Approach for ID Recognition in Person Identification Systems

Mariana T. Stoeva, Sivo V. Daskalov , Krasimir D. Dimitrov

***Abstract*—** The proposed approach enables automatic recognition from Identification Document (ID) images. The approach is invariant in terms of scaling and rotation of the input. It includes three main stages of data processing: initial image preparation, automated setup of the parameters of a chosen filter and main working mode of the system for ID identification.

***Keywords*—**Computer aided vision, Text recognition, Image processing.

I. Introduction

The development of identification and verification systems nowadays grows in importance. The requirements towards these systems have also significantly risen with the implementation of such in various spheres of our everyday life. The way towards satisfying these set requirements is searched for both in development of hardware and software.

This paper presents an approach that could be applied in identification and verification systems. It enables automatic recognition and data extraction from images of identification documents. The approach is invariant in regard to the transformations rotation and scaling. It includes three main stages of data processing – initial ID image preparation, automated setup of the parameters of a chosen filter and main working mode of the system for ID identification. In the development of the approach we have used filtering techniques in the processing of the images until an optimal state for text recognition has been reached. The idea of automatic filter tuning has been developed.

One of the less popular uses of the technology we have developed is that of text recognition in computer aided vision. However, a series of transformations needs to be applied to the original image in order to enable the use of recognition software. Furthermore, these modifications of the image are not universal, thus requiring the development of software to automatically adjust its filters in relation to the given task.

M. T. Stoeva, Associate Professor in Technical university of Varna, Department of Computer Science and Technology, 9010 Varna, Bulgaria, (phone: +39 894 651 799; e-mail: mariana_stoeva@abv.bg).

S. V. Daskalov, student in Technical Uiversity of Varna, Department of Computer Science and Technology, 9010 Varna, Bulgaria; (-mail: sivodaskalov@gmail.com)

K. D. Dimirov, student in Technical Uiversity of Varna, Department of Computer Science and Technology, 9010 Varna, Bulgaria (e-mail: kr.dimitrov93@gmail.com).

II. Theoretical base

The approach in this paper uses different theoretical base for its various stages.

The initial processing stage consists of transformation of the colored image to grayscale, as shown in [4] and extraction of the rectangle that includes all visual data that needs to be read. The rectangle parameters are then calculated and a composition of transformations including rotation and translation in order to provide invariance of the results in terms of possible positioning of the scanned ID.

The steps towards automated parameter tuning of a chosen filter and operations performed in the main working cycle of the system are based on the methods of image processing, and the text recognition techniques, explained in detail in [1] and [5]. Also used are tools for image quality improvement as described in [2]. A combination of Tikhonov quadratic filter and a total variational filter has been used as shown in [3]. These two filters can be automatically selected by a soft threshold function.

Digital image processing offers a wide variety of algorithms, of which a collection of filters will be used to remove impulse, white and Gaussian noise [1][5]. Among those are the Median, Adaptive Threshold and Bilateral filters. Automatic tuning of the threshold function for the adaptive threshold filter is shown in [3]. In our research to optimize performance speed the filter parameters will be selected by processing a sample group of images along with their text data and storing the sets of parameters that lead to a correct reading of all sample images in a database.

Optical Character Recognition (OCR), is the mechanical or electronic conversion of scanned or photographed images of typewritten or printed text into machine-encoded/computer-readable text. It is widely used as a form of data entry from a paper data source. It is a common method of digitizing printed texts so that they can be electronically edited, searched, stored more compactly, displayed on-line, and used in machine processes such as machine translation, text-to-speech, key data extraction and text mining. OCR is a field of research in pattern recognition, artificial intelligence and computer vision.

Euclidean distance is used for the identification of the picture in the input ID image when compared to data stored in the database for the same person.

III. Approach for ID text recognition

The approach consists of the following three phases:

A. Preprocessing of input ID image

The first phase includes next processes:

- Transform the color image to gray scale. Locating the bounding box of the ID's visual data and evaluating its vertices (P1, P2, P3, P4).
- Calculating the angle α, between the larger edge of the bounding-box and the X axis of the input device.
- Transformation of the input image using the composition given with equation (1), which includes rotation around the vertex with lowest Py coordinate at angle (-α).

$$Pnew = P_{Y\min} + \begin{Vmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{Vmatrix} (P - P_{Y\min}) \quad , \quad (1)$$

where P are the vertices (P1, P2, P3, P4) of the rectangle that includes the visual data of the ID, PYmin – the vertex with the lowest y coordinate, α - the angle between the larger edge of the bounding-box and the X axis of the input device.

- Invariance in scaling is reached with the usage of normalized coordinates, in relation to the length of the edges of the bounding-box rectangle. Normalized coordinates are also used for the extraction of the various sub areas with text of interest and the photograph from the ID.

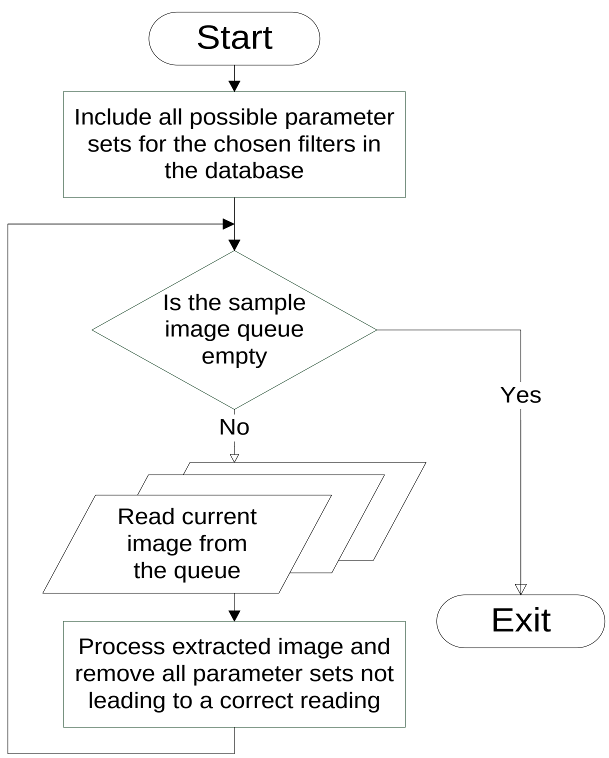

Fig. 1. Block diagram of the initial filter database creation

B. Initial filter database creation

A sample group of images is chosen including a wide variety of noise and contrast differences and their string data is manually input by the operator.

The filter database initially includes all possible sets of filter parameters for the chosen filters. Every sample image is processed by all each of the parameter sets in the database and then OCR is performed for all cases. If the results from the reading differ from the manually input data for the original image, the parameter set leading to such a reading is removed from the database Fig.1. This preparation stage of the approach is time consuming but improves significantly the performance of the processing.

C. Main cycle of ID processing

The processes performed by the algorithm are depicted in Fig. 2.

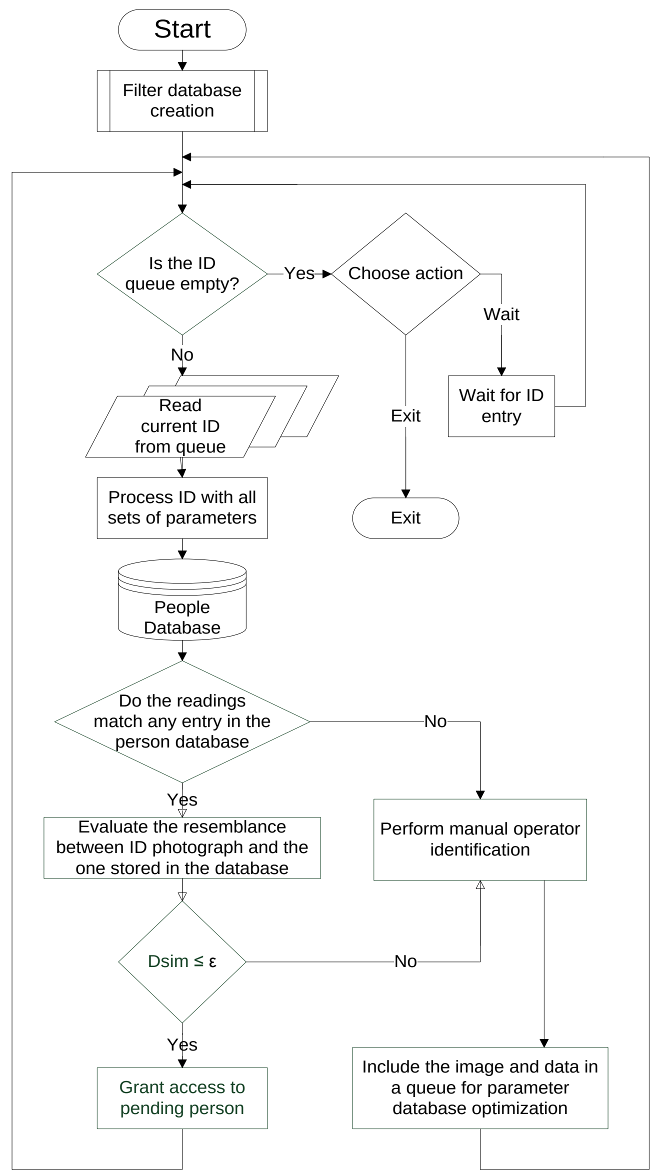

Fig. 2. Block diagram of the main cycle of person identification

Each of the sets in the parameter database is used to process the image and OCR is performed for each result. If the readings match an entry in the person database, the image is considered successfully read. However, if this is not the case identification is redirected to an operator and the data input by him along with the ID is stored in a queue for future parameter database optimization.

When the ID queue is empty or on demand, the process of parameter database optimization can be performed. Images and data from the PDO queue are processed with all possible sets of parameters and filters until the data is correctly read from the image. When that occurs, the parameter set leading to the correct reading is added to the parameter database for future reference. This process is tedious and should be performed during the inactive hours of the day.

Common inconsistencies could be observed, resulting from the various fonts and ID specifics. In order to reach a correct reading, the resulting collection of strings read by the OCR system needs to be passed to a correction function.

In case of a successful reading, the image corresponding to the read text data is extracted from the People database and compared to the current input photograph. The resemblance between the two photographs is evaluated by the distance measurement shown with (2)

$$Dsim(Q,I) = \frac{1}{n}\sqrt{\sum_{j=1}^{n}(q_j - i_j)} \quad (2)$$

where Dsim(Q,I) – similarity distance measure between two n-dimensional descriptions, Q={q1 , … qn) –description of face picture from input ID, I={i1 , … in) – description of photograph in the database, n –number of descriptions.

Dsim(Q,I) =0 in case of an exact match, Dsim(Q,I)< = ε, in case of an acceptable resemblance and Dsim(Q,I)> ε, when there is little similarity between the images. The error threshold ε is experimentally set for the current application.

IV. Experimental results

A. Experimental setup

The experiments have been conducted on Bulgarian ID cards Fig. 3.

Fig. 3. Image -example of a Bulgarian ID card

For image processing we've used OpenCV [6], an open source C++ library, which contains powerful tools for digital image editing. For optical character recognition we've used Tesseract OCR [6], also an open source C++ library. The images used are scanned with resolution of 300 dpi.

The experimental software was executed on a PC with the following characteristics: CPU - Intel Core i5 3570 @ 3.40GHz, RAM - 8.00GB Dual-Channel DDR3 @ 1.33GHz, GPU-NVIDA GeForce GTX770

B. Filter usage

The three filters of our choice are the median, adaptive threshold and bilateral filter [6]. The example results of applying Median filter with size of matrix = 3 are unreadable by Tesseract OCR. The results of applying the bilateral filter with parameters diameter of each pixel neighborhood that is used during filtering = 15, filter sigma in the color space= 60 and filter sigma in the coordinate space = 60 is readable by Tesseract OCR.

Fig. 4. Results of applying adaptive threshold with unsuitable parameters

The results of applying Adaptive Threshold Filter with size of matrix = 21 and calculation constant = 10 are shown on Fig.4. and are unreadable by Tesseract OCR.

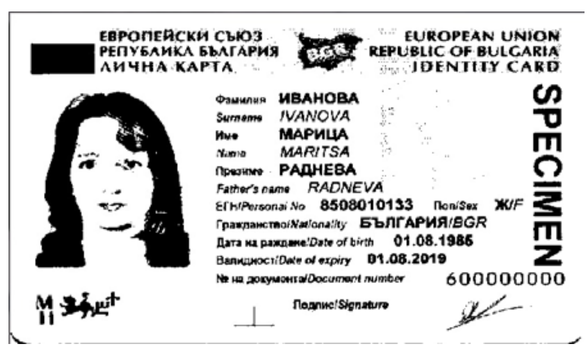

Fig. 5. Results of applying adaptive threshold with suitable parameters

The results of applying Adaptive Threshold Filter with size of matrix = 149 and calculation constant = 18 are shown on Fig.5. and are readable by Tesseract OCR. Best results for Bulgarian ID are reached by using the adaptive threshold and therefore it is the main focus of the majority of our experiments. The initial parameter database consists of all pairs of values for size of matrix (Size) and calculation constant (C) where Size is an odd number and between 1 and 255 and C is between 1 and 64.

C. Optimization of the filter database

A collection of 5 sample Bulgarian ID images has been chosen for the parameter database optimization. The initial count of the parameter sets was reduced from 16320 to 21 after processing all of the sample images. The time required to process each subsequent image was reduced significantly as depicted in Table I and Fig. 6.

Table I. Reduction in the number of entries in the parameter database

Image №	Number of parameter sets	Time eclapsed (min)
Start	16320	X
1	698	48
2	312	20
3	140	9
4	56	3
5	21	1

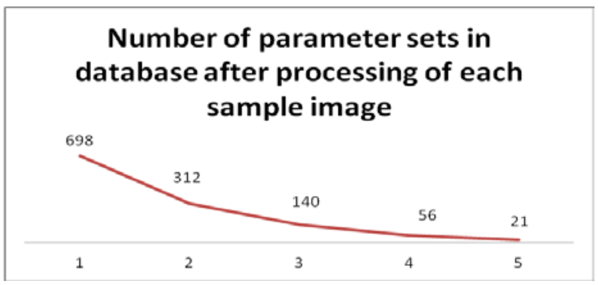

Fig. 6. Reduction in the number of entries in the parameter database

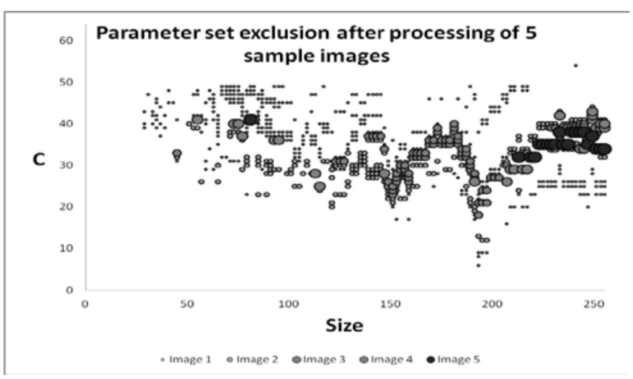

Fig. 7. Distribution of the remaining parameter sets after processing of each sample image

On Fig. 7. is shown the distribution of the parameter sets in the database for every round of optimization. The remaining values after the processing of each subsequent sample image are shown with a larger and darker circle.

D. Evaluation of resemblance between various images

The aim of this experiment is to calculate a suitable value for the error threshold ε for the similarity distance measurement Dsim when comparing the input to the database image. For the experiment we have used three modified versions of the original image and images of two different people. The results are shown in Table II.

As a result from this experiment we have evaluated the error threshold Dsim ε=0.1 in regards to the specifics of Bulgarian ID. The aforementioned threshold should be evaluated for each specific application since there is no optimal value.

Table II. Values of Dsim when comparing with the original image

Image	 original					
Distance Dsim	0	0.0783	0.0169	0.0165	0.2631	0.4135

V. CONCLUSION

The suggested approach is invariant in terms of rotation and scaling, and offers efficiency for usage in identification and verification systems. A well defined filter parameter database significantly reduces the time needed to process a random image from the ID queue for correct readability. Adaptive threshold seems to be most suitable for filtering the design specifics of Bulgarian ID. However, this is not necessarily the case for other types of ID. The aforementioned approach can be modified in order to be applied to various types of objects apart from ID. The approach is adaptive, which lowers the possibility of a misreading when working with bad quality or damaged images.

One of the aims towards improvement of readability is to create a custom filter in order to reach optimal results when manipulating a certain type of ID. The way parameter sets are added to or removed from the database can also be improved in the future. Each parameter set could be evaluated in a way as to secure or deny its position in the database. The expectations towards the approach need to be either confirmed or disproved by observation of the behavior of the created software when processing large amounts of data.

REFERENCES

[1] R. Gonzales., R. Woods, Digital Image Processing, Hardcover, 2007 Gatesmark Publishing A Division of Gatesmark,C www.gatesmark.com
[2] A. Wubuli, Zhen-Hong, Xi-Zhong Jia;, Y. Qin;Jie; N. Kasabov, Medical Image Enhancement Based on Shearlet Transform and Unsharp Masking, Journal of Medical Imaging and Health Informatics, Volume 4, Number 5, American Sc. Publishers, October 2014, pp. 814-818(5)
[3] H. Yan , X. Jianlou , B. Jian , H. Yu, Image decomposition combining a total variational filter and a Tikhonov quadratic filter, October 2013, Multidimensional Systems and Signal Processing, http://link.springer.com/ article/10.1007/s11045-013-0260-5
[4] J. Zhengmeng , K. Ng. Michael, A contrast maximization method for color-to-grayscale conversion, Multidimensional Systems and Signal Processing, August 2014, http://link.springer.com/ article/10.1007/s11045-014-0295-2
[5] И.С. Грузман, В.С. Киричук, *Цифровая обработка изображений в информационных системах*. – Новосибисрк: Изд-во НГТУ 2002.
[6] http://docs.opencv.org/