

Восстановление формы страницы текста для коррекции геометрических искажений

Ramiz Zeynalov, Alexander Velizhev, Anton Konushin
Department of Computational Mathematics and Cybernetics

Moscow State University, Moscow, Russia
ramiz.zeynalov@gmail.com, avelizhev@graphics.cs.msu.ru, ktosh@graphics.cs.msu.ru

Аннотация

В последнее время цифровые камеры получили широкое распространение. Поэтому их часто используют для быстрого фотографирования текста. При этом тексты на таких фотографиях плохо подходят для распознавания и не очень удобны для чтения. Основная проблема – это не прямые строки текста. Существующие на сегодняшний день системы оптического распознавания текста не могут дать удовлетворительного результата на подобных входных данных без соответствующей предобработки. Наибольшую сложность в этой предобработке представляет геометрическая коррекция. В данной статье предлагается развитие предложенного ранее метода исправления описанных выше искажений, основанного на восстановлении поверхности страницы по форме найденных на изображении строк текста.

Ключевые слова: строки текста, распознавание текста, улучшение качества текста.

1. ВВЕДЕНИЕ

В последнее время цифровые камеры получили широкое распространение – они встраиваются в сотовые телефоны и КПК. Поэтому их часто используют для быстрого фотографирования текста. При этом тексты на таких фотографиях плохо подходят для распознавания и не очень удобны для чтения.

Предлагается развитие метода из работ [18] и [19], который позволял решить частную задачу со следующими ограничениями:

1. Искажения формы происходят только в направлении, перпендикулярном строкам текста
2. Перспективных искажений нет
3. Главная оптическая ось камеры проходит через центр изображения текста по вертикали

В этой статье описано развитие метода, в котором удалось отказаться от всех вышеперечисленных ограничений, что существенно расширило область применения метода.

1.1 Постановка задачи

Дано изображение текста, которое может быть подвергнуто следующим типам искажений:

- Неравномерное освещение и блики
- Перспективные искажения
- Непланарность исходного текста

Требуется исправить перечисленные типы искажений, чтобы результат работы OCR был лучше и визуально строки текста стали прямыми, одного размера и ориентированы горизонтально.

2. ОБЗОР СУЩЕСТВУЮЩИХ МЕТОДОВ

2.1 Регрессия искривленных линий

Этот метод используется для коррекции геометрических искажений на сканированных изображениях, возникающих вследствие деформации поверхности книги в окрестности границы между страницами.

Здесь отдельно обрабатывается затененная область и незатененная область. Концы строк с затененной области аппроксимируются полиномами второй степени ([5]) и выпрямляются с учётом этого полинома. После выпрямления концов строк они объединяются с прямыми фрагментами строк с незатененной области. Более подробно о методе можно прочитать в [2].

Достоинства:

- Простота реализации
- Высокая скорость работы

Недостатки:

- Исправляются искажения только в окрестности границы страницы

2.2 Использование патча Безье

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. Подробнее о методе можно узнать из [8].

Достоинства:

- Высокая скорость работы
- Простота реализации

Недостатки:

- Реальное искажение не всегда можно приблизить плоской деформацией

2.3 Робастное приближение формы кривых строк текста

В данном подходе осуществляется поиск связанных компонент, затем для каждого символа находится локальная линия, на которой лежит строка, содержащая данный. Для каждого найденного символа вычисляется расстояние между строками, затем оно усредняется и считается постоянным

для всего текста. Используя найденное среднее расстояние между строками и расстояния между строками на изображении, строится поверхность книги, с использованием которой выполняется коррекция изображения текста. Более подробно узнать о методе можно в [3].

Достоинства:

- Корректируются и геометрические, и перспективные искажения

Недостатки:

- Высокая вычислительная сложность
- Необходимы данные о начертании символов (для использования алгоритма RAST)

2.4 Глобальная оптимизация формы поверхности

Для описания поверхности текста используется полигональная модель. Сначала строятся векторные поля – направление текста, направление расположения строк, нормали к поверхности, и учитываются параметры камеры. Затем строится функционал, который учитывает ограничения на ортогональность введенных полей в каждой точке поверхности и на гладкость результирующей поверхности. Этот функционал итерационно оптимизируется. Затем, каждая грань полигональной модели поверхности независимо отображается на скорректированное изображение, после чего результаты склеиваются с учётом условия непрерывности полученной плоской поверхности текста. Подробнее о методе можно узнать из [4].

Достоинства:

- Решает широкий спектр проблем (поиск областей с текстом, поворот, перспективная и геометрическая коррекция текста)

Недостатки:

- Сложность в использовании (много ручных параметров, задача определения оптимальных значений которых нетривиальна)
- Низкая производительность по сравнению с другими методами
- Хорошо исправляются лишь перспективные искажения

2.5 Выводы

Все рассмотренные методы обладают своими достоинствами и недостатками. Как правило, каждый метод, решающий поставленную задачу, находит компромисс между эффективностью и универсальностью. Из методов, которые обладают высокой производительностью, качественнее всех работает метод, использующий патч Безье для аппроксимации плоской деформации изображения.

3. КРИТЕРИИ ОЦЕНКИ КАЧЕСТВА РЕШЕНИЯ

С точки зрения системы оптического распознавания символов качество результата можно оценить через увеличение процента правильно распознанных символов после выполнения коррекции. Однако различные системы распознавания работают по-разному и используют разную информацию для своей работы. Таким образом, использовать

систему распознавания текста в качестве критерия качества некорректно. Поэтому в качестве критерия будем использовать визуальное качество результата, для этого формализуем это понятие, учитывая, что человеку удобнее читать прямой текст, ориентированный горизонтально и под прямым углом [17].

4. ОПИСАНИЕ МЕТОДА

Алгоритм коррекции деформаций состоит из нескольких последовательных шагов:

1. Подавление неравномерного освещения и бинаризация
2. Объединение символов в строки
3. Анализ геометрии строк и построение карты глубины
4. Построение поверхности по карте глубины
5. Проецирование с учётом полученной поверхности

4.1 Подавление неравномерного освещения и бинаризация

Целью данного шага является выравнивание освещенности на изображении в целом. Для этого используются два метода – деление изображения на его низкочастотную составляющую и ретинекс [10].

4.2 Объединение символов в строки

Для упрощения задачи поиска строк текста, отдельные символы и слова объединяются в связанные компоненты, соответствующие строкам текста. Это осуществляется путем последовательного применения операций математической морфологии ([18, 19]).

4.3 Анализ геометрии строк и построение карты глубины

Сначала производится поиск связанных компонент на бинарном изображении, полученном на предыдущем шаге. Для этого используется алгоритм [13]. Области классифицируются как шум, иллюстрации и строки. По строкам текста строятся кривые, описывающие их форму. Полученная линия хорошо описывает форму строки во внутренних точках, а окрестности границ строк обрабатываются отдельно. Далее, для каждой страницы вычисляются линии левой и правой границ. Для нахождения этих прямых используется RANSAC ([15]). По полученным границам рассчитывается матрица гомографии ([16]) из четырёхугольника, образованного пересечением этих границ с верхней и нижней границами изображения. После чего определяется область действия этого преобразования как четырёхугольник, в котором лежат все точки всех строк. В результате получается изображение без перспективных искажений. Кроме того, искажения из-за неплоской поверхности текста имеют место только в вертикальном направлении относительно плоскости страницы. После того, как область с текстом найдена, осуществляется поиск горизонтальной линии, в окрестности которой эти искажения минимальны. Полученные значения аппроксимируются прямой линией и находится значение ординаты, при которой эта прямая пересекает 0. Далее, с учётом формы строк, их желаемого расположения и оптического центра по вертикали, строится форма поверхности – глубина поверхности для каждого пиксела

входного изображения, соответствующего точек на строке. Затем находится зависимость между искажением формы строки и глубиной соответствующей точки строки. Подробное описание этих действий можно найти в статье ([18, 19]).

4.4 Построение поверхности по карте глубины

Необходимо построить поверхность во всех точках таким образом, чтобы полученная поверхность была гладкой. Кроме того, в точках, в которых она уже была определена, возможно, с ошибками, она была близка к точкам, полученным ранее. Определить поверхность в остальных точках надо для того, чтобы итоговое преобразование было непрерывным, чтобы нивелировать ошибки для отдельных строк и преобразовать все области, а не только области, соответствующие центрам строк.

4.5 Проецирование с учётом полученной поверхности

Для того чтобы полученное изображение было гладким, применяется обратное преобразование к каждой точке изображения-результата, в результате чего вычисляется цвет рассматриваемого пиксела.

5. РЕЗУЛЬТАТЫ

Тестирование производилось на нескольких наборах данных, которые отличались сложностью искажений и разрешением.

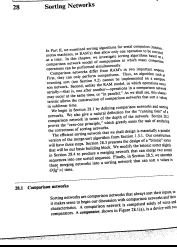

До коррекции	После коррекции
	
а)	б)

Рис. 1: Пример работы метода: а) исходное изображение, б) скорректированное изображение.

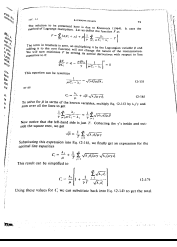

До коррекции	После коррекции
	
а)	б)

Рис. 2: Пример работы метода: а) исходное изображение, б) скорректированное изображение.

6. СРАВНЕНИЕ

Для сравнения с предложенным методом был выбран метод, основанный на построении патча Безье (из работы [8]). Выбор обусловлен тем, что метод, использующий патч Безье, согласно проведённому обзору, лучший на сегодняшний день. Сравнение производилось на стандартной для решаемой задачи тестовой базе IUPR [9], критерии сравнения были описаны в разделе “Критерии оценки качества решения”. Ниже на рисунках 3-5 изображены диаграммы с результатами сравнения ошибок неискривлённости, ориентации и перспективы для предложенного метода и метода, основанного на моделировании плоского искажения патчем Безье и скелете изображения.

Из рисунка 3 видно, что по критерию неискривлённости предложенный метод превосходит патч Безье.

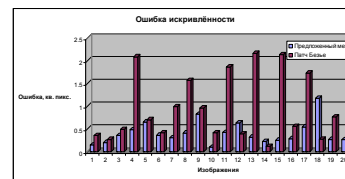


Рис. 3. Ошибка искривлённости (отклонение от прямой)

Из рисунка 4 видно, что ошибка ориентации чаще всего получается близкой к нулю.

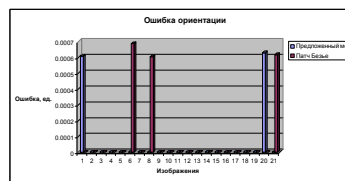


Рис. 4. Ошибка ориентации (отклонение от горизонтали)

Из рисунка 5 видно, что ошибка перспективы у обоих методов незначительная.

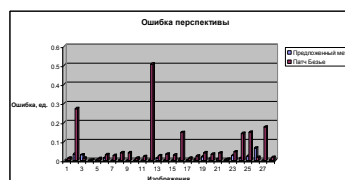


Рис. 5. Ошибка перспективы

7. ПРОИЗВОДИТЕЛЬНОСТЬ

Скорость работы приложения зависит от разрешения входного изображения и от некоторых параметров. Ниже приведена таблица (таблица 1) зависимости средней скорости работы от разрешения входного изображения. Измерения производились на компьютере средней конфигурации.

Разрешение, Мпикс	Время работы, секунды
1.3	1-1.5
6	5-6
8	7-10

Таблица 1. Производительность предложенного метода

8. ЗАКЛЮЧЕНИЕ

В данной работе предложен новый метод, позволяющий исправлять геометрические и перспективные искажения строк текста. Произведено сравнение с одним из существующих методов на стандартной тестовой базе. Метод для сравнения был выбран с тем же порядком производительности, что и предложенный.

9. ЛИТЕРАТУРА

- [1]Владимир Вежнев. Задача восстановления 3D формы по закраске [HTML] (<http://cgm.graphicon.ru/content/view/59/62/>)
- [2]Zheng Zhang, Chew Lim Tan. Correcting Document Image Warping Based on Regression of Curved Text Lines // Proc. ICDAR-2003, 2003, pp.589-563
- [3]Adrian Ulges, Christoph H. Lampert, Thomas M. Breuel. Document Image Dewarping using Robust Estimation of Curled Text Lines // Proc. ICDAR-2005, 2005, pp.1001- 1005
- [4]Adrian Ulges, Christoph H. Lampert, Thomas M. Breuel. Unwarping Images of Curved Documents Using Global Shape Optimization // Proc. International workshop on Camera-based Document Analysis and Recognition, 2005, pp.25-29
- [5]Н. С. Бахвалов, Н. П. Жидков, Г. М. Кобельков. Численные Методы. Бином, 2003, Москва
- [6]Т. М. Breuel. Robust least square baseline finding using a branch and bound algorithm // Proc. of SPIE/IS&T 2002 Document Recognition & Retrieval IX Conf. (DR&RIX), Jan. 2002, pp. 20–27
- [7]Li Zhang, Chew-Lim Tan. Warped Document Image Restoration Using Shape-From-Shading and Physically-based Modeling // Proc. WACV, 2007, pp.29-32
- [8]Антон Масалович, Леонид Местецкий. Использование патча Безье для аппроксимации искажения изображений текстовых документов // Proc. Graphicon-2007, 2007, pp.171-174
- [9]Faisal Shafait, Thomas M. Breuel. Document Image Dewarping Contest // 2nd Int. Workshop on Camera-Based Document Analysis and Recognition, CBDAR'07, Brazil. Sep. 2007. pp 181-188.
- [10]Zhixi Bian, Yan Zhang. Retinex Image Enhancement Techniques // Final Project Report for EE264
- [11]John W. Tukey. Exploratory Data Analysis // Addison-Wesley, Reading, MA. 1977.
- [12]The Open Computer Vision Library [HTML] (<http://www.opencv.org/>)
- [13]Кормен, Лейзерсон, Ривест, Штайн. Алгоритмы. Построение и анализ // Издательский дом “Вильямс”, Москва, 2005, стр. 635
- [14]Шапиро, Стокман. Компьютерное зрение // Бином, Москва, 2006, стр. 172
- [15]Антон Конушин. Устойчивые алгоритмы оценки параметров модели на основе случайных выборок [HTML] (<http://cgm.computergraphics.ru/content/view/47/>)
- [16]Антон Конушин. Геометрические свойства нескольких изображений [HTML] (<http://www.cgm.computergraphics.ru/content/view/141/>)

[17]Legge GE, Pelli DG, Rubin GS, Schleske MM. Psychophysics of reading. I. Normal vision. // Vision Res., 1985;25: pp.239–252

[18]Зейналов Р. Ш., Геометрическая коррекция фотографий документов // Proc. Ломоносов-2008, 2008, pp. 60

[19]Зейналов Р. Ш., Геометрическая коррекция фотографий документов // Proc. Graphicon-2008, 2008, pp. 258-261

10. ОБ АВТОРАХ

Зейналов Рамиз Шакирович – выпускник факультета вычислительной математики и кибернетики Московского Государственного Университета им. М. В. Ломоносова.

E-Mail: ramiz.zeynalov@gmail.com

Велижев Александр Брониславович – к. т. н., научный сотрудник Лаборатории Компьютерной Графики и Мультимедиа факультета вычислительной математики и кибернетики Московского Государственного Университета им. М. В. Ломоносова.

E-Mail: avelizhev@graphics.cs.msu.ru

Конущин Антон Сергеевич – к. ф.-м. н., научный сотрудник, руководитель группы компьютерного зрения Лаборатории Компьютерной Графики и Мультимедиа факультета вычислительной математики и кибернетики Московского Государственного Университета им. М. В. Ломоносова.

E-Mail: ktosh@graphics.cs.msu.ru

Document images geometrical distortions correction using text lines shape extraction

Abstract

By reason of PDA and Smartphones with integrated digital cameras became widespread, people have started to use it to capture images from broadside of a book. Naturally the text on this-way captured images is suffered by both geometrical and perspective distortions because of non-planar geometry shape of the book broadside. The OCR methods couldn't get text from such images with satisfactory accuracy. In this article an enhancement of previous method of geometric correction of such images is described. There are some suggestions in new method retracted. Using this method the recognition accuracy and text lines straightness are improved. This method is based on the surface shape from text lines shapes extraction.

About the authors

Ramiz Zeinalov is graduate student at Moscow State University, Faculty of Computational Mathematics and Cybernetics. Contact E-Mail: ramiz.zeynalov@gmail.com

Alexander Velizhev is researcher at Moscow State University, Faculty of Computational Mathematics and Cybernetics. Contact E-Mail: avelizhev@graphics.cs.msu.ru

Anton Konushin is researcher at Moscow State University, Faculty of Computational Mathematics and Cybernetics. Contact E-Mail: ktosh@graphics.cs.msu.ru