



HAL
open science

Ontologies et documents anciens: application aux images de lettrines

Mickaël Coustaty, Tsopze Norbert, Bouju Alain, Karell Bertet, Georges Louis,
Jean-Marc Ogier

► **To cite this version:**

Mickaël Coustaty, Tsopze Norbert, Bouju Alain, Karell Bertet, Georges Louis, et al.. Ontologies et documents anciens: application aux images de lettrines. CIFED, Mar 2012, Bordeaux, France. pp.231-246. hal-00715366

HAL Id: hal-00715366

<https://hal.science/hal-00715366>

Submitted on 6 Jul 2012

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.

Ontologies et documents anciens: application aux lettrines

**Mickaël Coustaty, Norbert Tsopze, Alain Bouju, Karell Bertet,
Georges Louis, Jean-Marc Ogier**

*Laboratoire L3i - Université de La Rochelle
Avenue Michel Crépeau - 17042 La Rochelle
{mcoustat,ntsopze,abouju,kbertet,glouis,jmogier}@univ-lr.fr*

RÉSUMÉ. Cet article présente une analyse d'image de documents anciens, et plus particulièrement d'images de lettrines, qui vise à réduire le fossé sémantique entre des traitements d'images et les mots-clés d'un domaine d'expertise. Extraire automatiquement de l'information d'images de documents anciens n'est pas chose aisée, et l'information extraite n'est pas directement interprétable par les utilisateurs finaux (les historiens). Nous proposons donc dans cet article des méthodes d'extraction d'information adaptées au cas des images graphiques de documents anciens, et un cadre formel basé sur des ontologies de représentation de ces connaissances. De manière analogue, nous proposons également une ontologie qui permette de représenter les connaissances issues du domaine d'expertise des images : les historiens. Le lien entre ces deux domaines repose alors sur des règles d'inférences qui créent un pont entre mots-clés du domaine et caractéristiques bas-niveau.

ABSTRACT. This paper proposes a novel method of historical documents image analysis, and more specially images of lettrines, that aims at reducing the semantic gap between automatic image processing algorithms, and keywords from experts's domain. Automatic extraction of information from historical document is a difficult task, and the extracted information is not directly usable by end users (the historians). We thus propose some new adapted algorithms to extract a relevant information from these graphical images, and a formal framework based on ontologies, to represent them. Similarly, we propose an ontology to represent historians's knowledge. A link between these two domains then relies on inference rules, that create a bridge between experts's keywords and low-level features.

MOTS-CLÉS : Ontologies, Documents anciens, Lettrines, Formes, Textures, Spatial

KEYWORDS: Ontologies, Historical documents, Lettrines, Shapes, Textures, Spatial

1. Introduction

Etablir un lien entre une information pixellaire (bas-niveau) et une information sémantique (haut-niveau) décrivant des images est une problématique importante en analyse d'images. Cette problématique, identifiée sous le terme de *fossé sémantique*, consiste à *représenter des données complexes* décrivant l'image dans un objectif de *recherche d'information* qui peut se décliner sous la forme de recherche par mot-clé, par requête, par indexation, par navigation, ou encore recherche par le contenu. Ces problématiques se distinguent des problématiques de reconnaissance ou de classification d'images à partir d'une signature de l'image, classiquement abordées en fouille de données et en apprentissage.

Les technologies de types *ontologies* ont récemment été introduites dans le domaine du web sémantique pour permettre d'organiser les données et connaissances d'un domaine (Uschold *et al.*, 1996). Ces technologies ont pour but de partager, diffuser et actualiser les connaissances du domaine, et de fournir un cadre méthodologique à la fois pertinent et novateur pour représenter des données complexes. Les ontologies peuvent être utilisées comme modèle formel prenant en charge des opérations d'ingénierie (extraction automatique d'une information de bas niveau, construction semi-automatique de l'ontologie, généricité, modularisation, etc), où l'interopérabilité joue un rôle majeur. Par ailleurs, les ontologies reposent sur le formalisme des logiques de description (Horrocks *et al.*, 2003), ce qui leur permet de proposer différents *mécanismes de raisonnement* permettant de déduire de la connaissance implicite à partir de la connaissance explicitement formalisée.

Dans le domaine de l'image, les ontologies semblent être une solution prometteuse pour définir une sémantique explicite : elles permettent de surmonter le problème de la représentation implicite et cachée d'une connaissance visuelle et peuvent être utilisées comme une référence commune pour échanger des contenus sémantiques. Plusieurs travaux (Liu *et al.*, 2007, Hanbury, 2008, Tousch *et al.*, 2012, Zhang *et al.*, 2012, Bannour *et al.*, 2012) proposent une étude des techniques de résolution du problème du fossé sémantique. Cinq catégories de techniques dans la résolution du fossé sémantique y sont identifiées, allant de l'utilisation des ontologies à des techniques d'apprentissage automatique. Dans (Hanbury, 2008), un typologie des métadonnées pour annoter les images est proposée : métadonnées indépendantes du contenu de l'image (nom de l'auteur, date,...) ; métadonnées liées aux caractéristiques de bas niveau de l'image (texture, couleur,...) ; métadonnées liées au contenu sémantique de l'image (voiture, arbre). Il est possible d'associer un vocabulaire à l'image globale, ou à des régions de l'image.

Dans (Zhang *et al.*, 2012) le processus d'extraction des caractéristiques et la représentation des images est décrit, puis un ensemble de méthodes pour les annoter est dénombré : ces méthodes basées sur l'apprentissage automatique (SVM, arbres de décision, réseaux de neurones artificiels, réseaux bayésiens) permettent de classer les images (ou les régions) suivant les concepts préalablement appris. Il note aussi que ce traitement est exposé aux problèmes liés à l'apprentissage : nécessité d'une

base d'apprentissage, temps d'apprentissage,... Dans (Tousch *et al.*, 2012), la place de l'utilisateur, son objectif ainsi que le contexte dans l'annotation des images sont soulignées. Une méthode de construction automatique de la hiérarchie entre les concepts est proposée dans (Bannour *et al.*, 2012) en combinant les informations visuelles, conceptuelles et contextuelles. Dans (Neumann *et al.*, 2008) les scènes (images) sont modélisées à l'aide de la logique de description, montrant ainsi son importance dans l'interprétation des scènes. Cette étude note que la logique formelle permet d'éviter des erreurs qui peuvent survenir quand les connaissances et l'inférence sont construites de manière intuitive. Elle propose un modèle formel d'interprétation d'une scène et souligne l'importance des contextes spatial et temporel dans la tâche d'interprétation.

L'interprétation d'une image dépend de l'utilisateur, du contexte et du degré de granularité de la connaissance. L'utilisation de l'ontologie a différents objectifs : description unifiée des caractéristiques de l'image, description visuelle des relations entre les caractéristiques (lignes, région,...), utilisation des informations contextuelles et enfin rapprochement du niveau visuel et du niveau sémantique.

La *représentation spatiale* de primitives issues d'une image est une problématique abordée en particulier dans les systèmes d'information géographique (Sharma *et al.*, 1994) où une large gamme de modèles et outils a été développée. Les problématiques abordées concernent avant tout une représentation de données complexes décrivant des images géographiques, où l'interopérabilité entre représentations joue un rôle majeur. La prise en compte de règles métier pour la modélisation et le raisonnement spatial et temporel est une problématique classique, citons à titre d'exemple les travaux de (Malki *et al.*, 2009).

Le domaine du document est principalement concerné par la problématique du fossé sémantique, les images de documents représentant des images issues d'un domaine souvent porteur d'une connaissance métier sémantiquement forte. Les archives de documents anciens représentent une vraie richesse qui n'ont pas seulement une valeur historique pour leur apparence physique mais également pour leur contenu. En effet, l'intérêt ici n'est pas seulement de numériser des documents mais également de créer une librairie numérique de documents enrichis par des annotations sémantiques permettant de décrire, classer et indexer des documents par leurs contenus.

En ce qui concerne les images de lettrines, images extraites de livres anciens, les historiens ont identifié quatre niveaux de décomposition les décrivant et pour lesquels une taxonomie a été définie : le fond (noir, blanc, hachuré, ou criblé), le motif (figuratif ou non figuratif), la lettre et le bord. De récents travaux au laboratoire ont permis d'en extraire certaines zones d'intérêts (Coustaty, 2011). Nous présentons dans cet article une ontologie des images de lettrines regroupant l'ensemble de ces données (taxonomie de mots-clé, zones d'intérêt extraites des images, relations spatiales) au sein d'une même représentation. Cette ontologie nous a permis de tester le mécanisme d'inférence de nouvelles connaissances. Nous avons ainsi pu identifier certaines zones selon qu'elles correspondent ou non à une lettre, à un élément du fond figuratif, ou qu'elles participent à la constitution d'un fond hachuré.

Les images de lettrines sont décrites dans la section 2. La section 3 décrit l'ontologie des lettrines que nous avons mis en place. Enfin, nous avons réalisé quelques expérimentations décrites dans la section 4.

2. Images de lettrines

Les images de lettrines sont des images graphiques que l'on retrouve dans des documents anciens des *XV^{ème}* et *XVI^{ème}* siècles. La Figure 1 en donne un exemple. Il s'agit d'une lettre décorée, en début des paragraphes des livres de cette époque, obtenue à l'aide de tampon en bois, sculpté à la main. Composée principalement d'une lettre en son centre, une lettrine se caractérise également par un arrière-plan qui peut être décoratif, ou bien représenter des scènes sociales de l'époque. Les nuances et les ombres étaient obtenues à l'aide de traits parallèles.

L'étude des lettrines fait l'objet de travaux au centre d'étude supérieur de la renaissance (CESR) de Tours. Les lettrines représentent une source d'information exploitée par les historiens pour situer les documents dans le temps, ou encore pour étudier les scènes sociales des fonds figuratifs. En effet, les tampons étant utilisés plusieurs fois, leur usure permet de situer les documents les uns par rapport aux autres. De plus, ils portent souvent des caractéristiques propres à celui qui les a sculptés.

La mise en place d'outils d'analyse d'images adaptés permet d'envisager des approches automatisées de recherche par le contenu, ou encore de navigation dans un document, lettrine par lettrine.



Figure 1. *Exemples de lettrines*

La sémantique d'une image de lettrine est décrite par un thesaurus issu de travaux historiens (Jimenes, 2008), qui respecte à la fois le standard TEI (Text Encoding Initiative) (Barnard *et al.*, 1995) - standard de représentation des textes - et la classification Iconclass (Waal, 1985) - système standard de classification des documents d'une collection par sujets. La description des lettrines y est décomposée en quatre couches, chaque couche fournit une information spécifique (voir Figure 2) :

Lettre (Letter) : positionnée au centre de l'image, la couche lettre caractérise en particulier la lettre qu'elle contient, sa couleur (noire ou blanche), l'alphabet (latin, grec, hébreu, ...) et la police utilisée (romain, gothique) ;

Motif (Pattern) : il est constitué des formes ornementales qui peuvent être décoratives ou figuratives ;

Fond (Background) : le fond peut être uniforme (noir ou blanc), hachuré ou criblé ;

Cadre (Frame) : enfin, le cadre correspond aux bords du tampon typographique. Il peut être composé de zéro, un ou deux traits.

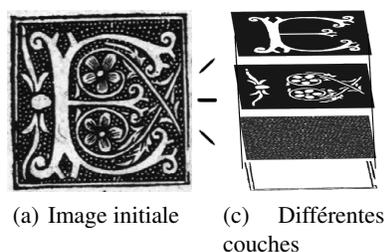


Figure 2. *Décomposition en couches des historiens*

Les images de lettrines sont des images particulièrement difficiles à traiter. En effet, ce sont des images qui se sont détériorées avec le temps - jaunissement du papier, pages abîmées, tâchées - et les techniques développées doivent être robustes à ces détériorations. Par ailleurs, il s'agit d'images binaires à base de traits sur lesquelles les approches texture classiques d'analyse d'images ne peuvent pas s'appliquer. C'est pourquoi une approche spécifique d'extraction d'information à partir d'images à base de traits a été proposée dans (Coustaty, 2011).

Tout d'abord, l'image est décomposée en plusieurs couches selon une décomposition décrite dans (Dubois *et al.*, 2008), résultat d'une série de projections (Hamidi *et al.*, 2010). La géométrie d'une image se retrouve alors principalement sur sa *couche forme*, alors que la *couche texture* permet d'analyser les traits qui la composent.

Zones de formes. Une méthode d'extraction adaptée à la couche forme, décrite dans (Coustaty *et al.*, 2011), permet d'en extraire certaines zones forme constitutives de l'image. En effet, la lettre, ou encore des éléments de scènes naturelles des lettrines à fond figuratif, correspondent à des formes de l'image. Ces zones de forme sont obtenues en deux étapes :

1) Les composantes connexes de la couche forme sont segmentées à l'aide d'une loi de Zipf (Pareti *et al.*, 2006), particulièrement adaptée par sa robustesse aux variations de niveaux de gris, et par son indépendance à la couleur des composantes.

2) Enfin, ces composantes sont extraites de l'image, formant ainsi des zones. Seules les zones dont l'aire est supérieure à 1% de l'image sont retenues, les autres zones étant considérées moins pertinentes.

Zones de traits. Un traitement de la couche texture a également été mis en place pour en extraire des zones constituées de traits similaires (Coustaty, 2011). Ces zones de traits sont obtenues en cinq étapes :



(a) Image d'origine (b) Couche forme (c) Couche texture

Figure 3. *Décomposition en couches d'une lettrine*

1) Un pré-traitement de binarisation est tout d'abord appliqué, suivi d'un débruitage pour ne conserver que les traits dans l'image binarisée.

2) Une squelettisation permet ensuite d'extraire le squelette des traits qui peuvent être d'épaisseurs variables.

3) Chaque trait est ensuite caractérisé à l'aide de caractéristiques prégnantes pour la vision humaine, à savoir sa longueur, son épaisseur, son orientation, son homogénéité. Une classification non supervisée à partir de ses caractéristiques permet ensuite de créer des classes de traits similaires.

4) Enfin, les traits similaires qui sont voisins dans l'image sont regroupés au sein d'une même zone de traits.

3. Ontologie des lettrines

Nous avons défini une représentation ontologique des images de lettrines en plusieurs étapes, chaque étape permettant l'intégration dans la base d'une description spécifique des images (cf Figure 4) :

Connaissance métier (historian knowledge) : la connaissance des historiens est décrite par un thesaurus que nous avons transcrit en une ontologie, i.e. en logique de description. Un fond de 4288 images de lettrines, annotées à la main par les historiens, est utilisé pour peupler la partie A-box de cette ontologie.

Zones forme et zones de traits des lettrines : nous avons défini une ontologie des zones d'une image de lettrine composée de zones formes extraites à partir de la couche forme, et de zones de traits extraites à partir de la couche texture. Cette ontologie est peuplée par 911 images de lettrines (lettrines à fond blanc, hachuré, noir, criblé et indéfini), décrites par 5588 zones forme et par 451711 zones de traits.

Combinaison des deux ontologies. Nous avons combiné l'ontologie métier et l'ontologie enrichie des zones en une seule *ontologie des lettrines*.

Complex Knowledge Database

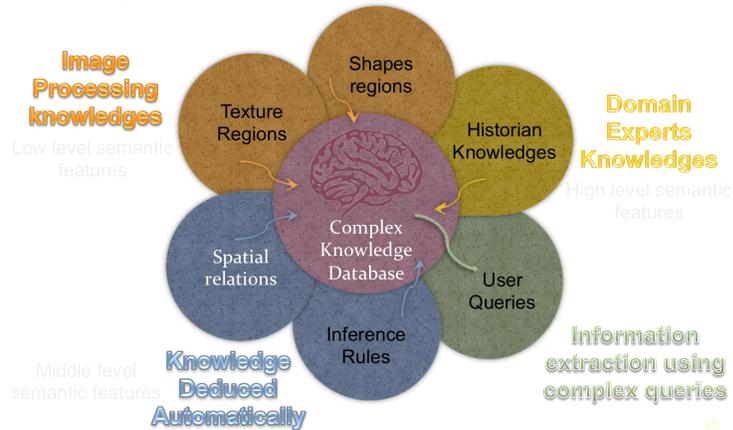


Figure 4. *Fleur ontologique*

Extraction des zones forme. Nous avons enrichi l'ontologie des zones avec un mécanisme de partitionnement de l'image et des relations spatiales, dans le but de pouvoir localiser chacune des zones forme selon qu'elle se situe au centre de l'image, ou sur un bord. Puis nous avons ajouté des descriptions logiques de propriétés permettant d'identifier et d'annoter certaines zones forme comme la lettre de la lettrine, ou encore une partie d'un personnage du fond de la lettrine

Extraction des zones de traits. En ce qui concerne les zones de traits, c'est l'ensemble des zones constitutives d'une image qui fournit une information pertinente. Nous avons ici résumé les caractéristiques de l'ensemble des zones de traits à l'aide de nouvelles propriétés au niveau de l'image elle-même, puis utiliser ces nouvelles propriétés pour déterminer si l'image est une lettrine à fond hachuré ou non.

Nous avons utilisé l'outil DLV¹, base de données déductive, pour construire et peupler ces ontologies. L'exécution en ligne de commandes basée sur la programmation logique déductive est simple à utiliser, DLV offre une facilité d'expression des requêtes, avec des temps de réponse relativement faibles.

1. <http://www.dlvsystem.com/dlvsystem/index.php>

3.1. *Ontologie métier des lettrines*

La T-box de l'ontologie métier est construite à partir du thesaurus déjà défini par les historiens. Un concept - vocabulaire de la logique de description - ou classe - vocabulaire des ontologies - est défini pour chaque couche sémantique d'une lettrine (*Letter, Background, Pattern, Frame*); ainsi que pour spécifier la lettre (*ColorLetter, TypeFont,*). La classe principale *Lettrine* est reliée à ces concepts par des propriétés :

- *Lettrine* **hasFrame** *Frame*
- *Lettrine* **hasBackground** *Background*
- *Lettrine* **hasPattern** *Pattern*
- *Lettrine* **hasLetter** *Letter*
 - *Letter* **hasIdentificationLetter** *IdentificationLetter*
 - *Letter* **hasColor** *ColorLetter*
 - *Letter* **hasAlphabet** *Alphabet*
 - *Letter* **hasTypeFont** *TypeFont*

3.2. *Ontologie des zones des lettrines*

Un concept *Image* permet de représenter chaque image de lettrine. Des informations globales relatives à l'image - unité de mesure, système de référencement utilisé - sont introduites par des propriétés :

- *Image* **hasImageMeasurementUnit** *MeasurementUnit*
- *Image* **hasImageReferenceSystem** *ImageReferenceSystem*

Les concepts *ROI* (Region Of Interest) et *ROISet* permettent d'associer un ensemble de zones à une lettrine :

- *Image* **hasROISet** *ROISet*
- *ROISet* **hasROI** *ROI*

Pour distinguer les zones forme des zones traits, nous avons créer deux sous-concepts du concept *ROI* :

- *ROI* **is-a** *ROIShape*
- *ROI* **is-a** *ROITexture*

Des caractéristiques communes à toutes les zones sont définies comme des propriétés du concept *ROI* :

- *ROI* **hasX** et *ROI* **hasY** : indique les coordonnées du centre de gravité de la zone, permettant ainsi de la situer dans l'image.
- *ROI* **hasAire** : indique l'aire de la zone, en nombre de pixels, utilisée pour trier

les zones selon leur taille.

Les zones forme sont caractérisées par trois propriétés supplémentaires :

– *ROIShape hasExcentricite* : indique l'excentricité Ecc de la zone, qui donne une indication de forme. L'excentricité est définie comme le rapport entre le rayon minimal r_m et le rayon maximal r_M de l'ellipse minimale englobant la zone (Pratt, 2007) :

$$Ecc = \frac{r_M - r_m}{r_M + r_m} \quad [1]$$

– *ROIShape hasGreyMean* et *ROI hasGreySTD* : indique la moyenne en niveaux de gris, ainsi que leur déviation, permettant d'estimer la couleur de la zone, ainsi que sa régularité.

– *ROIShape hasEuler* : indique le nombre d'Euler E_n de la zone, pour en estimer la compacité. Le nombre d'Euler introduit dans (Pratt, 2007), permet d'estimer le plus généralement la compacité d'un ensemble de zones. Pour une seule zone, il se réduit à $E_n = 1 - H$, où H représente le nombre de trous dans la zone.

Les zones de traits sont caractérisées par six propriétés supplémentaires :

– *ROITexture hasStrokesNumber* : indique le nombre de traits d'une zone.

– *ROITexture hasStrokesLenght* : indique la longueur moyenne des traits d'une zone.

– *ROITexture hasStrokesWirdh* : indique l'épaisseur moyenne des traits d'une zone.

– *ROITexture hasStrokesOrientation* : indique l'orientation moyenne des traits d'une zone.

– *ROITexture hasStrokesHomogeneity* : indique si les orientations des traits qui composent la région sont homogènes :

$$Hom = std(O_i) \quad [2]$$

où std représente la fonction *Ecart-Type* et O_i indique l'orientation du i_{eme} trait de la région traitée.

– *ROITexture hasStrokesCurvature* : indique la courbure moyenne des traits d'une zone. La courbure C_i permet d'indiquer si un trait est rectiligne ou courbe et est obtenue par un ratio entre l'orientation principale O_P et l'orientation secondaire O_S du trait :

$$C_i = \frac{O_P}{O_S} \quad [3]$$

3.3. Combinaison des deux ontologies

Nous avons ensuite combiné l'*ontologie métier* et l'*ontologie des zones enrichie* en une seule *ontologie des letrines*. En effet, ces deux ontologies portent chacune une

| | | |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 | 9 |

Figure 5. Partitionnement d'une image utilisée pour localiser les zones.

information pour une image de lettrine. Techniquement, la T-box de cette ontologie combinée s'obtient en unifiant le concept *Lettrine* de l'ontologie métier, et le concept *Image* de l'ontologie des zones.

3.4. Enrichissement de l'ontologie des zones forme

1) Nous avons mis en place un mécanisme de partitionnement d'une image de lettrine. Plus précisément, chaque image est divisée en neuf partitions (voir figure 5), une partition correspond au centre de l'image, 4 partitions correspondent aux différents bords, et les 4 dernières aux coins de l'image. Les partitions sont associées à l'image à l'aide d'un concept *Partition* et de la propriété suivante :

- *Image hasPartition Partition*

2) Nous avons ensuite enrichi l'ontologie des zones avec une ontologie standardisée des relations spatiales. L'information spatiale permet de préciser si chaque zone forme se positionne au centre de l'image, ou sur un bord. Les concepts *ROI* et *Partition* héritent de la classe abstraite *Surface* du standard SFS de l'OGC, les huit relations spatiales entre deux objets spatiaux définies par le standard RCC8 (Cohn *et al.*, 1997) (cf Figure 6) s'y définissent pas des propriétés. Ainsi, nous pouvons déterminer si une zone forme est située au centre de l'image à l'aide de la relation spatiale *contains* entre cette zone et la partition centrale :

- *ROI contains Part*

Il est ici à noter que les relations spatiales entre une zone forme et chacune des neuf partitions peuvent également se calculer lors de la phase d'extraction des zones, chaque relation entre une zone et une partition donnant lieu à une caractéristique booléenne.

3) Enfin, nous avons enrichi l'ontologie des lettrines avec de nouvelles propriétés qui intègrent expertise du domaine et expertise de traitement d'images. Il s'agit de propriétés calculées, définies par une formule ou une règle logique, permettant d'annoter sémantiquement certaines zones d'une lettrine extraites de la couche forme : propriétés *isLetter* et *isBody* (cf Figure 7) :

3.4.0.1. Propriété *isLetter* pour une zone forme :

ROIShape isLetter.

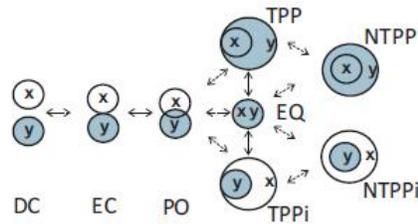


Figure 6. Relations spatiales de l'algèbre RCC-8

Cette propriété indique si une zone est identifiée comme la lettre d'une lettrine. Elle s'exprime en fonction de trois propriétés qui intègrent des caractéristiques des zones forme, mais également une information spatiale avec le partitionnement de l'image :

(isLetter) zone d'aire maximale
située au centre de l'image,
contenant peu de trous et contenant le centre de l'image de la lettrine dans son rectangle englobant.

a) zone située au centre de l'image : cette propriété se teste à l'aide des relations spatiales. Il s'agit de vérifier que la zone est contenue dans la partition centrale, et ne chevauche par les zones du bord.

b) Zone avec peu de trous : le nombre d'Euler, associé à chaque zone dans l'ontologie des zones, est ici utilisé pour tester cette propriété. Après quelques expérimentations, nous avons positionner le nombre d'Euler entre -2 et +2 pour caractériser une zone sans trous.

c) zone d'aire maximale : dans l'ontologie des zones, l'aire est indiquée pour chaque zone. Parmi toutes les zones vérifiant les deux propriétés précédentes, celle d'aire maximale est retenue pour être annotée par la propriété *isLetter*.

d) contenant le centre de l'image de la lettrine dans son rectangle englobant : Cette propriété permet de préciser que le centre de la lettrine doit appartenir au rectangle englobant de la zone choisie, avec une marge de 15 pixels.

3.4.0.2. Propriété *isBody* pour une zone forme :

ROIShape isBody.

Cette propriété indique qu'une zone est identifiée comme la partie d'un personnage du fond d'une lettrine. Elle s'exprime en fonction de cinq propriétés qui intègrent des caractéristiques des zones forme, une information spatiale avec le partitionnement, ainsi qu'une information de l'ontologie du domaine :

(isBody) zone d'une lettrine figurative
située au centre de l'image,
sans trous, de couleur gris-clair,
et qui n'est pas une lettre.

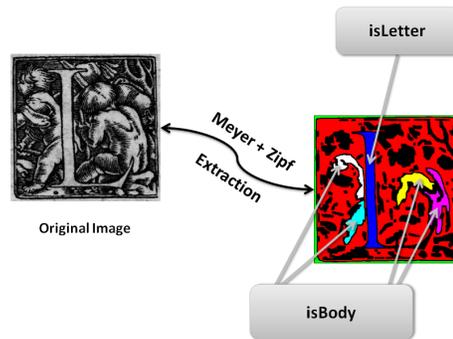


Figure 7. Extraction de formes d'une lettrine

a) zone d'une lettrine figurative : propriété de l'ontologie métier.

b) zone située au centre de l'image : comme pour la propriété *isLetter*, cette propriété se teste à l'aide des relations spatiales de l'ontologie des zones enrichie.

c) Zone avec peu de trous : comme pour la propriété *isLetter*, il s'agit de tester le nombre d'Euler associé à chaque zone dans l'ontologie des zones.

d) zone de couleur gris-clair : la caractéristique colorimétrique d'une zone, présente dans l'ontologie des zones, doit être supérieure à 90 pour indiquer une couleur gris-clair.

e) zone qui n'est pas une lettre : parmi toutes les zones vérifiant les propriétés précédentes, seules celles qui ne vérifient pas la propriété *isLetter* sont annotées.

3.5. Enrichissement de l'ontologie des zones de traits

1) En ce qui concerne les zones de traits, ce sont les caractéristiques de l'ensemble des zones de traits qui portent une information pertinente au niveau de l'image elle-même. Ainsi, une lettrine hachurée peut se définir par un ensemble de zones de traits correspondant à des traits hachurés, et dont la superficie est suffisamment grande pour couvrir une bonne partie de l'image. Nous avons enrichi la description d'une image avec la propriété suivante :

- Image *hasRatio* : indique la proportion de l'image couverte par les zones de traits :

$$Ratio = \frac{\sum_{\text{zone de traits } z} aire(z)}{aire \text{ image}}$$

2) Nous avons ensuite enrichi la description d'une image avec des propriétés décrivant l'ensemble des zones de traits qui la composent :

- Image *hasYYY-XXX* avec :

XXX : correspond à chacune des 6 propriétés spécifiques des zones de traits, à savoir **hasStrokesNumber**, **hasStrokesLenght**, **hasStrokesWirdh**, **hasStrokesOrientation** et **hasStrokesHomogeneity**.

YYY : correspond à la valeur minimale, maximale, moyenne, ou à l'écart-type (**Min**, **Max**, **Avg** et **Std**) de chacune des 6 caractéristiques de l'ensemble des zones de traits constitutives de l'image.

3) Enfin, nous avons enrichi l'ontologie des zones avec une propriété indiquant si une lettrine est constituée de suffisamment de zones de traits pour en déduire qu'elle a un fond hachuré :

3.5.0.3. Propriété *isHashed* pour une lettrine :

Image isHashed.

Cette propriété détermine si une lettrine possède un fond hachuré à partir des caractéristiques décrites ci-dessus qui résument celles de l'ensemble de ses zones de traits.

Cette propriété s'exprime seulement à partir de caractéristiques bas-niveau. C'est pourquoi nous avons pu utiliser une approche classique de classification supervisée (à savoir un arbre de décision C4.5) pour déterminer les seuils des différentes caractéristiques permettant de décider si la lettrine est hachurée ou non.

(isHashed) : *comparaison des valeurs minimale, maximale, moyenne, ou d'écart-type des caractéristiques des zones de traits par rapport à des seuil (7 cas) obtenus par apprentissage (cf table 1*

4. Résultats expérimentaux

Les règles ont été appliquées sur l'ensemble des 909 lettrines peuplant l'ontologie des zones (lettrines à fond blanc, hachuré, noir, criblé et indéfini), les images y étant décrites par 5588 zones forme et par 451711 zones de traits.

4.0.0.4. Règle pour la propriété *isLetter* :

La règle *isLetter* a pour objectif d'identifier la zone de forme correspondant à la lettre d'une image de lettrine. Cependant, cette règle ne garantit pas l'existence d'une zone annotée pour chaque lettrine. Après vérification manuelle, 816 zones de forme s'avèrent être correctement annotées, et 103 ne le sont pas, d'où un taux de reconnaissance de 89%, et un taux d'erreur de 11%.

Il serait ici possible de comparer la forme de la zone avec celle attendue de la lettre, information présente dans l'ontologie métier. Une telle extension permettrait de détecter les cas où la lettre se décompose en plusieurs zones, et ainsi réitérer la règle pour la propriété *isLetter* jusqu'à ce que toutes les zones de la lettre soient détectées.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-------------------|--------------|-------------|--------------|--------------|--------------|--------------|--------------|
| StLength seuil | Std > 27 | Std > 27 | Std > 27 | Std > 27 | Std > 27 | Std ≤ 27 | Std ≤ 27 |
| StNb seuil | Moy > 1 | Moy ≤ 1 | Moy ≤ 1 | Moy > 1 | Moy ≤ 1 | Moy > 1 | Std > 4 |
| StCurv seuil | Max ≤ 453 | | Max ≤ 400 | Max ≤ 453 | Max > 213 | Max > 213 | Max ≤ 213 |
| StWidth seuil | | | | Max ≤ 150 | Max ≤ 150 | | Avg ≤ 216 |
| StOrient seuil | | Moy ≤ 38 | Moy > 38 | | | | |
| StHom. seuil | | | Moy ≤ 824 | | Moy > 128 | | |
| X seuil | Max ≤ 574 | | Max > 238 | Max ≤ 57 | Max ≤ 5 | Max > 241 | Max > 241 |
| aire seuil | | Max ≤ 1 | Std ≤ 884 | | | | |
| ratio | ≤ 894 | ≤ 892 | | > 894 | > 894 | | |

Tableau 1. Seuils (7 cas) pour la règle *isHashed*

4.0.0.5. Règle pour la propriété *isBody* :

La règle *isBody* a pour objectif d'identifier les zones de forme des lettres à fond figuratif correspondant à des parties de personnages. Cette règle sélectionne 517 images sur 909 possédant un fond figuratif, puis annoté 927 zones de forme dans 397 images, d'où un taux d'images figuratives possédant une ou plusieurs zones forme annotées de 76,6%. Le nombre de zones annotées par image varie de 1 à 10 selon la répartition suivante :

| Nb zones annotées | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-------------------|--------|-------|-----|----|----|----|---|---|---|---|----|
| Nb images | 120 | 146 | 106 | 69 | 45 | 15 | 5 | 6 | 2 | 1 | 1 |
| | 23,21% | 76,6% | | | | | | | | | |

Cependant, certaines zones annotées ne correspondent pas à des parties de personnage. La figure 8 en fournit un exemple. On peut y observer que la zone jaune correspondant à une partie d'un personnage n'est pas annotée, son nombre d'Euler est trop grand. En revanche, la zone blanche est annotée alors qu'elle ne correspond pas à une partie d'un personnage. L'utilisation d'un partitionnement plus complexe de l'image, ou encore d'autres descripteurs statistiques des zones, pourrait permettre d'annoter correctement les zones de ce type.

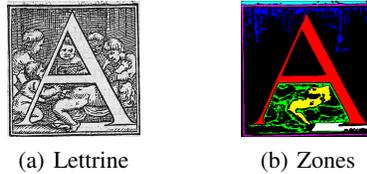


Figure 8. Exemple d'image (à gauche) où des zones extraites (à droite) ne sont pas correctement identifiées comme partie de personnage par *isBody*

4.0.0.6. Règle pour la propriété *isHashed* :

La règle *isHashed* permet d'identifier une lettrine comme possédant un fond hachuré à partir des zones de traits qui la composent. L'ensemble des 909 lettrines peuplant l'ontologie des zones est composé de 140 lettrines hachurées. La règle sélectionne 126 images. Après une vérification automatique sur l'ontologie métier (des historiens), 123 sont effectivement hachurées. D'où un rappel de $\frac{123}{140} = 87,8\%$, et une précision de $\frac{123}{126} = 97,8\%$.

5. Conclusion et perspectives

Ces travaux autour des lettrines sont des travaux récents, plusieurs perspectives sont d'ores et déjà envisagées, voire amorcées. L'amélioration des règles existantes est une première perspective : pour la règle *isLetter*, possibilité de comparer la zone lettre avec la forme attendue ; pour la règle *isBody*, mise en place d'un partitionnement plus sophistiqué, et utilisation de nouveaux descripteurs des zones. Sur le même modèle que la règle *isHashed*, une annotation des lettrines selon qu'elles soient lettrines à fond criblé, blanc, noir ou indéfini semble envisageable.

Notons également la possibilité de concevoir des règles intégrant des connaissances issues à la fois des couches zones et des couches de traits, ce qui laisse envisager des possibilités pertinentes d'extraction de nouvelles connaissances à partir d'informations bas-niveau hétérogènes.

Enfin, l'intégration d'une vérité terrain des zones d'une lettrine nous semble indispensable pour pouvoir vérifier automatiquement - et non manuellement - les règles mises en place, et ainsi les appliquer sur une plus grande quantité de zones et de lettrines. Des travaux en ce sens sont en cours.

6. Bibliographie

Bannour H., Hudelot C., « Combinaison d'information visuelle, conceptuelle, et contextuelle pour la construction automatique de hiérarchies sémantiques adaptées à l'annotation d'images », *Actes de la conférence RFIA 2012 - Reconnaissance des Formes et Intelligence Artificielle*, p. 460 -469, juin, 2012.

- Barnard D. T., Burnard L., Gaspart J.-P., Price L. A., Sperberg-McQueen C. M., Varile G. B., « Hierarchical Encoding of Text : Technical Problems and SGML Solutions », *Computers and the Humanities*, vol. 29, n° 3, p. 211-231, 1995. <http://www.tei-c.org/index.xml>.
- Cohn A., Bennet B., Gooday J., Gotts N., « Representing and reasoning with qualitative spatial relations about regions », *Spatial and temporal reasoning*, Kluwer, p. 97-134, 1997.
- Coustaty M., Contribution à l'analyse complexe de documents anciens. Application aux lettrines, PhD thesis, Université de La Rochelle, La Rochelle, France, 2011.
- Coustaty M., Pareti R., Vincent N., Ogier J.-M., « Towards historical document indexing : extraction of drop cap letters », *IJDAR*, vol. 14, n° 3, p. 243-254, 2011.
- Dubois S., Lugiez M., Péteri R., Ménard M., « Adding a Noise Component To A Color Decomposition Model For Improving Color Texture Extraction », *proc. of the 4th European Conference on Colour in Graphics, Imaging, and Vision (CGIV'08)*, p. 394-398, 2008.
- Hamidi A. E., Ménard M., Lugiez M., Ghannam C., « Weighted and extended total variation for image restoration and decomposition », *Pattern Recognition*, vol. 43, n° 4, p. 1564 - 1576, 2010.
- Hanbury A., « A survey of methods for image annotation », *Journal of Visual Languages and Computing*, vol. 19, p. 617-627, October, 2008.
- Horrocks I., Patel-Schneider P., Harmelen F., « An introduction to description logics », *The description logic handbook : theory, implementation, and applications*, vol. , p. 1-40, 2003.
- Jimenes R., Les Bibliothèques Virtuelles Humanistes et l'étude du matériel typographique, Technical report, Centre d'Etude Supérieur de la Renaissance, 2008.
- Liu Y., Zhang D., Lu G., Ma W.-Y., « A survey of content-based image retrieval with high-level semantics », *Pattern Recognition*, vol. 40, n° 1, p. 262 - 282, 2007.
- Malki J., Mefteh W., Bouju A., « Une approche ontologique pour la modélisation et le raisonnement sur les trajectoires. Prise en compte des règles métiers, spatiales et temporelles », *3ème édition des journées francophones sur les ontologies (JFO'09)*, 2009.
- Neumann B., Möller R., « On scene interpretation with description logics », *Image Vision Comput.*, vol. 26, p. 82-101, January, 2008.
- Pareti R., Vincent N., « Ancient Initial Letters Indexing », *Proc. of the 18th International Conference on Pattern Recognition (ICPR'08)*, IEEE Computer Society, Hong Kong, China, p. 756-759, 2006.
- Pratt W., *Digital Image Processing : PIKS Scientific Inside*, 4 edn, Wiley-Interscience, February, 2007.
- Sharma J., Flewelling D.-M., Egenhofer M.-J., « A qualitative spatial reasoner », *Proc. of the 6th International Symposium on Spatial Data Handling (SDH'94)*, p. 665-681, 1994.
- Tousch A.-M., Herbin S., Audibert J.-Y., « Semantic hierarchies for image annotation : A survey », *Pattern Recognition*, vol. 45, p. 333-345, 2012.
- Uschold M., King M., « Ontologies : principles, methods and applications », *Knowledge Engineering Review*, vol. 11, n° 2, p. 93-155, 1996.
- Waal H., *ICONCLASS, an Iconographic Classification System*, Completed and edited by L. D. Couprie et al, 1985. <http://www.iconclass.nl/>.
- Zhang D., Islam M. M., Lu G., « A review on automatic image annotation techniques », *Pattern Recognition*, vol. 45, p. 346-362, 2012.