente un bref aperçu des différents ensembles de descripteurs de texture étudiés. Par la suite, un bref aperçu de l'approche proposée pour la segmentation basée sur l'analyse de texture est décrite dans la section 3. Ensuite, le protocole expérimental et les résultats des tests effectués sont détaillés dans les sections 4 et 5 respectivement. Enfin, les conclusions de cette étude et nos perspectives sont présentées dans la section 6.

2. Descripteurs de texture

De nombreux algorithmes d'extraction de descripteurs ont été proposés dans la littérature du traitement d'images. Reed et DuBuf (Reed et DuBuf, 1993) les catégorisent en des approches à base de descripteurs, de modèles et de structures. Feddaoui et Hamrouni (Feddaoui et Hamrouni, 2006), quant à eux, les classifient en trois approches : structurelles, statistiques et spatio-fréquentielles. Toyoda et Hasegawa (Toyoda et Hasegawa, 2005) les classent en deux types : locales (par exemple, motif binaire local (LBP) (Ojala *et al.*, 2002)) et fréquentielles (les transformées en ondelettes (Mallat, 1989), filtres de Gabor (Manjunath et Ma, 1996),...). Les méthodes d'extraction et d'analyse à base de texture peuvent être divisées en quatre classes (Chen *et al.*, 1998 ; Tuceryan et Jain, 1998) :

– *Les méthodes statistiques* analysent la distribution spatiale des valeurs de niveaux de gris par le calcul des indices locaux dans l'image et déduisent par la suite un ensemble de statistiques. GLCM (Haralick *et al.*, 1973) est l'une des méthodes statistiques de segmentation à base de texture fréquemment citée.

– *Les méthodes géométriques* sont utilisées pour décrire les motifs complexes et déduire les propriétés des textures. Les descripteurs de texture peuvent être extraits, par exemple, par une différence de gaussiennes (Tuceryan et Jain, 1990). Ces méthodes permettent de caractériser les propriétés géométriques des textures et trouver les règles qui régissent leur organisation spatiale.

– *Les méthodes à base de modèle* estiment un modèle paramétrique en fonction de la distribution d'intensité des descripteurs de texture calculés. Les méthodes à base de modèles probabilistes sont largement utilisées telles que les champs aléatoires conditionnels (CRF) (Lafferty *et al.*, 2001), les champs aléatoires Markoviens (MRF) (Nicolas *et al.*, 2005), les champs aléatoires Markoviens gaussiens (GMRF) (Chellappa et Chatterjee, 1984), les fractales (Ferrell *et al.*, 2003), LBP,...

– *Les méthodes fréquentielles* analysent les fréquences de l'image. Les méthodes fréquentielles les plus utilisées sont les filtres de Gabor (Jain *et al.*, 1992), la transformée de Fourier (Sabharwal et Subramanya, 2001), les transformées en ondelettes (Mallat, 1989) et les méthodes de segmentation à base de calcul de moments (Tuceryan, 1994). Certaines approches étudient les propriétés locales de l'image analysée (GMRF, LBP,...). D'autres méthodes sont basées sur des représentations statistiques et/ou spatiales et/ou fréquentielles (les transformées en ondelettes, les filtres de Gabor,...).

Un aperçu complet des dernières techniques d'extraction d'indices de texture pour la segmentation d'images est présentée dans (Reed et DuBuf, 1993) incluant les filtres de Gabor, GLCM, fractales,... Qiao *et al.* (Qiao *et al.*, 2006) combinent les filtres de Gabor, les ondelettes et les méthodes à base de noyau pour la segmentation d'images de documents. Nourbakhsh *et al.* (Nourbakhsh *et al.*, 2006) évaluent deux catégories de descripteurs de texture extraits des filtres de Gabor et des ondelettes pour séparer les zones textuelles des régions non-textuelles dans des images de documents. Dans cet article, trois catégories de descripteurs de texture sont extraites : la fonction d'auto-corrélation, les matrices de co-occurrence des niveaux de gris et les filtres de Gabor.

Dans notre travail, nous avons choisi d'analyser ces trois ensembles de descripteurs de texture pour les raisons suivantes. Tout d'abord, nous avons fait une étude comparative pour avoir le meilleur choix des descripteurs de texture à analyser. Ce choix est déterminé en visant le meilleur compromis entre les meilleures performances obtenues dans la littérature, un nombre réduit de paramètres et seuils à fixer et un temps de calcul des indices de texture raisonnable (Mehri *et al.*, 2013a). D'autre part, l'extraction des indices de texture devrait être faite avec le moins possible de paramètres à fixer. En effet, sans aucune information sur la structure du document (le modèle de document, la mise en page, son contenu,...), le choix de nombreux seuils et paramètres appropriés est une tâche très difficile à gérer. En outre, la pertinence des expériences de segmentation réalisées dans le domaine d'analyse d'image de document basées sur la fonction

d'auto-corrélation (Journet *et al.*, 2008), les matrices de co-occurrence des niveaux de gris (Busch *et al.*, 2005) et les filtres de Gabor (Jain et Farrokhnia, 1991), nous amène à évaluer et comparer ces trois exemples de descripteurs de texture dans l'objectif de segmenter les documents anciens. Enfin, les descripteurs de texture sont principalement étudiés séparément dans le domaine de l'analyse d'images de documents. À notre connaissance, il n'existe pas d'étude comparative des descripteurs de texture pour la segmentation de documents anciens. Le présent article propose une contribution dans cette direction en évaluant, sur une tâche de segmentation des images de documents anciens, ces trois ensembles de descripteurs largement utilisés et répandus. Tous les tests sont élaborés sans aucune connaissance *a priori* ni sur la mise en page du document analysé, ni sur son contenu. Nous présentons dans la suite de cet article, l'évaluation de ces trois ensembles de descripteurs de texture dans une tâche de segmentation sur plus de 300 images de documents. Ces descripteurs issus de la fonction d'auto-corrélation, de la matrice de co-occurrence des niveaux de gris et des filtres de Gabor sont détaillés dans (Mehri *et al.*, 2013a).

3. Approche proposée de segmentation

Afin d'évaluer et de comparer les trois ensembles de descripteurs de texture, une approche de segmentation de documents anciens basée sur l'analyse de texture est proposée dans cet article. Notre système de segmentation basée sur l'analyse de texture est composé de quatre modules : une phase de pré-traitement, l'étape d'extraction de descripteurs de texture, le processus de segmentation en régions homogènes sur la base des valeurs d'indices de texture extraits en utilisant une technique de regroupement de pixels "classification non-supervisée des pixels" et une étape de post-traitement.

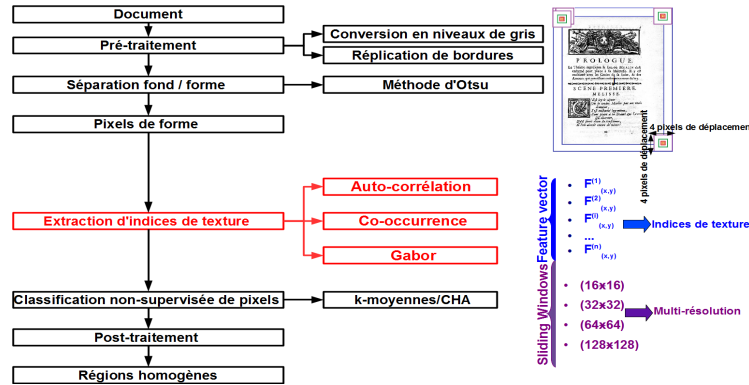


Figure 1. Approche proposée de segmentation de documents anciens basée sur l'analyse de descripteurs de texture.

Tout d'abord, les descripteurs de texture sont extraits uniquement sur les pixels de forme des documents en niveaux de gris. En effet, nous considérons qu'il n'est pas

nécessaire de caractériser les pixels représentant le fond du document. Par ailleurs, le fait de n'extraire les descripteurs que sur les pixels associés au tracé permet de réduire drastiquement le volume de pixel traités et par conséquent l'occupation mémoire et le temps de traitement des algorithmes d'extraction. La sélection des pixels de forme (bruits, champs de texte, illustrations...) est effectuée en utilisant une approche classique non-paramétrique de binarisation : la méthode d'Otsu (Otsu, 1979). Même si d'autres choix auraient pu être opérés, l'utilisation d'une approche globale de seuillage, comme celle proposée par Otsu, offre un moyen rapide de sélection de pixels de forme de document. Par ailleurs, la méthode proposée par Otsu s'est avérée donner de bons résultats dans (Busch *et al.*, 2005).

Suite à la sélection des pixels de forme, afin d'adopter une approche multi-résolution permettant de s'abstraire de la résolution des images, les descripteurs sont calculés sur quatre fenêtres d'analyse de tailles différentes et centrées sur les pixels à caractériser. Un algorithme de classification non-supervisée est ensuite exécuté sur les vecteurs d'indices de texture normalisés en fixant le nombre maximum de régions homogènes égal à celui défini dans notre vérité terrain (Mehri *et al.*, 2013a). La vérité terrain utilisée dans ce travail a été saisie manuellement à l'aide de l'environnement GEDI¹. L'extraction des différents indices de texture : auto-corrélation, co-occurrence et Gabor à l'aide de 4 tailles différentes de fenêtres d'analyse donne un total de 20 (5 auto-corrélation indices $\times$ 4 fenêtres glissantes pour la multi-résolution), 72 (18 co-occurrence indices $\times$ 4 tailles de fenêtres glissantes) et 192 (48 Gabor indices $\times$ 4 tailles de fenêtres glissantes) indices extraits respectivement (*cf.* Table 2). Deux algorithmes classiques de classification non-supervisée (k-moyennes (MacQueen, 1967) et classification ascendante hiérarchique (CAH) (Lance et Williams, 1967)) ont été utilisés pour la segmentation du document analysé en régions homogènes sur la base des valeurs d'indices de texture extraits. La phase de classification non-supervisée de pixels est exécutée sur les vecteurs de textures calculés afin de partitionner l'espace des vecteurs de texture en des regroupements compacts et dont les caractéristiques de texture sont distinctes d'un regroupement de pixels à l'autre. Il est à noter que cette étape de segmentation n'intègre pas d'information spatiale. Seuls les descripteurs de texture sont extraits sur des fenêtres d'analyse centrées et intègrent donc une partie d'information contextuelle. Dans ce travail, nous introduisons une technique de vote majoritaire postérieure à la classification des pixels en tant que post-traitement. La technique de vote mise en place met en œuvre la multi-résolution en pondérant le vote des pixels voisins en fonction de leur appartenance ou non à des fenêtres de voisinage de tailles différentes.

Pour affiner les résultats de segmentation, de nombreux chercheurs ont mis en place une étape de post-traitement afin de prendre en considération les informations liées à la localisation ou au voisinage des pixels. Chang et Kuo (Chang et Kuo, 1992) justifient l'usage de la phase de post-traitement pour les deux raisons suivantes : tout d'abord, en intégrant les relations de voisinage (filtre médian ou dilatations morphologiques), l'information spatiale qui n'a pas été prise en considération lorsque les indices

1. <http://gedigroundtruth.sourceforge.net/>

de texture sont extraits et analysés, est introduite dans la phase de post-traitement. D'autre part, les auteurs confirment que l'extraction des indices de texture à partir de fenêtres d'analyse de tailles pré-définies n'est pas un choix pertinent car cette technique peut générer des artefacts. L'étape de post-traitement permet ainsi d'améliorer la segmentation. Etemad *et al.* (Etemad *et al.*, 1994) utilisent la technique de vote majoritaire pondéré dans le schéma d'intégration de décision afin d'identifier le texte et les illustrations sur une page de document. Pour la segmentation de document, Jain *et al.* (Jain *et al.*, 1992) utilisent une analyse en composantes connexes afin d'obtenir des blocs de texte rectangulaires et lissés après l'extraction des indices de Gabor. Pour la segmentation de pages de documents non structurés, les auteurs (Etemad *et al.*, 1997) intègrent les opérations de morphologie mathématique (fermeture) sur les régions extraites afin d'éliminer le bruit ou les valeurs aberrantes dans les régions segmentées. Palfray *et al.* (Palfray *et al.*, 2012) utilisent la technique de vote majoritaire pour affiner la segmentation à base de MRF de journaux anciens. Les auteurs (Charrada et Amara, 2012) combinent plusieurs opérations de morphologie mathématique (érosion) et techniques d'analyse en composantes connexes en tant que post-traitement après l'utilisation des filtres de Gabor pour extraire différents types de filets des périodiques arabes anciens.

Dans notre travail, les relations de voisinage entre les pixels sont introduites en intégrant la technique de vote majoritaire appliquée par une analyse multi-résolution. Cette technique non-paramétrique est largement utilisée dans la segmentation de documents anciens (Etemad *et al.*, 1994 ; Palfray *et al.*, 2012). Elle vise à éliminer les pixels isolés dans des régions homogènes. En appliquant une approche multi-résolution dans la technique de vote majoritaire, la décision sur chaque pixel est prise en tenant compte du nombre maximal ou de la majorité des étiquettes de pixels effectuées à quatre tailles différentes de fenêtres d'analyse.

4. Protocole expérimental

Afin d'évaluer la robustesse des indices de texture extraits pour la segmentation de documents anciens, nous cherchons à analyser le pouvoir discriminant de chaque ensemble de descripteurs de texture et de déterminer la meilleure catégorie, en testant deux algorithmes classiques de classification non-supervisée : k-moyennes et CAH. Par la suite, nous présentons une étude comparative de trois ensembles de descripteurs texturaux décrits précédemment dans (Mehri *et al.*, 2013a). Enfin, la contribution d'une étape de post-traitement est analysée afin de souligner le rôle des relations de voisinage entre pixels.

Dans le cadre du projet de DIGIDOC (Document Image diGitisation with Interactive DescriptiOn Capability)², nous nous intéressons à la simplification et à l'améliora-

2. Le projet de DIGIDOC est référencé sous ANR-10-CORD-0020. Pour plus de détails, voir [http://www.agence-nationale-recherche.fr/en/anr-funded-project/?tx_lwmsuivibilan_pi2\[CODE\]=ANR-10-CORD-0020](http://www.agence-nationale-recherche.fr/en/anr-funded-project/?tx_lwmsuivibilan_pi2[CODE]=ANR-10-CORD-0020)

tion de l’archivage, le traitement, la comparaison et l’indexation des anciens ouvrages numérisés et collectés par la bibliothèque numérique Gallica³. Ainsi, dans nos expériences, 314 pages ont été sélectionnées. Ces pages sont extraites d’ouvrages Français publiés sur une période de six siècles (1200-1900). Notre corpus expérimental est composé de :

- 89 pages contenant uniquement deux polices de caractères
- 60 pages contenant uniquement trois polices de caractères
- 108 pages contenant illustrations et une seule police de caractères
- 57 pages contenant illustrations et deux polices de caractères

Pour évaluer quantitativement les différents résultats obtenus, des mesures d’évaluation de classification non-supervisée et de classification supervisée sont calculées : la largeur moyenne de silhouette (*SW*), la pureté par bloc (*PPB*), la F-mesure (*F*) et le taux de bonne classification (*CA*) (Mehri *et al.*, 2013a ; Mehri *et al.*, 2013c). Plus les valeurs de ces mesures sont élevées, meilleurs sont les résultats. Dans le Tableau 2, nous avons distingué deux valeurs nommées “*Moyenne*”. Les valeurs des cases du tableau appartenant à la ligne “*Moyenne**” sont obtenues en moyennant toutes les valeurs des différentes colonnes du tableau à l’exception des valeurs issues de la catégorie “*Illustrations et deux polices de caractères***”. Tandis que les valeurs des cases du tableau appartenant à la ligne “*Moyenne***” sont obtenues en moyennant toutes les valeurs des différentes colonnes du tableau à l’exception des valeurs issues de la catégorie “*Illustrations et deux polices de caractère du texte**”. La catégorie des documents “*Illustrations et deux polices de caractère du texte**” représente le cas où toutes les polices de caractères sont associées à une étiquette distincte dans la vérité terrain. La classification non-supervisée des pixels est alors appliquée en fixant à 3 le nombre de types de contenu (illustrations et texte avec deux polices de caractères distinctes). La catégorie des documents “*Illustrations et deux polices de caractère du texte***” considère le cas où toutes les polices du texte ont la même étiquette dans la vérité de terrain. La classification non-supervisée est appliquée en fixant à 2 le nombre de types de contenu (illustrations et texte). Cette distinction permet une analyse plus fine. Elle permet notamment de distinguer une tâche de séparation texte/graphique d’une tâche visant non seulement la séparation texte/graphique mais également la distinction simultanée des différentes polices de caractère du texte.

L’analyse des trois ensembles de descripteurs de texture sur notre corpus par un seul algorithme de classification non-supervisée, avec et sans intégration de la phase de post-traitement, donne un total de 2226 images de document analysées (314 + 57 images × 3 textures × 2). Les expériences ont été exécutées sur un serveur de calcul de type SGI Altix ICE 8200 (1 seul CPU et 2Gb de mémoire allouée dans Quad-Core X5355@2.66GHz sous Linux).

3. <http://gallica.bnf.fr>

5. Évaluation et résultats

La première série d'expériences est réalisée afin de déterminer lequel des deux algorithmes classification non-supervisée (k-moyennes et CAH) est le plus approprié pour les données traitées. Pour l'algorithme k-moyennes, lorsqu'on utilise "*Illustrations et deux polices de caractère du texte***" dans le calcul de "*Moyenne***", nous observons que les performances des descripteurs d'auto-corrélation extraits sont meilleures que lorsqu'on utilise "*Illustrations et deux polices de caractère du texte**" dans le calcul de "*Moyenne**". En effet, des gains globaux de performance 7%(SW), 3%(PPB), 4%(F) et 2%(CA) sont obtenus. De même, pour CAH, nous observons des gains globaux de performance : 8%(SW), 1%(PPB), 5%(F) et 2%(CA).

En comparant les deux méthodes de classification non-supervisée : k-moyennes et CAH, des meilleures performances sont notées en utilisant CAH, *i.e.* des gains de performance sont obtenus en utilisant CAH pour les valeurs "*Moyenne**" et "*Moyenne***" respectivement : 4% ; 2%(PPB), 2% ; 3%(F) et 4% ; 4%(CA). Mais, une légère baisse de performance pour la largeur moyenne de silhouette est observée (-2% et -1% pour "*Moyenne**" et "*Moyenne***" respectivement). Ce phénomène peut être expliqué par le processus de fusion progressive de la CAH, où à des niveaux supérieurs de la hiérarchie, deux points distants dans l'espace des caractéristiques peuvent être associés au même regroupement. Ceci engendre une faible valeur de la largeur de la silhouette. Cette justification peut être appuyée par la particularité de la largeur moyenne de silhouette comme étant une mesure d'évaluation interne ou non-supervisée qui vise à estimer la cohérence d'une solution de regroupement en mesurant la compacité des regroupements. En conclusion, la plupart des mesures d'évaluation calculées s'accordent à indiquer que l'algorithme de CAH est plus performant que celui des k-moyennes.

La deuxième série d'expériences vise à déterminer quel ensemble de descripteurs de texture est le mieux adapté pour discriminer les régions graphiques des régions textuelles d'une part, et d'autre part, pour séparer les régions textuelles de différentes tailles et de différentes polices de caractères, en comparant les trois catégories de descripteurs de texture. En outre, un bref aperçu sur le temps de calcul et la complexité de chaque catégorie de descripteurs de texture est détaillé dans le Tableau 2. Le temps de traitement dépend fortement de la résolution, de la taille de l'image de document analysée et finalement du nombre de pixels de forme présents sur l'image. L'exemple d'une page d'un ouvrage ancien numérisé à 300 ppp et de taille 1965*2750 pixels est donné dans le Tableau 2. Le traitement d'une page de taille 1965*2750 pixels est maximum avec l'utilisation des filtres de Gabor tandis qu'il est minimum avec les descripteurs de co-occurrence (14s). Les valeurs de temps de traitement observées sont en accord avec la complexité théorique des algorithmes permettant l'extraction des descripteurs. Nous observons que l'approche basée sur l'analyse des filtres de Gabor a la plus forte complexité alors que la plus basse est observée pour celle basée sur les descripteurs de co-occurrence (*cf.* Tableau 2). Par conséquent, cette étude conclut que l'approche basée sur les descripteurs de co-occurrence est la meilleure en termes de temps de traitement. Les approches basées sur la co-occurrence et les filtres de

Gabor sont les plus consommatrices en termes d'occupation mémoire. Cela s'explique notamment par le fait que les descripteurs à base de filtre de Gabor correspondent à des vecteurs de grande dimension (192).

Une comparaison des résultats obtenus suite à l'analyse des trois ensembles de descripteurs de texture dans notre approche de segmentation de documents anciens est présentée dans le Tableau 2. Les résultats caractérisés via différentes mesures d'évaluation de classification non-supervisée et de classification supervisée sont concordants. Ces bons résultats justifient le choix d'utiliser les indices de texture pour effectuer une segmentation par approche ascendante sans connaissances *a priori*. Les résultats visuels illustrant la qualité de la segmentation d'images sont présentés sur la Figure 2. Nous observons également que l'approche basée sur les filtres de Gabor (*cf.* Figure 2(c)) a donné les meilleures performances pour la majorité des mesures d'évaluation de classification non-supervisée et de classification supervisée, et ce, sans prendre en considération les relations de voisinage des pixels. Toutefois, lorsque la complexité numérique (temps de calcul) est prise en compte, l'approche à base des filtres Gabor est la plus consommatrice. Nous notons que les performances de l'approche à base des filtres Gabor sont largement supérieures (18%(*F*) gain moyen noté de différence de performance entre les deux approches à base co-occurrence et les filtres de Gabor) dans le cas des documents ne contenant que des régions textuelles avec trois distinctes polices (*cf.* Figures 2(i)). Nous observons que les valeurs les plus basses des mesures d'évaluation de classification non-supervisée et de classification supervisée sont obtenues pour les documents contenant uniquement trois polices de caractères distinctes par l'utilisation de descripteurs de co-occurrence. Nous concluons que les descripteurs de co-occurrence ne sont pas les plus performants pour séparer différentes polices bien qu'ils présentent l'avantage de présenter un temps de traitement plus faible que les autres. Les performances obtenues dans ce travail sur plus de 300 documents confirment celles obtenues dans nos travaux présentés dans (Mehri *et al.*, 2013a), *i.e.* les performances de l'approche à base des filtres Gabor sont les plus élevées.

Finalement, en introduisant la technique de vote majoritaire appliquée par une analyse multi-résolution comme une étape de post-traitement, nous cherchons à affiner les résultats de segmentation en intégrant les informations relatives au voisinage des pixels. Il est important de souligner que notre approche de segmentation basée sur l'analyse de texture est robuste. En effet, elle affiche des résultats globalement prometteurs. Nous obtenons 80% et 86% de F-mesure pour les deux catégories "*Moyenne**" et "*Moyenne***" respectivement sans prendre en compte les informations du voisinage. Néanmoins, nous notons que l'approche à base des filtres de Gabor présente les meilleures performances après l'introduction de la phase de post-traitement dans notre approche de segmentation des documents anciens basée sur la texture. Les filtres de Gabor qui affichent les meilleurs résultats offrent des descripteurs robustes. On constate en effet qu'il n'existe pas d'amélioration notable des performances suite à l'ajout de la phase de post-traitement par vote majoritaire (*cf.* Figures 2(f) et 2(l)).

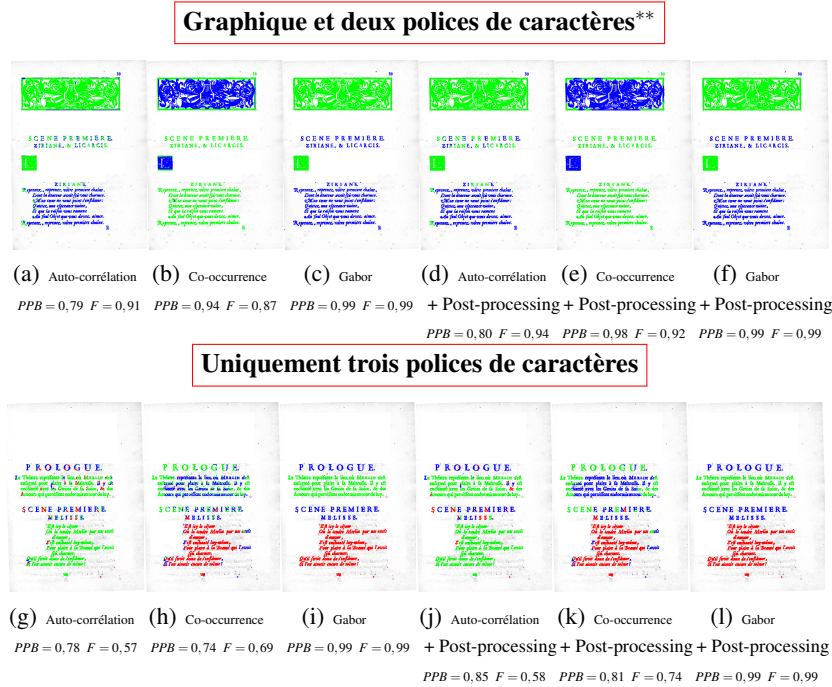


Figure 2. Exemples de résultats de segmentation de documents anciens. Puisque l'approche proposée est non-supervisée, les couleurs attribuées aux pixels texte/illustration peuvent être différentes d'une page à l'autre.

6. Conclusions et perspectives

Cette étude présente une approche pour évaluer et comparer trois ensembles de descripteurs de texture, comprenant l'auto-corrélation, les matrices de co-occurrence des niveaux de gris et les filtres de Gabor. Ce travail a prouvé la robustesse d'une approche à base d'analyse de texture pour segmenter des documents anciens. Sur la base de nos expériences, nous concluons que les descripteurs à base des filtres de Gabor sont les meilleurs choix pour discriminer les régions textuelles de ceux graphiques en ne tenant pas en compte les relations de voisinage entre pixels. En outre, nous avons également montré que l'approche basée sur les filtres de Gabor est la meilleure pour la segmentation d'images de documents ne contenant que des régions textuelles ayant plusieurs polices de caractères distinctes. Cependant, lorsque la complexité numérique est prise en compte et la contribution d'une étape de post-traitement, l'approche à base de la fonction d'auto-corrélation pourrait s'avérer un meilleur choix. D'autres travaux pourraient s'intéresser à l'étude de la complémentarité des descripteurs et pourraient aboutir en la proposition d'un vecteur de caractéristiques optimal combinant certains des descripteurs étudiés.

Tableau 1. Évaluation de l'analyse de descripteurs de texture et de l'apport de l'étape de post-traitement.

		Auto-corrélation			Co-occurrence			Gabor		
		$A^{\dagger}$	$B^{\ddagger}$	$C^{\lrcorner}$	$A^{\dagger}$	$B^{\ddagger}$	$C^{\lrcorner}$	$A^{\dagger}$	$B^{\ddagger}$	$C^{\lrcorner}$
Illustrations et une seule police de caractère du texte	PPB	0,91	0,96	0,05	0,89	0,94	0,05	0,97	0,97	0
	CA	0,88	0,92	0,04	0,83	0,89	0,06	0,92	0,92	0
	F	0,89	0,92	0,03	0,84	0,89	0,05	0,93	0,93	0
Illustrations et deux polices de caractère du texte*	PPB	0,80	0,88	0,08	0,86	0,91	0,05	0,93	0,93	0
	CA	0,71	0,82	0,11	0,69	0,75	0,06	0,72	0,72	0
	F	0,62	0,69	0,07	0,73	0,69	-0,04	0,72	0,72	0
Illustrations et deux polices de caractère du texte**	PPB	0,88	0,93	0,05	0,94	0,97	0,03	0,99	0,99	0
	CA	0,88	0,92	0,04	0,85	0,89	0,04	0,94	0,94	0
	F	0,88	0,93	0,05	0,86	0,91	0,05	0,93	0,93	0
Uniquement deux polices de caractère du texte	PPB	0,89	0,93	0,04	0,84	0,88	0,04	0,95	0,95	0
	CA	0,85	0,89	0,04	0,69	0,69	0	0,88	0,88	0
	F	0,79	0,85	0,06	0,73	0,73	0	0,91	0,91	0
Uniquement trois polices de caractère du texte	PPB	0,75	0,83	0,08	0,79	0,86	0,07	0,89	0,89	0
	CA	0,73	0,79	0,06	0,66	0,69	0,03	0,70	0,70	0
	F	0,61	0,67	0,06	0,63	0,67	0,04	0,65	0,65	0
Moyenne*	PPB	0,84	0,90	0,06	0,84	0,90	0,06	0,93	0,94	0,01
	CA	0,79	0,86	0,07	0,72	0,76	0,04	0,81	0,81	0
	F	0,73	0,78	0,05	0,71	0,75	0,04	0,80	0,80	0
Moyenne**	PPB	0,86	0,91	0,05	0,86	0,91	0,05	0,95	0,95	0
	CA	0,84	0,88	0,04	0,76	0,79	0,03	0,86	0,86	0
	F	0,79	0,84	0,05	0,77	0,8	0,03	0,86	0,86	0

$A^{\dagger}$, $B^{\ddagger}$ et $C^{\lrcorner}$ représentent l'évaluation de l'analyse de descripteurs de texture dans les cas sans étape de post-traitement, avec étape de post-traitement et la différence entre $A^{\dagger}$ et $B^{\ddagger}$ respectivement. Les cellules du tableau dont les fonds sont de couleur jaune et vert indiquent les meilleures performances de l'analyse des descripteurs de texture sans étape de post-traitement ($A^{\dagger}$) et avec étape de post-traitement ($B^{\ddagger}$) respectivement. Les valeurs des cases du tableau appartenant à la ligne "Moyenne*" sont obtenues en moyennant toutes les valeurs des différentes colonnes du tableau à l'exception des valeurs issues de la catégorie "Illustrations et deux polices de caractères**". Tandis que les valeurs des cases du tableau appartenant à la ligne "Moyenne**" sont obtenues en moyennant toutes les valeurs des différentes colonnes du tableau à l'exception des valeurs issues de la catégorie "Illustrations et deux polices de caractère du texte*".

Tableau 2. Complexité et temps de calcul : exemple sur une image de document ancien numérisé de taille 1965*2750 pixels.

	Auto-corrélation	Co-occurrence	Gabor
Durée	00 : 02 : 33	00 : 00 : 14	00 : 06 : 05
Mémoire	$\approx 48\text{Mo}$	$\approx 587\text{Mo}$	$\approx 552\text{Mo}$
Complexité	$O(M(\theta_a W \log_2 W))$	$O(M d_c G^2)$	$O(f_g \theta_g (N^2 \log_2 N))$
Taille du vecteur d'indices de texture	$20 = I_a \times W$	$72 = I_c \times W$	$192 = I_g \times W$
Nombre d'indices de texture	$I_a = 5$	$I_c = 8d_c + 2 = 18$	$I_g = 2f_g \theta_g = 48$

Tel que I_a , I_c et I_g sont les indices d'auto-corrélation, de co-occurrence et de Gabor respectivement. W représente la taille de la fenêtre d'analyse qui est égale à 4. M est le nombre de pixels de forme. N est la dimension de l'image $N \times N$ analysée. G est le nombre de niveaux de gris, *i.e.* 255 niveaux de gris. θ_a est le nombre d'orientation de la rose des directions, *i.e.* 180 valeurs d'orientation. d_c est la distance particulière des matrices de co-occurrence des niveaux de gris définie par la probabilité de paires des niveaux de gris. Dans cette étude, d_c est égale à 2. f_g et θ_g représentent la fréquence et l'orientation spatiales des filtres de Gabor respectivement. "Mo" désigne Méga-octet. La durée d'exécution est indiquée selon le format "hh : mm : ss" (Mehri *et al.*, 2013a).

7. Bibliographie

- Agam G., Bal G., Frieder G., Frieder O., « Degraded document image enhancement », *DRR*, 2007.
- Busch A., Boles W. W., Sridharan S., « Texture for script identification », *PAMI*, vol. 27, n° 11, p. 1720-1732, 2005.
- Chang T., Kuo C. C. J., « Texture segmentation with tree-structured wavelet transform », *TFTSA*, p. 543-546, 1992.
- Charrada M. A., Amara N. E. B., « Texture approach for nets extraction application to old Arab newspapers images structuring », *IPTA*, p. 212-216, 2012.
- Chellappa R., Chatterjee S., « Classification of textures using Markov random field models », *ICASSP*, p. 694-697, 1984.
- Chen C. H., Pau L. F., Wang P., *Texture analysis in the handbook of pattern recognition and computer vision*, second edn, World Scientific, 1998.
- Crasson A., Fekete J. D., « Structuration des manuscrits : du corpus à la région », *CIFED*, 2004.

- Eglin V., Bres S., Rivero C., « Hermite and Gabor transforms for noise reduction and hand-writing classification in ancient manuscripts », *International Journal of Document Analysis and Recognition*, vol. 9, n° 2-4, p. 101-122, 2007.
- Etemad K., Doermann D., Chellappa R., « Page segmentation using decision integration and wavelet packets », *ICPR*, p. 345-349, 1994.
- Etemad K., Doermann D., Chellappa R., « Multiscale segmentation of unstructured document pages using soft decision integration », *PAMI*, vol. 19, n° 1, p. 92-96, 1997.
- Feddaoui N., Hamrouni K., « Personal identification based on texture analysis of Arabic hand-writing text », *ICIT*, p. 1302-1307, 2006.
- Ferrell R., Gleason S., Tobin K., « Application of fractal encoding techniques for image segmentation », *QCAV*, p. 69-77, 2003.
- Gabor D., « Theory of communication. Part 1 : The analysis of information », *Journal of the Institution of Electrical Engineers - Part III : Radio and Communication Engineering*, vol. 93, n° 26, p. 429-441, 1946.
- Grana C., Borghesani D., Cucchiara R., « Automatic segmentation of digitalized historical manuscripts », *Multimedia Tools and Applications*, vol. 55, n° 3, p. 483-506, 2011.
- Haralick R. M., Shanmugam K., Dinstein I., « Textural features for image classification », *SMC*, vol. 3, n° 6, p. 610-621, 1973.
- Jain A. K., Bkattacharjee S. K., Chen Y., « On texture in document images », *CVPR*, p. 677-680, 1992.
- Jain A. K., Farrokhnia F., « Unsupervised texture segmentation using Gabor filters », *PR*, vol. 24, n° 12, p. 1167-1186, 1991.
- Journet N., Ramel J., Mullot R., Eglin V., « Document image characterization using a multi-resolution analysis of the texture : application to old documents », *IJDAR*, vol. 11, n° 1, p. 9-18, 2008.
- Julesz B., « Visual Pattern Discrimination », *IT*, vol. 8, n° 2, p. 84-92, 1962.
- Lafferty J., McCallum A., Pereira F., « Conditional Random Fields : probabilistic models for segmenting and labeling sequence data », *ICML*, p. 282-289, 2001.
- Lance G. N., Williams W. T., « A general theory of classificatory sorting strategies 1. Hierarchical systems », *The Computer Journal*, vol. 9, n° 4, p. 373-380, 1967.
- MacQueen J. B., « Some methods for classification and analysis of multiVariate observations », *Berkeley Symposium on Mathematical Statistics and Probability*, p. 281-297, 1967.
- Mallat S. G., « A theory for multiresolution signal decomposition : The wavelet representation », *PAMI*, vol. 11, n° 7, p. 674-693, 1989.
- Manjunath B. S., Ma W. Y., « Texture features for browsing and retrieval of image data », *PAMI*, vol. 18, n° 8, p. 837-842, 1996.
- Mehri M., Gomez-Krämer P., Héroux P., Boucher A., Mullot R., « Texture feature evaluation for segmentation of historical document images », *HIP*, p. 102-109, 2013a.
- Mehri M., Gomez-Krämer P., Héroux P., Mullot R., « Old document image segmentation using the autocorrelation function and multiresolution analysis », *DRR*, 2013b.
- Mehri M., Héroux P., Gomez-Krämer P., Mullot R., « A pixel labeling approach for historical digitized books », *ICDAR*, p. 817-821, 2013c.

- Nagy G., « Twenty years of document image analysis in PAMI », *PAMI*, vol. 22, n° 1, p. 38-62, 2000.
- Nicolas S., Kessentini Y., Paquet T., Heutte L., « Handwritten document segmentation using hidden Markov random fields », *ICDAR*, p. 212-216, 2005.
- Nourbakhsh F., Pati P. B., Ramakrishnan A. G., « Text localization and extraction from Complex gray images », *ICVGIP*, p. 776-785, 2006.
- Ojala T., Pietikäinen M., Mänpää T., « Multiresolution gray-scale and rotation invariant texture classification with local binary patterns », *PAMI*, vol. 24, n° 7, p. 971-987, 2002.
- Otsu N., « A threshold selection method from gray-level histograms », *SMC*, vol. 27, n° 11, p. 62-66, 1979.
- Ouji A., Leydier Y., Bourgeois F. L., « Chromatic / achromatic separation in noisy document images », *International Conference on Document Analysis and Recognition*, IEEE, p. 167-171, 2011.
- Palfray T., Tranouez D. H. P., Nicolas S., Paquet T., « Segmentation logique d'images de journaux anciens », *CIFED*, 2012.
- Petrou M., Sevilla P. G., *Image processing : dealing with texture*, John Wiley & Sons, 2006.
- Qiao Y., Lu Z., Song C., Sun S., « Document image segmentation using Gabor wavelet and kernel-based methods », *ISSCAA*, p. 450-455, 2006.
- Reed T. R., DuBuf J. M. H., « A review of recent texture segmentation and feature extraction techniques », *CVGIP : Image Understanding*, vol. 57, n° 3, p. 359-372, 1993.
- Ro Y. M., « Matching pursuit : contents featuring for image indexing », *MSAS III*, p. 89-100, 1998.
- Sabharwal C., Subramanya S., « Indexing image databases using wavelet and discrete Fourier transform », *SAC*, p. 434-439, 2001.
- Toyoda T., Hasegawa O., « Texture classification using extended higher order local autocorrelation features », *TAS*, p. 131-136, 2005.
- Tuceryan M., « Moment based texture segmentation », *PRL*, vol. 15, n° 7, p. 659-668, 1994.
- Tuceryan M., Jain A. K., « Texture segmentation using Voronoi polygons », *PAMI*, vol. 12, n° 2, p. 211-216, 1990.
- Tuceryan M., Jain A. K., *Texture analysis*, The Handbook of Pattern Recognition and Computer Vision (2nd Edition), by C. H. Chen, L. F. Pau, P. S. P. Wang (eds.), World Scientific Publishing Co, 1998.

Remerciements

Ce travail s'appuie sur le support de l'Agence Nationale de la Recherche Française (ANR) dans la cadre du projet DIGIDOC sous contrat *ANR-10-CORD-0020*. Les auteurs tiennent également à remercier Geneviève Cron de la Bibliothèque Nationale de France (BnF).