SEGMENTATION OF ANCIENT NEUMATIC MUSICAL NOTATION

***Lasko Laskov, and **Dimo Dimov**

**PhD student, Informatics Dep., New Bulgarian Univ., llaskov@nbu.bg, Humanities Informatics IMI – BAS*
***Assoc. Prof., Ph.D., Institute of Information Technologies, Bulgarian Academy of Sciences,*
Tel: (+359 2) 8706493, E-mail: dtdim@iinf.bas.bg

Abstract: During the last decade a lot of effort is put in preservation of historical heritage which is connected with the digitalization of ancient documents of different kinds to build digital databases and libraries. This brings up the problem of automatic (or automated) Document processing and Optical Character Recognition (OCR) of non-standard ancient documents. Unfortunately, the commercial OCR software fails to handle this kind of documents due to a number of difficulties, typical for them ˜ bad quality, old alphabet, ancient language, etc.) In this paper we propose a method for symbol segmentation of ancient documents, containing neumatic musical notation, based on simple horizontal and vertical projection profiles, which is an important stage of an OCR system, designed to handle ancient neumatic musical notation.

Key words: historical document processing, optical character recognition, text segmentation, neumatic musical notation

INTRODUCTION

Neumatic musical notation is a type of musical notation used by the Medieval Orthodox and Gregorian Christian Church which is used until nowadays in its contemporary form. As explained in [6], there is a vast variety of old manuscripts containing ancient neumatic notation which are of interest to the scientist in the fields of history, music and theology. A software tool which could assist the studies of this type of ancient manuscripts is an OCR-based system which can help either building a digital database or searching and comparison of data concerning neumatic musical notation.

Nowadays, there are a number of commercial OCR systems which successfully process standard printed documents of different ˜ languages, written in different contemporary alphabets ˜ Latin, Cyrillic, Greek, etc.. Unfortunately these kind of systems fail to process old historical documents due to a number of difficulties, typical for the ancient manuscripts, as the bad quality, old alphabet, handwritten symbols, ancient language. These conditions also concern the case of the ancient documents containing neumatic notation and together with the big material for investigation is the reason to develop an OCR based software which will process ancient documents of neumatic musical notations.

Because of the specific characteristics of the problem there are very few attempts to create such a system.

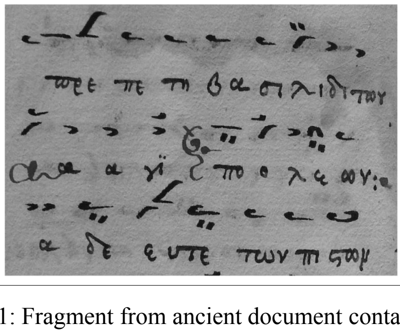

Fig. 1: Fragment from ancient document containing neumatic notation.

Gezerlis proposed in [3] OCR algorithm for contemporary neumatic notation but it was designed to recognize individual neumes and not to process whole document pages. In the more recent work [1] a complete OCR system for neumatic notation is proposed in which all stages of a recognition system are covered, like image preprocessing, page segmentation, symbol classification and semantic analysis. Even though this system is reported to give good results, it is still designed to recognize the contemporary notation and works with printed documents. Our goal is to create OCR system for ancient manuscripts which contain neumatic musical notation.

PROBLEM DESCRIPTION

Text (and symbol) segmentation is an important stage in the automatic document processing. This is the stage in which the document structure (lines, paragraphs, etc.) is discovered and characters, or symbols, syllabi, words etc. are segmented. There are number of methods for symbol and text segmentation, many of them based on projection techniques. The so called (ρ,θ) version of Hough Transform (HT) [3] is often adopted as a method for line skew and symbol skew detection by evaluation of the all projections along enough large number of directions θ, $\theta \in [0, 2\pi)$.

For the paper simplicity we suppose that the lines of our documents are mostly horizontal, i.e. we will apply only two projections, horizontally and vertically for the aims of segmentation. The neume manuscripts have a specific structure – each odd line contains neumatic notation and each even – Greek text. The neumatic lines show how the Greek text below them has to be interpreted in musical form. In some documents, the neume symbols and the text are written in different colors, or there are subsidiary notes written in different color. In this case color segmentation is applicable, which is described in [5].

Nevertheless neumes and Greek letters are often written in different colors, we will presume here the heavier case of they are written in one and the same colour, see Fig.1. Thus we can consider that input document images are already binarized, e.g. following [5], and consist only of black pixels corresponding to symbols of interest (neumes and/or Greek letters), and white pixels corresponding the background, see Fig.2. Although the documents are handwritten, the lines in general are almost straight and the symbols are written in a calligraphic manner (Fig. 1). This gives us the opportunity to apply a comparatively simple approach for document line and symbol segmentation, based on horizontal and vertical projection profiles of the document image.

For symbols' separators we shall consider the gaps (empty spaces) between symbols. For this reason there will be two types of neumes – simple neumes (composed out of a symbol only) and composite ones (composed by a group of symbols). As for the handwritten Greek it will be brought as usual to three types of objects: letters, syllables and words. At segmentation, the neumatic lines will be considered dominating over Greek text lines that will be very usable for the cases of neumes and (Greek) text overlapping (vertical collisions). The cases of horizontal collisions are not interesting at this time. This is intuitively clear for the Greek lines which will be processed only to evaluate some time intervals. The horizontal collisions in a neume line will be interpreted like composed neumes. Besides, we shall treat both the neumatic symbols and Greek letters as unknown symbols.

PROBLEM SOLUTION

The method proposed has three general stages: (i) approximate line segmentation by horizontal projection profile of the image, (ii) approximate symbol segmentation by separate vertical projection profile for each discovered document line; and (iii) collision solve for overlapping symbols or symbol parts belonging to different lines.

Approximate Line Segmentation. In this stage of segmentation, we calculate the horizontal projection profile histogram A(x), x = 0, 1, 2,..., H, where H is the image height. A(x) is accumulated by summing the pixel values along each bitmap

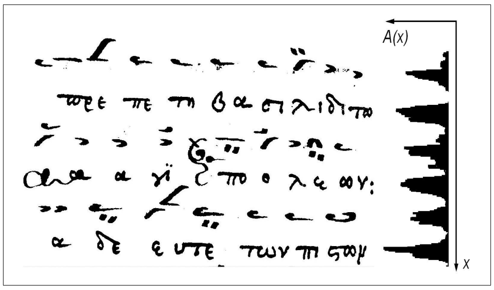

Fig. 2: A binarized (binary threshold) image of a neume notation and its horizontal projection profile.

row separately (Fig. 2). The local maxima of the resulting projection histogram correspond to the document line baselines and the local minima correspond to the gaps between lines. We find the minima using a data-driven approach known as swimming mean filtering for smoothing of A(x). Initially, when the horizontal projection histogram is accumulated, it has a lot of "false" local extrema (minima and maxima) caused by different reasons natural handwriting irregularities and/or noise. Suppose that we smooth *A(x)* with a 1D mean filter $g_s(.)$:

$$g_s(x) = \frac{1}{s}\sum_{i=-s/2}^{s/2} A(x+i), \quad x \in [0, H-1]$$

where s is the filter size and A(x) is expanded with zeros outside the interval [0,H-1]. A number of the minima of A(x) will dissipater depending on the filter size, see Fig. 3. If we start with a small filter size, for example with s=3, and continue to filter A(x) while increasing s, the number of minima will decrease until we reach the point when A(x) is nearly flat.

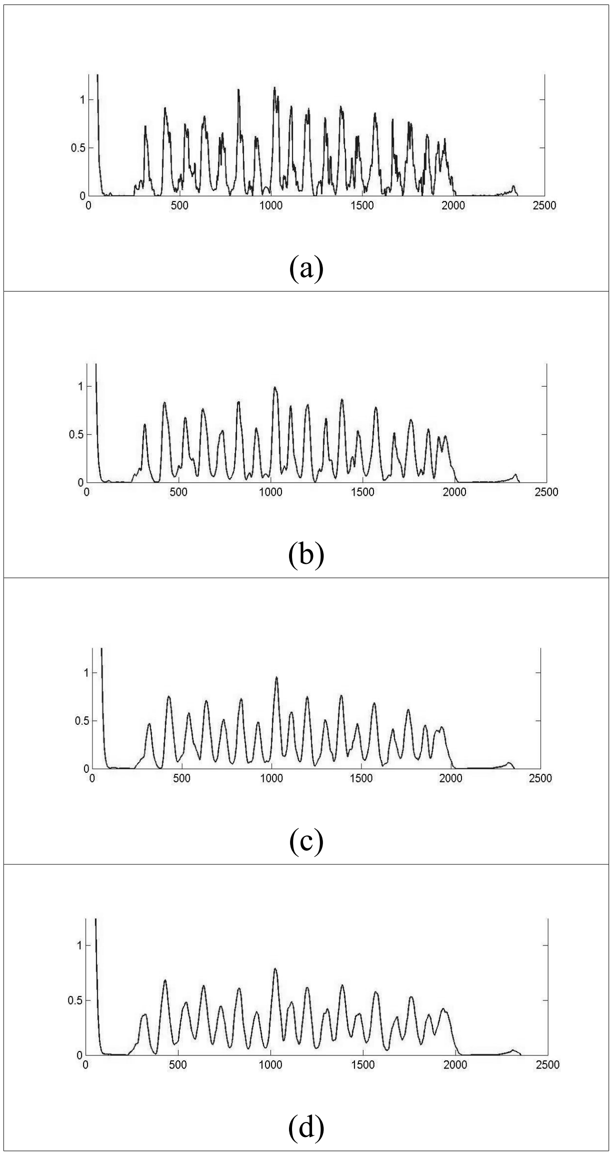

Fig. 3: The projection profile A(x) filtered with mean filter of size: (a) 3, (b) 17, (c) 31 and (d) 45; (the vertical document size H is 2500).

In this way we define a function $n=\Phi(s)$, where n is the number of local minima of $A(x)$, and we suppose that Φ is monotonously decreasing for $s=3, 5, 7,...,m$, where m is the filter size for which $A(x)$ is nearly flat. Since the aimed extrema which correspond to the document lines and the gaps between the lines dominate over the other "false" extrema we can expect that when $A(x)$ becomes smooth enough, Φ will not change for a number of successive filter sizes s. Such a plateau in Φ corresponds to the filter size(s) which will smooth the projection profile histogram $A(x)$ and will leave only the aimed minima corresponding to the gaps between the document lines, and maxima that are not interesting at the time being. A shape of function Φ is shown on Fig. 4.

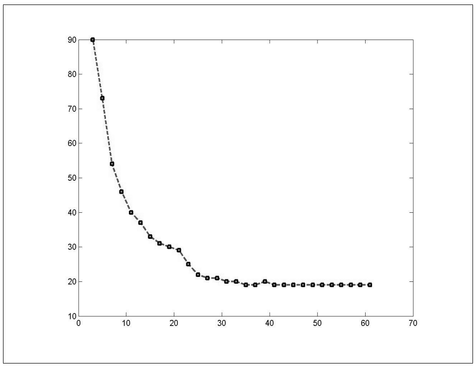

Fig. 4: Illustration of the correspondence between filter sizes s and number $\Phi(s)$ of the minima in the (horizontal) projection profile histogram.

Approximate Symbol Detection. After the document lines have been allocated, for each of them separately we apply vertical projection profile. The resulting vertical projection profile histogram (one for each line) is calculated in a similar way like in the horizontal case, but this time by pixel accumulation by columns of the respective line bitmap. Each vertical projection histogram is looked for the gaps between the symbols (group of symbols or symbol parts) which give an approximation of the symbols' location. Since the document image is announced binarized, we consider as a minimum each location of the scanned histogram where there is no object pixel detected, i.e. where the histogram value is zero. This leads to detection of gap areas between the symbols which we reduce to the middle coordinate of the respective gap area. In the last stage of the image segmentation, the area between the local minima in each document line is used to calculate the minimal wrapping window of containing a symbol (or a cluster of symbols). Similar technique is used to isolated possible noise-pixel errors.

Segmentation Collision Solve. The result of this stage will be a detected symbol, or group of connected symbols determined by the coordinates of the respective minimal wrapping window in the whole bitmap. First, for each image area, bounded between horizontal and vertical minima, the minimal wrapping window is calculated (Fig. 5). At this point all almost empty windows are detected and discarded as noise. After that the broken symbols which cross the line boundaries are detected by finding the minimal wrapping windows, which have a common horizontal boundary. Once a vertical collision is found, the broken symbols (or symbol parts) are marked and their own minimal wrapping window is calculated. For the remaining of the initial wrapping windows, the minimal windows are recalculated without taking into account the broken (conflicting) symbol parts. If some of them appear to be empty, they are discarded. In contrary, if all the three wrapping windows, the one containing the broken parts and the two recalculated, are not empty, the conflicting window is merged with the closest wrapping window but giving priority to the lines which contain neumatic notation (Fig. 6).

EXPERIMENTAL RESULTS

In this section an example of a result of our method is given. Fig. 5 shows a fragment of neumatic manuscript, after the stages of approximate line segmentation and symbols detection. Each detected symbol is surrounded by its minimal wrapping window. There are three vertical collisions in this example, which are solved in the collision solve stage. The result is shown in Fig. 6, where the conflicting symbols are shown in grey color. The neumes are correctly segmented, even though some their parts cross the Greek text line.

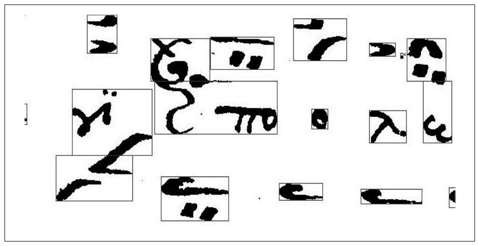

Fig. 5: Initially detected symbols in minimal wrapping windows.

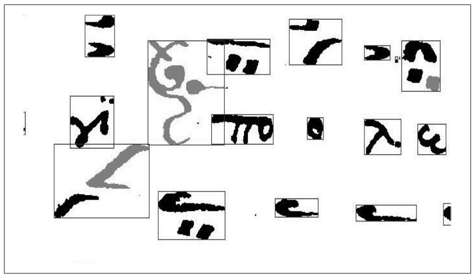

Fig. 6: Merged minimal wrapping windows with the collisions shown in grey.

CONCLUSIONS AND FUTURE WORK

In this paper we propose a method for text/symbol segmentation for ancient documents containing neumatic musical notation. The method is based on the horizontal and vertical projection profiles (histograms). A data-driven approach has been proposed for local extrema (minima and maxima in the projection histograms) detection, namely by a swimming mean filter that is successfully applied in our experiments.

The goal of the neume segmentation is to separate simple symbols and composed symbols (group of closed and/or connected symbols) and to store the coordinates of their minimal bounding boxes together with some additional information, like the number of line, line type (neumatic or Greek text), etc.. This gives us the opportunity to treat each symbol as a separate bitmap extracted from the original image or from its binarized version. Our next objective will be to use the extracted information to create a database of the symbols which will be composed from two main separate parts – one for the neumatic symbols and one for the Greek letters. The first one will be our priority, since our efforts are focused on the neumatic notation classification. The database should contain the following information for each detected symbol:

- document from which it was extracted,
- minimal bounding box coordinates,
- document line number to which it belongs,
- type of symbol (neumatic or Greek),
- pointer to detected symbol bitmap, etc.

The proposed method represents segmented symbols as separate bitmaps of minimal size that is very suitable for next classification/recognition step to be designed using neural networks and/or 2D Fourier transform. This representation is suitable to compare the symbols and to classify them in a number of distinct classes – one class for each essential neumatic symbol, and after that to organize an automatic recognition.

Also as a future work we will represent the detected neume symbols by a 1D Fourier transform on their contours. This representation has a number of advantages. First, it allows us to represent the symbols in more compressed form. Secondly a fast image retrieval method for fast access into the database can be applied as the one proposed in [2].

REFERENCES

1. Dalitz C., G. K. Michalakis, and C. Pranzas: Optical Recognition of Psaltic Byzantine Chant Notation, International Journal of Document Analysis and Recognition (submitted 06 March 2008).
2. Dimov D.: Fast Image Retrieval by the Tree of Contours' Content, Cybernetics and Information Technologies, Bulgarian Academy of Sciences, Vol.4, 2, pp. 15-29 (2004).
3. Duda, R. O., and P. E. Hart: Use of the Hough Transformation to Detect Lines and Curves in Pictures, Comm. ACM, Vol. 15, pp. 11–15 (January, 1972).
4. Gezerlis, V., and G., S. Theodoridis: Optical character recognition of the orthodox Hellenic Byzantine Music notation. Pattern Recognition 35 (4), pp. 895-914 (2002).
5. Laskov, L., and D. Dimov: Color Image Segmentation for Neume Note Recognition, In: Proceedings of the Int.Conf. A&I'07, 03-06.10.07, Sofia, 2007, pp. III.37-41.
6. Laskov, L.: Classification and Recognition of Neume Note Notation in Historical Documents, Proceedings of CompSysTech'06, the International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing, June 15-16, 2006, V.Tyrnovo, Bulgaria, pp. V.4.1-4.