High-speed OCR algorithm for portable passport readers

Victor Bessmeltsev, Evgeny Bulushev, Nikolay Goloshevsky
Institute of Automation and Electrometry SB RAS
Novosibirsk, Russia
e.d.bulushev@gmail.com

Abstract

The modern portable passport readers are able to acquire high quality images (up to 450 dpi), but have poor performance embedded computational facilities, because of the requirement on long battery life. For the extraction and recognition of machine-readable document codes the high-speed and accurate projection-based OCR algorithm is developed. Profile projection is calculated for middle columns of document image and is smoothed using the moving average filter. The slant compensation is conducted, using the angle between the horizontal line and the line connecting centers of the extreme individual codes. For the embedded processor (AMD Geode 500 MHz 128 RAM) the algorithm shows 100% accuracy and processing time less than 1 second for 300 and 450 dpi images.

Keywords: *MRZ, recognition of passports, OCR, high-speed segmentation, slant compensation, portable passport readers, profile projection*

1. INTRODUCTION

A biometric passport and visa contains picture and user information areas for visual inspection and the machine-readable zone (MRZ) with individual codes, specified in the standard ICAO 9303, which can be captured and processed using optical character recognition (OCR) systems. Currently the relevance of the development of portable passport readers for rapid examination of documents on board ships, road and railway transport, border crossings increases. Basic requirements for such devices – light weight (less than 3 kg), several hours of battery life, short time of document processing (less than 1 second) and wireless data transmission to the server for additional inspection. Such devices include embedded computing facilities, sensor module to acquire document image and modules for reading a contactless chip and data transmission. A document image is processed using both OCR methods and face recognition methods. In order to improve the reliability of the recognition and the determination of forged documents images are acquired in high-resolution (up to 450 dpi) and in several spectral ranges (UV, visible, IR).

Portable passport readers exploit relatively poor performance processors in order to remain low-power and have long battery life. According to our experience the commercial solutions for mobile passport readers recognize MRZ in more than 3 seconds. The efficiency of standard solutions for recognition of individual codes was estimated, as with the specialized software, based on the algorithm [7], and the general-purpose software – CuneiForm OCR version 12. The average recognition time of high resolution (450 dpi) document image is respectively 7 and 3 seconds at the portable passport reader with the embedded processor AMD Geode 500MHz and 128 MB RAM. Such performance is unacceptable, and therefore the development of algorithm with the accuracy higher than 99% and the processing time less than 1 second is required.

2. OVERVIEW

The document image processing consists of the following steps: detection of MRZ and code sequence blocks, extraction and slant compensation of document codes, individual codes recognition. In the specialized algorithm [7] code sequence blocks are extracted from the document image using Sobel masking, horizontal smearing and contour tracking. Template matching, morphological methods [8], fuzzy logic, artificial neural networks [9] and crosscheck with the visual inspection area [9] are used for the individual code recognition. The implementation [6] of these algorithms is efficient (accuracy above 98%), robust to noisy and low detail images (resolution below 300 dpi, skew angle up to 10 degrees), but demanding to the computational resources. Generally such high robustness is excess for modern mobile passport readers, due to their ability to acquire high resolution images (up to 450 dpi) with high signal to noise ratio [2] and small maximum image skew (less than 1 degree).

Printed and hand-written documents are segmented using projection profiles method [11]. The vertical projection profile is obtained by summing intensity of pixels along the horizontal axis for each image row. The profile can be used to segment text lines as average intensity of background and symbols differs. The projection-based method is applied to the black-and-white image to extract MRZ and picture area [10]. The projection-based method is less resource-intensive than method, based on horizontal smearing, due to applying simple algorithms of image transformation and analyzes. Meanwhile, the smearing of vertical profile, caused by image skewing and noise, makes this technique less reliable and accurate.

The accuracy of individual codes recognition can be improved for skewed document images using slant compensation. The Hough transform is performed to determine skew angle [16]. The accuracy of 1-2 degrees was achieved for the low resolution images (150 dpi) with irregular illumination, and maximum skew angle up to 10 degrees.

In Karateev et al. [5] is argued that individual codes recognition is the most resource-intensive step of the document image processing, because of the numerous template matching operations. To accelerate this step the rapid template-based matching is performed using font image, generated with special software. According to our experience the speed of individual code recognition can be boosted, using parallelization, both at the CPU and at the hardware level (e.g. using FPGA). Meanwhile for the high resolution images MRZ detection and code sequence blocks extraction takes longer time and is not easily parallelized.

So, methods, based on the horizontal smearing, and the profile projection exist for MRZ detection and individual codes extraction. The first is highly robust but relatively slow, the second is efficient but less robust in case of noise and document slanting. In this paper the high-speed and accurate algorithm is developed for individual codes extraction from high resolution (more than 300 dpi) document images with small slant angle (less than 1 degree). The algorithm is based on the transformation of vertical profile projection and skew angle detection, using the individual codes position.

3. EXTRACTION AND SLANT COMPENSATION OF DOCUMENT CODES

The MRZ is detected using the vertical projection method. Unlike the standard methods, which project the intensity of source image or black-and-white image, we use the projection of red channel of source image. In Figure 1 the lower part of the document image is depicted. The vertical projection profile of red channel is shown in Figure 2. The right extreme sharp gap corresponds to the bottom border of the document; the two next recessions match the centers of the code sequence blocks. The bottom part of gap corresponds to the center of text line.

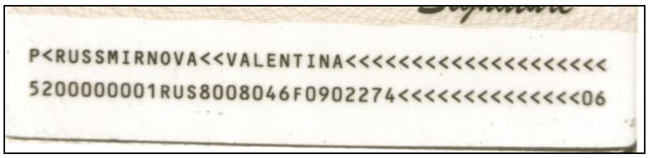

Figure 1: MRZ and individual codes.

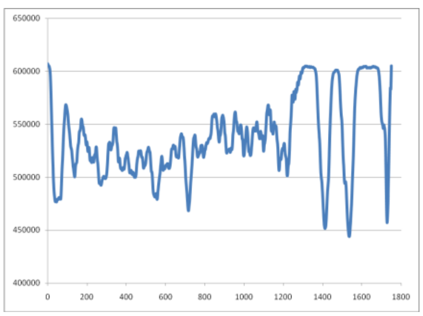

Figure 2: Projection of red channel. In horizontal direction - row number, in vertical direction – the total sum of red color intensity for row. Three rightmost drops correspond to code sequence blocks and border of the document.

Two gaps, corresponding to code sequence blocks, intersect a little for small skew angles (less than 1-2 degree) of document image and can be detected independently. The complexity of automatic detection of the required minima is caused by plenty of local extrema, caused by skew or moderate fluctuations of intensity of text lines. The criterion of the predefined depth of profile gap is not adaptive, because of the various intensity distributions and features of different countries documents. The statistical criterion, calculated based on the histogram of distance between adjacent local minima, is used in [1] to determine potential segmentation points in printed documents.

In order to increase robustness of segmentation an image may be divided into vertical strips and profiles sought inside each strip [18]. The Gaussian filtering can be used to reduce the number of local minima of profile [12]. We applied both approaches with small time-saving modifications. In order to increase computational speed the moving average filter (MAF) is applied instead of the Gaussian smoothing. The vertical projection is computed only for middle strip, that reduces the volume of processed information and increases the speed of document analyze. Generally the duration of data analyze can be reduced up to one order.

The resulting profile has less local extrema (Figure 3). Code sequence blocks are extracted using the criterion of monotonicity. First two profile points, for which the previous values monotonically decrease in more than N consecutive image rows, starting from the bottom of image, are the required extrema. For the demonstrated profile the following coefficients were used (N = 20, the size of the smoothing window = 41, coefficients of constraints of strip are 0.4 and 0.6). Experiments show that the developed algorithm for code sequence blocks extraction is adaptive to small image slanting, change of illumination and diversity of document features.

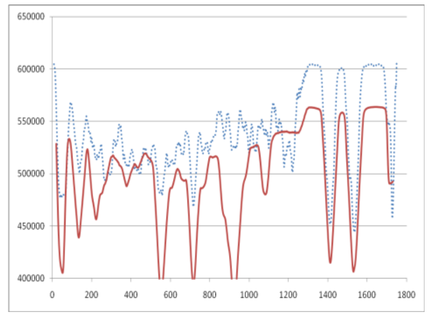

Figure 3: The dotted blue line - the projection of the red channel. The solid red line - the projection of the red channel, calculated for the middle columns, smoothed by MAF, and scaled by “y”

The MRZ image is formed, after the code sequence blocks position is determined. In order to simplify the individual codes extraction the threshold conversion of the MRZ image is performed, using the threshold value, computed automatically using the Otsu method [15]. This method is robust to the average intensity change of document image. Figure 4 shows the result of image conversion.

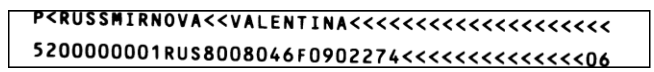

Figure 4: The result of MRZ image binarization

The horizontal position of individual passport codes is determined using horizontal projection of black-and-white image. The vertical position of codes is refined using vertical projection in the vicinity of codes location. The MRZ image and zoom image with red rectangles, allocating the result of individual codes extraction, are depicted on Figure 6.

Figure 5: Individual codes extraction. Red rectangles correspond to document code positions.

Black-and-white MRZ image is adjusted, using slant compensation, in order to improve the efficiency of recognition of individual codes. The Hough transform [4] or image moments

[17] are usually used to estimate a skew angle. However these methods are inefficient in case of geometric distortions of image, which can be significant and caused by imperfection of optical registration system. Even in the case of correction of geometrical distortions the presented algorithms are relatively slow, because a lot of statistical data is analyzed (more than a thousand of points). For this reason the efficient and fast slant compensation method is proposed. The slant of each code sequence block is determined as the angle between the horizontal line and the line between centers of the extreme rectangles of individual codes. Image skew is calculated as the average of two values.

For small skew angles (angle is less than 0.1 degree), image is not adjusted since the recognition accuracy does not change. In this case the algorithm demonstrates high productivity, as a skew angle is defined based on the statistical data reduced in several orders. If skew angle exceeds the preset threshold a document image is adjusted and positions of individual codes are refined using projection method in the vicinity of previous locations. The result of image slant compensation is shown on Figure 6 for the MRZ image and zoom image.

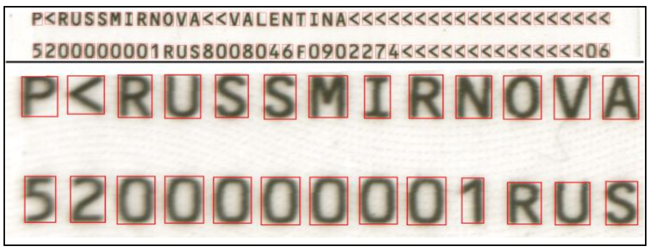

Figure 6: Image slant compensation. Red rectangles correspond to document code positions.

4. RECOGNITION

The recognition step is performed after document codes are extracted, that includes feature extraction and classification. The template matching and line intersections methods can be used to get features of the individual codes [3]. In [14] the general case of feature extraction from the monochrome characters image is considered. We use template-based matching, which is efficient in case of high resolution and low noise document images. The template was generated on the basis of the standard font OCR-B (Figure 6).

Figure 6: Font Template

5. RESULTS

The algorithm is tested on the portable passport reader [2], consisting of image sensor module, single-board industry computer and storage device. Image sensor module is intended to preprocess data acquired by Foveon F13 (14 million elementary photodetectors, organized in matrix 2688 x 1792 x 3 layers with up to 50MHz data output frequency) [13]. Field-programmable gate array (FPGA) controls the matrix and carries out real-time preprocessing of data stream. The main preprocessing stages are linearization, dark frame subtraction, column filtering, bad pixel replacement and color space transformation. Module has external memory connected to FPGA to store intermediate data, calibrating tables and images. Converted data is transferred to processing module connected to image sensor module by parallel 32-bit interface. Image recognition, visualization and storing are performed by industry computer (AMD Geode based). In order to output and transfer information, system is completed with display and wireless LAN.

The algorithm is implemented in C++ and its code is optimized using the static analyzer PVS-Studio and VS 2010 profiler. For the performance evaluation we experimented on the CPU Intel Pentium Dual-Core 2.50 GHz 1.96 GB of RAM and the portable passport reader with embedded computational facilities AMD Geode low power LX800 500 MHz 128 MB RAM. The results of the testing are shown in Table 1.

	PC. Time (ms)	Portable passport reader. Time (ms)
450 dpi	31	460
Image skew. 450 dpi	47	801
300 dpi	25	360
Image skew. 300 dpi	37	530

Table 1: Testing the performance of the algorithm

The first row shows the duration of recognition of high quality images (450 dpi) from the test database of 30 documents. The performance higher than one document in half of second and 100% accuracy was achieved. The second row shows the time span of processing of document images rotated by 1 degree. The required performance and the 100% accuracy were obtained. However the necessity of recalculating of profile projections significantly reduces the calculation speed. The testing has shown that the main time-saving algorithmic improvement is the computation of the vertical projection for the middle strip. It saves about 50 ms or 10% of calculating time. For the described portable passport reader the algorithm is robust to image skew angle less than 2 degree. For bigger skew angles the criterion of monotonicity is not effective, because of the smearing of vertical profile.

For downsampled to 300 dpi images 100% accuracy was achieved and calculation speed increased. So the reduction of resolution of document image may be applied, in order to increase the productivity. For 200 dpi images the accuracy of recognition dropped, because of using template-based approach with general font image. Black-and-white individual codes, extracted by the proposed algorithm on utilized portable passport reader, may be used to specify the template in order to increase the accuracy of recognition.

6. CONCLUSION

This paper proposes the high-speed and accurate algorithm for individual codes extraction from document image. The algorithm is based on segmentation of vertical projection profile, supplemented by MAF smoothing and data reduction to middle strip. The vertical projection profile is computed for the red channel of source image. In addition, the efficient slant compensation method is proposed, using the angle between the horizontal line and the line, connecting centers of the extreme individual codes. The algorithm is 100% accurate for high quality document images (300 dpi and 450 dpi, high signal to noise ratio, skew angle less than 2 degree). For low resolution images (less than 200 dpi), high skew angles (more than 2 degree) or low signal to noise ratio algorithm is inaccurate.

For the embedded processor (AMD Geode 500 MHz 128 RAM) the implementation of the algorithm processes document images in less than 1 second. It revealed at least double performance increase over the algorithm implementation, based on the horizontal smearing. The boost of the document analyze is caused by the algorithmic reduction of the processing data volume and by applying general OCR approaches for the specified task of extraction of individual codes. The performance of the developed software can be improved by at least twice in the case of implementing the algorithm on the hardware level (e.g., FPGA) and recognition of the individual codes in parallel.

7. REFERENCES

[1] A. Antonacopoulos, D. Karatzas, “Document image analysis for World War II personal records”, *First International Workshop on Document Image Analysis for Libraries, DIAL’04*, Palo Alto, pp. 336–341.

[2] V. Bessmeltsev and N. Goloshevsky, “System of high-resolution machine vision with stacked photodiodes structure of photosensitive sensor,” *Proceedings of the IASTED International Conference, ACIT*, Novosibirsk: 2010.

[3] A.V. Bondarenko, V.I. Galaktionov, V.A. Goremychkin, A.V. Yermakov, and S.Y. Zgelto, *Research of the Approaches to Construction of Systems of Automatic Reading of the Symbolical Information Preprint, Inst. Appl. Math., the Russian Academy of Science*, Moscow: 2003.

[4] G. Kapogiannopoulos, “A fast high precision algorithm for the estimation of skew angle using moments,” *Proceeding of IASTED*, 2002.

[5] S.L. Karateev, I.V. Beketova, M.V. Ososkov, V.A. Knyaz, Y.V. Vizilter, A.V. Bondarenko, and S.Y. Zheltov, “Software-Hardware System for Digital Face Imagery Acquisition and Testing for Biometric Documents,” *Вестник компьютерных и информационных технологий*, vol. 2, 2008, pp. 9-14.

[6] K. Kim, “Intelligent immigration control system by using passport recognition and face verification,” *Advances in Neural Networks–ISNN 2005*, 2005, pp. 147–156.

[7] K. Kim, “Passport Recognition Using Enhanced ART2-based RBF Neural Networks,” *IJCSNS*, vol. 6, 2006, pp. 12-17.

[8] K. Kim, “A passport recognition and face verification using enhanced fuzzy ART based RBF network and PCA algorithm,” *Neurocomputing*, vol. 71, 2008, pp. 3202–3210.

[9] K.B. Kim, J.H. Cho, and C.K. Kim, “Recognition of passports using FCM-based RBF network,” *AI 2005: Advances in Artificial Intelligence*, 2005, pp. 1241–1245.

[10] T. Kim and Y. Kwon, “Crosscheck of Passport Information for Personal Identification,” *Graphics Recognition. Ten Years Review and Future Perspectives*, 2006, pp. 162–172.

[11] L. Likforman-Sulem, A. Zahour, and B. Taconet, “Text line segmentation of historical documents: a survey,” *International Journal of Document Analysis and Recognition (IJDAR)*, vol. 9, Sep. 2006, pp. 123-138.

[12] R. Manmatha and N. Srimal, “Scale space technique for word segmentation in handwritten documents,” *Scale-Space Theories in Computer Vision*, 1999, pp. 22–33.

[13] R. Merrill, “Color separation in an active pixel cell imaging array using a triple-well structure,” *U.S. Patent 5,965,875*, 1999

[14] I.A. Mikhaylov, “Image recognition using a radial neighborhood method,” *Computer Optics*, vol. 34, 2010, pp. 399-407.

[15] N. Otsu, “A threshold selection method from gray-level histograms,” *Automatica*, vol. 11, 1975, p. 285–296.

[16] Y.V. Visilter, S.Y. Zheltov, and A.A. Lukin, “Development of OCR system for portable passport and visa reader” *Proceedings of SPIE*, 1999, pp. 194-199.

[17] B. Yu and A. Jain, “A robust and fast skew detection algorithm for generic documents,” *Pattern Recognition*, vol. 29, 1996, pp. 1599-1629.

[18] A. Zahour, B. Taconet, P. Mercy, and S. Ramdane, “Arabic hand-written text-line extraction,” *Proceedings of Sixth International Conference on Document Analysis and Recognition*, 2001, pp. 281–285.