

A METHOD FOR MACHINE-READABLE ZONES LOCATION BASED ON A COMBINATION OF THE HOUGH TRANSFORM AND THE SEARCH FOR FEATURE POINTS

B.I. Savelyev^{1,2}, *N.S. Skoryukina*^{1,2}, *V.V. Arlazarov*^{2,3}

¹Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Moscow, Russian Federation

²Smart Engines Service LLC, Moscow, Russian Federation

³Institute for Information Transmission Problems RAS, Moscow, Russian Federation

E-mail: bsaveliev@smartengines.com, skleppy.inc@smartengines.com, vva@smartengines.com

This article describes a method for machine-readable zones location in document images based on a combination of the Hough transform and the search for feature points. The search for feature points, filtering, and clustering using the Hough transform are described step-by-step. In addition to the machine-readable zone location, we develop a solution for determining the orientation of the zone. This method is designed to meet the requirements for real-time operation on mobile devices. The paper presents the results of measuring the quality of the method on an open synthetic dataset and the operating time on mobile devices. An experimental study on an artificial dataset show that the proposed algorithm allows to achieve a quality of 0,82 in terms of the mean value of the Jaccard indices. The operating time of the proposed algorithm for machine-readable zone location on a mobile device is 6 ms on the iPhone SE 2.

Keywords: machine-readable zone; image analysis; mobile OCR; recognition algorithms.

Introduction

Modern technologies in document processing systems allow to speed up the work of access control points and reduce the number of errors in comparison with manual input. In order to implement recognition systems, a group of experts from the International Civil Aviation Organization (ICAO) developed a standard for passports and other travel documents [1]. On documents developed according to the standard, in addition to the usual data fields, there is a special machine-readable zone (MRZ), which contains complete information about the type of document, the owner, and checksums.

At the moment, in addition to the ICAO, there are several standards for machine-readable documents, and within each of them there are several types of MRZ. Each of the types has fixed characteristics, including geometric ones, such as the aspect ratio of the zone. Some countries adopted its own MRZ standard, which is identical in geometric characteristics to the MRZ. Examples of MRZ types are shown in Figure 1.

The basic solution for standardized document recognition is special passport scanners. They allow to get an image in significantly less time compared to conventional scanners, but the image quality can be significantly lower. Depending on the type of scanner, the resulting images contain either a machine-readable zone directly, or, like a conventional scanner, a document in its entirety without scene elements [2, 3].

The next stage was the use of smartphones with mobile software for scanning documents, including documents with the MRZ recognition function. This allowed to

the same reason, we can not use another popular method that is horizontal morphological blurring on a binarized image in combination with contour analysis [11, 12]. Although some studies [3, 12, 13] considered the case of a slight tilt of the document in the image and its evaluation using the Hough and Fourier transforms for subsequent correction, their applicability was evaluated based on the following assumptions:

- 1) MRZ is the key source of straight lines in the image;
- 2) Binarization successfully turns the text to black and the background to white.

The camera frame can capture various elements of the environment and significant lighting differences, so both of these assumptions are incorrect for the problem in question.

The closest approach to the problem under consideration is the work [14], which uses the approach to the search for MRZ through the selection of text blocks using the morphological filter Toggle Mapping, but the method under consideration shows low frame-by-frame search efficiency even on data synthesized without taking into account the natural bends of the document and lighting.

2. Proposed Approach

We propose an algorithm for localizing a machine-readable zone in an image, which is based on the selection and clustering of feature points, followed by analysis, transformation, and clipping of clusters based on geometric features.

The main steps of the algorithm are shown in Figure 2.

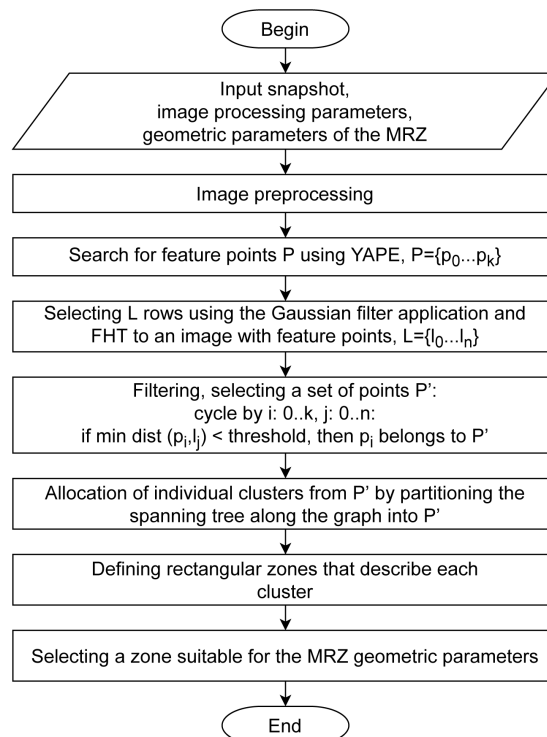


Fig. 2. Algorithm diagram

3. Preprocessing and Searching for Feature Points

The following requirements were formulated for the method of selecting feature points:

- speed;
- high density of text coverage, especially machine-readable (at least one point per symbol);
- ignoring areas with straight borders [15].

In this paper, the YAPE method is used [16,17] (Figure 3).

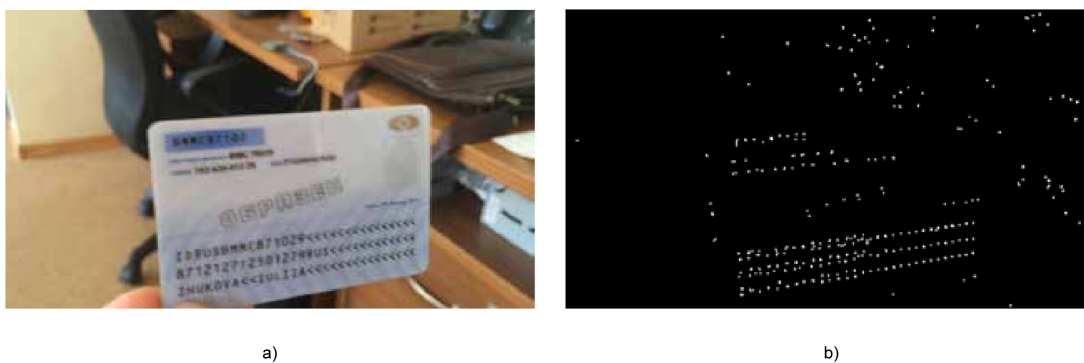


Fig. 3. a) input snapshot; b) search for feature points using YAPE

The input image is pre-scaled to a fixed size (800 px in width), converted to gray, and smoothed using a Gaussian filter with $\sigma = 1$. Scaling, in addition to additional noise reduction, is associated with one of the YAPE parameters that is the maximum radius of the area analyzed for each point. For a comparable result of the algorithm on images of different resolutions, the radius must correlate with the resolution (the higher the resolution the larger the radius), which, unlike pre-scaling, means a significant slowdown (quadratic dependence) on large images with minimal impact on the further algorithm.

4. Filtering Points

In order to discard the points that most probably do not belong to the desired zone we use the assumption that a) the lines of the machine-readable zone are the longest text object on document b) the selected feature point detector does not select points belonging to continuous boundaries (straight lines in an image) and apply the following scheme:

- 1) calculate n candidate lines in an image containing feature points using the fast Hough transform [18];
- 2) for each point p_i , determine the nearest straight line among the candidates l_i . The point is discarded if the distance d_{min} to the nearest straight line is greater than the threshold T . In this paper, we use $T = 0.5 \cdot h_{sym}$, where h_{sym} is the maximum possible height of the MRZ symbol in the fixed-size image.

5. Point Clustering

Next, we combine the points into clusters as follows.

1. Construct a complete graph in which the points are vertices, and the weight of the edge between the points p_i and p_j is calculated using the following formula:

$$w_{ij} = f(\text{angle}) \cdot \text{dist}_{ij}, \quad (1)$$

where dist_{ij} is the distance between each two points, f is some monotone function, angle is the minimal angle between the lines l_i and l_j .

Further this weight will be used to penalize cases that are not similar to parallel MRZ strings.

2. Define a minimal spanning tree of the graph.
3. Divide the tree into several parts (clusters), throwing out the edges whose weight is greater than the threshold (for example, $T = 2 \cdot w_{sym}$, where w_{sym} is the maximum possible width of the MRZ symbol in the image of the fixed size). If the cluster size (the number of points in it) is less than the minimum allowed number of MRZ characters in the string, then discard it.

6. Cluster Analysis & Selection of Resulting Rectangle

For each of the obtained clusters, the angle of inclination to the sides of the image is found as the average of the angles of the lines corresponding to the points included in the cluster. After that, there is a rectangle circumscribed around the cluster points, located at this angle (Figure 4).

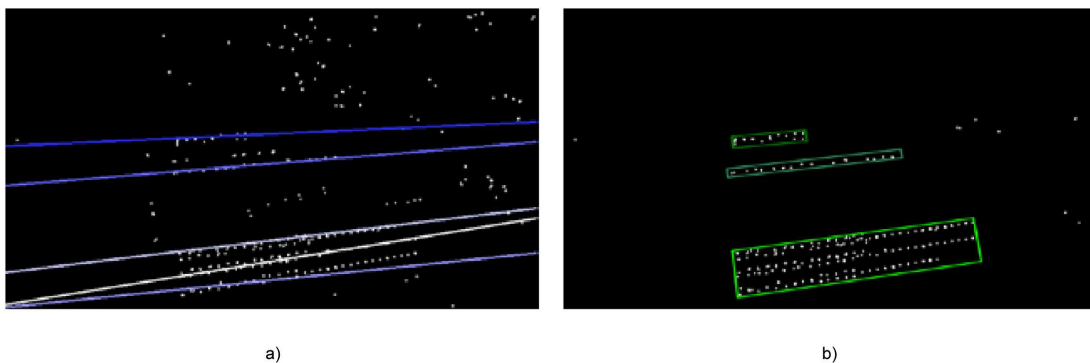


Fig. 4. a) straight lines corresponding to text strings; b) clusters and circumscribed rectangles

From the resulting rectangles, the final one is selected that best matches the known geometric constraints of any of the MRZ types.

In the case where several clusters are structurally identical, we analyze the raster within each cluster. The character “<” is an MRZ-specific separator, i.e. it occurs at least 2 times in the vast majority of cases and usually is not presented in other text fields.

The separator is stably detected as a feature point, so it is sufficient to analyze the local neighborhoods of the points at the beginning and end of the cluster.

To do this, we represent the reference image of the character “<” as a local descriptor. In this paper, we use the binary descriptor RFD [19].

We can not use all 37 characters as a local descriptor that can be used in the MRZ alphabet, because the OCR-B font can occur not only in the MRZ, but also in the document’s fill-in text and such cases are not too rare.

We cut out the neighborhoods around the points from the cluster from the original image, taking into account the size and rotation angle of the cluster, and then calculate the RFD descriptors for them. For each such descriptor, we calculate the Hamming distance with the reference descriptor. We consider the character detection to be successful if the distance is less than the threshold. The weight for the cluster is calculated as follows:

$$\left(1 - \frac{\text{number of mappings}}{\text{number of points}}\right) \cdot \text{mean Hamming distance.} \quad (2)$$

The smaller the weight, the more similar the cluster is to the MRZ.

Due to the fact that the stability of the RFD to rotation is limited (about 15 degrees), and the rotation of the text is unknown at this stage, we calculate “<” descriptors in two rotations: 0 and 180 degrees. The check is performed for each turn independently, and from the answers we choose the one with the greater weight.

7. Results

The developed algorithm was tested using the MRZ recognition module on an image containing an machine-readable zone [20]. To evaluate the quality of the algorithm, the synthetic dataset described in [14] was used. The dataset consists of 456000 artificial images, with a document containing the MRZ. To annotate all images, a different detection algorithm from SmartIDReader was used and results were converted to ground truth. MRZ was not found on all images, so only 422338 images were annotated. Of course, the algorithm evaluation with such ground truth cannot be considered to be accurate, but can estimates its approximate value.

Figure 5 illustrates the example of MRZ detection (red rectangle) and ground truth (green rectangle). The quality was evaluated using the Jaccard metric. The final quality is a mean of Jaccard indices and equal to 0,82 (Figure 6).

The average MRZ detection time is 40 ms and 6 ms on the iPhone 5s and the iPhone SE 2, respectively.

Conclusion

An algorithm for machine-readable zones location in a document image is proposed and tested. According to the test results, the proposed algorithm allows to achieve a quality of 0,82 in terms of the mean value of the Jaccard indices, and the speed of execution of the algorithm on mobile devices allows to use it in real time. The main problems of the algorithm include unstable operation on documents, the content of which generates many feature points on non-text areas, which in turn can incorrectly expand the cluster area. Further research includes the search for solutions to stabilize the algorithm in such cases.

References

1. ICAO Doc 9303 Machine Readable Travel. *International Civil Aviation Organization*, Seventh Edition, Parts 2–7, 2015.
2. Bessmeltsev V., Bulushev E., Goloshevsky N. Highspeed OCR Algorithm for Portable Passport Readers. *Graphicon*, no. 11, 2011, pp. 25–29.
3. Visilter Y.V., Zheltov S.Y., Lukin A.A. Development of OCR System for Portable Passport and Visa Reader. *Proceedings of SPIE*, 1999, no. 3651, pp.194–199.
4. Bulatov K.B., Polevoy D.V., Ilin D.A., Chernyshova Y.S. [Problems of Machine-Readable Zone Recognition Captured with Digital Mobile Cameras]. *Proceedings of ISA RAN*, 2015, vol. 65, no. 3, pp. 85–93. (in Russian)
5. Arlazarov V.V., Zhukovsky A., Krivtsov V., Nikolaev D., Polevoy D. [Analysis of Using Stationary and Mobile Small-Scale Digital Video Cameras for Document Recognition]. *Information Technologies and Computation Systems*, 2014, no. 3, pp. 71–78. (in Russian)
6. Bulatov K., Matalov D., Arlazarov V.V. MIDV-2019: Challenges of the Modern Mobile-Based Document OCR. *Proceedings of SPIE*, 2020, no. 11433, pp. 717–722.
7. Konovalenko I.A., Shemiakina J.A., Faradjev I.A. [Calculation of a Vanishing Point by the Maximum Likelihood Estimation Method]. *Bulletin of the South Ural State University. Mathematical Modelling, Programming and Computer Software*, 2020, vol. 13, no. 1, pp. 107–117. (in Russian)
8. Xiangrong C., Yuille A.L. Detecting and Reading Text in Natural Scenes. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, vol. 2, pp. II-366–II-373.
9. Petrova O., Bulatov K. Methods of Machine-Readable Zone Recognition Results Post-Processing. *Proceedings of SPIE*, 2019, no. 11041, pp. 387–393.
10. Kwon Y., Kim J. Recognition Based Verification for the Machine Readable Travel Documents. *International Workshop on Graphics Recognition*, Curitiba, Brazil, 2007.
11. Kwang-Baek K., Sungshin K. A Passport Recognition and Face Verification Using Enhanced Fuzzy ART Based RBF network and PCA Algorithm. *Neurocomputing*, 2008, pp. 3202–3210.
12. Martin-Rodriguez F. Automatic Optical Reading of Passport Information. *2014 International Carnahan Conference on Security Technology*, Rome, 2014, pp. 1–4.
13. Lee H., Kwak N. Character Recognition for the Machine Reader Zone of Electronic Identity Cards. *2015 IEEE International Conference on Image Processing*, Quebec City, 2015, pp. 387–391.
14. Hartl A., Arth C., Schmalstieg D. Real-Time Detection and Recognition of Machine-Readable Zones with Mobile Devices. *VISAPP 2015 – 10th International Conference on Computer Vision Theory and Applications*, 2015, pp. 79–87.
15. Harris C., Stephens M. A Combined Corner and Edge Detector. *Alvey Vision Conference*, Manchester, 1988, 50 p.
16. Lepetit V., Fua P. Towards Recognizing Feature Points using Classification Trees. *Technical Report IC/2004/74*, 2004, 13 p.
17. Lukoyanov A., Nikolaev D., Konovalenko I. Modification of YAPE Keypoint Detection Algorithm for Wide Local Contrast Range Images. *Proceedings of SPIE*, 2018, vol. 10696, pp. 305–312.
18. Nikolaev D.P., Nikolaev I.P., Nikolaev P.P., Karpenko S.M. Hough Transform: Underestimated Tool in the Computer Vision Field. *European Conference on Modelling and Simulation*, 2008, pp. 238–243.

19. Fan B., Kong Q., Trzcinski T., Wang Z., Pan C., Fua P. Receptive Fields Selection for Binary Feature Description. *IEEE Transactions on Image Processing*, 2014, vol. 23, no. 6, pp. 2583–2595.
20. *Smart 3D OCR MRZ v.1.0*. Available at: <https://smartengines.ru/smart-mrzreader/> (accessed 24 May 2021).

Received March 29, 2022

УДК 51-77:001.895

DOI: 10.14529/mmp220208

МЕТОД ЛОКАЛИЗАЦИИ МАШИНОЧИТАЕМЫХ ЗОН, ОСНОВАННЫЙ НА КОМБИНАЦИИ ПРЕОБРАЗОВАНИЯ ХАФА И ПОИСКА ОСОБЫХ ТОЧЕК

Б.И. Савельев^{1,2}, *Н.С. Скорюкина*^{1,2}, *В.В. Арлазаров*^{2,3}

¹Федеральный исследовательский центр «Информатика и управление» РАН,
г. Москва, Российская Федерация

²ООО «Смарт Энджинс Сервис», г. Москва, Российская Федерация

³Институт проблем передачи информации имени А.А. Харкевича РАН, г. Москва,
Российская Федерация

В данной статье описан метод локализации машиночитаемых зон на изображениях документов, основанный на комбинации преобразования Хафа и поиска особых точек. Описаны поэтапно поиск особых точек, фильтрация и кластеризация с помощью преобразования Хафа. Помимо локализации машиночитаемой зоны разработано решение определение ориентации зоны. Данный метод разработан с учетом требования к работе в режиме реального времени на мобильных устройствах. В работе представлены результаты замеров качества работы метода на открытом синтетическом датасете и времени работы на мобильных устройствах. Экспериментальное исследование на искусственном датасете показало, что предложенный алгоритм позволяет достичь качества 0.82 в терминах среднего значения коэффициента Жаккара. Время работы предложенного алгоритма локализации машиночитаемой зоны на мобильном устройстве составляет 6 мс на iPhone SE 2.

Ключевые слова: машиночитаемая зона; анализ изображений; мобильное распознавание; алгоритмы распознавания.

Литература

1. ICAO Doc 9303 Machine Readable Travel // International Civil Aviation Organization. – 2015. – Edition 7. – Parts 2–7.
2. Bessmeltsev, V. Highspeed OCR Algorithm for Portable Passport Readers / V. Bessmeltsev, E. Bulushev, N. Goloshevsky // Graphicon. – 2011. – № 11. – P. 25–29.
3. Visilter, Y.V. Development of OCR System for Portable Passport and Visa Reader / Y.V. Visilter, S.Y. Zheltov, A.A. Lukin // Proceedings of SPIE. – № 3651. – 1999. – P. 194–199.
4. Булатов, К.Б. Проблемы распознавания машиночитаемых зон с использованием малоформатных цифровых камер мобильных устройств / К.Б. Булатов, Д.А. Ильин, Д.В. Полевой, Ю.С. Чернышова // Труды ИСА РАН. – 2015. – V. 65, №. 3. – С. 85–93.

5. Арлазаров, В.В. Анализ особенностей использования стационарных и мобильных мало-размерных цифровых камер для распознавания документов / В.В. Арлазаров, А. Жуковский, В. Кривцов, Д. Николаев, Д. Полевой // Информационные технологии и вычислительные системы. – 2014. – № 3. – С. 71–78.
6. Bulatov, K. MIDV-2019: Challenges of the Modern Mobile-Based Document OCR / K. Bulatov, D. Matalov, V.V. Arlazarov // Proceedings of SPIE. – 2020. – № 11433. – P. 717–722.
7. Коноваленко, И.А. Оценка точки схода отрезков методом максимального правдоподобия / И.А. Коноваленко, Ю.А. Шемякина, И.А. Фараджев // Вестник ЮУрГУ. Серия: Математическое моделирование и программирование. – 2020. – Т. 13, № 1. – С. 107–117.
8. Xiangrong, C. Detecting and Reading Text in Natural Scenes / C. Xiangrong, A.L. Yuille // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2004. – № 2. – P. II-366–II-373.
9. Petrova, O. Methods of Machine-Readable Zone Recognition Results Post-Processing / O. Petrova, K. Bulatov // Proceedings of SPIE. – 2019. – № 11041. – P. 387–393.
10. Kwon, Y. Recognition Based Verification for the Machine Readable Travel Documents / Y. Kwon, J. Kim // International Workshop on Graphics Recognition. – Curitiba, 2007.
11. Kwang-Baek, K. A Passport Recognition and Face Verification Using Enhanced Fuzzy Art Based RBF Network and PCA Algorithm / K. Kwang-Baek, K. Sungshin // Neurocomputing. – 2008. – P. 3202–3210.
12. Martin-Rodriguez, F. Automatic Optical Reading of Passport Information / F. Martin-Rodriguez // 2014 International Carnahan Conference on Security Technology. – Rome, 2014. – P. 1–4.
13. Lee, H. Character Recognition for the Machine Reader Zone of Electronic Identity Cards / H. Lee, N. Kwak // 2015 IEEE International Conference on Image Processing. – Quebec City, 2015. – P. 387–391.
14. Hartl, A. Real-time Detection and Recognition of Machine-Readable Zones with Mobile Devices / A. Hartl, C. Arth, D. Schmalstieg // VISAPP 2015 – 10th International Conference on Computer Vision Theory and Applications. – 2015. – P. 79–87.
15. Harris, C. A Combined Corner and Edge Detector / C. Harris, M. Stephens // Alvey Vision Conference, Manchester. – 1988. – P. 50.
16. Lepetit, V. Towards Recognizing Feature Points using Classification Trees / V. Lepetit, P. Fua // Technical Report IC/2004/74. – 2004. – 13 p.
17. Lukoyanov, A. Modification of YAPE Keypoint Detection Algorithm for Wide Local Contrast Range Images / A. Lukoyanov, D. Nikolaev, I. Konovalenko // Proceedings of SPIE. – 2018. – № 10696. – P. 305–312.
18. Nikolaev, D.P. Hough Transform: Underestimated Tool in the Computer Vision Field / D.P. Nikolaev, I.P. Nikolaev, P.P. Nikolaev, S.M. Karpenko // European Conference on Modelling and Simulation. – 2008. – P. 238–243.
19. Fan, B. Receptive Fields Selection for Binary Feature Description / B. Fan, Q. Kong, T. Trzcinski, Z. Wang, C. Pan, P. Fua // IEEE Transactions on Image Processing. – 2014. – V. 23, №6. – P. 2583–2595.
20. Smart 3D OCR MRZ v.1.0. Available at: <https://smartengines.ru/smart-mrzreader/> (accessed 24 May 2021).

Борис Игоревич Савельев, научный сотрудник – программист, ООО «Смарт Энджинс Сервис» (г. Москва, Российская Федерация); ведущий программист, Федеральный исследовательский центр «Информатика и управление» Российской академии наук (г. Москва, Российская Федерация), bsaveliev@smartengines.com.

Наталья Сергеевна Скорюкина, научный сотрудник – программист, ООО «Смарт Энджинс Сервис» (г. Москва, Российская Федерация); программист первой категории, Федеральный исследовательский центр «Информатика и управление» Российской академии наук (г. Москва, Российская Федерация), skleppy.inc@smartengines.com.

Владимир Викторович Арлазаров, кандидат технических наук, генеральный директор, ООО «Смарт Энджинс Сервис» (г. Москва, Российская Федерация); и.о. ведущего научного сотрудника, Институт проблем передачи информации имени А.А. Харкевича РАН (г. Москва, Российская Федерация), vva@smartengines.com.

Поступила в редакцию 29 марта 2022 г.