Identification of authors of documents based on offline signature recognition

Tončo Marušić*, Željko Marušić, Željko Šeremet**
*Faculty of Science and Education, Mostar, BiH
**Faculty of Mechanical Engineering and Computing, Mostar, BiH
tonco.marusic@gmail.com

Abstract - Handwritten signature is used in various applications on daily basis. Whether one signs a contract, work documents, petition, or wants to approve a check payment, one will use personal signature to do all those things. In this paper we use this daily based biometric characteristic for identification and classification of students' papers and various exam documents used at University of Mostar. In this paper we used OpenCV library as an image processing tool for feature extraction. As regards to classification method, we used Support Vector Machine.

I. INTRODUCTION

In the teaching process, teaching staff (professors, assistants) must organize and manage a lot of documentation. Exam lists, quizzes, written assignments, tests, attendance lists are just some of them. When we bear in mind that these materials must be archived and stored for some time, it is clear that it is necessary to digitize. This heap of documentation can then be used for various analyses, as well as for formation of a knowledge base. In the process of digitizing, we can use different types of scanners or even cameras from smartphones. Once the documents are scanned, they should be grouped, sorted and stored in individual folders and directories. This entire process can be automated.

Generally, signatures and student ID numbers are used for identification of these documents. Handwritten signatures or handwritten ID numbers represent non-invasive and inexpensive identification methods.

Recognition of individual digits of a number is a problem that is relatively well resolved. A much difficult problem is to recognize digits when some of the handwritten digits are connected. A simple solution is keeping digits separated during enrollment, and to write them to bounding boxes. This is illustrated in Figure 1 that represents proposed header of documents used in various student exams.

![Figure 1. Proposed header of test documents with place for signature and student ID.](image)

During this enrollment, students often type the wrong digit by mistake or they simply cross the bounding box reserved for the individual digits, which will inevitably result in the wrong classification of digits, and thus the misidentification of the author of the document. It would be desirable that the system is resistant to these situations, and potential candidate is to use signatures as a way of verification and/or identification of the author. In this paper, we will check what are the possibilities for identification of various scanned exam documents based on students’ signatures. After recognition of the authors, the documents can be automatically classified in the individual directories and linked to document management system (DMS).

As regards to literature, there is much more to deal with the offline verification of signatures in relation to the offline identification of signatures. Terms are similar in that they rely on the database which contains records about user’s biometric characteristic features. Then we compare given biometric characteristic features with those stored in the database. The main difference is that verification answers the question: “Is the user really who he or she says to be?” On other hand, biometric identification answers the question: “Who is the person?” As regards to classification point of view, verification is often referred to as „one-to-one” comparison. Biometric identification is known as „one-to-many” comparison [1]. Since they are similar, papers that deal with verification can be used in our analysis.

Various off-line techniques have been proposed in literature for the signature verification problem. These techniques include: Hidden Markov Models (HMM), Neural classifiers, Wavelets, Dynamic Time Warping [2], Graph matching, Support Vector Machines and so on. Generally, we can say that SVM outperform HMM [5] and Neural networks [3], therefore, we chose SVM as a classifier.

Offline signature verification and identification is a pattern recognition problem and a typical pattern recognition system has the following steps [4]:

- Data Acquisition
- Preprocessing
- Feature Extraction
- Classification
- Performance Evaluation
II. ACQUISITION OF SIGNATURES

There are inevitable variations in the signature patterns produced by the same person (intrapersonal variability) [6]. Some factors that affect our signatures are: a person’s physical and psychological state, body position, writing material, purpose of signing (formal vs. informal). In order to build a reliable system, we must obtain a database of signatures extensive enough to capture possible individual variations. We can address this issue by modeling acquisition system that will include different size rectangle areas, pens with different stroke width, different color of imprint, different angle and so on [7].

There are two types of acquisitions in our approach. First, we collected the signatures from 30 students. Each student wrote 22 signatures on a printed sheet A4 paper with the same size rectangles. They used different pens with different color (blue, black, red) and with different stroke width. During the acquisition, we emphasized to students that during the signing the boundary lines should not be touched. With the development of technology for the purpose of acquisition, we can use cameras and mobile phones with additional image processing. After that, the whole A4 document was photographed with a camera from mobile phone Sony Xperia T with a resolution of 4128 * 3096 (13Mpx). For purpose of photography we built improvised tripod with led lamps that illuminate the document. The signatures were manually extracted using an external tool for image processing Gimp. These signatures were then used for our database.

Another type of acquisition included collecting various tests, exam papers of students to whom we performed our classification. In total, we collected 145 documents. We photographed them in an analogous manner with the same resolution as above. However, in this case we did not extract the signatures by manual cropping, but we used computer vision algorithms for doing this automatically. We used mobile phone camera just to capture the documents and all processing was done on PC.

III. PREPROCESSING

Preprocessing is a set of subsequent operations applied for the improvement of quality of signature image and locating ROI (Region of Interest) of signature.

The various sub-processes which can be considered as image preprocessing are: Cropping, Binarization, Noise removal, Morphological Operations, Image normalization, Skeletonization, etc.

A. Binarization

Binarization represents an initial step in most systems for the analysis of document images. It refers to conversion of the gray-scale image into a binary image. It is crucial, because it is the basis for successful segmentation and recognition of characters, words or zones. Generally, we distinguish global and local methods. Global methods use one threshold value to classify the pixels in the image object or to classify the background. On other hand, local schemes can use multiple values based on the information in the local area. Global techniques are not suitable for binarization of documents that are delivered via camera, because such documents suffer from non-uniform lighting, low contrast, shadows [8]. We used local adaptive technique instead.

B. Image denoising

After image acquisition and adaptive threshold, noises like salt and pepper, Gaussian noise were generated during painting. In the documents it is possible to use a series of filters as the median, Gaussian and Wiener filter. We used the median filter which reduces noise with a very small degree of image degradation an also preserves sharp edges unlike the mean filter and considerably less blurs the image in relation to the Gaussian filter.

C. Data Area Cropping and rotation

At the top of each test sheet there is a header containing basic information about the exam. This information consists of the name of the exam, the exam type, number of exam period, the date of the exam, the student's signature, index number, etc. If we are dealing with documents, we must first segment the header, and then in the header we locate the region of interest of signature. For this purpose, the algorithm used is based on the topological structural analysis of the image that has been implemented in OpenCV [9]. This is illustrated in Figure 2 and Figure 3.

In Figure 3 we can see that base line is extracted with a signature. This line is of great importance to us, we can very easily determine and correct the angle of image acquisition to zero with rotation and using contour segmentation. In this way we improve accuracy of feature extraction.
D. Morphological operations

After binarization, cropping and rotation, we continue processing resulting images in order to improve the quality of text regions and preserve connectivity within certain characters, and meet some possible gaps, holes within the character. For this purpose, we used mathematical morphology operations such as: dilation and image erosion. These operations are usually applied to binary images. Furthermore, in this process we removed the base line with dilation and bitwise operation. After cropping, we found bounding box of signature like in Figure 4.

![Figure 4. Bounding box of a signature](image)

E. Normalization width

Width of the photographed document is 3096 pxs. Region which the signature occupies is approximately one-third of the header. Once we determine the signature region, each signature will be scaled using bicubic interpolation to constant width of 384 pixels keeping fixed aspect ratio.

F. Skeletonization

When signing the test sheet, a student does not always use the same pen. Thick imprint of pens can greatly affect the value of extracted features. Precise determination of values for certain features in this paper can only be obtained on the skeleton image. Furthermore, this speeds up the execution of algorithms in gathering values for some features. This is illustrated in Figure 5.

![Figure 5. Skeleton signature image](image)

After preprocessing, we can move to determine values of global, grid and SIFT features.

IV. EXTRACTING FEATURES

Features are a key process in achieving high accuracy in identifying signatures. The ideal separation techniques use minimal sets of features that are used to maximize the difference among different people’s signatures while minimizing the difference of the same person’s signature [10].

Global structural and statistic features of signatures are extracted from the entire image signatures. Grid features are extracted from the grid formed by the parts of the signature [7]. Also we used SIFT feature descriptor in our work.

A. Global features

In literature, there are numerous possible global features [7][10][11][12][13][14][15][16]. After the analysis and testing, we chose 16 global features that proved their efficiency.

1. Aspect ratio - represents the proportionality rate of the skeleton signature image. This is calculated by dividing the height with the width of the signature;

2. Signature occupancy represents the number of pixels that belong to the signature;

3. Baseline shift is a difference between the vertical center of gravity of the left and right side skeleton signature image. At first, counting divided the image signature vertically in two halves and calculated the center of gravity for each half of the CL and CR. Then baseline shift (baseline shift) was defined as BSL = CL – CR;

4. Global Slant Angle - It is the overall direction of line strokes in the skeleton signature. The original signature was rotated from -45degrees to +45 degrees in steps of 5 degrees. For each rotation, the original signature was first pre-processed followed by counting the number of vertical 3-pixels connections from the rotated skeleton image. The global slant angle is the angle which has the maximal number of vertical 3-pixels connections. This is illustrated in Figure 6.

![Figure 6. Rotated and skeletonized images for procedure of finding Global slant](image)

5. Number of Edge Points - According to [15] an edge point is a black pixel which has only one 8-neighbour. This is illustrated in Figure 7.

![Figure 7. Number of edge points](image)

6. Number of Cross-Points and Spatial Symbols [14]

Every person uses some spatial symbols in his signature, such as: some ‘x’ marks (cross marks), star
marks or other symbols. The total number of spatial symbols of a person’s signature is unique. If we find that one pixel has more than two neighbors, then those pixels will form a spatial symbol. Such types of pixels are shown in Figure 8.

Figure 8. Number of Cross-Points and Spatial Symbols

7. and 8. Horizontal and vertical center of gravity

Center of gravity is pair (X,Y) represented with formula (1)

\[ X = \frac{1}{N} \left( \sum_{i=1}^{n} x_i \right) \quad Y = \frac{1}{N} \left( \sum_{i=1}^{n} y_i \right) \]  

where \( x_i \) and \( y_i \) represent numbers of white pixels in columns and rows respectively on binarized image. Part X represents horizontal centre of gravity, and part Y represents vertical centre of gravity.

9. and 10. Length name, Length surname

After analysis of students’ papers, we can say that students identify themselves with full first and last name. Generally, there is a considerably larger gap between first and last name in relation to the distance between the individual characters that are often completely connected. This property can be utilized to segment first name and last name, and then retrieve the width for each part. For this purpose, we iterated morphological dilation operation until we got only two segmented parts. During the acquisition we found no instance where all pixels of a signature were completely connected. However in case that there is only one connected contour left during this process, both features would be set to value 0. In case that person has multiple names the largest gap between names is used for division.

11. and 12. Maximal horizontal and vertical histogram

The horizontal histogram is calculated for each row. Row with the highest value is taken as the maximum. Similarly, we calculate the vertical histogram for each column. Column with highest value is taken as maximal vertical histogram.

13. The length and ratio of Adjacency Columns [14]

In this stage, we firstly computed the length of the adjacency columns from top and from bottom of the Sample signature image and stored it in a 1D array \( L_T \) and \( L_B \). Then we computed the sum of all elements of \( L_T \) and all elements of \( L_B \). Overall procedure for this feature extraction is shown in the following Figure 6. Upper half of Figure 6 shows a portion of a signature image. Lower half of Figure 6 shows the length of the corresponding adjacency columns from the top and from the bottom.

Now Adjacency Ratio can be calculated (2) as follows:

\[ \text{Adjacency Ratio} = \frac{\text{sum} \left( L_T \right)}{\text{sum} \left( L_B \right)} \times \text{Signature Occup.} \]  

Adjacency ratio will be more or less the same for every signature of a person. This is shown in Figure 9.

Figure 9. Illustration of finding the length and ratio of Adjacency Columns [14]

14. “Heaviness of signature”

Calculation of this feature is based on previous feature. We calculated the absolute difference between the top and the bottom adjacency columns (Figure 10).

Figure 10. Illustration of “heaviness” procedure

Then we summed up the difference. After that, we calculated the ratio of difference between sum and background area (black pixels). We used formula (3) based on previous figure.

\[ \text{ALR} = \frac{\text{sum(height) - (L_T + L_B)}}{\text{number of black px}} \]  

15. Ratio of first vertical pixel to height of image.

In the bounding box image we calculated distance between top left pixel and the y coordinate of the signature first pixel. Then, this distance is divided with height of bounding box. In this way we can find out where the first vertical pixel in our picture is located in relation to height.

16. Number of Closed Loops

It represents the number of closed regions in a signature. We first find the closed region using [9] and then we subtracted with original image and counted contours that left out. In figure we can see 2 closed loops (Figure 11).
There are a significant number of other global features in literature that we can use, such as: height, pure width, pure height, local slant angle. Since we normalized the width of signature area, some features like width and height were redundant. Instead of a pure width, we used width of two main parts of a signature (name and surname).

B. Grid Features

Each signature image is divided into 18 equal grid regions. From these grid regions, grid characteristics as local features are estimated, such as: width, height, area of signature pixels of each grid region, center of gravity, horizontal and vertical projections, etc. [13]. The global features can also be considered as local features for each grid region. In this implementation we only used number of pixels that belong to a signature for every particular region as a grid feature. We divided signature bounding box to 18 regions so in total we collected 18 features per signature.

C. SIFT Features

SIFT operator is an image local feature descriptor introduced by David G. Lowe in 2004[17]. It is one of the most popular local features. SIFT algorithm detects features in the scale space and confirms the location and scale of key points. Then, it sets the direction of gradient as the direction of the point. Thus the scale and direction invariance of the operator are realized. SIFT is a local feature, which is invariant to rotation, scaling and change of light and stable in a certain extent of changes in visual angle, affine transformation and noise. It ensures specificity and abundance, so it is applicable to fast and accurate matching among mass feature data. Its large quantity ensures that even a few objects can generate a number of SIFT features, high-speed satisfies the requirement of real-time, and extensibility makes it easy to combine with other feature vectors.

We used SIFT features using Bag of Words Model. This model method is extensively used in image classification and retrieval [18]. The main steps are as following:

1. Detect key points through image division or random sampling etc.
2. Extract the local features (SIFT) of the image and generate the descriptor.
3. Cluster these feature related descriptor (usually via K-means) and generate visual vocabulary, in which each clustering center is a visual word.
4. Summarize the frequency of each visual word in a histogram.

This method was applied to grayscale images. After extracting the features, descriptor was clustered using k-means algorithm where k = 48. In this way we obtained 48 features.

When we obtained a set of global, grid and SIFT features, we combined these set of feature sets into a feature vector of 82 elements.

V. DATA TRAINING

In this paper, offline identification of signatures was made with the support of vector machine (SVM). The SVM, a learning method introduced by Vapnik et al. [19][20], tries to find an optimal hyper plane for separating two classes. Therefore, the misclassification error of data both in the training set and the test set is minimized. Basically, SVM was used for linear separation of the two classes. When data are non linearly separable, a kernel function is used as polynomial function, radial basis function (RBF) or multi layer perception. The classification based on SVM involves training and testing stages. The training stage was used to find the optimal parameters. Hence, the two parameters should be determined: the kernel parameter and the regularization parameter. These two parameters were found experimentally depending on the dataset. The testing stage allowed evaluating the robustness of the classifier parameters [21]. For finding parameters and for evaluation of robustness of classifier we used signatures from acquisition. That is, we used 15 collected signatures per student for training (15*30=450) and 5 signatures for testing (5*30=150). Various kernels were tried, and the best results were achieved using a RBF kernel. Parameters C and γ for RBF kernel were determined through a grid search in Weka, machine learning software. Overall prediction accuracy for learning model was 98%.

VI. PERFORMANCE EVALUATION

In order to test proposed system with real exam documents, we collected 145 various exam documents of 30 students (from 5 students, we collected 4 documents and for rest it was 5 per student). After feature extraction and learning model on signatures from database, we started identification/ prediction of the authors using SVM as classifier. For measuring performance concerning of the identification of signatures and its affiliation in a particular set of students we used accuracy measurement. The testing was carried out individually for each set of features as well as combination of three selected feature sets.

Results of evaluation are shown in Table 1.
TABLE 1. RESULTS OF IDENTIFICATION OF AUTHORS ON EXAM DOCUMENTS BASED ON SIGNATURE

<table>
<thead>
<tr>
<th>Method</th>
<th>Correct Prediction</th>
<th>Error Prediction</th>
<th>Average Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Features</td>
<td>104</td>
<td>41</td>
<td>71.72%</td>
</tr>
<tr>
<td>Grid Features</td>
<td>90</td>
<td>55</td>
<td>62.97%</td>
</tr>
<tr>
<td>SIFT Features</td>
<td>84</td>
<td>61</td>
<td>57.93%</td>
</tr>
<tr>
<td>Global + Grid + SIFT features</td>
<td>129</td>
<td>16</td>
<td>88.97%</td>
</tr>
</tbody>
</table>

From the table, it can be concluded that global structural and statistic characteristics of signature were the best discriminators with 71.72% accuracy. Set of SIFT and Grid features had a significantly worse accuracy. However, when these individual feature set are joined into a single vector of features we gained significantly in performance and get the best accuracy of the proposed model with 88.97% correctly identified signatures and consequently authors of exam documents.

VII. CONCLUSION AND FUTURE WORK

In this paper we examined the problem of classification of exam documents based on the vector of global, local and SIFT features of students’ signatures. The best performance was achieved by using local, grid and SIFT features combination, where average accuracy was 88.97%. Although the learning model showed excellent performance, when we tested the model on real documents it showed significantly lower results. We believe that problem was during acquisition where we emphasized to students that during the signing on rectangular areas, the boundary lines should not be touched. But on real documents they generally crossed the baseline with lower parts of signature.

However the results are promising, considering that our approach is relatively simple. Certainly, the results can be facilitated by using some other feature set (Histogram of oriented gradients, Dynamic Time Warping, Sparse Coding, Features using graph theory) and with some other classifier or with ensemble of classifiers.

In general, we can say that the proposed system will be of great help when used with the system for the recognition digits of the students’ ID numbers.

REFERENCES

[1] Fotak T., Baća M., Koruga P., Handwritten signature identification using basic concepts of graph theory , WSEAS TRANSACTIONS on SIGNAL PROCESSING