AUTOMATIC SEARCH AND DELIMITATION OF FRONTISPIECES IN ANCIENT SCORES

Cristian Segura [1], *Isabel Barbancho* [2], *Lorenzo J. Tardón* [3], *Ana M. Barbancho* [4]

Departamento de Ingeniería de Comunicaciones, Universidad de Málaga
Campus de Teatinos s/n, 29071, Málaga, Spain
[1] cseguraguerrero@gmail.com, [2] ibp@ic.uma.es, [3] lorenzo@ic.uma.es [4] abp@ic.uma.es

ABSTRACT

In this paper, an automatic system to locate frontispieces in ancient scores is presented. The system is able to identify images containing two pages and separate them. Then, the algorithm performs a correction of the rotation. The next step is to delimit the area of interest which will be used along the next steps. After that, a mask containing the shape of staves is calculated to identify the areas with information different from the staff related musical information. With this, an indetermined amount of regions are extracted. Between these candidates, some restrictions are applied to obtain the final candidates to frontispieces. From the final candidates, the possible frontispieces are selected. The algorithm obtains good results and constitutes an important piece in a global OMR system.

1. INTRODUCTION

The aim of Optical Music Recognition (OMR) is to provide a computer the ability to get music information from scanned scores. A particular case of OMR focuses on ancient scores, which present several additional difficulties compared with classic OMR [1], [2]. Some of these difficulties are due to the quality of the conservation of the original scores (marks and blots hiding the original symbols, heterogeneous illumination, variable color of the ink inside the same score, presence of fungi or mold, irregular leveling of pages, etc.) [3]. Also the digitalization process itself may introduce more degradation to the digital image.

In this context, the goal of the OMR system presented, in this paper, is to extract all the information available in the score about idenfitication of the author, style, origin, age, etc. The most part of this information can be extracted from the analysis of the frontispieces (decorative illustrations and fonts) that can be found in ancient scores. In the literature, other OMR systems focused on ancient scores [4],[5],[6] can be found but none of them extract automatically the frontispieces in the scores. To that end, the system presented in this paper will identify the frontispieces in order to analyze them to extract information that will be compared with a database. This database includes different ancient scores preserved in the Archivo de la Catedral de Málaga (ACM), with scores from different authors.

In this paper, the method implemented to automatically detect and extract the frontispieces from the scores is presented. It will be able to work with a large amount of different books from different authors and ages and the scanned images can contain one or two pages. In section 2, the algorithm is presented including the different steps, like preprocessing of the image, generation of the staves mask, searching for frontispieces candidates, and the selection over the candidates. In section 3, an overview of the algorithm performance is given. Finally, in section 4, some conclusions are drawn.

(a)

(b)

Figure 1: Examples of the type of images to work with. (a) One page image. (b) Two pages image.

2. ALGORITHM FOR AUTOMATIC SEARCH AND DELIMITATION OF FRONTISPIECES

In this section, all the stages of the algorithm will be presented: preprocessing of the image, generation of the staves mask, search of potential frontispieces and, finally, selection of the frontispieces among all the candidates.

 ISSN 2076-1465

Figure 2: (a) Portion of the image employed to calculate the rotation of image shown in Fig. 1(a). (b) Image after the rotation correction

Figure 3: (a) Image after delimitation of Area of Interest. (b) Image after binarization.

2.1 Image Preprocessing

The starting point of the algorithm is a RGB image as the ones shown in Fig. 1. The steps followed at the preprocessing stage are similar to those in [8]:

- *Correction of Rotation.* To estimate the global angle of rotation of the image, the central part of the image will be used in order to reduce computation time. Let $[X_{rows}, Y_{columns}]$ represent the dimensions of the image, then the portion of the image with coordinates (x, y) such that $x \in [0.4X_{rows}, 0.6X_{rows}]$ and $y \in [0.4Y_{columns}, 0.6Y_{columns}]$ define the portion of the image employed to estimate the rotation angle. The linear Hough transform is employed to estimate the angle of rotation [7]. In Fig. 2(a), the portion employed to calculate the rotation of image shown in Fig. 1, together with the image after the correction of the rotation is presented.
- *Delimitation of Area of Interest.* Now, the useful parts of the images should be selected and the borders of the page and also the parts of the page that doesn't contain any information should be removed. In [8], the Area of Interest is extracted in three different ways: the first method uses an eroded version of the Sobel edges of the image [9]. The second method uses the horizontal detail of the LeGall 3/5 Wavelet transform [10]. The last method is a modification of the first one and combines the eroded version of the Sobel edges with the eroded binary image. In this paper, the first algorithm for Automatic Delimitation of Area of Interest described in [8] is employed. An example of its performance is shown in Fig. 3(a). This algorithm also identifies images with two pages and divides them when needed.
- *Conversion to grayscale.* Due to the high heterogeneity of the scores database, the color information isn't considered useful in our system. Then all the process is performed using grayscale or binary images. The RGB is converted to grayscale using [11]:

$$I(grayscale) = 0.30R + 0.59G + 0.11B \qquad (1)$$

  where R, G and B are the color coordinates of each pixel in the RGB space.
- *Contrast-limited adaptive histogram equalization.* This equalization enhances the contrast of the image by transforming the values in the intensity image I by operating on small data regions (tiles). Each tile's contrast is enhanced, so that the histogram of the output region approximately matches the specified histogram.
- *Illumination compensation.* The image is filtered using a two dimensional Gaussian filter to simulate the background illumination of the page. The filter used is a low-pass Gaussian filter of size $[0.1X_{rows}, 0.1Y_{columns}]$ (the size is fixed heuristicly to 10% of the image size to assure it gets enough background information to obtain a good estimation) and standard deviation $0.05X_{rows}$ (the 5% of the image's height, considering the image has more rows than columns). The aim is to obtain an estimation $P(x,y)$ of the ilumination $L(x,y)$ to obtain a new image $C(x,y)$ as show in eq. 2 where the image corrected $C(x,y)$ shows a reflectance $R(x,y)$ cleaned for its heterogeneous lighting $L(x,y)$, by its compensated lighting $P(x,y)$.

$$C(x,y) = \frac{I(x,y)}{P(x,y)} = \frac{R(x,y)L(x,y)}{P(x,y)} \approx R(x,y) \qquad (2)$$

- *Binarization.* The image is binarized employing Otsu method [12].

2.2 Removal of staff regions

At this stage, the goal is to generate a mask that will allow the removal of all the staff regions in the image, so that only the other elements remain in the image: frontispieces, lyrics and any other elements outside the staves. The procedure employed performs the following steps:

- *Erosion of the binary image.* The binary image is eroded with a rectangular structure to remove all the elements except the staff lines. Due to the irregularity of the staff lines along all the books, since they were handwritten or because of irregularities on the pages, a large rectangular structure can't be used. Instead, the structure dimensions are heuristicly obtained: a length of 4% of the image number of columns and only one pixel of height due to the thickness of the staff lines and their irregularity.
- *Dilation of the staff lines.* Once the staff lines are isolated, the next step is to dilate those lines to transform

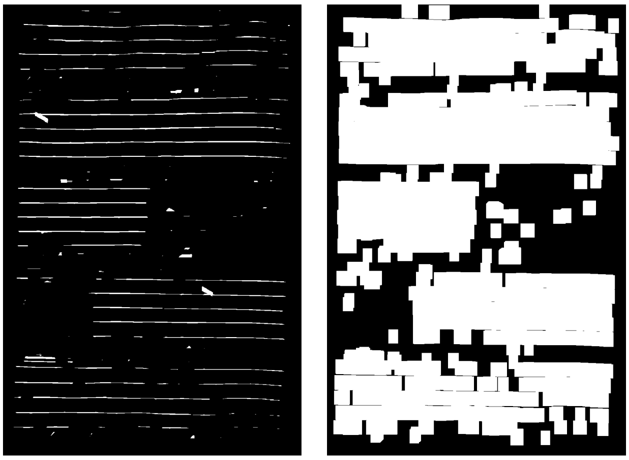


Figure 4: (a) Erosion with a rectangular structure of image shown in Fig. 3(b). (b) Dilation of the staff lines.

them into blocks, so one block for each staff is obtained. To that end, we need to fill the spaces between the 5 lines of the staff. To get a structure able to perform that task in all the scores of the available database, an estimation for the number of staves in the image has to be performed. The database has books with different number of staves per page (from 5 to 12 staves per page), so the mean of staves for page is calculated and a result of approximately 8 staves per page is obtained. Thus, the dilation structure will be a rectangle with dimensions $[X_{rows}/(8staves \times 4holes), Y_{columns}/(8staves \times 4holes)]$. This mask is also used to calculate the mean stave height that will be used lately to discriminate between all the possible candidates.

- *Application of the mask.* Once the mask is available, it is applied to the binary image, so that only the elements out of the staves remain in the new image.

2.3 Search for Candidates to Frontispieces

Once the staves are removed, the candidates to frontispieces are searched through the next steps:

- *Search of division points in row histogram.* The image obtained at the end of the stages descripted in the previous section is employed to calculate its row histogram. After that, a lowpass filter is applied to smooth the high variation in the histogram. A threshold is defined to be $0.25\mu_n$, with μ_n being the histogram mean. Then, all the points under the threshold are set to zero.
- *Selection of horizontal sections.* With the thresholded histogram, all the separate blocks with non-zero values are extracted to cells (a kind of sepparated groups). Let be μ_i the mean value of cell i, max_i its max value and $\mu[\ldots]$ the mean value of what is inside. Among all j cells, only the ones satisfying max_i to be over $0.25\mu[max_{i=1toj}]$ and μ_i to be over $0.25\mu[\mu_{i=1toj}]$ are selected. The selected horizontal sections of Fig. 3(a) are shown in Fig. 6.
- *Obtaining the column histogram of horizontal sections.* Once the horizontal sections of the image are extracted, the column histogram of each section is calculated and a lowpass filter is applied but, this time, no threshold is applied. The reason for that is the nature of frontispieces,

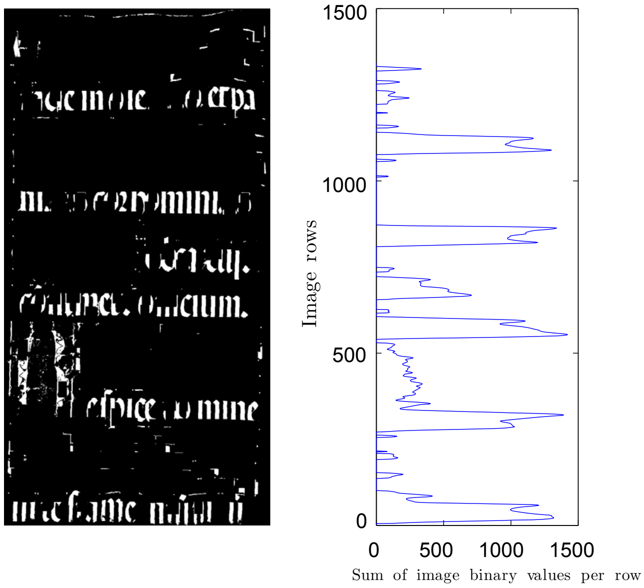


Figure 5: (a) Image after application of staves mask. (b) Row histogram.

that will usually have very thin parts that could be easily removed from the histogram using a threshold; this could break the frontispiece in different parts or remove it completely in the next steps.

- *Selection of vertical sections.* With the thresholded colum histograms of every horizontal sections, the same selection process used to select horizontal sections is applied to each column histogram.
- *Extraction of first candidates.* Once the horizontal and vertical selections are made, the image is cut, generating an indetermined amount of candidates.

2.4 Selection of Frontispieces between Candidates

Among all the candidates to frontispieces, a lot of them will be just little parts of lyrics in the score, so a restriction in the size of the candidates is applied. Only the candidates that fulfill the next restrictions will be taking into account. The height of the candidate needs to be at least the 90% of the average height of staves (calculated with the staves mask) and its width needs to be at least the 10% of the image width (this restriction assures that any thin part of the image that went through the histograms phase will be removed). The results can be observed in Fig. 8 and Fig. 9, being Fig. 8(b), Fig. 9(b), Fig. 9(c) and Fig. 9(d) the selected candidates.

3. RESULTS

In order to evaluate the performance of the developed algorithm, the following data has been obtained:

- Probability of detection off all frontispieces in the score.
- Probability of detection of 75% of the frontispieces in the score.
- Probability of detection of 66% of the frontispieces in the score.
- Probability of detection of 50% of the frontispieces in the score.
- Probability of no detection of frontispieces in the score.

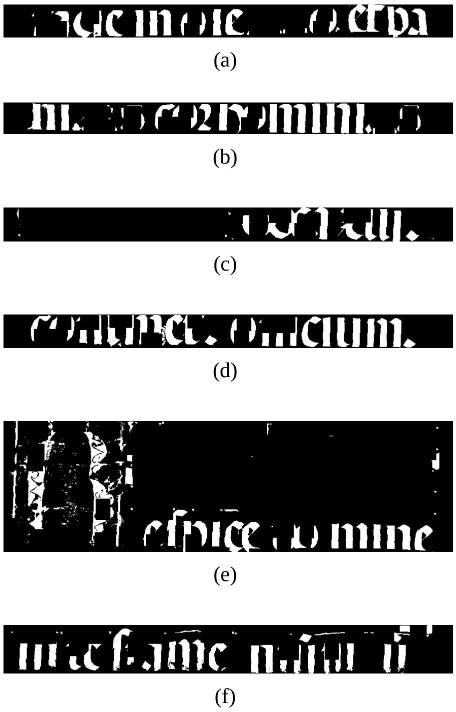


Figure 6: Extracted horizontal sections of the image.

- Probability of any amount of false positives.
- Probability of detection between 1 and 3 false positives.
- Probability of detection between 4 and 6 false positives.
- Probability of detection between 7 and 9 false positives.

To calculate these statistics, images from 15 different books have been considered. Some of the books where captured at one page per image and others at two pages per image. A set of 100 images were ramdonly chosen among the different books, assuring each of them a minimum representation. A variable number of frontispieces per page is presented, varying from 1 to 6 frontispieces per page in the sample set.

Table 1 shows the algorithm offers good results at finding possible frontispieces, not finding any of possible frontispieces only in the 5% of cases. In Table 2, the probability of false positives is given. The heterogeneity of the book samples considered in the tests is the reason of the large probability of false positives, mainly due to: the irregular orientation of the pages and the different sizes of the staff lines, which are used to obtain the Fig. 4(a), cause the mask for staff removal to be badly generated, eroding many lines on the binary image.

4. CONCLUSIONS AND DISCUSSION

In this paper, a method to extract the frontispieces from ancient scores has been presented. The algorithm is able to separate the image into two pages when needed, performing a correction of rotation and determinating the region where the frontispieces could be.

This algorithm is part of a global OMR system which goal is to determinate all the information available of an ancient score using only its frontispieces, by comparing a series of descriptors with a huge database. To this end, a process of

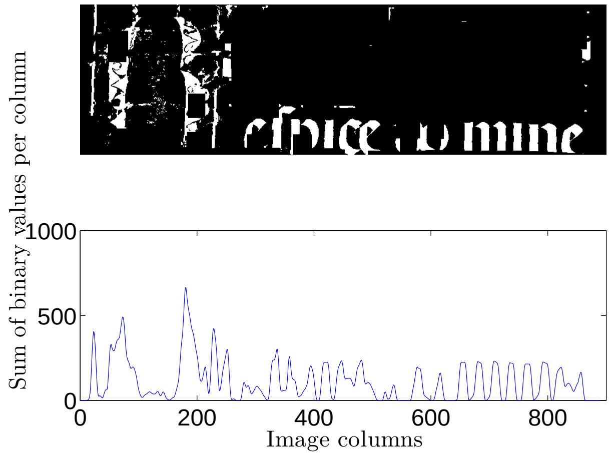


Figure 7: Sample column histograms.

Table 1: Performance of the developed algorithm.

Probability of detection of	**Global**	**Only pages with several frontispieces**
100% of frontispieces	0.82	0.82
75% of frontispieces	0.02	0.02
66% of frontispieces	0.01	0.01
50% of frontispieces	0.10	0.13
0% of frontispieces	0.05	0.02

Table 2: Probability of false positives.

Number of false positives	**Global**	**Only pages with several frontispieces**
Any	0.42	0.44
[1–3]	0.33	0.34
[4–6]	0.05	0.05
[7–9]	0.04	0.05

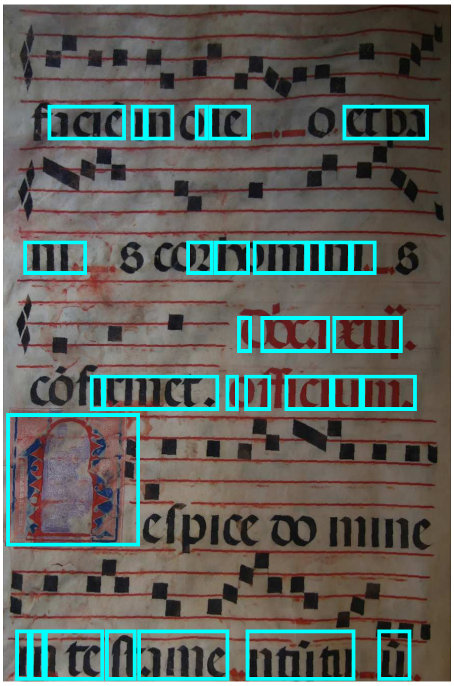
tracking needs to be implemented in the future, to correct the possible partial cut suffered by some frontispieces in this process, and the use of some descriptors to distinguish the real frontispieces between other false positive cases, reducing the amount of false positives.

Acknowledgment

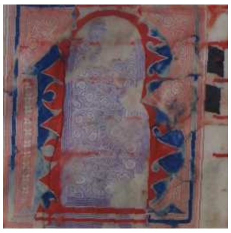
This work has been funded by the Ministerio de Educación y Ciencia of the Spanish Government under Project No. TSI2007-61181 and by the Junta de Andalucía under Project Number P07-TIC-02783. The authors are grateful to the person in charge of the Archivo de la Catedral de Málaga, who allowed the utilization of the data sets used in this work.

REFERENCES

[1] J. Caldas Pinto, P. Vieira, M. Ramalho, M. Mengucci, P. Pina and F. Muge. "Ancient music recovery for digital libraries." *Lecture Notes in Computer Science - Proceedings of the 4th European Conference on Research and*

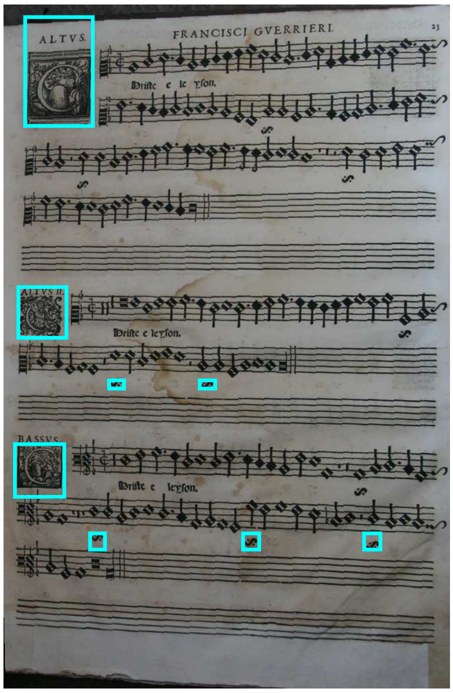


(a)

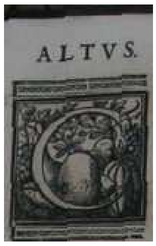


(b)

Figure 8: (a) All the candidates to frontispieces over the image outlined in boxes. (b) The candidates after the selection step.



(a)



(b)

(c)

(d)

Figure 9: (a) All the candidates in a page from Fig. 1(b). (b), (c), (d) Selected candidates.

Advanced Technology for Digital Libraries, pp. 24–34, Jan. 2000.

[2] D. Bainbridge and T. Bell,“The challenge of optical music recognition.” Journal *Computers and the Humanities*, vol. 35, number 2, pp. 95–121, May 2001

[3] L.J. Tardón, S. Sammartino, I. Barbancho, V. Gomez and A. Oliver. “Optical Music Recognition for Scores Written in White Mensural Notation,” *EURASIP Journal of Imagen and Video Processing*, vol. 2009, Article ID 843401, 23 pages, 2009.

[4] L. Pugin, J. Hockman, J. A. Burgoyne, and I. Fujinaga. “GAMERA versus ARUSPIX two optical music recognition approaches,” *International Conference on Music Information Retrieval (ISMIR 2008)*, vol. 9, pp. 419–424, September 2008.

[5] J. C. Pinto, P. Vieira, and J. Sousa. “A new graph-like classification method applied to ancient handwritten musical symbols,” *International Journal on Document Analysis and Recognition, Springer Berlin / Heidelberg*, pp. 10–22, August 2003.

[6] A. Fornes, J. Llados, and G. Sanchez. “Primitive segmentation in old handwritten music scores,” *Lecture Notes in Computer Science, Springer Berlin / Heidelberg*, pp. 279–290, October 2006.

[7] W.K. Pratt. *Digital Image Processing*. John Wiley & Sons, Feb. 2007.

[8] I. Barbancho, C. Segura and L.J. Tardón and A.M. Barbancho. “Automatic Delimitation of Area of Interest in Ancient Scores,” *IEEE Mediterranean Electrotechnical Conference* (MELECON), Valletta, Malta, pp. 326–331, April 2010.

[9] R.O. Duda and P.E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, June 1973.

[10] C. Christopoulos, A. Skodras and T. Ebrahimi. “The JPEG2000 Still Image Coding System: An Overview,” *IEEE Transactions on Consumer Electronics*, vol. 46, number 4, pp. 1103–1127, 2000.

[11] R.C. González and R.E. Woods. *Digital Image Processing*. Prentice-Hall, Aug. 2007.

[12] N. Otsu. “A threshold selection method from gray-level histograms.” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, number 1, pp. 62–66, Jan. 1979.