2009 10th International Conference on Document Analysis and Recognition

ORNAMENTAL LETTERS IMAGE CLASSIFICATION USING LOCAL DISSIMILARITY MAPS

J. Landré
Univ. of Reims-Champagne-Ardenne - CReSTIC
IUT, 9 rue de Québec, BP 396
10 026 Troyes cedex, France
jerome.landre@univ-reims.fr

F. Morain-Nicolier
Univ. of Reims-Champagne-Ardenne - CReSTIC
IUT, 9 rue de Québec, BP 396
10 026 Troyes cedex, France
frederic.nicolier@univ-reims.fr

S. Ruan Univ. of Reims-Champagne-Ardenne - CReSTIC
IUT, 9 rue de Québec, BP 396
10 026 Troyes cedex, France
su.ruan@univ-reims.fr

Abstract

This article describes a new method for ancient books ornamental letters segmentation and recognition. The purpose of our work is to automatically determine the letter represented in an ornamental letter image. Our process is divided in two parts: a segmentation step of the ornamental letter is followed by a recognition step. The segmentation process uses multiresolution analysis to filter background decorations followed by a binarisation step and a morphologic reconstruction of the expected letter. The recognition process use the previously obtained reconstruction and compares it with capital letters images used as a dictionary of shapes with the Local Dissimilarity Map (LDM) distance.

1 Introduction

Nowadays, cultural heritage preservation [8] is a very important task for every nation in order to save knowledge and literature for future generations. Ancient books from the hand-press period often have beautiful page layouts because at this time, books were precious and rare objects. Ornamental letters (or decorated initials) were used to start new chapters to illustrate the content of the chapter. On the web, many databases of ornamental letter images are available for online search [5]. Efficient tools are needed to browse these databases.

Figure 1 shows several ornamental letters images out of the BVH [6] database.

Figure 1. Examples of ornamental letters from the BVH database.

Our article is divided in seven parts. Section 2 describes related work in the field. The local dissimilarity map tool is presented in section 3. In section 4, a description of our method is proposed. Detailed information on our technique is developed in section 5. Experimental results are proposed in section 6. At last, section 7 gives a conclusion and future directions concerning our work.

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.226

2 RELATED WORK

With the growing of interest in cultural heritage preservation in the 1990s, numerous works of digitization of historical collections have been carried out. Nowadays, large databases of ornaments of the hand-press period are available on the Internet [5] and will go on growing in the future. These ornaments are extracted from the whole digitized pages, using full automatic [4] or user-driven [10] segmentation methods, or recorded independently.

Previous works of the Calypod research group [5] have emphasized that the key problem is now to make these images available for research to history specialists and for general users (like artists, designers, publishers, printers, students, etc.). All of these individuals have different needs which require varied and sophisticated means of searching and accessing information.

An overview of existing methods for image segmentation and recognition is given by Garain at al. [3]. The main problem of ornamental letter images recognition is to separate the image in two regions: the object (the letter) and decorations (background). The problem is to find a binary representation of the image where ideally the object is alone and decorations are removed.

Mathematical morphology is a very good theory offering interesting tools to work on binary images. Our method uses wavelets and mathematical morphology operators to obtain a segmented image of the ornamental letter. The obtained segmented images are compared with a dictionary of shapes using the Local Dissimilarity Map (LDM) distance map to determine the letter represented on the image. The LDM algorithm was proposed by Baudrier et al. [1] and gives good results on "Renaissance" wooden stamps. The LDM method uses the Hausdorff distance to build a map of local dissimilarities between two images in order to compare them.

3 LOCAL DISSIMILARITY MAP

The local dissimilarity (or distance) map LDMap [1] is an image processing transform which computes dissimilarities between two images. It is based on the windowed Hausdorff distance and the distance transform of the two images. It is a growing window algorithm which differs from the simple bit to bit comparison of two images.

For two non-empty finite images $F \in \mathbb{R}^2$ and $G \in \mathbb{R}^2$, for each point $x \in \mathbb{R}^2$, the local dissimilarity map is defined as: $LDMap(x) = \mid 1_F - 1_G \mid \times max(d(x,F), d(x,G))$ where d is the Hausdorff distance, 1_F is equal to 1 when $x \in F$ and to 0 elsewhere.

An important property of the LDMap is its symmetry: $LDMap(A,B) = LDMap(B,A)$ due to the definition. But an interesting modification allows to build a directed LDMap. By definition, the directed LDMap (DLDMap) is: $DLDMap(x) = \mid 1_F - 1_G \mid \times d(x,G)$.

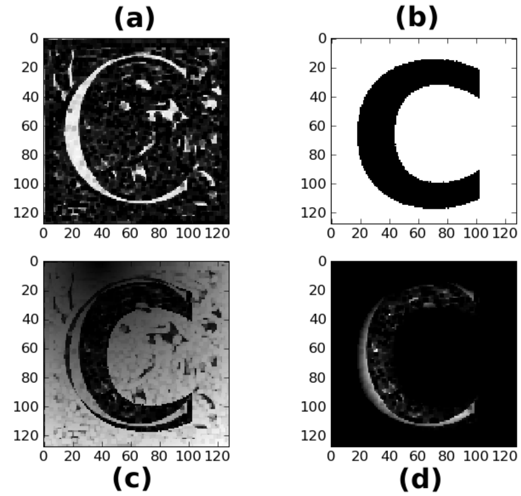

Figure 2. Examples of LDM and DLMD. Image C (a), synthetic image C' (b), LDM(C,C') (c) and DLDM(C,C') (d).

On Figure 2, the local dissimilarity map and the directed local dissimilarity map are computed between an ornamental letter image and a synthetic image. The directed LDM between C (a) and C' (b) offers to consider pixels only in $C \cap C'$ (c). DLDM (d) avoid the computation of the LDM outside the synthetic letter C'.

4 OVERVIEW OF SEGMENTATION

The main principle of our method is to clean the image of its ornamental decorations to obtain a good representation of the letter. Figure 3 illustrates the principle of our segmentation method.

At first, a multiresolution analysis of the ornamental letter image (a) is performed at scale $s = 1$, leading to a multiresolution representation (b). The approximation image (c) is extracted from (b) by reconstruction with no details at scale $s = 0$. The image (c) is binarized to get binary image (d) in which the biggest region with connected pixels is computed and represented in a new image (e). Then, binary reconstruction is used between images (a) and (e) to get the final image (f). Each step of this method is described in details in the next section.

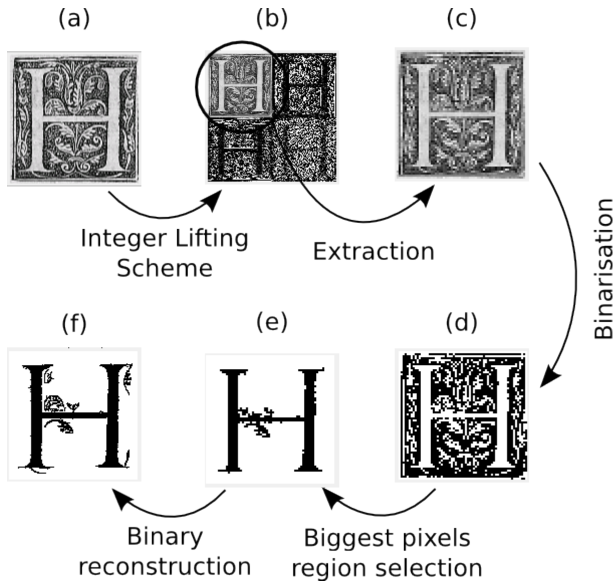

Figure 3. Overview of our segmentation method.

5 DETAILED SEGMENTATION

5.1 Multiresolution analysis

Our method uses an integer lifting scheme algorithm to decompose the original image (defined at scale $s = 0$) at multiple scales. This method was proposed by Calderbank at al. [2]. The Haar wavelets S-transform was used in our method at scale $s = 1$ because it is easy to compute and sufficient for cutting unimportant details of the background of the ornamental letter image.

After this lifting scheme decomposition, details are set to zero value and the approximation of the obtained image is used alone for reconstruction at scale $s = 0$ (original size of the ornamental letter image). An example of this process is proposed on Figure 4.

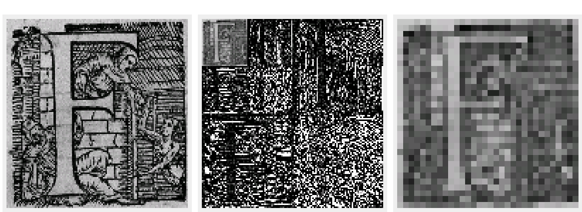

Figure 4. An original image, its integer lifting scheme decomposition at scale $s = 2$ and the reconstructed image without details at scale $s = 0$.

5.2 Binarization

In this step, an adaptative binarization process is applied to the aproximation image obtained in the previous step. Our binarization technique is based on the Otsu [9] method. Otsu's method consist in minimizing the inter-class variance of the image histogram. However, because a lot of images were not well binarized using Otsu's optimal threshold t, an adaptative threshold is defined.

A new threshold T is computed as $T = \alpha(M - t) + t$, where $0 \leq \alpha \leq 1$ is a shifting coefficient for the threshold, M is the maximum gray level in the image histogram. If $\alpha = 0$, $T = t$ Otsu's optimal threshold is kept unchanged and if $\alpha = 1$, $T = M$ and no pixel is left in the resulting binarization process.

α was experimentally fixed to the value $\alpha = 0.5$. Results on ornamental images are shown on Figure 5.

Figure 5. Examples of Otsu's adaptive thresholding for ornamental letters. From left to right: original image, Otsu optimal thresholding ($\alpha = 0$) and our method ($\alpha = 0.5$).

5.3 Biggest region detection

The computed binary image contains letter and decoration pixels. In a letter, pixels composing the letter are supposed to be connected together to form the shape of the letter. In this step, connected pixels (in 4-connectivity) of the image are grouped into regions. From all the connected pixels region, the biggest in size (number of pixels) is chosen and supposed to represent the object. The found region is defined as a mask for further processing.

5.4 Binary reconstruction

Reconstruction by dilation [12] (chapter 6, Geodesic transformations) allows reconstruction of all objects of an image X marked by an image Y. This a very powerful morphological operation that reconstruct (retains) connected particles in an image (called mask) based on markers present in another image (called seed). Morphological reconstruction consists in dilating the seeds inside the mask (so particles that do not have seeds are not reconstructed).

It is thus needed to have a mask image and a seed image. The mask image will be the binarized image obtained from 5.2. The seed image needs to be idealy an image containing only pixels from the letter. As we make the hypothesis that the initial spatial information is mainly coarse, the decoration pixels can be removed from the binarized image thanks to an erosion operator.

The final filtering algorithm is an erosion, followed by a binary reconstruction of the binarized image based on the eroded image as the seed. An application of this filtering is given in Figure 6.

Figure 6. Example of the binary reconstruction operator. From left to right: original image, binary mask image, seed (obtained by erosion of the binary mask image) and the obtained reconstruction (used for image recognition).

6 EXPERIMENTAL RESULTS

In order to test the recognition performance of our method, a Java plugin running under the ImageJ [11] image processing application was developed. the main part of our process was developed using Sage [7]. To demonstrate the efficiency of our method, the segmentation was performed on a set of images and was compared to a dictionnary of shapes using the DLDM and using a binary bit-to-bit comparison.

Our working database contains 823 ornamental letter images. Several examples of this set are shown on Figure 7.

A subset of 60 images representing capital letters (only C, H, M, O, Q and S because of their interesting shapes) from several families was used to test our method. Each image of this subset was segmented using our algorithm. After that, each image was compared to a dictionary of shapes

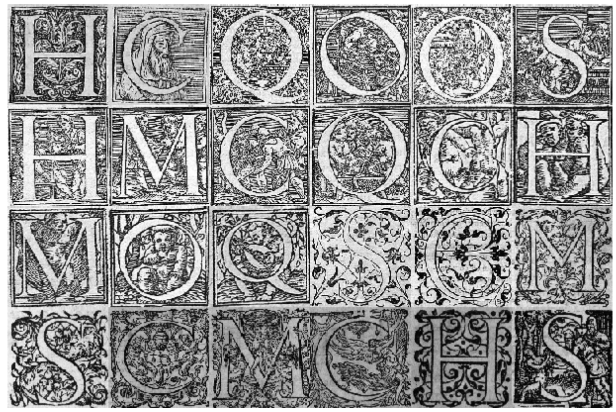

Figure 7. Examples of images from our ornamental letter images database.

(only letters C, H, M, O, Q and S) using two methods: binary AND distance and LDM distance. The dictionary of shapes is given on Figure 8.

Figure 8. Our reduced dictionary of shapes with C, H, M, O, Q and S capital letters.

Results are shown on table 9. For the 60 images database, the binary AND distance is not sufficient to match dictionary elements. Several letters were misclassified due to their bad segmentation. But this rate increases to 89 % if only good segmented letters are took in account for statistical results.

On several families of images, segmentation was not good because of the presence of big regions of background textures. These images did not respect our model in which the letter was the biggest object made of connected pixels in the images. Several examples of bad results are given on Figure 10.

Another problem of this approach is that several capital letters have very close shapes. For instance, a lot of misclassification were between the M letter and the H shape (and the opposite), also between O, C and Q letters and shapes which are very close to each others.

For the 823 images collection, the success rate is about 62% which is less than the 60 images case. Of course, because there are many letters in the database, the success rate decreases. Many images were misclassified, but errors are still of the form “O” for “Q”, “M” for “H”, … Visually similar images are the main error type for the 823 images database.

There are many images for which the classification process took place with success. Several examples of good segmentation leading to good recognition are shown on Figure 11.

	60 images	823 images
Binary AND distance	68.2 %	31.1%
Local Dissimilarity Map	89.4 %	62.7%

Figure 9. Results of segmentation and recognition.

Figure 10. Bad segmentation (and recognition) examples.

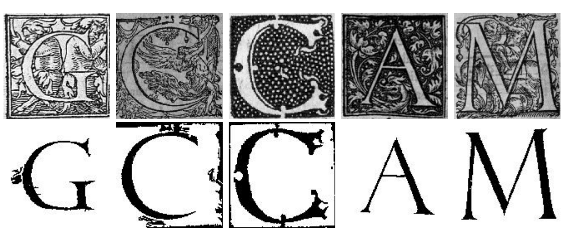

Figure 11. Good segmentation (and recognition) examples.

7 CONCLUSION

In this article, a new method for ornamental letters segmentation is presented. It is applied to ornamental letters gray level images recognition using a dictionnary of shapes and a distance defined by the Local Dissimilarity Map. This approach is new and offers good performances in terms of recognition based on a dictionary of shapes. Several images are not well segmented due to the determination of the object as the biggest connected pixels region.

LDM offers better results than a simple bitwise AND distance between ornamental letter images and dictionary elements. It leads to a good segmentation and recognition rate without the use of learning classification methods. However there are still problems with certain families of images due to the presence of big textured area.

Future research directions will concern the enhancement of the segmentation process with difficult ornamental letter families (i.e. with many textured information) by considering not only the biggest family of connected pixels but the three biggest families (for instance) and choose the one representing the object to be the largest in terms of spatial repartition in the image.

Another approach would be to use neural networks or SVM (Support Vector Machines) learning methods for better segmentation and classification results. This approach could lead to better results.

References

[1] E. Baudrier, F. Nicolier, G. Millon, and S. Ruan. Binary-image comparison with local-dissimilarity quantification. *Pattern Recognition*, 41(5):1461–1478, 2008.

[2] R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo. Wavelet transforms that map integers to integers. *Applied and Computational Harmonic Analysis (ACHA)*, 5(3):332–369, 1998.

[3] U. Garain, T. Paquet, and L. Heutte. On foreground-background separation in low quality document image. *International Journal on Document Analysis and Recognition (IJDAR)*, 8(1):47–63, 2006.

[4] N. Journet, J. Ramel, R. Mullot, and V. Eglin. A proposition of retrieval tools for historical document images librairies. In *International Conference on Document Analysis and Recognition (ICDAR)*, volume 2, pages 1053–1057, 2007.

[5] Collaborative work. Calypod. Working group on ancient books, http://calypod.free.fr.

[6] M.-L. Demonet et al. Bibliothèques virtuelles humanistes. Centre d'Études Supérieures de la Renaissance, CNRS UMR 6576, http://www.bvh.univ-tours.fr.

[7] Many contributors. Sage: Open source mathematics software. http://www.sagemath.org.

[8] J.-M. Ogier and K. Tombre. Madonne: Document image analysis techniques for cultural heritage documents. In *International Conference on Digital Cultural Heritage*, 2006.

[9] N. Otsu. A threshold selection method from grey scale histogram. *IEEE Trans. on Syst. Man and Cyber.*, 1:pp. 62–66, 1979.

[10] J. Ramel, S. Leriche, M.-L. Demonet, and S. Busson. User-driven page layout analysis of historical printed books. *International Journal on Document Analysis and Recognition (IJDAR)*, 9(2–4):pp. 243–267, 2007.

[11] W. Rasband. *ImageJ*. U. S. National Institutes of Health, Bethesda, Maryland, USA, 1997-2008. http://rsb.info.nih.gov/ij/.

[12] P. Soille. *Morphological image analysis: principles and applications*. Springer, 2003.