

УДК 681.5

ОПТИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ФИНАНСОВЫХ ДОКУМЕНТОВ ПО ИЗОБРАЖЕНИЯМ, ПОЛУЧЕННЫМ С ПОМОЩЬЮ ФОТОКАМЕРЫ В МОБИЛЬНЫХ УСТРОЙСТВАХ

Э.В. МИХАЙЛОВ*(Полоцкий государственный университет)*

Рассмотрены основные методы оптической идентификации финансовых документов с учетом особенностей изображений, полученных с помощью встроенных фотокамер мобильных устройств. Анализируются основные алгоритмы предварительной обработки таких изображений, а также проблемы внедрения современных методов оптической идентификации документов в финансовые институты и организации.

Введение. Эффективность использования систем OCR (OpticalCharacterRecognition, оптическое распознавание текстов) в сфере банковского обслуживания давно подтверждена на практике. Современные специализированные решения, поставляемые такими известными компаниями, как ABBYY или Nuance [1] позволяют организовать комплексный оптический ввод финансовых документов с помощью специальных сканеров или повседневных средств получения цифровых изображений (бытовой сканер, фотоаппарат и т.д.). Такие сложные системы поставляются в финансовые учреждения, специфика работы которых подразумевает постоянную работу с бумажными носителями информации. Как правило, это филиалы банков, ежедневно обрабатывающих большое количество платежных поручений и справок, предоставляемых клиентами. Несмотря на то, что практически все финансовые институты стараются избавиться от «бумажного» документооборота, внедряя системы штрихового кодирования и дистанционного обслуживания, для соблюдения легитимности при работе с клиентами необходимо отражать часть информации на бумажных носителях (чеках, счетах и т.д.). В этом случае для оптического ввода финансовой информации самим клиентом актуальным является развитие систем OCR на базе мобильных устройств, оснащенных портативными фотокамерами. Изображения, полученные такими «мобильными» системы, имеют ряд особенностей, связанных с ограниченными возможностями оптического устройства ввода (фотокамеры): шумы [2], низкое разрешение, невысокая резкость, неточности цветопередачи [3].

В данной работе рассмотрены методы идентификации платежных документов как один из первых этапов оптического ввода финансовой информации по изображениям, полученным с помощью фотокамеры мобильного телефона.

Рассмотрим особенности финансового документа в контексте систем оптического ввода сложных структурных документов [4]. К таким особенностям можно отнести:

- обязательное наличие табличных данных;
- документы, относящиеся к одной финансовой организации и имеющие одинаковое назначение, имеют одинаковую структуру для всех клиентов. Таким образом, для всех клиентов фотография документа одного и того же типа будет иметь примерно одинаковое количество структурных примитивов;
- наличие большого объема бесполезной информации (реклама, справочная информация и т.д.);
- полезная информация в основном представлена монохромными участками;
- наличие специфических структурных примитивов (печати, логотипы, штрихкоды). Данные структуры хорошо выделены на исходном изображении и могут быть использованы для идентификации типа документа.

Под идентификацией финансового документа понимают выделение множества признаков (маркеров) на изображении документа, которые позволяют соотнести бумажную версию документа с его электронной копией [5]. Идентификация не является обязательным этапом в цикле обработки исходного изображения, а имеет место лишь в том случае, если обрабатываемые документы имеют неоднородную структуру: различное расположение полезной информации, различие в табличной структуре документа и т.д.

Примерами идентификации могут выступать следующие случаи:

- получение детальных данных для оплаты услуг ЖКХ по фотографии расчета за предоставленные услуги, с последующей автоматической оплатой данных услуг при помощи банковского ПО;
- получение детальной информации о предоставленной услуге по фотографии квитанции от поставщика услуг (например, квитанция из магазина) с последующим автоматическим занесением данной информации в домашнюю бухгалтерию.

1. Метод идентификации с помощью штрихкодов

Рассматриваемый метод является наиболее простым и распространенным методом идентификации финансовой информации. На каждый «бумажный» документ наносится специальный штрихкод, который

кодирует номер электронного документа в базе данных финансового учреждения. В этом случае с помощью специального программного обеспечения клиент имеет возможность, сфотографировав штрихкод, получить всю доступную ему информацию из базы данных соответствующей организации. Наиболее распространенные мобильные операционные системы – Android и iOS – имеют встроенные средства и инструменты разработчика для считывания различных видов штрихкодов.

К достоинствам данного метода следует отнести [6], такие как:

- высокая информационная плотность, или высокое разрешение. Очень маленькие коды могут быть отпечатаны и использованы в любом месте документа, при этом они занимают небольшую площадь. Таким образом, при фотографировании штрихкода можно получить достаточно качественное изображение кода в большом разрешении;

- оптимальное расположение данных, когда возможность ошибок чтения практически нулевая;

- легкость дешифровки. Штрихкоды, например, используемые в банковской сфере, имеют точно определенное содержание данных. Они структурируются для обеспечения удобства большого количества пользователей. Некоторые штрихкоды разработаны с поддержкой значительного количества наборов символов, тогда как другие поддерживают только цифровые данные.

К недостаткам использования штрихового кодирования относят:

- штрихкод не содержит полезной для клиента информации. В финансовой сфере штрихкод на бумажном носителе используется как дополнение к полезной информации;

- без связи с базой данных организации, которая выдала штрихкодированный документ, невозможно получить полезную информацию и, следовательно, использовать данную информацию;

- для банков и электронных платежных систем необходимо заключать договора для взаимодействия с поставщиками услуг при использовании клиентами банка штрихкодов этих поставщиков. В то же время у клиентов есть возможность оплатить те или иные услуги путем ввода реквизитов счетов от поставщиков вручную.

На рисунке 1 представлены примеры штрихкодов, полученных с помощью фотокамер мобильных устройств. Из приведенных трех изображений только первое сверху может быть распознано как штрихкод [7].



Рис. 1. Примеры изображений штрихкодов, полученных с помощью фотокамер мобильных устройств

2. Метод идентификации по структурным элементам изображения документа. Данный метод является наиболее трудоемким и его использование в мобильных устройствах в первую очередь обусловлено ростом производительности вычислений мобильных систем, а также улучшением технических характеристик фотокамер мобильных телефонов. Основным достоинством данного метода является возможность узнать тип документа по его изображению без взаимодействия с организацией-поставщиком данного документа. Зная тип документа, становится возможным получить его структуру (например, по

заранее сформированному шаблону), а с применением технологий OCR – полезную информацию для ее последующей обработки.

Основные этапы идентификации документа с бумажного носителя представлены на рисунке 2.

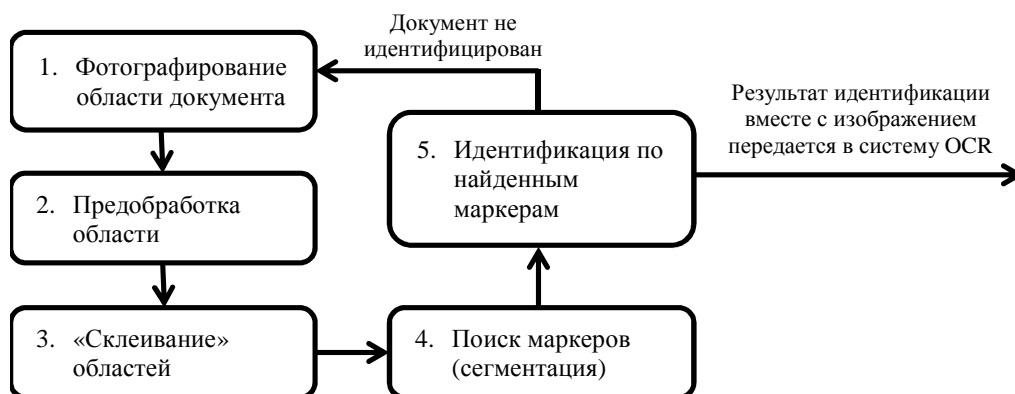


Рис. 2. Основные этапы идентификации финансового документа

Первым этапом идентификации является фотографирование области документа. Фотографирование документа целиком в силу низкой разрешающей способности матрицы фотокамеры мобильного устройства, как правило, не дает нужного результата. Данный метод идентификации предполагает **ручное** или **автоматическое** получение целого изображения из его частей. На данном этапе также выполняется проверка качества изображения. Наиболее интересными методами предварительной оценки качества в системах OCR являются:

- использование набора детекторов дефектов изображения (системы на базе ABBYY Mobile Imaging SDK предлагают использовать детектор расфокусировки и размытости изображения, детектор бликов, детектор шумов [8]);

- использование OCR детектора. В детекторе применима система OCR для распознавания текста на некоторой выборке из блоков изображения. В случае успешного распознавания символов выбранных блоков всё изображение считается допустимым для последующей обработки;

- методы оценки качества, основанные на расчете отклонения слоев изображения от некоторой идеальной маски [9].

Как правило, изображения, полученные с мобильного телефона, делаются при недостаточном освещении. Исходный фрагмент документа представлен на рисунке 3.



Рис. 3. Фрагмент изображения квитанции, полученный с помощью встроенной фотокамеры мобильного телефона

В этом случае для достижения приемлемой яркости изображения поднимают светочувствительность матрицы портативного фотоаппарата, что приводит к появлению **яркостного и хроматического**

шума на изображении. Отсутствие достаточной освещенности бумажного документа и специфические для мобильного устройства условия съемки («дрожание» фотокамеры и т.д.) формируют изображение размытым и неконтрастным. Для сохранения изображение подвергается сжатию в специальные форматы с потерей качества, например JPEG, что зачастую приводит к появлению «артефактов». Еще одной особенностью мобильных фотоснимков является наличие геометрических и перспективных искажений.

Предварительная обработка изображения направлена на улучшение качества изображения и включает фильтрацию изображения от шумов, повышение резкости и контрастности изображения, выравнивание и преобразование в используемый системой формат [10]. Типовой стек алгоритмов предварительной обработки изображения представлен на рисунке 4.

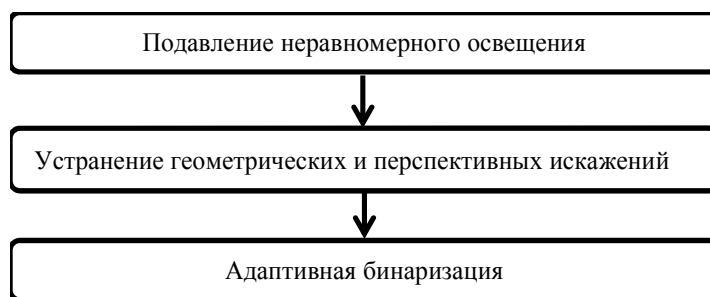


Рис. 4. Схема применения алгоритмов предварительной обработки изображения

Для подавления неравномерного освещения в условиях мобильной съемки, а также нормализации локальных уровней освещенности хорошо подходит алгоритм SSR (Single Scale Retinex) [11]:

$$R_i x, y = \log I_i x, y - \log F_1 x, y \cdot I_i x, y, \tag{1}$$

где $R_i x, y$ – выходное изображение; $I_i x, y$ – исходное изображение в i -том спектральном диапазоне; $F_1 x, y$ – функция Гаусса.

К преимуществам алгоритма SSR можно отнести: высокую скорость обработки входных данных, параметризуемое качество результата, отсутствие размывания в ходе работы, и как следствие, «чуткость» к деталям, нормализация локальных уровней освещенности [12].

Для **устранения геометрических и перспективных искажений** пригоден метод, основанный на использовании патча Безье [13]. В этом методе используется понятие скелета многоугольной фигуры. Для решения задачи строится внешний скелет для межстрочных интервалов, после чего производится фильтрация скелета, затем строится патч Безье. С помощью полученного патча итерационным методом производится выпрямление строк. К достоинствам данного метода относят: простоту реализации и высокую скорость работы, к недостаткам: исправление искажений только в окрестности границы страницы. Недостатки метода учтены в его модификациях [14].

В качестве **алгоритма адаптивной бинаризации** следует применять модифицированный алгоритм Кристиана (Christian) [15; 16]. Порог бинаризации для скользящего окна вычисляется согласно формуле (2) [17]:

$$T = (1 - \alpha_1) \cdot \mu + \alpha_2 \cdot \left(\frac{\sigma}{R_\sigma} \cdot [\mu - M] + \alpha \cdot M \right), \tag{2}$$

где T – порог бинаризации для заданного окна; R_σ – устанавливается как максимальное значение среднеквадратического отклонения, рассчитанное для окрестности заданного окна; M – минимальное значение яркости в окрестности заданного окна; $\mu = \frac{1}{N} \sum_{i=1}^N x_i$ – математическое ожидание значения яркости в заданном окне;

$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$ – среднеквадратическое отклонение яркости в заданном окне;

$\alpha_1 = k_1 \cdot \left(\frac{\sigma}{R_\sigma} \right)^\gamma$, $\alpha_2 = k_2 \cdot \left(\frac{\sigma}{R_\sigma} \right)^\gamma$, где γ , k_1 и k_2 – положительные константы. При $\gamma = 2$ параметры k_1 и k_2 рекомендуется брать в диапазонах 0,1...0,2; 0,15...0,25 и 0,01...0,05.

Данный алгоритм в совокупности с предыдущими методами позволяет устранить шумы на изображении.

Автоматическое «склеивание» изображений из уже обработанных частей происходит путем сопоставления SIFT (Scale Invariant Feature Transform) дескрипторов. SIFT является наиболее эффективным алгоритмом поиска характерных точек на двух или более изображениях, инвариантным к масштабу изображений и изменениям яркости. SIFT также достаточно устойчив к изменениям ракурса изображения [18]. В результате работы алгоритма получается множество особенных точек на каждой паре изображений с набором своих дескрипторов. Эти точки сравниваются попарно на соседних изображениях, и если дескрипторы двух выбранных точек совпадают в пределах заданной точности, две точки принимаются как сопряженные. Можно считать, что эти точки были получены как проекция одной и той же трехмерной точки на плоскости двух камер.

Выделение и поиск маркеров происходит по найденному множеству SIFT дескрипторов и по биаризированному изображению документа. Основные типы маркеров можно разделить на 3 основные группы относительно их значимости для идентификации документа:

- к самым значимым признакам можно отнести: печати, логотипы, штрихкоды, стилизованные надписи;

- ко второй группе относят: линии, пересечения (углы), текстовые блоки и отдельные надписи;

- наименее значимыми признаками являются отдельные символы.

Описанный алгоритм позволяет не только идентифицировать документ, а также представить документ в виде упорядоченного множества хорошо различимых признаков, а промежуточный результат работы алгоритма может быть использован системами OCR без предварительной обработки. Сложность такого подхода к идентификации документа обусловлена в первую очередь качественными характеристиками исходного изображения и способом его получения.

Заключение. В отличие от метода идентификации документов по штрихкоду, который активно сегодня используется и внедрен в основные финансовые институты, метод идентификации по структурным элементам изображения практически не имеет внедрения, однако широко востребован электронными платежными банковскими системами. Внедрение данного метода позволяет развиваться таким банковским системам изолированно, без договорных отношений с поставщиками услуг, и взаимодействовать напрямую с клиентами. Как следствие, такие системы становятся более привлекательными. Однако сдерживающим фактором для внедрения метода является задача выделения четких структурных элементов из изображений, получаемых с помощью портативных фотокамер, что само по себе является наукоемкой задачей, требующей использования самых новых исследований алгоритмов машинного зрения и, следовательно, существенных финансовых затрат. Выходом из данной ситуации является использование ряда ограничений, налагаемых на процесс получения исходного изображения и сам процесс идентификации. Практическим примером таких ограничений может служить ситуация, когда программное обеспечение мобильного устройства сначала предлагает клиенту пошагово сфотографировать логотипы и/или печати документов, далее, если удастся распознать организацию и тип документа, также в пошаговом режиме, предлагает сфотографировать полезную информацию с указанием мест, где эта информация может находиться.

ЛИТЕРАТУРА

1. Вся правда об OCR [Электронный ресурс]. – 2014. – Режим доступа: <http://internetno.net/category/obzoryi/ocr/>. – Дата доступа: 10.08.2014.
2. Шумы на фотографии [Электронный ресурс]. – 2014. – Режим доступа: <http://fotoprimer.ru/chto-k-chemu/shumyi-na-fotografii>. – Дата доступа: 11.08.2014.
3. Улучшение качества фото с мобильного телефона [Электронный ресурс]. – 2014. – Режим доступа: <http://www.photoshop-master.ru/lessons.php?rub=3&id=587>. – Дата доступа: 11.08.2014.
4. Структурные методы распознавания сложноорганизованных исторических табличных форм [Электронный ресурс]. – 2014. – Режим доступа: http://www.ipages.ru/index.php?ref_item_id=20269&ref_dl=1. – Дата доступа: 14.08.2014.
5. Быстрая идентификация документа DIRECTUM RapID [Электронный ресурс]. – 2014. – Режим доступа: <http://www.directum.ru/1960173.aspx>. – Дата доступа: 14.08.2014.
6. Технологии штрихового кодирования [Электронный ресурс]. – 2014. – Режим доступа: http://www.tppkuban.ru/services.php?sub_id=104. – Дата доступа: 23.08.2014.
7. Ender Tekin, James Coughlan. Proceeding CRV'09 Proceedings of the 2009 Canadian Conference on Computer and Robot Vision, IEEE Computer Society Washington, DC, USA, 2009.

8. ABBYY Mobile Imaging SDK [Электронный ресурс]. – 2014. – Режим доступа: <http://www.abbyy.ru/mobile-imaging-sdk/description/>. – Дата доступа: 23/08/2014.
9. Николаев, Д.П. Метод автоматической оценки качества цветовой сегментации в задаче упаковки изображений печатных документов / Д.П. Николаев, Д.В. Полевой, Т.С. Чернов // Труды ИСА РАН. – 2013. – Т. 63.
10. Методы распознавания текстов [Электронный ресурс]. – 2014. – Режим доступа: <http://habrahabr.ru/post/112442/>. – Дата доступа: 26.08.2014.
11. Bian, Z. Retinex Image Enhancement Techniques / Z. Bian, Y. Zhang // Final Project Report for EE264.
12. Построение системы оптического распознавания структурной информации на примере Imago OCR [Электронный ресурс]. – 2014. – Режим доступа: <http://habrahabr.ru/post/172651/>. – Дата доступа: 26.08.2014.
13. Масалович, А. Использование патча Безье для аппроксимации искажения изображений текстовых документов / А. Масалович, Л. Местецкий // Proc. Graphicon-2007. – 2007. – P. 171–174.
14. Zeynalov, R. Восстановление формы страницы текста для коррекции геометрических искажений / R. Zeynalov, A. Velizhev, A. Konushin // Department of Computational Mathematics and Cybernetics. Moscow: Moscow State University, Russia.
15. Wolf, C. Extraction and recognition of artificial text in multimedia documents / C. Wolf, J.M. Jolion // Patt. Anal. Appl. – 2004. – Vol. 6, No. 4, Feb. – P. 309–326.
16. Meng-Ling Feng and Yap-Peng Tan. Contrast adaptive binarization of low quality document images // IEICE Electronics Express. – Vol. 1, No.16. – P. 501–506.
17. Адаптивная бинаризация [Электронный ресурс]. – 2014. – Режим доступа: <http://www.heclub.ru/2013/12/03-yy.html>. – Дата доступа: 01.09.2014.
18. Lowe, David G. Object recognition from local scale-invariant features / David G. Lowe // Proceedings of the International Conference on Computer Vision 2: 1150–1157, 1999.

Поступила 05.09.2014

**OPTICAL IDENTIFICATION OF FINANCIAL DOCUMENTS BY IMAGES,
OBTAINED BY EMBEDDED MOBILE DEVICES CAMERA**

E. MIHAILOV

This paper reviews the basic methods of optical identification for financial documents, taking into account the characteristics of document images obtained by embedded mobile camera. The basic pre-processing algorithms for such images are also reviewed, as well as the problems of practical implementation of new optical identification methods for paper documents in different financial institutions and organizations.