

Распознавание гильоширных элементов: определение страниц паспорта РФ

В. В. АРЛАЗАРОВ, Д. П. НИКОЛАЕВ, С. А. УСИЛИН, Д. Л. ШОЛОМОВ

Аннотация. В работе описан алгоритм распознавания гильоширных элементов на примере определения номеров страниц паспорта РФ. Алгоритм использует морфологическую фильтрацию для предобработки изображения и метод поиска объектов Виолы и Джонса для непосредственного распознавания гильоширных объектов. Несмотря на то, что описанный алгоритм настроен для определения номеров страниц паспорта РФ, после незначительных модификаций он может быть использован для распознавания других документов.

Ключевые слова: распознавание изображений документов, гильош, математическая морфология, AdaBoost.

Введение

Гильош — это специальная технология защиты банкнот, документов, ценных бумаг и других видов полиграфической продукции (билетов, акцизных марок, сертификатов и многих других документов государственного масштаба) [1]. Защита документов обеспечивается путем нанесения на бланки сложных композиций различных гильоширных элементов. Гильоширный элемент представляет собой замысловатый рисунок из множества многократно пересекающихся тончайших кружевных линий (рис. 1). Обычно такие элементы представлены разного рода защитными сетками, розетками, бордюрами, виньетками и уголками. Гильош может быть как симметричным, так и асимметричным по своему дизайну. Согласно существующим нормативам [2], гильоширные элементы должны занимать не менее 70 % площади ценных бумаг. Причем из этой площади большая часть должна содержать многоцветные гильоширные композиции.

При распознавании изображений документов часто приходится иметь дело с различными гильоширными элементами. При этом, являясь обычно частью фона, гильош только усложняет тем самым процесс распознавания значимой информации (усложняя алгоритм «подавления» фона, что является неотъемлемой частью любой системы распознавания [3]). Однако в ряде задач гильоширные элементы сами по себе являются значимыми объектами и должны быть распознаны. Например, такая ситуация возникает при распознавании страниц паспорта гражданина РФ (рис. 1б). Для обеспечения дополнительной защиты номер страниц паспорта печатается в виде гильоширной розетки (рис. 1б). Однако,

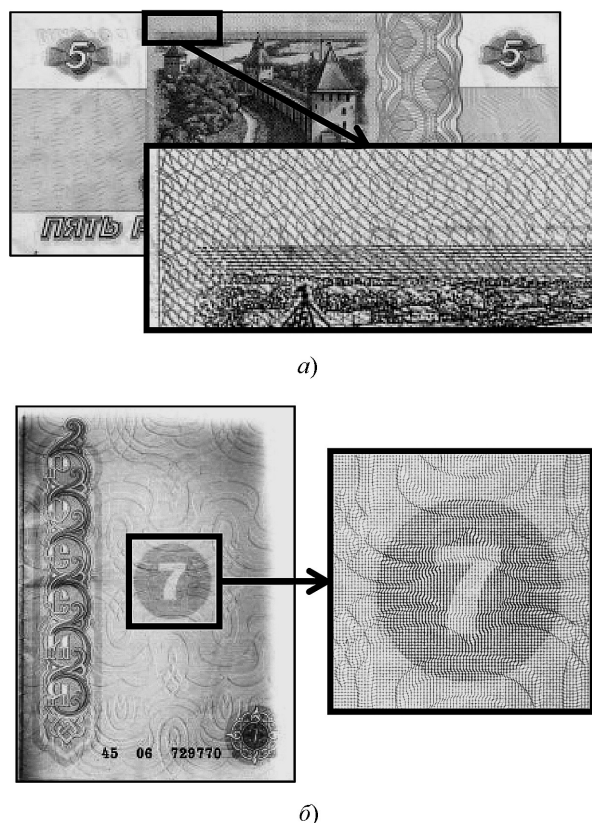


Рис. 1. Примеры документов, содержащих гильоширные элементы: а) банкнота номиналом 5 рублей образца 1997 года; б) изображение страницы паспорта гражданина РФ

при настройке алгоритмов распознавания такого рода графики возникает ряд трудностей. При попытке построить вектор признаков распознаваемого объ-

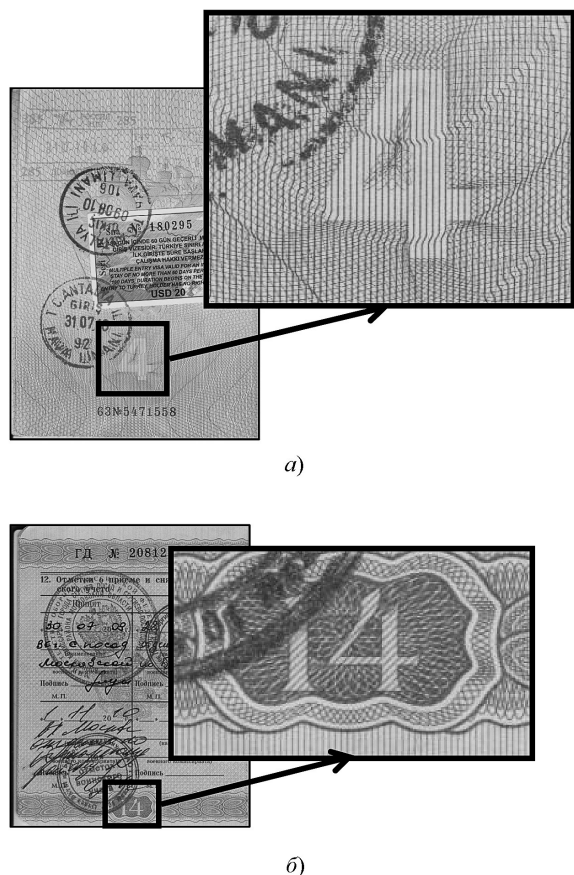


Рис. 2. Примеры документов, содержащих гильоширные элементы: а) заграничный паспорт гражданина РФ
б) Удостоверение личности военнослужащего РФ

екта по изображению высокого разрешения, в силу особенности гильоширных элементов (которые, напомним, представляют собой рисунок из тончайших линий), значение его будет существенно изменяться при незначительных геометрических искажениях (таких как, например, сдвиг или поворот). С другой стороны, попытка построить вектор признаков по изображению с пониженным разрешением оказывается неудачной из-за недостаточной контрастности распознаваемого объекта. Кроме того, дополнительные трудности для распознавания создают возможные дополнительные помехи. Например, номера страниц паспорта, и без того слабо поддающиеся распознаванию, часто перекрываются различными штампами и печатями.

В данной работе описана техника распознавания гильоширных элементов на примере номеров страниц паспорта гражданина РФ. Описанная ниже техника после незначительных модификаций может быть использована для обработки и распознавания гильоша, наносимого на другие документы (рис. 2).

1. Алгоритм распознавания страницы паспорта

Итак, на изображении страницы паспорта требуется определить номер. При этом точная ориентация страницы неизвестна, а также возможно наличие помех в виде штампов и печатей (рис. 3а).

Непосредственно перед распознаванием необходимо «подготовить» исходное изображение: подавить помехи и определить правильную ориентацию изображения. Заметим, что как распознаваемые объекты (номера страниц паспорта), так и возможные помехи преимущественно представимы в виде композиции линий, геометрические характеристики которых известны (в нашем случае важна только толщина). При работе с такими объектами часто используют морфологическую фильтрацию.

Морфология (математическая морфология) — это инструмент для выделения и анализа на изображении графических элементов с известной геометрической структурой [4–8]. В общем случае морфология является ресурсоемкой операцией, поэтому на практике часто используют фильтры с прямоугольными примитивами, для которых существуют быстрые алгоритмы вычисления [9]. Основными морфологическими операциями являются дилатация и эрозия. Дилатация изображения $f(x, y)$ по прямоугольному структурному элементу

$$e = \{(x, y) : |x| \leq a, |y| \leq b\}$$

обозначается как $f \oplus e$ и определяется следующим образом:

$$(f \oplus e)(x, y) = \max_{|s| \leq a, |t| \leq b} \{f(x + s, y + t)\}. \quad (1)$$

Аналогично, эрозия изображения $f(x, y)$ по структурному элементу e обозначается как $f \otimes b$ и определяется следующим образом:

$$(f \otimes e)(x, y) = \min_{|s| \leq a, |t| \leq b} \{f(x + s, y + t)\}. \quad (2)$$

Рассмотрим еще две важные морфологические операции: размыкание и замыкание. Размыкание изображения $f(x, y)$ по структурному элементу e обозначается $f \circ e$ и определяется равенством $f \circ e = (f \otimes e) \oplus e$. Аналогично, замыкание изображения $f(x, y)$ по примитиву e обозначается $f \bullet e$ и определяется как $f \bullet e = (f \oplus e) \otimes e$. Замыкание сглаживает контуры объекта, «обрывает» узкие перешейки и ликвидирует выступы небольшой ширины. Размыкание также проявляет тенденцию к сглаживанию участков контуров, но, в отличие от размыкания, «заливает» узкие разрывы и длинные углубления малой ширины, а также заполняет промежутки контура.

Пусть характерная толщина кружевных линий гильоша t_g , характерное расстояние между двумя



Рис. 3. Распознавание страницы паспорта гражданина РФ: а) исходное изображение (в перевернутом виде); б) морфологическая фильтрация типа «размыкание»; в) выделенное изображение фона; г) выделенное изображение печатей; д) определение ориентации страницы; е) распознавание номера страницы

линиями гильоша d_g , а толщина линий штампов t_s . Заметим, что для страниц паспорта справедливо соотношение:

$$t_g < d_g < t_s \quad (3)$$

Шаг 1. С помощью операции opening «сольем» друг с другом линии образующие гильош. Это позволит увеличить контраст распознаваемых объектов, при этом практически не изменит вид помех (так как $d_g < t_s$). Введем структурный элемент $e_1 = \{(x, y) : |x| \leq d_g, |y| \leq d_g\}$. Тогда контрастированное задается следующим образом (рис. 3б):

$$f_1(x, y) = (f \circ e_1)(x, y) \quad (4)$$

Шаг 2. Зная толщину линий штампов, с помощью замыкания подавим возможные помехи. Введем структурные элементы для «снятия» горизонтальных и вертикальных линий штампов как $e_{2h} = \{(x, y) : |x| \leq 0, |y| \leq t_s\}$ и $e_{2v} = \{(x, y) : |x| \leq t_s, |y| \leq 0\}$ соответственно. Тогда изображение фона (то есть изображение страницы паспорта «без помех») получается следующим образом (рис. 3в):

$$\begin{aligned} f_{2h}(x, y) &= (f_1 \bullet e_{2h})(x, y) \\ f_{2v}(x, y) &= (f_1 \bullet e_{2v})(x, y) \\ f_2(x, y) &= \max\{f_{2h}(x, y), f_{2v}(x, y)\} \end{aligned} \quad (5)$$

Шаг 2*. После получения фоновое изображение в качестве бонуса легко получить изображение $f_{2*}(x, y)$, содержащее только помехи (рис. 3в). Оно получается путем вычитания фона из исходного изображения и может быть полезно для анализа заполнения страницы (есть ли какие-либо штампы на текущей странице паспорта).

После снятия помех изображение считается подготовленным к определению ориентации и номера страницы. Заметим, что преимущественная ориентация страницы может быть найдена путем детектирования и анализа расположения углового метки страницы (рис. 3е), которая для паспорта РФ должна быть всегда снизу. Отметим, что алгоритм детектирования должен быть устойчив к незначительным искажениям объекта и шуму. Кроме того, настройка алгоритма должна проводиться на небольшом обучающем наборе. Пожалуй, одним из самых известных алгоритмов, удовлетворявших указанным требованиям, является метод Виолы и Джонса [10].

Алгоритм детектирования объектов Виолы и Джонса изначально разрабатывался для поиска лиц на изображениях в режиме реального времени. Он рассматривает задачу детектирования как задачу бинарной классификации в каждой точке, то есть для каждого возможного сдвига на изображении при помощи предварительно обученного распознавателя (классификатора) проверяется гипотеза о присутствии искомого объекта в данном положении. Ключевыми особенностями алгоритма являются:

- Использование прямоугольных признаков Хара [11] в качестве базовых признаков при построении классификатора.
- Использование так называемого интеграла изображения [10, 12], позволяющего быстро вычислять элементарные признаки.
- Использование эффективного алгоритма машинного обучения AdaBoost [13] для обучения классификаторов.
- Использование каскадной структуры классификаторов [10] с целью уменьшения суммарного времени обработки изображения.

Этот же алгоритм используется и для распознавания номера страницы паспорта путем предварительного обучения индивидуальных классификаторов для каждой гильоширной розетки, соответствующей номерам страниц.

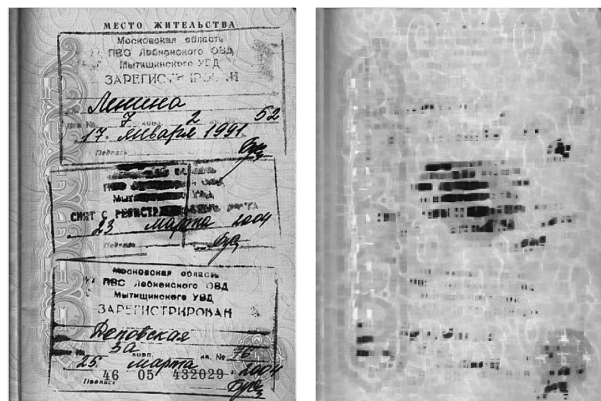
Шаг 3. Используя предварительно обученный классификатор, определим возможное расположение угловых меток на изображении фона $f_2(x, y)$. Если найденная метка расположена сверху, то повернем изображение на 180 градусов вокруг центра (рис. 3д).

Шаг 4. Определив ориентацию страницы и повернув при необходимости изображение, применим каждый из обученных классификаторов гильоширных розетт-номеров к изображению фона $f_2(x, y)$. Классификатор, обеспечивающий большее число срабатываний в центральной области изображения, определят искомый номер страницы (рис. 3е).

2. Экспериментальные результаты

Для тестирования изложенного алгоритма было подготовлено два набора данных. Первый использовался для обучения распознавателей угловых элементов номеров страниц паспорта (состоял из 333 изображений, то есть по 37 различных изображений каждой страницы паспорта). Второй набор данных использовался для проверки работоспособности и оценки качества детектирования и состоял из 616 различных изображений (распределение по типам страниц неравномерное, точная информация приведена в табл. 1).

В результате среднее качество детектирования составило 92.86 %. Заметим, что много ошибок было получено при распознавании страницы № 5 (несмотря на достаточно большую обучающую выборку). Это связано в первую очередь с наибольшим заполнением данной страницы по сравнению с другими в среднем. Некоторые из нераспознанных страниц к тому же с большим трудом могут быть распознаны даже человеком (рис. 4).



а)

б)

Рис. 4. Примеры нераспознаваемых изображений страниц паспорта: а) исходное изображение; б) изображение фона

Таблица 1

Статистика распознавания

	Количество изображений	Качество детектирования
Стр. 4	34	88.86 %
Стр. 5	129	89.15 %
Стр. 6	90	91.11 %
Стр. 7	88	97.73 %
Стр. 8	58	91.38 %
Стр. 9	59	94.92 %
Стр. 10	57	91.23 %
Стр. 11	53	100.00 %
Стр. 12	48	91.67 %
Итого	616	92.86 %

Заключение

При распознавании документов иногда возникает задача распознавания объектов, представимых в виде композиции гильоширных элементов. Сложность обработки таких объектов может быть к тому же увеличена за счет дополнительных помех на изображении (например, печати или штампы). В данной работе предложен подход к распознаванию такого рода объектов. Описанный алгоритм предназначен для определения номеров страниц внутреннего паспорта РФ, однако после незначительных модификаций может быть использован для распознавания других документов.

Литература

1. [-0.2pt]Warner R. D., Adams R. M. Introduction To Security Printing. Graphic Arts Center Publishing Company, 2005. P. 115.
2. ГОСТ Р 54109–2010. Защитные технологии. Продукция полиграфическая защищенная. Общие технические требования. 2010.
3. Арлазаров В. Л., Славин О. А. Алгоритмы распознавания и технологии ввода текстов в ЭВМ // Информационные технологии и вычислительные системы. 1996. Т. 1. Вып. 6. С. 48–54.
4. Soille P. Morphological Image Analysis: Principles and Applications // Morphological Image Analysis Principles and Applications. Springer-Verlag, 1999. P. 391.
5. Курпентев А. В. и др. Выделение графических примитивов и текстовых блоков на изображениях документов с помощью морфологических операций // Труды 51-й научной конференции МФТИ «Современные проблемы фундаментальных и прикладных наук». Часть IX. Инновации и высокие технологии. М.: МФТИ, 2008. С. 29–32.
6. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2006. С. 1072.
7. Asadchikov V. E. et al. Morphological analysis and reconstruction techniques for computer tomography // International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM). 2011. Vol. 3. P. 19–25.
8. [-0.2pt]Николаев Д. П. Методы автоматической обработки изображений // Межд. науч. школа-семинар «Современные методы анализа дифракционных данных». Великий Новгород, 12–16 ноября, 2011. С. 92.
9. Van Herk M. A fast algorithm for local minimum and maximum filters on rectangular and octagonal kernels // Pattern Recognition Letters. New York, NY, USA: Elsevier Science Inc., 1992. Vol. 13. № 7. P. 517–521.
10. Viola P., Jones M. Robust Real-time Object Detection // International Journal of Computer Vision. 2002.
11. Papageorgiou C. P., Oren M., Poggio T. A general framework for object detection // Sixth International Conference on Computer Vision. 1998. P. 555–562.
12. Crow F. C. Summed-area tables for texture mapping // SIGGRAPH '84: Proceedings of the 11th annual conference on Computer graphics and interactive techniques. New York, NY, USA: ACM, 1984. P. 207–212.
13. Freund Y., Schapire R. E. A decision-theoretic generalization of on-line learning and an application to boosting // Journal of Computer and System Sciences. 1997. Vol. 55. № 1. P. 119–139.

Арлазаров Владимир Викторович. С. н. с. ИСА РАН. К. т. н. Окончил в 1999 г. МИСиС. Количество печатных работ: 6. Область научных интересов: распознавание образов, обработка изображений, системы массового обслуживания. E-mail: vva777@gmail.com

Николаев Дмитрий Петрович. Зав. сектором ИППИ РАН. К. ф.-м. н. Окончил в 2000 г. МГУ им. М. В. Ломоносова. Количество печатных работ: 103. Область научных интересов: быстрые алгоритмы обработки изображений. E-mail: dimonstr@iitp.ru

Усилин Сергей Александрович. Научный сотрудник ИСА РАН. Окончил в 2009 году МФТИ. Количество печатных работ: 16. Область научных интересов: цифровая обработка изображений и распознавание образов. E-mail: usilin.sergey@gmail.com

Шоломов Дмитрий Львович. С. н. с. ИСА РАН. К. т. н. Окончил в 1997 году МГУ им. М. В. Ломоносова. Количество печатных работ: 15. Область научных интересов: распознавание образов. E-mail: sholomov@list.ru