

Метод локализации машиночитаемых зон, инвариантный к условиям съёмки*

Н.С. СКОРЮКИНА

Аннотация. В данной статье описан метод локализации машиночитаемых зон на изображениях документов, полученных с малоформатных цифровых камер. Предложенный метод основан на поиске особых точек с их последующей фильтрацией с кластеризацией с помощью преобразования Хафа. Метод разработан с учётом требования к работе в режиме реального времени на мобильных устройствах. В работе приведены результаты тестирования алгоритма на реальных данных.

Ключевые слова: *машиночитаемая зона, анализ изображений, мобильные устройства.*

Введение

Современные технологии в системах обработки документов позволяют ускорять работу контрольно-пропускных пунктов и уменьшать число ошибок в сравнении с ручным вводом. В интересах внедрения систем распознавания группой экспертов Международной Организации Гражданской Авиации (ICAO) был разработан стандарт для паспортов и прочих проездных документов [1]. На документах, разработанных согласно стандарту, помимо обычных полей данных расположена специальная машиночитаемая зона (MRZ – machine readable zone), которая содержит полную информацию о типе документа, владельце, а также контрольные суммы.

На данный момент помимо ICAO существует несколько стандартов машиночитаемых документов, а в рамках каждого из них – несколько типов MRZ. Каждый из типов обладает фиксированными характеристиками, в т.ч. геометрическими, такими как соотношение сторон зоны. В России принят собственный стандарт МЧЗ [15], который по геометрическим характеристикам идентичен MRZ. Примеры типов MRZ приведены на рисунке 1.

Базовым решением для распознавания стандартизованных документов являются специальные сканеры паспортов. Они позволяют получать изображение за существенно меньшее время в сравнении с обычными сканерами, но качество изображения может быть существенно ниже. В зависимости от типа сканера, получаемые изображения содержат либо непосредственно машиночитаемую зону,

либо по типу обычного сканера документ целиком без элементов сцены [2][3].

Следующим этапом стало использование смартфонов с мобильным ПО для сканирования документов, в т.ч. с функцией распознавания MRZ. Это позволяет существенно удешевить и ускорить оснащение пунктов контроля, но предъявляет дополнительные требования к программной части системы. Малобюджетные мобильные устройства позволяют получать поток изображений невысокого качества и ограничены собственной вычислительной мощностью в использовании существующих алгоритмов детекции и распознавания, поскольку данные машиночитаемой зоны являются персональной информацией, передача которой по сети регулируется законодательством. Таким образом, разрабатываемое ПО должно работать в режиме реального времени даже на маломощных устройствах и поддерживать высокое качество распознавания на изображениях, полученных с камер низкого разрешения.

В статьях [4][5] авторы подробно разобрали проблемы, возникающие при съёмке машиночитаемых документов с малоформатных цифровых камер, такие как:

- проявления «цифрового шума» и артефактов алгоритмов сжатия
- перепады яркости, блики и цветовые искажения
- поворот документа и проективное искажение
- изгиб строк документа

При этом

1. МЧЗ отчётливо видна на изображении, занимает его значимую часть (ширина полной строки МЧЗ составляет не менее $\frac{1}{3}$ ширины кадра) и все буквы полностью различимы;

* Работа выполнена при финансовой поддержке РФФИ (проекты №№17-29-03370 и 17-29-03161)

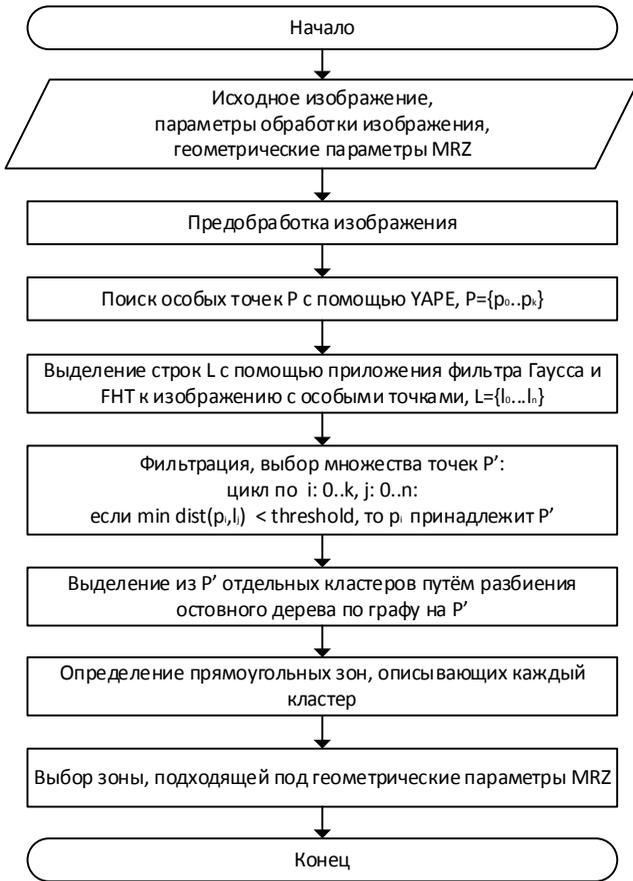


Рис. 2. Схема алгоритма

Входное изображение предварительно масштабируем к фиксированному размеру (800 px по ширине), переводим в серое и сглаживаем с помощью фильтра Гаусса. Масштабирование, помимо дополнительного подавления шумов, связано с одним из параметров YAPE - максимальным радиусом области, по которой идёт анализ для каждой точки. Для сопоставимого результата работы алгоритма на изображениях разных разрешений радиус должен коррелировать с разрешением, что, в

отличие от предварительного масштабирования, означает значимое замедление на больших изображениях при минимальном влиянии на дальнейший алгоритм.

3. Фильтрация точек

Отбросим точки, заведомо не принадлежащие искомой зоне. Для этого используем знание о том, что а) строки машиночитаемой зоны - наиболее протяженный текстовый объект на документе б) выбранный детектор особых точек не выделяет точки, принадлежащие непрерывным границам (прямым на изображении) и применим следующую схему:

1. Вычислим на изображении, содержащем особые точки, n прямых-кандидатов, используя быстрое преобразование Хафа [13].
2. Для каждой точки p_i определяем ближайшую прямую среди кандидатов l_i . Точка отбрасывается, если расстояние d_{min} до ближайшей прямой больше порога T . В данной работе используется $T = 0.5 * h_{sym}$, где h_{sym} - максимально возможная высоты символа MRZ на изображении зафиксированного размера.

4. Кластеризация точек

Объединим точки в кластеры следующим образом:

1. Построим полный граф, в котором точки будут вершинами, а вес ребра между точками p_i и p_j будет рассчитываться по следующей формуле:

$$w_{ij} = f(\text{angle}) * \text{dist}_{ij}$$

где dist_{ij} - расстояние между каждыми двумя точками,

f - некоторая монотонная функция,

angle - угол между прямыми l_i и l_j

Таким образом мы оштрафуем случаи, не похожие на параллельные строки MRZ.



а)



б)

Рис. 2. а) Исходное изображение б) выделение точек с помощью YAPE

2. Определим на графе минимальное остовное дерево
3. Разобьём дерево на несколько частей (кластеров), выкинув рёбра, вес которых больше порога (например $T = 2 * w_{sym}$, где w_{sym} – максимально возможная ширина символа MRZ на изображении зафиксированного размера). Если размер кластера (количество точек в нём) меньше, чем минимально допустимое количество символов MRZ в строке, то отбрасываем его.

5. Анализ кластеров, выбор результирующего прямоугольника

Для каждого из полученных кластеров находится угол наклона к сторонам изображения как среднее из углов прямых, соответствующих точкам, входящим в кластер. После этого находится описанный вокруг точек кластера прямоугольник, расположенный под этим углом.

Из полученных прямоугольников выбирается тот, который лучше всего соответствует известным геометрическим ограничениям какого-либо из типов MRZ.

6. Результаты

Тестирование разработанного алгоритма проходило с помощью модуля распознавания MRZ на изображении, содержащем исключительно машиночитаемую зону [14]. В работе модуля выделены следующие шаги: 1) сегментация зоны на строки; 2) сегментация строк на символы; 3) распознавание символов; 4) исправление результата с учётом контекста.

Будем считать, что зона найдена успешно в том случае, когда оба этапа сегментации прошли успешно (а значит, успешно определён тип MRZ).

Для тестирования было составлено 2 набора данных:

- а) Позитивная выборка, состоящая из 3300 изображений машиночитаемых документов, снятых в различных условиях и с использованием камер разных мобильных устройств. Разрешение изображений варьируется от 3761x2353px до 560x540px. Каждому изображению сопоставлена разметка - настоящие координаты окаймляющего прямоугольника для MRZ и значения строк.
- б) Негативная выборка (для оценки, состоящая из 3000 изображений документов, с синтетически затёртой машиночитаемой зоной либо не содержащей её вовсе. Для данных документов прямоугольники в разметке были описаны вокруг произвольных текстовых полей.

Табл. 1.

Доля успешных сегментаций

	Зона из разметки	Зона, полученная с помощью алгоритма
Позитивная выборка	99.67%	95%
Негативная выборка	0%	0%

Результаты тестирования, приведённые в таблице 1, следует интерпретировать следующим образом:

- 1) На негативной выборке отсутствуют успешные сегментации, т.е. у модуля, использованного для тестирования, нет ложноположительных ответов.
- 2) На положительной выборке выборке успешно отсегментировано 99.76% реальных зон и 95% зон, найденных с помощью предложенного метода. Таким образом, реальное качество работы предложенного метода $\geq 95\%$.

В таблице 2 приведены результаты оценки времени работы предложенного алгоритма.

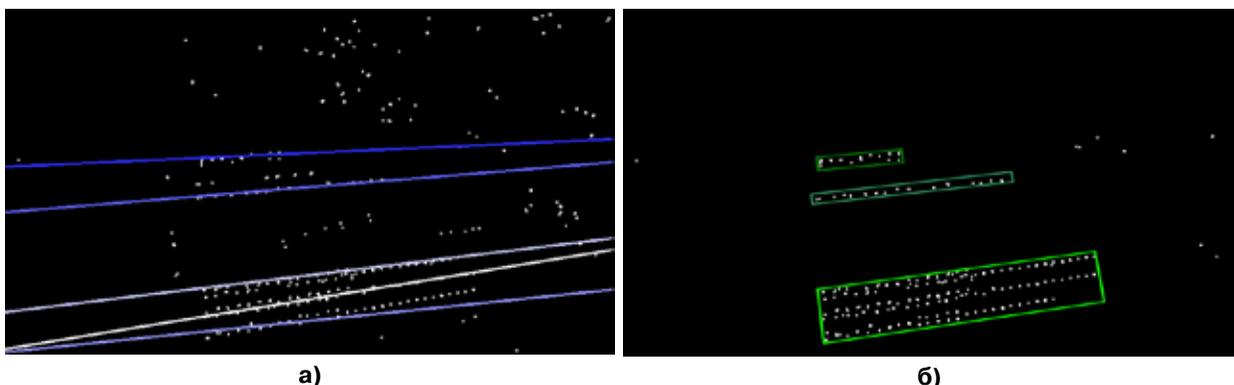


Рис. 3. а) прямые, соответствующие текстовым строкам б) кластеры и описывающие прямоугольники

Табл. 2.
Время выполнения алгоритма на мобильных устройствах. Замеры включают затраты на масштабирование

Тестируемое устройство	Среднее время выполнения, сек	Максимальное время выполнения, сек
iphone 5s	0,034	0,080
samsung galaxy s6	0,058	0,160

Заключение

В работе предложен и протестирован алгоритм поиска MRZ на изображениях машиночитаемых документов. Согласно результатам тестирования, предложенный алгоритм позволяет корректно детектировать MRZ в ~95% случаев, а скорость выполнения алгоритма на мобильных устройствах позволяет использовать его в режиме реального времени. К основным проблемам алгоритма относится неустойчивая работа на документах, содержимое которых порождает множество особых точек на нетекстовых областях, что в свою очередь приводит к ошибкам кластеризации. Дальнейшие исследования включают поиск решений для стабилизации работы алгоритма в подобных случаях.

Литература

1. ICAO Doc 9303 Machine Readable Travel. Seventh Edition. Parts 2-7. 2015. International Civil Aviation Organization.
2. Bessmeltsev V., Bulushev E., Goloshevsky N. High-speed OCR algorithm for portable passport readers // *Graphicon* '11. 2011. P. 25–29.
3. Visilter Y.V., Zheltov S.Y., and Lukin A.A., “Development of OCR system for portable passport and visa reader”. *Proceedings of SPIE*, 1999, pp. 194-199.4.
4. Булатов К.Б., Ильин Д.А., Полевой Д.В., Чернышова Ю.С., Проблемы распознавания машиночитаемых зон с использованием малоформатных цифровых камер мобильных устройств // *Труды ИСА РАН*, 2015. Т. 65. № 3. С. 85–93.
5. Арлазаров В.В., Жуковский А., Кривцов В., Николаев Д., Полевой Д. Анализ особенностей использования стационарных и мобильных малоразмерных цифровых видео камер для распознавания документов // *Информационные технологии и вычислительные системы*. 2014. № 3. С. 71–78.
6. Xiangrong Chen and Yuille A. L., “Detecting and reading text in natural scenes,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, pp. II-366-II-373 Vol.2.
7. Kwon Y., Kim J., “Recognition based Verification for the Machine Readable Travel Documents,” in *International Workshop on Graphics Recognition (GREC 2007)*, Curitiba, Brazil, September 20-21, 2007
8. Kwang-Baek Kim and Sungshin Kim. A passport recognition and face verification using enhanced fuzzy ART based RBF network and PCA algorithm // *Neurocomputing*. 2008. 71, 16-18 (October 2008), 3202-3210.
9. Martín-Rodríguez F., “Automatic optical reading of passport information,” 2014 International Carnahan Conference on Security Technology (ICCST), Rome, 2014, pp. 1-4.
10. Lee H. and Kwak N., “Character recognition for the machine reader zone of electronic identity cards,” 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, 2015, pp. 387-391.
11. 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; Proceedings*. 3. 79-87.
12. Lepetit V., Fua P. Towards Recognizing Feature Points using Classification Trees // *Technical Report IC/2004/74*. École polytechnique fédérale de Lausanne. 2004.
13. 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* 22. 2008.
14. *Smart 3D OCR MRZ v.1.0* <https://smartengines.ru/smart-mrzreader/>
15. Приказ Федеральной миграционной службы от 30 июня 2011 г. N 279 «Об утверждении Правил и способа формирования машиночитаемой записи в паспорте гражданина Российской Федерации, удостоверяющем личность гражданина Российской Федерации на территории Российской Федерации»

Скорюкина Наталья Сергеевна. Научный сотрудник-программист ООО «СМАРТ ЭНДЖИНС СЕРВИС». Окончила в 2013 г. НИТУ МИСИС. Количество печатных работ: 5. Область научных интересов: анализ изображений, компьютерное зрение. E-mail: skleppy.inc@gmail.com

Machine-readable zones localization method robust to capture conditions

N.S. Skoryukina

Abstract. This paper describes a method for machine-readable zones localization on document images captured using mobile digital cameras. The proposed method is based on feature points detection, filtration and clustering using Hough transform. The algorithm is designed in accordance of real-time requirements for running on mobile devices. Test results for real data are presented.

Keywords: machine-readable zone, image analysis, mobile devices

References

1. ICAO Doc 9303 Machine Readable Travel. Seventh Edition. Parts 2-7. 2015. International Civil Aviation Organization.
2. Bessmeltsev V.; Bulushev E., Goloshevsky N. High-speed OCR algorithm for portable passport readers // *Graphicon* 11. 2011. P. 25–29.
3. Visilter Y.V., Zheltov S.Y., and Lukin A.A., “Development of OCR system for portable passport and visa reader”. Proceedings of SPIE, 1999, pp. 194-199.4.
4. Bulatov K.B., Polevoy D.V., Ilin D.A., Chernyshova Y.S. Problems of machine-readable zone recognition captured with digital mobile cameras // *Труды ИСА РАН*, 2015. Т. 65. № 3. С. 85–93.
5. Arlazarov V.V., Zhukovskiy A., Krivtsov V., Nikolaev D., Polevoy D. Analiz osobennostey ispolzovaniya stacionarnykh i mobilnykh malorazmernykh tsifrovyykh video kamer dlya raspoznavaniya dokumentov // *Informatsionnye tekhnologii i vychislitelnye sistemy*. 2014. № 3. С. 71–78.
6. Xiangrong Chen and Yuille A. L., “Detecting and reading text in natural scenes,” Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004, pp. II-366-II-373 Vol.2.
7. Kwon Y., Kim J., “Recognition based Verification for the Machine Readable Travel Documents,” in International Workshop on Graphics Recognition (GREC 2007), Curitiba, Brazil, September 20-21, 2007
8. Kwang-Baek Kim and Sungshin Kim. A passport recognition and face verification using enhanced fuzzy ART based RBF network and PCA algorithm // *Neurocomputing*. 2008. 71, 16-18 (October 2008), 3202-3210.
9. Martin-Rodríguez F., “Automatic optical reading of passport information,” 2014 International Carnahan Conference on Security Technology (ICCST), Rome, 2014, pp. 1-4.
10. Lee H. and Kwak N., “Character recognition for the machine reader zone of electronic identity cards,” 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, 2015, pp. 387-391.
11. 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; Proceedings. 3. 79-87.
12. Lepetit V., Fua P. Towards Recognizing Feature Points using Classification Trees // Technical Report IC/2004/74. École polytechnique fédérale de Lausanne. 2004.
13. 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 22. 2008.
14. *Smart 3D OCR MRZ v.1.0* <https://smartengines.ru/smart-mrzreader/>
15. *Order of the Federal Migration Service of the Russian Federation of June 30, 2011 No 279* “About approval of the rules and method of forming machine-readable zone in the passport of a citizen of the Russian Federation, main identity document of the citizen of the Russian Federation in the territory of the Russian Federation.”

Skoryukina N.S. Place of work: OOO «Smart Engines Service», research and development department specialist. She graduated NUST MISIS in 2013. Number of publications: 5. Research field: image analysis, computer vision. E-mail: skleppy.inc@gmail.com