



**HAL**  
open science

## Reconnaissance et classification de lettres à base des descripteurs de bas niveau et de représentation structurelle

Maroua Mehri, Pierre Héroux, Mickaël Coustaty, Petra Gomez-Krämer, Julien  
Lerouge, Rémy Mullot

### ► To cite this version:

Maroua Mehri, Pierre Héroux, Mickaël Coustaty, Petra Gomez-Krämer, Julien Lerouge, et al.. Reconnaissance et classification de lettres à base des descripteurs de bas niveau et de représentation structurelle. Semaine du Document Numérique et de la Recherche d'Information (SDNRI 2016) - Colloque International Francophone sur l'Écrit et le Document (CIFED 2016), Mar 2016, Toulouse, France. pp.465-480. hal-01294406

**HAL Id: hal-01294406**

**<https://hal.science/hal-01294406>**

Submitted on 29 Mar 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

# Reconnaissance et classification de lettrines à base des descripteurs de bas niveau et de représentation structurelle<sup>1</sup>

Maroua Mehri\* — Pierre Héroux\* — Mickaël Coustaty\*\*  
— Petra Gomez-Krämer\*\* — Julien Lerouge\* — Rémy Mullot\*\*

\* *Laboratoire d'Informatique, du Traitement de l'Information et des Systèmes - LITIS, Université de Rouen, France*

{maroua.mehri, pierre.heroux, julien.lerouge}@univ-rouen.fr \*\* *Laboratoire Informatique, Image et Interaction - L3i, Université de Rochelle, France*

{mickael.coustaty, petra.gomez, remy.mullot}@univ-lr.fr

---

*RÉSUMÉ. Cet article s'intéresse à un des éléments importants dans l'analyse de graphiques sur les images de documents anciens que sont les lettrines. Nous proposons dans ce travail des méthodes génériques pour la reconnaissance et la classification de lettrines. Tout d'abord, une méthode ascendante de segmentation à base de descripteurs de bas niveau est présentée, permettant d'assurer la séparation des éléments de la couche de forme de ceux de la couche du fond de la lettrine. Ensuite, une représentation structurelle est proposée pour la caractérisation de lettrines à l'aide d'une analyse des descripteurs de bas niveau (texture, forme, géométriques et topologiques). Cette représentation structurelle est basée sur l'extraction et la sélection automatique des régions homogènes représentatives d'une lettrine pour générer une signature structurelle. Elle fournit ainsi une description riche et holistique permettant une catégorisation à partir du contenu (capturé par les descripteurs de bas niveau) et de la structure (représentée par des graphes). Cette signature permet la mise en œuvre de nombreuses applications pour gérer efficacement une base de lettrines. En comparant les différentes signatures structurelles par le biais de la distance d'édition entre graphes, les similitudes entre les lettrines en termes de style, structures (le motif décoratif du fond de la lettrine) et contenu (la lettre représentée sur la lettrine) peuvent être déduites. Enfin, sur une base d'images de plus de 300 lettrines, une étude expérimentale avec des observations qualitatives et numériques et une comparaison des performances obtenues avec celles de l'état de l'art, est présentée pour illustrer l'efficacité et la robustesse des solutions proposées.*

1. Traduction d'un article publié en anglais à ICDAR 2015 (Mehri *et al.*, 2015b)

*ABSTRACT.* This article tackles some important issues relating to the analysis of a particular case of complex ancient graphic images, called “letrines”, “drop caps” or “ornamental letters”. Our contribution focuses on proposing generic solutions for lettrine recognition and classification. Firstly, we propose a bottom-up segmentation method, based on auto-correlation features, ensuring the separation of the letter from the elements of the background in an ornamental letter. Secondly, a structural representation is proposed for characterizing a lettrine. This structural representation is based on filtering automatically relevant information by extracting representative homogeneous regions from a lettrine to generate a graph-based signature. The proposed signature provides a rich and holistic description of the lettrine style by integrating varying low-level features (e.g. texture). Then, to categorize and classify lettrines with similar style, structure (i.e. ornamental background) and content (i.e. letter), a graph-matching paradigm has been carried out to compare and classify the resulting graph-based signatures. Finally, to demonstrate the robustness of the proposed solutions and provide additional insights into their accuracies, an experimental evaluation has been conducted using a relevant set of lettrine images. In addition, we compare the results achieved with those obtained using the state-of-the-art methods to illustrate the effectiveness of the proposed solutions.

*MOTS-CLÉS :* Lettrine, reconnaissance, classification, indexation, texture, graphe.

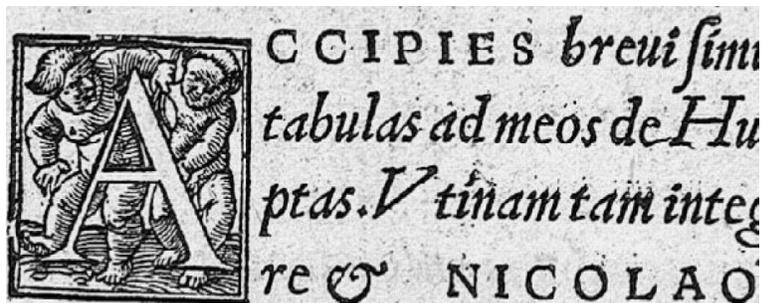
*KEYWORDS:* Lettrine, recognition, classification, indexing, texture, graph.

---

## 1. Introduction

Afin de renforcer l'exploitation et la valorisation des collections de documents patrimoniaux, d'une part, et d'autre part, afin de garantir une conservation durable et de fournir un accès plus large aux documents anciens, de nombreux projets de recherche ont été mis en place avec le soutien des fonds publics fourni par les gouvernements Européens et Américains. Les objectifs de ces projets se focalisent sur (i) la mise en ligne des bases consultables d'images de documents patrimoniaux, (ii) le développement d'outils efficaces d'indexation multilingue et des systèmes fiables de recherche documentaire par navigation libre, mots-clés et contenu dans un corpus de documents anciens et aussi le développement de solutions informatiques assurant l'analyse automatique et rapide des documents historiques. Certains projets visent à proposer des méthodes analysant des images de pages entières d'ouvrages anciens ou toutes les pages d'un ouvrage ancien (Journet *et al.*, 2008 ; Mehri *et al.*, 2015a), alors que d'autres projet se sont centrés essentiellement sur l'étude et l'analyse des parties de documents historiques (e.g. texte, styles/polices de caractères du texte, annotations manuscrites (Eglin *et al.*, 2007)) ou graphiques (e.g. ornements, portraits, illustrations, lettrines (Uttama *et al.*, 2006)). Les images de documents historiques comportent souvent des images graphiques particulières représentées par des lettrines (*cf.* figure 1). Les lettrines correspondent à des images graphiques très utilisées dans les ouvrages anciens des *XV<sup>ème</sup>* et *XVI<sup>ème</sup>* siècles. Une lettrine représente une lettre ornementale qui débute un chapitre ou un paragraphe. Elle est généralement composée de deux couches principales : la lettre et le motif décoratif peint au fond de la lettrine. Ainsi, trois composantes principales de la lettrine peuvent être déduites :

- 1) **Le fond** : illustre l'arrière plan de la lettrine. Il peut-être plein (noir ou blanc), hachuré ou criblé.
- 2) **Le motif** : correspond aux ornements. Il peut être décoratif ou figuratif.
- 3) **La lettre** : est un élément clef de la lettrine puisqu'il précise la lettre qui débute un chapitre ou un paragraphe. Il peut-être noir ou blanc et de différentes polices.



**Figure 1.** Exemple d'une lettrine.

En analysant des lettrines, plusieurs avantages, utilisations et intérêts pourraient se présenter pour l'analyse d'images de documents patrimoniaux. Tout d'abord, l'indexa-

tion des ouvrages anciens numérisés aux premiers temps de l'imprimerie pourrait être assurée. Un autre intérêt de l'étude et de l'analyse des lettrines consiste à enrichir sémantiquement les ouvrages anciens en ajoutant des méta-données ou des annotations sémantiques. En outre, les lettrines peuvent être caractérisées, comparées et classifiées par le biais des signatures statistiques, spatiales ou structurelles déduites suite à l'analyse de leur contenu. Ainsi, des ouvrages anciens pourraient être historiquement datés, authentifiés ou caractérisés en s'appuyant sur les différences des signatures ou des annotations sémantiques déduites des lettrines analysées. D'autres utilisations des lettrines consistent à développer des systèmes pertinents de recherche d'images de lettrines par le contenu, connus dans la littérature sous le nom de CBIR, pour "content-based image retrieval". L'idée consiste à fournir une image requête d'une lettrine au système CBIR développé pour les lettrines afin de récupérer à partir d'une base de lettrines celles qui sont similaires à l'image requête. D'autres besoins ont été identifiés par la communauté des chercheurs dans l'analyse d'images de documents historiques. Par exemple, en analysant les lettrines, la police, la couleur ou l'alphabet qui caractérise l'imprimerie pourraient également être déterminés. Dans d'autres cas, l'analyse des lettrines permet de regrouper l'alphabet utilisé dans les lettrines, étudier l'usure des tampons utilisés pour imprimer les lettrines, analyser le motif peint au fond de la lettrine, examiner le progrès des techniques d'impression utilisées, *etc.*

Néanmoins, l'analyse des lettrines est considérée comme complexe, car il y a une grande variété de styles et de modèles de lettrine, et ils contiennent beaucoup d'informations (e.g. texture, lettre et le motif décoratif du fond de la lettrine). Coustaty *et al.* (Coustaty *et al.*, 2011) stipule que l'analyse des lettrines n'était pas aussi simple que ce que l'on pensait en raison de la grande variabilité de la représentation de l'information en termes de couleurs, de formes et de connectivités aux différentes illustrations, *etc.* Dans le cadre de deux projets Français de recherche suivants : MADONNE<sup>1</sup> et NaviDoMass<sup>2</sup>, nombreuses études ont été menées pour extraire des parties spécifiques de ces images graphiques anciens complexes (Pareti et Vincent, 2005 ; Pareti et Vincent, 2006 ; Chouaib *et al.*, 2009) et de calculer des signatures pour l'indexation (Jouili *et al.*, 2010 ; Coustaty *et al.*, 2011 ; Coustaty et Ogier, 2011).

Par exemple, Coustaty *et al.* (Coustaty *et al.*, 2011) ont proposé une méthode d'extraction de la lettre représentée sur la lettrine en décomposant l'information contenue dans une lettrine en plusieurs couches (*i.e.* en séparant la lettre des éléments de l'arrière plan de la lettrine) dans l'objectif de générer automatiquement une signature pertinente permettant de caractériser une lettrine. Leur méthode est composée de trois principales étapes : (*i*) simplification des images de lettrines par le biais des techniques de décomposition en couches, (*ii*) extraction de formes à partir des couches déterminées dans la première étape, et (*iii*) sélection de la forme correspondant à la lettre à extraire à partir de la lettrine. Elle est basée sur la décomposition de Meyer et la loi de Zipf. La décomposition de Meyer a été utilisée pour filtrer le bruit, extraire les fréquences spatiales des images de lettrines et de les segmenter en différentes couches

---

1. <http://madonne.univ-lr.fr>

2. <http://navidomass.univ-lr.fr>

(forme, texture et bruit). Ensuite, la loi de Zipf a été appliquée sur les niveaux de gris de la couche forme afin de détecter les larges régions homogènes contenant la lettre dans la lettrine. Toutefois, le principal inconvénient de leur méthode consiste à identifier la méthode d'encodage la plus appropriée lors de l'application de la loi de Zipf pour définir les indices capables de distinguer chaque élément constituant une lettrine. En plus, il est assez difficile d'analyser la courbe résultante de la loi de Zipf afin de décomposer la couche de forme de lettrine. En outre, une phase d'apprentissage est nécessaire, qui est aussi basée sur la définition des critères et des règles complexes (e.g. la lettre est la grande composante connexe (CC), qui est forcément centrée dans la lettrine analysée, et elle ne touche pas les frontières de la lettrine).

Récemment, un autre intérêt d'analyse de lettrines a été révélé par les experts historiens, qui consiste à la discrimination assistée par ordinateur de lettrines en fonction de leur style pour répondre aux besoins d'assurer une indexation automatique et pertinente de lettrines. Par exemple, Coustaty et Ogier (Coustaty et Ogier, 2011) ont combiné la loi de Zipf avec le sac de motifs. En extrayant un ensemble de modèles les plus pertinents (avec des tailles pré-définies de masques) à partir d'images de lettrines et leur triant selon les critères de fréquences de type TF-IDF, ils ont proposé une méthode supervisée pour identifier les lettrines en fonction du style du motif décoratif peint au fond de la lettrine. Même s'ils ont obtenu un taux moyen de rappel égal à 80% avec une bonne précision, une phase d'apprentissage sur un ensemble considérable d'images de lettrines est nécessaire afin de générer le modèle d'apprentissage. En outre, ils ont utilisé une étape de quantification afin de réduire le nombre de motifs possibles.

Par ailleurs, l'extraction automatique d'une description détaillée et exhaustive des caractéristiques d'une lettrine est devenue un enjeu majeur afin de répondre aux besoins des experts historiens en termes d'assurer une classification et une catégorisation non-supervisée de lettrines. Par exemple, Jouili *et al.* (Jouili *et al.*, 2010) ont proposé une approche structurelle basée sur la technique d'appariement de graphes pour la classification de lettrines. Ils ont proposé une représentation structurelle en graphe pour caractériser une lettrine. Par la suite, ils ont comparé leur représentation avec une autre statistique basée sur le descripteur de Fourier générique, connu sous le nom de GFD, pour "generic Fourier descriptor". Ils ont noté que les performances de la représentation statistique basée sur le descripteur GFD sont les plus faibles et ils ont conclu que la représentation structurelle en graphe est meilleure que la représentation statistique pour catégoriser les images de lettrines. En se référant à la méthode proposée par Coustaty *et al.* (Coustaty *et al.*, 2009) pour l'extraction de la lettre représentée sur la lettrine, Jouili *et al.* (Jouili *et al.*, 2010) a utilisé les régions les plus représentatives, qui ont été déduites des six plus grandes composantes connexes extraites de la lettrine afin de définir les nœuds de la représentation structurelle en graphe. Ensuite, chaque nœud du graphe a été décrit par un vecteur de quatre descripteurs représentant les caractéristiques topologiques (les coordonnées du centre de gravité, la surface, le rapport entre la largeur et la hauteur de chaque région sélectionnée). Par la suite, les graphes obtenus, caractérisant les lettrines analysées, ont été comparés en utilisant le paradigme d'appariement de graphes pour la classification non-supervisée de lettrines

en fonction du style de leur fond ornemental. Même si la méthode proposée par Jouili *et al.* (Jouili *et al.*, 2010) s'est révélée plus pertinente que celle proposée par Riesen et Bunke (Riesen et Bunke, 2009), la non-utilisation des relations topologiques entre les différentes régions extraites (*i.e.* nœuds du graphe), demeure un inconvénient majeur et a un effet négatif sur les performances de la classification non-supervisée de graphes.

Dans le cadre de nos travaux de recherche, nous proposons des solutions génériques permettant l'indexation de lettrines et plus particulièrement la reconnaissance et la classification non-supervisée de lettrines. Tout d'abord, une méthode ascendante de segmentation à base de descripteurs de bas niveau est proposée pour l'extraction et la sélection automatique de régions homogènes représentatives d'une lettrine. La méthode proposée vise à extraire la lettre représentée sur la lettrine et aussi les éléments du motif décoratif peint au fond de la lettrine afin de faciliter l'indexation de lettrine. Basé sur les composantes représentatives extraites d'une lettrine, deux applications peuvent être mises en œuvre pour gérer efficacement un corpus ou une collection de lettrines. La première application vise à séparer la lettre représentée sur la lettrine des éléments du motif décoratif peint au fond de la lettrine dans le but de reconnaître une lettrine. Et, la seconde application vise à définir une représentation structurelle afin de caractériser une lettrine. Cette représentation structurelle en graphe est construite à partir des régions homogènes extraites et reflétant le contenu et la structure de la lettrine. À l'aide des descripteurs de bas niveau (texture, forme, géométriques et topologiques), une description riche et holistique de la lettrine permettant une catégorisation à partir du contenu (capturé par les descripteurs de bas niveau) et de la structure (représentée par des graphes) est générée. Ainsi de suite en comparant les différentes signatures structurelles par le biais de la distance d'édition entre graphes, les similitudes entre les lettrines en termes de leurs structures et/ou contenus peuvent être déduites. Par conséquent, les lettrines ayant un style similaire sont catégorisées ou regroupées.

Cet article est structuré de la façon suivante. La section 2 décrit la méthode proposée d'extraction et de sélection automatique de régions homogènes représentatives. Dans la section 3, deux applications de la méthode proposée sont présentées pour la reconnaissance et la classification non-supervisée de lettrines (*cf.* sections 3.1 et 3.2). La section 4 détaille tout d'abord les protocoles expérimentaux mis en œuvre pour évaluer la robustesse de la méthode proposée et de ses deux applications. Ensuite, les différentes étapes de l'approche proposée ont donné lieu à des évaluations par le biais d'expérimentations menées sur une base d'images de plus de 300 lettrines. En outre, des observations qualitatives et numériques et une comparaison des performances obtenues avec celles de l'état de l'art, sont présentés pour illustrer l'efficacité et la robustesse des solutions proposées. Enfin, nos conclusions et nos perspectives sont présentées dans la section 5.

## 2. Méthode proposée

Les historiens mettent en lumière le fait que l'analyse de lettrines dans les ouvrages imprimés anciens est une étape importante qui leur permet par exemple d'identifier l'imprimeur de l'ouvrage ou de dater relativement les impressions de deux ouvrages en examinant l'usure des tampons. Parmi les différentes tâches s'appliquant au traitement des lettrines, la détermination de la lettre de la lettrine et l'identification du style de la lettrine sont considérées comme deux étapes importantes pour la recherche d'occurrences de plusieurs impressions à partir du même tampon de lettrine. L'extraction automatique de ces informations par une analyse d'image serait un outil d'une grande assistance pour les historiens.

Du point de vue de l'analyse d'image, les techniques similaires à celles appliquées par les OCR visant à reconnaître les lettres formées par des composantes connexes se révèlent inopérantes, notamment en raison de la grande variabilité dans les styles de lettrines, de la présence de bruit et de dégradation qui perturbe de façon non-uniforme leurs contours. Par ailleurs, une lettrine est constituée de deux couches d'information principales. D'une part, un fond ornemental, qui est une des caractéristiques du style de la lettrine, est caractérisé par des fréquences régulières. D'autre part, la lettre en elle-même et d'autres motifs, sont caractérisés par de faibles fréquences spatiales et des faibles variations des niveaux de gris. Il est de ce fait de combiner une analyse de texture pour traiter le fond de la lettrine et une analyse en composantes connexes pour traiter la lettre représentée. L'approche que nous proposons consiste en une segmentation ascendante combinant l'analyse de texture et de composantes connexes pour déterminer les principales régions homogènes des lettrines.

### 2.1. *Prétraitements et extraction de la caractéristique d'auto-corrélation*

Dans un premiers temps, les lettrines sont représentées comme des images numériques en niveaux de gris. L'extraction des informations relatives à la texture est appliquée directement à l'image en niveaux de gris afin d'éviter le biais qu'introduirait une phase de binarisation. L'analyse de texture s'opère sur les pixels de l'image en examinant son voisinage. Pour permettre le traitement de tous les pixels de l'image, notamment ceux situés en limite, par une technique de fenêtre d'analyse centrée sur le pixel à caractériser, nous introduisons une étape de réplique des bords de l'image. Dès lors, il nous est possible d'extraire cinq caractéristiques de la matrice d'auto-corrélation : l'orientation principale de la rose de direction, l'intensité de la fonction d'auto-corrélation pour l'orientation principale, la variance des intensités de la rose de direction, l'épaisseur moyenne et la hauteur moyenne selon les directions spécifiques). Ces caractéristiques ont été présentées dans (Mehri *et al.*, 2013). Ces caractéristiques sont extraites de l'image numérique en niveau de gris pour 4 tailles de fenêtre d'analyse ((16 × 16), (32 × 32), (64 × 64) et (128 × 128)) de façon à adopter une approche multi-résolution qui puisse s'affranchir de la résolution des images. En fait, dans ce travail, les tailles de fenêtres d'analyse varient seulement de (16 × 16)

jusqu'à  $(128 \times 128)$ , car au-delà la taille  $(128 \times 128)$ , l'étape extraction de la caractéristique d'auto-corrélation serait coûteuse en termes de temps de traitement. En plus, les dimensions des images de lettres évaluées dans ce travail ne dépassent pas  $(128 \times 128)$ . Donc, suite à l'étape d'extraction de la caractéristique d'auto-corrélation, chaque pixel est alors caractérisé par un vecteur à 20 dimensions (5 caractéristiques pour chacune des 4 tailles de fenêtre d'analyse).

## **2.2. Classification automatique des pixels**

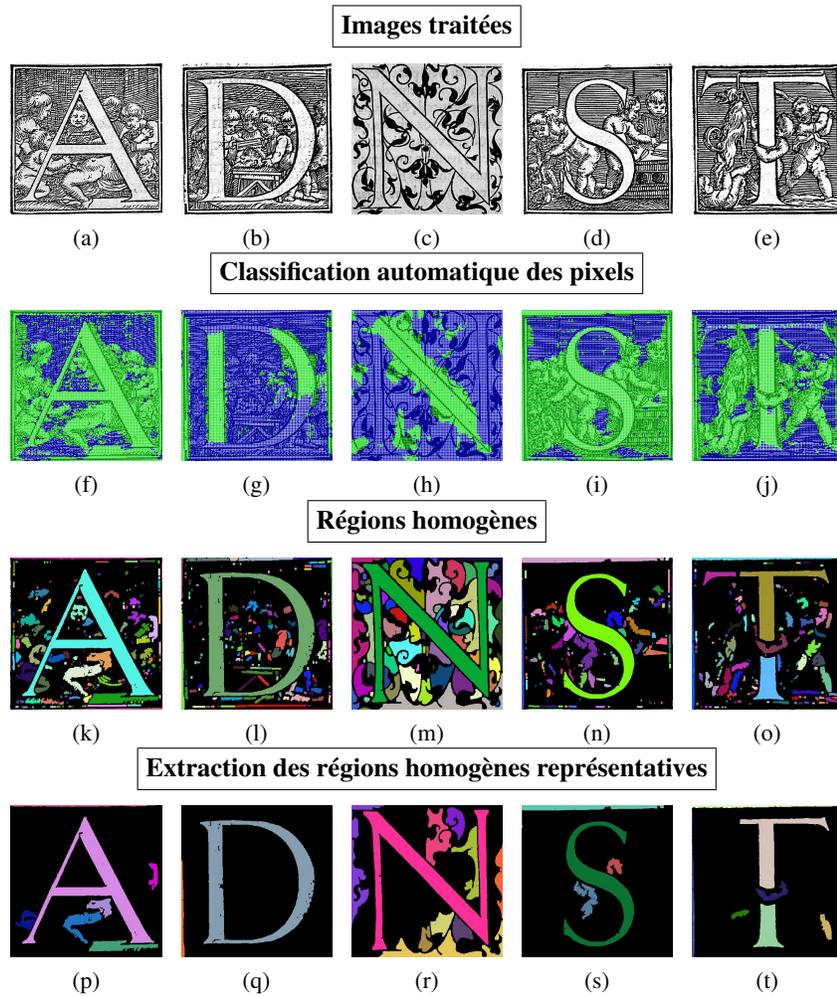
Chacune des dimensions du vecteur de caractéristique des pixels est ensuite normalisée. Ces vecteurs normalisés sont donnés en entrée d'un algorithme de classification non-supervisée pour lequel nous fixons le nombre de classes à deux. L'idée sous-jacente est d'appliquer une classification non-supervisée qui discrimine les pixels du font de ceux des formes. L'algorithme de classification utilisé est celui de la classification ascendante agglomérative. Les figures 2(f), 2(g), 2(h), 2(i) et 2(j) montrent le résultat de ce traitement sur quelques images de lettres.

## **2.3. Post-traitement**

Après la classification automatique (*cf.* section 2.2), nous appliquons un traitement visant à identifier des masses de pixels disposant de caractéristiques similaires du point de vue de la texture. Pour ce faire, nous procédons à une extraction et une analyse des composantes connexes par le biais de l'algorithme ARLSA (adaptive run-length smoothing algorithm) qui vise à combler les interstices entre les composantes connexes extraites (Mehri *et al.*, 2015c). Le résultat de ce traitement est illustré sur les figures 2(k), 2(l), 2(m), 2(n) et 2(o).

## **2.4. Sélection des régions homogènes représentatives**

L'objectif de cette étape consiste à ne considérer, parmi l'ensemble des composantes connexes identifiées par l'étape précédente, que les plus représentatives. Dans un premier temps, nous attribuons une étiquette à chaque composante connexe produite par l'étape précédente (*cf.* section 2.3). Cette étiquette est déterminée par un vote majoritaire des pixels qui la composent. L'étape ultérieure de sélection consiste à ne conserver que les composantes connexes dont l'étiquette est identique à celle de la plus grande composante connexe (il s'agit le plus souvent d'une composante associée à la lettre) et dont l'aire est supérieure à 5% de l'aire de la plus grande composante connexe. Les composantes connexes ainsi sélectionnées correspondent à des parties des lettres représentées ou à des éléments caractéristiques du fond ornemental tel qu'illustré sur les figures 2(p), 2(q), 2(r), 2(s) et 2(t).



**Figure 2.** *Illustration des différentes phases de traitement*

### 3. Reconnaissance et classification de lettrines

Nous proposons d'utiliser les régions extraites tel qu'indiqué précédemment pour deux applications : la reconnaissance de lettrines et la classification non-supervisée de lettrines.

### 3.1. Reconnaissance de letrines

La reconnaissance automatique de la lettre représentée sur la letrine est un moyen de fournir une clé d'index sur ces letrines. Selon Coustaty *et al.* (Coustaty *et al.*, 2011), la plus grande composante connexe au sein d'une letrine couvre une partie significative de la lettre représentée. Partant de ce constat, nous proposons d'appliquer un logiciel de reconnaissance optique de caractères à la plus grande composante connexe issus des traitements précédents.

### 3.2. Classification non-supervisée de letrines

Dans cette autre application, nous mettons à profit les nombreux avantages associés à l'utilisation de représentations structurales (Conte *et al.*, 2004 ; Jouili *et al.*, 2010). Nous proposons de représenter l'image de la letrine par un graphe modélisant les principales régions de l'image issues des traitements précédents et les relations topologiques respectives entre ces régions.

Dans ce graphe, les nœuds représentent les régions homogènes et sont étiquetés par un vecteur de 66 caractéristiques. Parmi ces 66 caractéristiques, on retrouve 20 caractéristiques associées à la texture. Il s'agit des valeurs moyennes des caractéristiques d'auto-corrélation prises sur les quatre tailles de fenêtre d'analyse, calculées sur l'ensemble des pixels associé à la composante connexe. Les 46 caractéristiques restantes sont des caractéristiques associées à la forme de la composante connexe. On y retrouve notamment les coordonnées du centre de gravité, la longueur du contour et l'aire de la composante, la moyenne et l'écart-type du niveau de gris, les coordonnées du rectangle englobant, sa largeur, sa hauteur, son aire, le rapport largeur sur hauteur, la hauteur (resp. largeur) de la composante normalisée par rapport à la hauteur (resp. largeur) de l'image de la letrine, 10 moments spatiaux, 7 moments centrés, 7 moments centrés normalisé et 7 moments de Hu.

Les arcs du graphe sont utilisés pour modéliser les relations topologiques entre les régions représentatives. Un arc est construit entre deux nœuds, si une interaction  $F$  (cf. equation 1) entre les deux régions atteint un niveau suffisant ( $F_e^{s,d} \geq Th_e$ ).

$$F_e^{s,d} = \frac{N_{G_v^d}}{(ED_{G_v^{s,d}})^2} \quad [1]$$

où

$ED_{G_v^{s,d}}$  représente la distance Euclidienne entre deux nœuds : source ( $G_v^s$ ) et destination ( $G_v^d$ );

$N_{G_v^d}$  représente le nombre de pixels du nœud destination ( $G_v^d$ );

$Th_e$  représente le seuil de la force d'attraction.

Nous proposons de construire cette interaction sur le modèle gravitationnel qui veut qu'il n'y ait pas d'interaction entre deux petites régions à moins qu'elles ne soient proches et que les grandes régions puissent être liées aux autres. Les arcs sont étiquetés par la valeur de cette force gravitationnelle et par les distances horizontale et verticale entre les centres de gravité des deux nœuds : source ( $G_v^s$ ) et destination ( $G_v^d$ ).

Afin de regrouper et catégoriser les lettrines disposant d'un style ornemental identique ou d'une même lettre, nous construisons les représentations structurelles des images et nous les comparons par un calcul de distances d'édition entre graphes (Bunke et Riesen, 2012). Cette distance est déduite au coût de la séquence d'édition de coût minimal permettant de transformer un des deux graphes vers l'autre, les éditions étant des substitutions, des insertions ou des suppressions de nœuds ou d'arcs. La catégorisation d'un ensemble d'images de lettrines est réalisée sur la base des distances entre les représentations structurelles calculées deux à deux. Le nombre de classes est une entrée donnée par les experts historiens et correspond au nombre de styles.

#### 4. Évaluation et résultats

Afin d'évaluer les méthodes proposées pour la reconnaissance des lettres et la classification, nous avons mené une évaluation sur une base de 301 images de lettrines. Cette base avec sa vérité terrain ont été fournies par le centre d'études supérieur de la Renaissance (CESR) de Tours, France<sup>3</sup>.

Tout d'abord, la première application de la méthode proposée d'extraction et de sélection automatique de régions homogènes représentatives, qui consiste à identifier la lettre représentée sur la lettrine pour la reconnaissance de lettrines, a été évaluée en catégorisant notre corpus expérimental selon le caractère alphabétique représenté sur la lettrine. Par conséquent, notre base d'images de lettrines est composée de 18 différents caractères alphabétiques.

En second lieu, les images de lettrines ont été étiquetées par des historiens afin de définir une vérité terrain plus appropriée permettant de catégoriser les lettrines disposant d'un style ornemental identique. Cette vérité terrain définie, vise à évaluer la seconde application de la méthode proposée d'extraction et de sélection automatique de régions homogènes représentatives, qui consiste à classifier les lettrines d'une manière non-supervisée. Elle a aussi pour objectif d'illustrer l'efficacité de la signature structurelle en graphe proposée pour la caractérisation de lettrines. La vérité terrain spécifie le nombre de classes de lettrines qui est une entrée donnée par les experts historiens et correspond au nombre de styles. Le nombre de classes de lettrines fourni par les experts historiens est égal à 3. Cette catégorisation de lettrine est basée sur la caractérisation des différents styles de lettrines en fonction de leurs fonds ornementaux, leurs apparences et leurs propriétés. Les trois styles d'images de lettrines, utilisés dans nos expérimentations, sont illustrés dans la figure 3.

3. <http://cesr.univ-tours.fr>

Par conséquent, notre corpus expérimental est composé de :

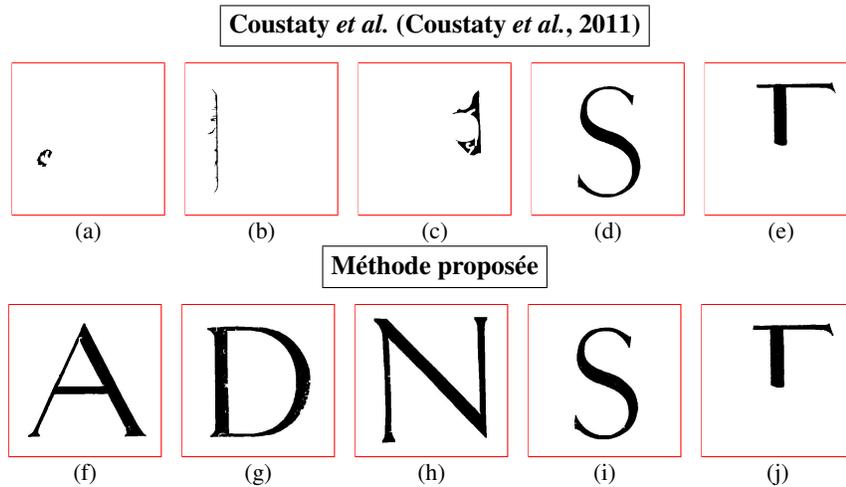
- 191 images de lettrines ayant le premier style “Style 1” (cf. figure 3(a)),
- 95 images de lettrines ayant le second style “Style 2” (cf. figure 3(b)),
- 15 images de lettrines ayant le troisième style “Style 3” (cf. figure 3(c)).



**Figure 3.** Illustration de quelques exemples d’images de lettrines utilisés dans notre corpus expérimental.

#### 4.1. Reconnaissance de lettrines

L’évaluation de la première application de la méthode proposée d’extraction et de sélection automatique de régions homogènes représentatives repose sur l’utilisation d’un logiciel OCR libre (Tesseract) sur les images de lettrines résultantes de l’extraction de la lettre représentée sur la lettrine (cf. section 3.1 et figures 4(f), 4(g), 4(h), 4(i) et 4(j)). En conséquence, le taux de reconnaissance de caractères (*RA*) est utilisé en tant qu’un critère d’évaluation des performances de la méthode proposée pour la reconnaissance de lettrines. En comparant nos résultats avec ceux obtenus avec une méthode proposée par Coustaty *et al.* (Coustaty *et al.*, 2011), un gain de *RA* égal à 17.6% est noté, en appliquant Tesseract comme un système d’OCR pour la reconnaissance de lettrines (cf. tableau 1). Une moyenne de *RA* égale à 84.7% a été obtenue. Nous concluons que notre méthode donne de meilleures performances pour l’application d’extraction de la lettre représentée sur la lettrine que la méthode proposée par Coustaty *et al.* (Coustaty *et al.*, 2011). Les résultats visuels illustrant la qualité des performances obtenues de notre méthode d’extraction de lettres et de reconnaissance de lettrines et celles de la méthode proposée par Coustaty *et al.* (Coustaty *et al.*, 2011) indique que les meilleurs résultats sont obtenus avec notre méthode (cf. figure 4). Néanmoins, même si les résultats montrent un plus grand pouvoir discriminant de notre méthode pour séparer la lettre représentée sur la lettrine en comparant la méthode proposée par Coustaty *et al.* (Coustaty *et al.*, 2011), les deux méthodes ont échoué à extraire la lettre “T” d’une lettrine (cf. figure 2(e)), ce qui a engendré de la confusion avec la lettre “r” lors qu’on a appliqué Tesseract (cf. figures 4(j) et 4(e)). Ceci peut être expliqué par la grande complexité de la lettrine analysée (*i.e.* des chevauchements des limites des zones de la lettre avec le motif décoratif peint au fond de la lettrine).



**Figure 4.** Illustration des images de lettrines résultantes de la première application de la méthode proposée d'extraction et de sélection automatique de régions homogènes représentatives : la reconnaissance de lettrines en identifiant la lettre représentée sur la lettrine.

**Tableau 1.** Évaluation de la méthode proposée d'extraction et de sélection automatique de régions homogènes représentatives pour la reconnaissance de lettrines.

	Méthodes évaluées	Tesseract (%)
Taux de reconnaissance de caractères (RA)	Coustaty et al. (Coustaty et al., 2011)	67.1
	Méthode proposée	84.7

#### 4.2. Classification non-supervisée

Afin d'examiner la performance de la seconde application qui consiste en une classification non-supervisée des images de lettrines et pour illustrer le pouvoir de représentation de la signature structurale proposée, nous présentons une matrice de confusion (cf. tableau 2). La matrice de confusion illustre la performance de la classification non-supervisée au regard de la distribution en classes fournie par les historiens. Les mesures de performances calculées sont la précision ( $P$ ), le rappel ( $R$ ), la  $F$ -mesure ( $F$ ) et le taux de bonne classification ( $CA$ ).  $P_i$  et  $R_i$  désignent respectivement la précision et le rappel sur le cluster  $i$ . Les résultats globaux sont relativement encourageants puisque nous obtenons 65% de précision, 82% de rappel, une  $F$ -mesure de 73% et un

taux de bonne classification de 72%. Cependant, il est remarqué une plus faible précision et un plus fort rappel sont observés sur la classe associée au “Style 1”, alors qu’on observe une tendance opposée pour la classe associée au “Style 2”. Il semblerait que l’application proposée tende à confondre les deux classes. Ce phénomène peut s’expliquer par une grande similitude entre les styles des fonds ornementaux. Cependant, en comparant les résultats obtenus à ceux de Coustaty *et al.* (Coustaty *et al.*, 2011), nous remarquons un gain prometteur avec un temps de traitement meilleur. Un rappel plus fort est obtenu en utilisant notre signature structurelle en graphe pour caractériser une lettrine.

**Tableau 2.** *Évaluation de la seconde application de la méthode proposée d’extraction et de sélection automatique de régions homogènes représentatives : la classification non-supervisée de lettrines en utilisant une représentation structurelle en graphe.*

Vérité terrain				
	Style 1	Style 2	Style 3	
Cluster 1	124	67	0	$\leftrightarrow P_1 = 0.65$
Cluster 2	9	86	0	$\leftrightarrow P_2 = 0.91$
Cluster 3	3	6	6	$\leftrightarrow P_3 = 0.4$

$\updownarrow$	$\updownarrow$	$\updownarrow$
$R_1 = 0.91$	$R_2 = 0.54$	$R_3 = 1$

## 5. Conclusions et perspectives

Cet article aborde les questions relatives à l’analyse de graphiques sur les images de documents anciens et plus particulièrement les lettrines. En effet, une approche ascendante de segmentation exploitant des descripteurs de texture mesurés à différentes échelles est tout d’abord proposée pour l’extraction des régions homogènes. Dès lors que les zones homogènes représentatives ont été extraites, la seconde étape de l’approche sépare la lettre des éléments de la couche du fond de la lettrine. Ensuite, une signature structurelle de la lettrine est générée sur la base de l’extraction et de la sélection automatique de régions homogènes représentatives. Elle fournit ainsi une description riche et holistique permettant une catégorisation à partir du contenu (capturé par les descripteurs de bas niveau) et de la structure (représentée par des graphes). Cette signature structurelle en graphe associant représentation des éléments de contenu et description structurelle, assure la caractérisation d’une lettrine offrant ainsi plusieurs modalités de reconnaissance et catégorisation de lettrines. Les nœuds du graphe ainsi produits sont associés aux zones homogènes et sont étiquetés par les attributs caractérisant les régions. Les arcs, quant à eux, caractérisent les liens topologiques entre les différentes régions. Ensuite, en comparant les différentes signatures structurelles par le biais de la distance d’édition entre graphes, les similitudes entre

les lettrines en termes de style, structures (le motif décoratif du fond de la lettrine) et contenu (la lettre dans la lettrine) peuvent être déduites. Par la suite, les lettrines ayant un style similaire sont regroupées. Les résultats d'expérimentations obtenus sur une base d'images de plus de 300 lettrines montrent clairement que notre méthode ascendante de segmentation de lettrines présentent de meilleures performances que celles de l'état de l'art pour les deux applications suivantes : la reconnaissance de lettrines et la classification non-supervisée de lettrines. Nos travaux futurs dans la lignée de ceux décrits dans cet article viseront à évaluer les solutions proposées avec d'autres ensembles de descripteurs de texture, par exemple le motif binaire local (LBP). En outre, nous envisageons d'évaluer les solutions proposées pour la reconnaissance et la classification de lettrines sur une variété plus large de styles des lettrines.

## 6. Bibliographie

- Bunke H., Riesen K., « Towards the unification of structural and statistical pattern recognition », *PRL*, vol. 33, p. 811-825, 2012.
- Chouaib H., Cloppet F., Vincent N., « Graphical drop caps indexing », *GREC*, p. 212-219, 2009.
- Conte D., Foggia P., Sansone C., Vento M., « Thirty years of graph matching in pattern recognition », *IJPRAI*, vol. 18, p. 265-298, 2004.
- Coustaty M., Ogier J. M., « Discrimination of old document images using their style », *ICDAR*, p. 389-393, 2011.
- Coustaty M., Ogier J. M., Vincent R. P. N., « Drop caps decomposition for indexing : a new letter extraction method », *ICDAR*, p. 476-480, 2009.
- Coustaty M., Pareti R., Vincent N., Ogier J. M., « Towards historical document indexing : extraction of drop cap letters », *IJDAR*, vol. 14, p. 243-254, 2011.
- Eglin V., Bres S., Rivero C., « Hermite and Gabor transforms for noise reduction and handwriting classification in ancient manuscripts », *IJDAR*, vol. 9, p. 101-122, 2007.
- Jouili S., Coustaty M., Tabbone S., Ogier J. M., « NaviDoMass : structural-based approaches towards handling historical documents », *ICPR*, p. 946-949, 2010.
- Journet N., Ramel J., Mullot R., Eglin V., « Document image characterization using a multiresolution analysis of the texture : application to old documents », *IJDAR*, vol. 11, p. 9-18, 2008.
- Mehri M., Gomez-Krämer P., Héroux P., Boucher A., Mullot R., « A texture-based pixel labeling approach for historical books », *PAA*, vol. 1, p. 1-40, 2015a.
- Mehri M., Gomez-Krämer P., Héroux P., Coustaty M., Lerouge J., Mullot R., « A bottom-up method using texture features and a graph-based representation for lettrine recognition and classification », *ICDAR*, p. 226-230, 2015b.
- Mehri M., Gomez-Krämer P., Héroux P., Mullot R., « Old document image segmentation using the autocorrelation function and multiresolution analysis », *DRR*, SPIE, 2013.
- Mehri M., Héroux P., Sliti N., Gomez-Krämer P., Amara N. E. B., Mullot R., « Extraction of homogeneous regions in historical document images », *VISAPP*, p. 47-54, 2015c.
- Pareti R., Vincent N., « Global discrimination of graphic styles », *GREC*, p. 120-130, 2005.
- Pareti R., Vincent N., « Ancient initial letters indexing », *ICPR*, p. 756-759, 2006.

Riesen K., Bunke H., « Approximate graph edit distance computation by means of bipartite graph matching », *IVC*, vol. 27, p. 950-959, 2009.

Uttama S., Loonis P., Delalandre M., Ogier J. M., « Segmentation and retrieval of ancient graphic documents », *GREC*, p. 88-98, 2006.