COLOR IMAGE SEGMENTATION FOR NEUME NOTE RECOGNITION

Lasko Laskov[1], and Dimo Dimov[2]

[1] Computer Science Department, New Bulgarian University,
21, Montevideo Str., 1618 Sofia, Bulgaria, tel. +359 2 811 06 11, e-mail: llaskov@nbu.bg
[2] Institute of Information Technologies, Bulgarian Academy of Sciences,
Bl.29-A, Acad.G.Bonchev Str., 1113 Sofia, Bulgaria, tel. +359 2 870 64 96, e-mail: dtdim@iinf.bas.bg

Abstract: Neume note notation is a specific type of writing notes and music used by the Christian Orthodox Church from ancient times until now-a-days. The historical documents containing neume notation provide vast material for research in many fields of the cultural sciences. This is our motivation to perform efforts towards creating a software tool which will help scientists in their investigations of neume notation. As a first step of automated/automatic processing of neume notation images, we will discus the separation of the objects of interest, neumes and text, from the document background. In this paper we try covering the problem by two algorithms applying classical thresholding for image segmentation in HSV color scheme.

Keywords: image processing, neume note segmentation, segmentation by HSV thresholding, periodic histogram thresholding.

INTRODUCTION

Neume note notation is a specific type of writing used by the Christian Orthodox Church to denote music and musical forms in the sacred documents from ancient times until now-a-days. Usually the neume symbols are written above the psalm text and denote how the text has to be performed, specifying tempo, tone, intonation, etc. Since the first documents, which contain neume notation, date from ancient times, the variety of such documents is vast and is an important source of information for scientists in the fields of history, music and cultural sciences. An example of a research in this field is the attempt to track down how the musical forms sounded in a given historical period, comparing the way neume notation developed and changed during the ages and different schools of copyists. Such kind of research involves a lot of comparison and searching for similar patterns in different historical documents. This inspires the creation of a software system which can help researchers by automating most of the technical activities involved in the investigation of historical documents, containing neume note notation.

The Optical Character Recognition (OCR) approach should play a key role in research and development of such system. Besides, the system should tend from automated to an automatic computer analysis and recognition of neume notation in historical documents. In other words, we try to construct a specific OCR for automated/automatic neume notation reading. Even only from text recognition viewpoint, OCR of historical documents is by itself a challenging problem because of: (i) the text is handwritten, (ii) the parchment or paper is often of bad quality or destroyed, (iii) the language is ancient, etc. For these reasons, computer processing of historical documents requires special techniques for document image enhancement, character recognition and so on, and is also the reason why the standard commercial software fails to solve these problems. Besides the problems caused from the fact that the documents are ancient, we have problems which are caused from the specific characteristics of the neume note notation itself. First of all, in the different periods of history the notation was permanently modifying, and also the notation varies in the different copyists schools. These make the variety of different neume notations very big and thus difficult to analyze and classify. It is also important to note that for the oldest neume notation, even the scientist/researchers are not sure how to interpret most of the neume symbols and/or their compositions (notations).

As the first stage of automatic computer processing of images, containing neume note notation, we will discuss the segmentation of the meaningful objects from the image background. In the literature this process is often referred to as image *binarization*, if we consider that we have two classes of pixels only – object pixels and background pixels. However, in our case we have objects of interest in different color, namely: (1) neumes and *accompanying text*, written by dark ink that seems almost black or brown and (2) subsidiary notes, most often written by red ink (Fig.1). In some rare cases, other ink inscriptions (green, blue, etc.) are added to the original neume documents in more recent time. The background color highly depends on the paper/parchment type and the age of the document, and may vary from gray-blue to yellow or light-brown.

Our current task will be to separate the objects of interest (neumes and texts) from the background using predominantly histogram-thresholding approaches to color characteristics of the documents.

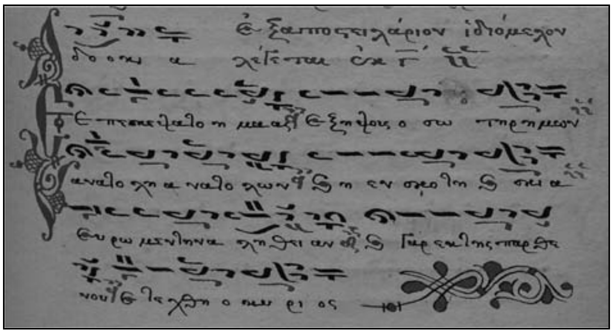

Fig.1. A fragment containing a relatively clean neume notation.

IMAGE HISTOGRAM ANALYSIS

To separate the objects of interest from the background of the document image we will use histogram based methods [2, 6, 7, 8]. We will also use Otsu's approach for histogram optimal thresholding or multi-thresholding [6, 7]. However, we will use these approaches most of all for histogram thresholding in HSV color space instead of the classical RGB color space.

In our case, we have to separate three types of objects:
(i) neumes and accompanying black text,
(ii) red or other-colored subsidiary notes, and
(iii) the background which can be achromatic or of specific color.

One approach of the segmentation could be to create a 3D color histogram for the RGB color space. Than we will expect to observe three "clouds" of RGB entities, one for each of the three classes (types of objects). Thus, our problem would be reduced to a separation of the first two classes from the third one (the background) in the RGB color space.

In fact, the experiments (see Fig 13÷16 below) show that after gray-level transform of the image the traditional approaches for statistically optimal binarization [6, 7] give relatively satisfactory results. As we can see from Fig.2 the two classes, "dark" and "light", are very good distinguishable. When this is not fulfilled we can modify those approaches to locally-adaptive ones, for example using [2].

However, the result of the binarization has to be applied as a *mask* over the original image for an additional separation of both classes of binarization, into the three classes defined for our case. The former will obviously require an extra color processing in some color space, for example in HSV color space instead of the traditional RGB.

The HSV (hue, saturation, value) color schema [4, 5, 8] is often preferred to RGB. It better represents the viewpoint of painters while the RGB was primary designed as a model of color generating techniques in display monitors. The HSV schema is definitely preferred recently in the biometrics area, for skin segmentation, e.g. in face or palm recognition, [1, 9].

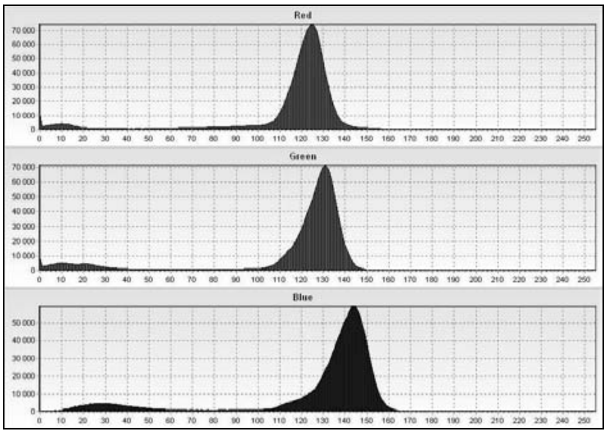

Fig.2. RGB color histograms for the image from Fig. 1.

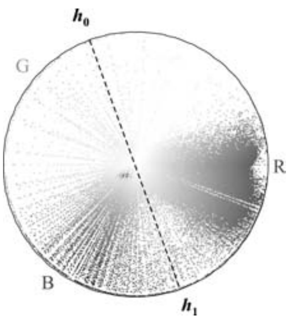

Fig.3. The HS color histogram of the image of neume notation from Fig.1. Both H-thresholds, h_0 = 110°, h_1 = 291°, are visible.

We consider the HSV adequate for our case because of the instances of the red notations seem to be simply segmented from the color of the background, as well as the achromatic (gray) neumes – from the gray of the background (see Fig 3).

Additionally, we could use an approach similar to Fisher's linear discriminant [3, 8], e.g. to search in the chosen color space (HSV, or RGB, etc.) for an optimal axis, or plane, on which the projected three classes will be most easily distinguishable. But this approach will lead to a clumsy iterative procedure to the solution, since the classes are preliminary unknown, for a given image.

Therefore it is reasonable to search for a specific approach to the observed problem solution which should unite the simplicity of the well-known approaches for histogram optimal thresholding with the unavoidable color processing, for which we chose the HSV color scheme.

HSV HISTOGRAM ANALYSIS

The HSV color space is usually represented as a cone [4, 5, 8], see also Fig.5 and 6, where:
1. *Hue,* the color parameter h, measured in angle degrees, $h\in[0°,360°)$, and sequentially passing the colors from red, through yellow, green, cyan, and blue, to magenta.
2. *Saturation s*, a normalized value $s\in[0, 1]$ that varies from unsaturated shades of gray (s=0) to fully saturated colors with no white component (s=1).
3. *Value v*, also normalized in the interval [0, 1] and representing the brightness of the image.

Using HSV color space we propose a histogram-based method to separate the three classes of pixels in the image – the dark colored letter and *neume* pixels, the red letter pixels and the background pixels (colored and/or achromatic ones).

The both algorithms we propose below are based on the following underlying operations:

Operation 1. Separation by the S-histogram

We apply the Otsu's method on the S-histogram, accumulated along the S-axis of HSV space for the neume image, i.e. to calculate a statistically optimal threshold s_0. In this way we separate the entire HSV-cone in two parts:
(i) an (almost) achromatic sub-cone $S_0(s)$ of height v=1, and $s< s_0$, where s_0 is the basis radius of S_0, and
(ii) the rest $S_1(s)$ of the HSV-cone ($s_0\leq s\leq 1$), containing the definitely color part of the image.

Operation 2. Separation by the V-histogram

Applying again the Otsu's approach we calculate a threshold v_0 for the V-histograms and so we can separate the respective area in two parts: the "dark" $V_0(v)$, where $0 \leq v \leq v_0$, and the "light" $V_1(v)$, where $v_0 < v \leq 1$.

Operation 3. Separation by the periodic H-histogram

To perform a color separation, e.g. in S_1 after operation 1, we can use the H-histogram. We can apply again the Otsu's approach but considering that the H-histogram is a periodic one (Fig. 3). In this connection we propose the following:

Modification of Otsu's approach for the periodical case:

The next text concerns the H-histogram but can play as a proof for the common case of periodic 1D-histograms:

- Obviously there are two thresholds, h_0 and h_1, necessary to separate two continuous areas in a periodic histogram, in our case – the HS-histogram (Fig.3) resulting in a periodic H-histogram (Fig.4).

- Let us suppose that the histogram start-point coincides with the threshold h_0. Then we have to calculate the threshold h_1 by the relatively simple Otsu's method which maximizes the Otsu's criterion $\eta_{ho}(h_1)$.

- As h_0 is a priory unknown we have to repeat the above procedure for each (integer or rational) h_0, $0 \leq h_0 < 360°$, and to get as result this couple (h_0, h_1) which maximizes the criterion $\eta_{ho}(h_1)$.

For the sake of the software optimization we extend the H-histogram over two periods, i.e. $h \in (0° \div 720°)$, see Fig.4.

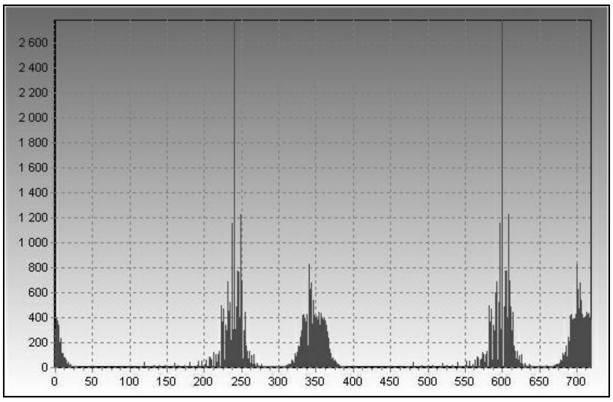

Fig.4. The H-histogram of the colored part $S_1(s)$ of the neume image HSV-cone, $s_0 \leq s \leq 1$. Two optimal thresholds (h_0, h_1), because of periodicity of the H-histogram, $h_0 = 110°$, $h_1 = 291°$.

TWO CASES OF SEGMENTATION

The two algorithms, *A* and *B*, for neume and text segmentation that we propose are based on the three operations, already described. Their segmentation strategies are illustrated on Fig. 5 and Fig. 6, respectively. Both algorithm descriptions are closely tied to both columns of experiments, see Fig.7÷16.

Algorithm A

Step 1: Apply Operation 1 on the given document image. The result is a threshold s_0 which separates the HSV cone for the image in two parts – the almost achromatic sub-cone $S_0(s)$ and the color rest $S_1(s)$ of the cone.

Step 2: Apply Operation 2 on the achromatic sub-cone $S_0(s)$. The threshold v_a we obtain, separates the dark symbols from the achromatic component of the background, i.e. restricts $S_0(s)$ to its upper part $S_0(s,v \mid v_a < v \leq 1)$, see Fig. 5.

Step 3: Apply Operation 2 again but this time on the color rest $S_1(s,v \mid v_a < v \leq 1)$ of the HSV-cone. The resulting threshold v_c separates the dark colored symbols (neumes and text) from the almost same (red) color of subsidiary notes, Fig. 5.

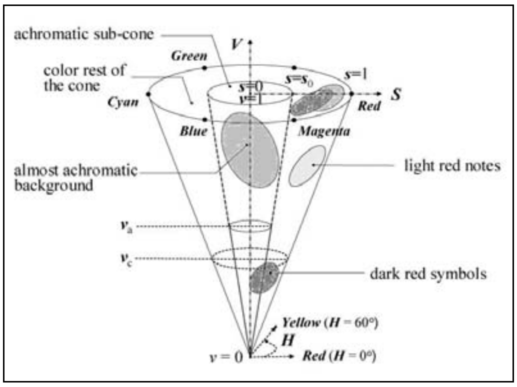

Fig.5. The HSV scheme of the significant volumes considered by the algorithm *A*.

Algorithm B

Step 1: The same as the Step 1 of Algorithm *A*.

Step 2: It is similar to Step 2 of Algorithm *A* but this time the interpretation of v_a-thresholding of the achromatic sub-cone $S_0(s)$ is to separate the dominating light part from the rare dark part. So this segmentation appears not very important, because of both parts of $S_0(s)$ are interpreted as background. By regular light conditions, the low part $S_0(s,v \mid 0 \leq v \leq v_a)$ often appears almost empty, see Fig. 6.

Step 3: Apply Operation 3 on the color part $S_1(s)$ of the HSV-cone. The obtained two thresholds of the periodic H-histogram separate the two essential classes of the document, the light neumes/symbols (in red) and the darker-colored psalm text, see Fig. 6.

Additionally, we chose for the light neumes/symbols this part of H-histogram that is more close to red ($h = 0°$). Another semantic rule for this could be – the light neumes/symbols space is smaller than the one of psalm text.

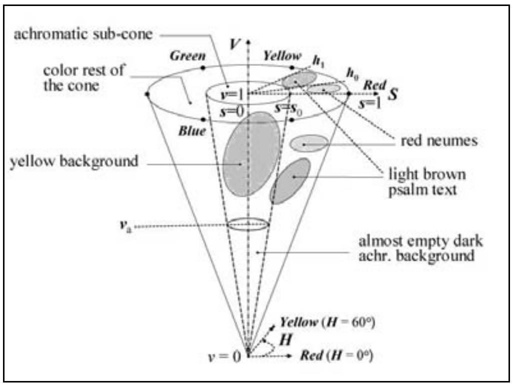

Fig.6. The HSV scheme of the significant volumes considered by the algorithm *B*.

EXPERIMENTS

We provide two examples to illustrate the two algorithms described in the previous section, see Fig.7÷16. The left column of figures (odd numbers) corresponds to algorithm *A*, while the right one – to algorithm *B*. A horizontal correspondence among the figures is also ensured.

The example images are given on Fig. 7 and Fig. 8 respectively. They both represent neume notation fragments. The first fragment consists of dark colored letters and neumes, red notes and gray-brownish background, while the second – of light-brown text, red neumes and yellowish background.

Fig.9 and 10 illustrate the both algorithms, by steps. On Fig. 11 and 12 we provide the obtained results. Fig.13 and 14 as well as their subsidiary Fig. 15 and 16, allow a comparison of our results with the results obtained by the Otsu's globally optimal thresholding of the example images of Fig.7 and 8, where the YIQ color scheme was applied for conversion of color images to gray.

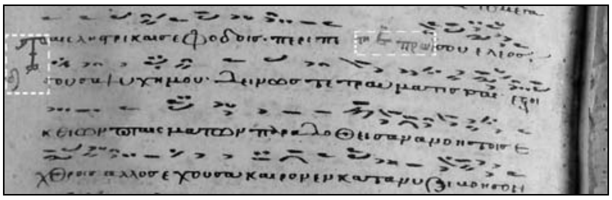

Fig.7. An original image fragment with blue-gray background, dark-brown neumes (and texts), and some text notes in red. The both dashed squares are shown for comparison with Fig.17 below.

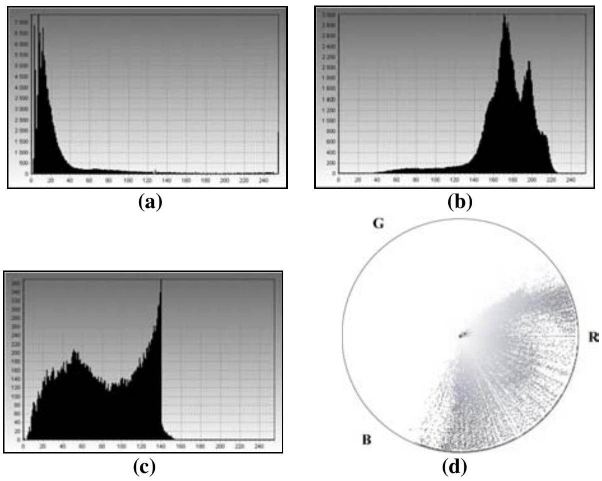

Fig.9. The algorithm *A* sequence of histograms (cf. Fig.5) applied for the image fragment on the left (Fig.7):

(a) S-histogram of the whole HSV space for the image, s_0=98;
(b) V-histogram for the achromatic sub-cone S_0, threshold v_a=140;
(c) V-histogram for the color cone part S_1, determined by s_0 and v_a; a new threshold v_c is obtained, v_c=81;
(d) HS-histogram for the color cone part S_1, thresholds h_0=148° and h_1=324°; this histogram is not very significant for the algorithm *A*.

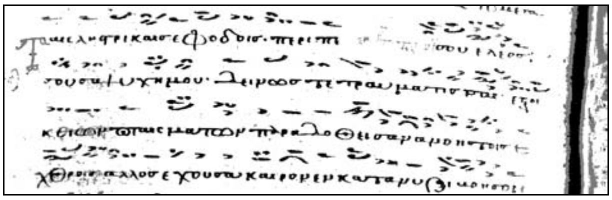

Fig.11. The image fragment (on the left) segmented by the algorithm *A* of the proposed method.

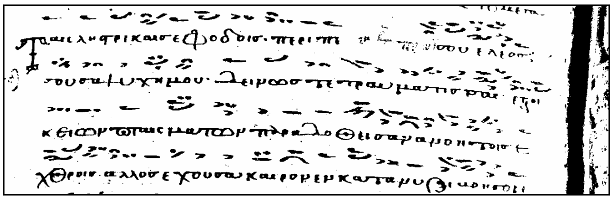

Fig.13. The image fragment (on the left), but preliminary converted to gray using YIQ color scheme and then binarized by Otsu's global optimal threshold (to be compared with Fig.11).

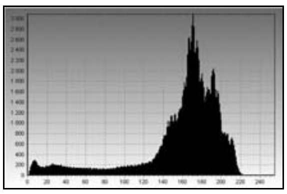

Fig.15. The Y-component histogram of the image fragment on the left; the Otsu's global optimal threshold obtained, y_0=113.

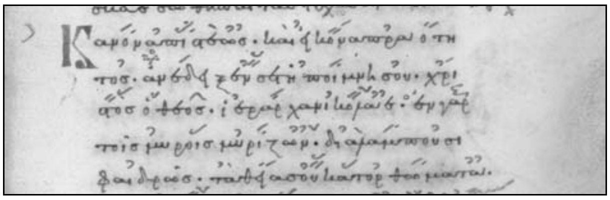

Fig. 8. Another original fragment with yellow background, neume notes in red and light-brown psalm texts.

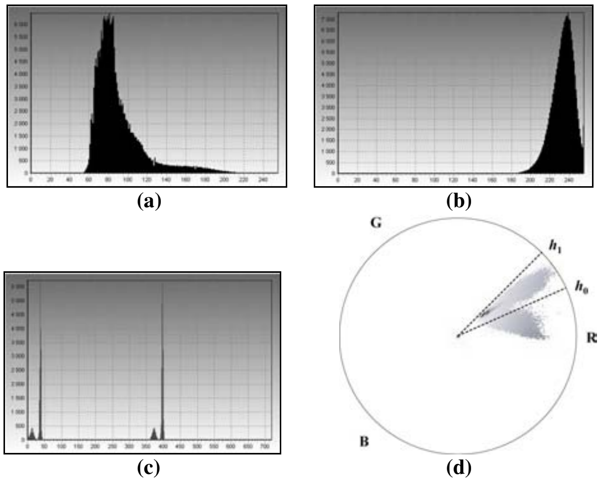

Fig.10. The algorithm *B* sequence of histograms (cf. Fig.6) applied for the image fragment on the right (Fig.8):

(a) S-histogram of the whole HSV space for the image, s_0=114;
(b) V-histogram for the achromatic sub-cone S_0, threshold v_a=230; this histogram is not very significant for the algorithm *B*;
(c) H-periodic-histogram for the color cone part S_1, determined by s_0; two thresholds h_0 and h_1 are calculated, h_0=24° and h_1=45°;
(d) HS-histogram for the color cone part S_1; the thresholds h_0 and h_1 are additionally shown.

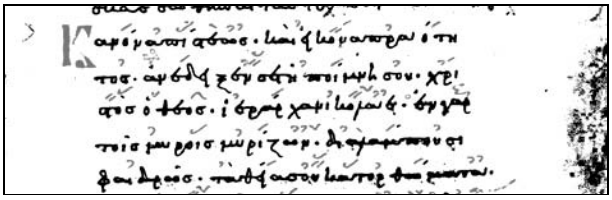

Fig.12. The image fragment (on the right) segmented by the algorithm *B* of the proposed method.

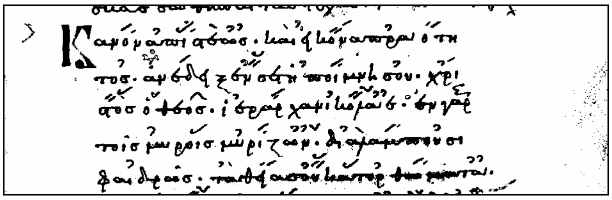

Fig.14. The image fragment (on the right), preliminary converted to gray using YIQ color scheme and then binarized by Otsu's global optimal threshold (to be compared with Fig.12).

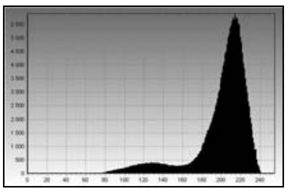

Fig.16. The Y-component histogram of the image fragment on the right; the Otsu's global optimal threshold obtained, y_0=169.

DISCUSSION AND CONCLUSION

Historical documents containing neume notation are often noisy and corrupted. Apart from this, the background, as well as the neumes and texts (psalm texts and/or subsidiary notes) have their specific color characteristic, besides the noise caused either by the time of archive keeping or by the inappropriate illumination of snapshot or scan. For these reasons the standard global thresholding methods are not applicable in this cases.

In this paper we have proposed a HSV color scheme based segmentation method which takes in advantage not only from the intensity characteristics of the image but also by the color information. As it has been analyzed in the paper second paragraph, we had to search for a specific solution of the problem that unites the simplicity of the well-known approaches for histogram optimal thresholding with the unavoidable color processing of images. Meanwhile, we have had to prove a modification of Otsu's approach [7] to a periodic 1D histogram as the Hue-histogram is.

We have proposed two algorithms that cover the most of the cases of our data pull of more than a few hundreds of neume notation images. We have not to exclude existence of "difficult" cases matching neither of both algorithms. For this reason we have in mind the following directions for future work in the topic:

• Eventual difficult cases are expected to be solved applying the similar strategy of H-S-V-histograms thresholding because of many combinations remaining not used by the proposed two algorithms.

• Observed illumination non regularities as well as non regularities in HSV-characteristics in the images of interest (see Fig.17) are foreseen to solution by extension of the proposed algorithms towards locally adaptive approaches, e.g. like [2].

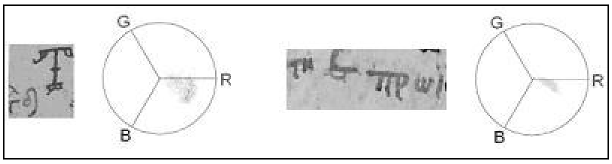

Fig.17. Two parts of the first image example (cf. Fig.7) and their HS-histograms to be compared with the HS-histogram of the entire image (cf. Fig.9d).

• A next step to an automated/automatic neume recognition will be a geometrical approach for contours' detection of the segmented neumes and symbols. Application of harmonic analysis (wavelets, Fourier, etc.) methods either on 1D contours or on 2D areas of interests could additionally help the segmentation and final recognition as well.

• A specialized database is also foreseen to be developed either for the experiments keeping or for the semantic studying of neume classes recognized/mined by the outer help of neume notation specialists/archivists too.

• Feedback approaches well-known from the control theory are also foreseen to research in the terms of considered problem, namely to analyze the behavior of the entire neume recognition system by variations of its inner parameters, including those of the proposed two algorithms.

It is expected that the considered system for neume recognition should not be trivially reproduced by on-the-shelf OCR system, nevertheless the large opportunities of learning they currently have recently.

ACKNOWLEDGMENTS

This work was partially supported by the following grants: Grant # 010088/2007 of Bulgarian Academy of Sciences, and Grant # I-1306/2003 of the Bulgarian Ministry of Education & Science.

REFERENCES

1. Bradski, G.R., and S. Clara. Computer Vision Face Tracking For Use in a Perceptual User Interface, *Intel Technology Journal*, Q2'98, (1998).
2. Dimov, D. T., and G. Y. Gluhchev. A Locally Adaptive Binary-Tree Method for Binarization of Text Images. In: Downton, A.C., and S. Impedovo (Eds.): Progress in Handwriting Recognition. World Scientific, London (1997) 575-580.
3. Duda, R. O., P. E. Hart, and D. G. Stock. Pattern Classification (2d ed.), John Wiley & Sons, Inc., 2001.
4. Foley, J. D., and A. van Dam. Fundamentals of interactive Computer Graphics, Vol. 2. Addison-Wesley Publ. Co., London Amsterdam Sydney, 1982.
5. http://en.wikipedia.org/wiki/HSV_color_space
6. Kurita T., N. Otsu, and N. Abdelmalek. Maximum Likelihood Thresholding Based on Population Mixture Models, *Pattern Recognition*, Vol. 25, No. 10 (1992) 1231-1240.
7. Otsu, N. A Threshold Selection Method from Gray-Level Histograms, *IEEE Trans. on Systems, Man, and Cybernetics*, Vol.9, No.1 (1979) 62-66.
8. Sonka, M., V. Hlavac, and R. Boyle. Image Processing, Analysis, and Machine Vision (2d ed.), Brooks/Cole Publishing Co., CA, 1998.
9. Zhu X., J. Yang, and A. Waibel. Segmenting hands of arbitrary color, In: Proceeding of the 4th International Conference on Automatic Face and Gesture Recognition, Grenoble, France, March 2000, 446-453.