

Аннотация

В данной работе рассматриваются различные автоматические методы исправления искажений на изображениях бумажных документов. Анализируются искажения, наиболее часто встречающиеся на фотографиях документов, и существующие методы борьбы с ними. Приводятся описание алгоритмов коррекции перспективных искажений и нормализации освещенности и их программная реализация. Предлагается метрика качества работы преобразований по исправлению искажений. Приводятся результаты работы описанных алгоритмов на примере изображений из собранной репрезентативной выборки.

Содержание

1 Введение	5
1.1 Постановка задачи	5
1.2 Актуальность работы	5
1.3 Цель и задачи работы	6
1.4 Набор данных	6
1.5 Анализ существующих решений	9
1.5.1 TurboScan	10
1.5.2 Abbyy FineScanner	10
1.5.3 Другие приложения	10
2 Исправление перспективных искажений	12
2.1 Описание искажения	12
2.2 Идея метода и ограничения	12
2.3 Детали реализации	14
2.3.1 Уменьшение входного изображения	15
2.3.2 Применение медианного фильтра	15
2.3.3 Детектирование границ	16
2.3.4 Преобразование Хафа	18
2.3.5 Выделение углов документа	20
2.3.6 Оценка соотношения длин сторон документа	22
2.3.7 Перспективное преобразование	27
2.4 Примеры работы и предложения по улучшению	27

3 Нормализация неравномерной освещенности	30
3.1 Описание искажения	30
3.2 Идея метода и ограничения	30
3.3 Детали реализации	32
3.3.1 Представление данных	32
3.3.2 Оценка фонового изображения	32
3.3.3 Вычисление исправленного изображения	34
3.4 Примеры работы	35
4 Измерение качества работы преобразований	37
4.1 Обоснование необходимости метрики	37
4.2 Описание метрики	38
4.3 Примеры измерения качества	39
5 Итоги и результаты работы	41
Список использованных источников	42
А Примеры работы алгоритма исправления перспективных искажений	45
Б Примеры работы алгоритма нормализации освещенности	61
В Примеры измерения качества работы предложенных алгоритмов	68
Г Примеры работы предложенных алгоритмов	73

Глава 1.

Введение

1.1. Постановка задачи

На сегодняшний день в связи с активно идущей цифровизацией экономики повышается потребность в переводе печатных документов в цифровой вид. Раньше основным средством получения цифрового изображения документа являлся сканер. Однако в последнее время доступность цифровых камер послужила залогом того, что именно они стали зачастую использоваться для получения цифрового изображения документа. При этом изображения документов на получающихся фотографиях страдают от различных видов неточностей, таких как перспективные искажения, неравномерное освещение, наличие бликов и ярких теней, геометрические искажения, вызванные изгибами и деформациями листа бумаги и другие. Все это затрудняет последующую работу, предполагающую оптическое распознавание информации с документа. Следовательно, требуется дополнительная предобработка изображений, направленная на подавление и исправление вышеперечисленных искажений.

1.2. Актуальность работы

Проблема устранения искажений на цифровых изображениях печатных документов достаточно давно беспокоит умы людей. Существует множество исследований, нацеленных на изучение конкретных видов искажений и способов их исправления. Также существует немало

программных продуктов, решающих данную задачу с той или иной долей успеха. Тем не менее, до сих пор не существует идеального решения, способного отлично справляться со всеми видами искажений на изображении произвольного документа. Необходимость разработки новых более универсальных подходов коррекции искажений, возникающих при фотографировании произвольных печатных документов, определяет актуальность темы данной работы.

1.3. Цель и задачи работы

Целью данной работы является разработка механизма удаления искажений с изображений печатных документов.

Задачами работы являются:

1. Сбор репрезентативной выборки фотографий печатных документов, содержащих различные виды искажений
2. Выявление искажений, наиболее часто встречающихся на изображениях
3. Анализ существующих инструментов коррекции выявленных искажений, их достоинств и ограничений
4. Разработка и реализация в программном коде методов автоматического устранения искажений
5. Разработка способа оценки качества работы созданных алгоритмов
6. Тестирование разработанных методов на собранных данных и систематизация полученных результатов

1.4. Набор данных

Для данной работы был собран набор фотографий печатных документов со всевозможными видами искажений. Цифровые

изображения были получены фотокамерами разнообразных смартфонов при различающихся внешних условиях. Фотографии также имеют разное разрешение. Всего было собрано порядка нескольких сотен изображений. Примеры фотографий, содержащихся в собранном наборе, приводятся в табл. 1.1.

Стоит отметить разнообразие документов, которые подвергались оцифровке. Документы могут иметь разный формат и соотношение сторон. Тексты на документах могут быть произвольного размера, на различных языках, шрифт надписей может быть как печатным, так и рукописным. В документах, помимо текста, могут содержаться и более сложные структурные элементы, например печати, таблицы, диаграммы, рисунки. Некоторые типы документов могут иметь строго фиксированную общеизвестную структуру, как, например, паспорт гражданина РФ или определенный тип бланков, в то время как структура других может быть произвольной. Документы могут быть как цветными, так и одноцветными, причем цвет бумаги документа также может быть различным. При этом, еще до оцифровки сам документ может обладать некоторыми дефектами, вызванными, например, плохим качеством печати или используемой бумаги.

Как видно из примеров, содержащихся в табл. 1.1, получающиеся фотографированием цифровые изображения бумажных документов страдают от различных видов искажений и поэтому не всегда лучшим образом передают информацию о содержании оригинального документа. Условия, при которых получают фотографии документов всегда разные, что, в частности, сказывается на освещенности получаемых изображений. Хорошой иллюстрацией описанной проблемы являются изображения 1d и 4b. Нередко люди хранят важные документы в пластиковых прозрачных файлах и перед фотографированием забывают достать из них документы, забывая о том, что эти файлы имеют прекрасно отражающую поверхность, что приводит к ярким бликам, которые мешают увидеть часть информации

№	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
1	An open passport showing two pages with official seals and a stamp page.	An open stamp book with multiple pages containing various official stamps.	A document with a large blue circular stamp.	A large document with a prominent blue circular stamp.
2	A receipt from a medical facility and a printed document with a QR code.	A receipt from a medical facility.	A document with handwritten text and several signatures.	A large document with a detailed table and a stamp at the bottom.
3	A receipt from a medical facility and a printed document with a QR code.	A receipt from a medical facility.	A document with a large blue circular stamp.	A large document with a detailed table and a stamp at the bottom.
4	A receipt from a medical facility and a printed document with a QR code.	A receipt from a medical facility.	A document with a large blue circular stamp.	A large document with a detailed table and a stamp at the bottom.

Таблица 1.1 – Примеры изображений печатных документов, содержащихся в собранной выборке

на документе, как на изображениях 3a и 4c. Часто на фотографиях, помимо документов, присутствует однородный или неоднородный, контрастный или сливающийся с документом внешний фон, как на изображениях 2b и 3c, что тоже не способствует распознаванию информации с документа. Иногда, стараясь избавиться от внешнего фона, фотограф не только не захватывает фон, но и достигает обратного эффекта - на изображении отсутствуют границы документа, а иногда и, возможно, часть ценной информации с документа, как на изображениях 2c и 4a. Так как при фотографировании бывает достаточно сложно расположить документ так, чтобы его плоскость была перпендикулярна оси фотокамеры, то неизбежно возникновение перспективных искажений, как на изображениях 4c, 2c и 2b. При фотографировании не всегда под рукой есть ровная поверхность, на которую можно было бы положить документ, поэтому часто приходится держать его в руках, и поэтому на изображениях присутствуют "пальцы", как на изображениях 3c, 3d, 4a. При таком стечении обстоятельств также могут возникнуть геометрические искажения, вызванные деформациями листа бумаги, как на изображении 3d или 4a. Наконец, нередки случаи, когда перечисленные искажения комбинируются на одном изображении.

1.5. Анализ существующих решений

В настоящее время существует множество программных продуктов, целью которых является получение цифровой копии документа в максимально приближенном к оригиналу виде. Самый частый формат таких систем оцифровки – это приложение для смартфонов. Многие из этих приложений заточены под задачу распознавания текста на документе при помощи различных технологий оптического распознавания символов (optical character recognition, OCR), а не под задачу получения цифровой копии документа без искажений в общем виде. Из-за закрытого исходного кода бывает сложно оценить ограничения и методы работы конкретной

системы и сравнить их между собой. Далее рассмотрим несколько таких приложений.

1.5.1. TurboScan

Приложение TurboScan[1] предназначено для определения границ документа на изображении и устраниении перспективного искажения. Определение границ документа происходит в полуавтоматическом режиме, то есть пользователь может вручную скорректировать местонахождение углов и сторон документа. Отличительной особенностью данного программного обеспечения является режим, в котором программа делает три различных фотографии одного документа и комбинирует их для получения четкого изображения и более эффективного устранения искажений.

1.5.2. Abbyy FineScanner

Программа для смартфонов Abbyy FineScanner[2] автоматически распознает документ на фотографии, определяет его границы, убирает внешний фон, выполняет коррекцию перспективы и выравнивает строки, чтобы устранить геометрические искажения, вызванные изгибами бумаги. Присутствует специальный режим для обработки текстов на разворотах книг. Также есть возможность произвести коррекцию освещения, выбрав один из нескольких доступных вариантов для черно-белого или цветного документа. Данная программа имеет функционал распознавания печатного текста на 193 языках и предлагает экспортовать распознанный текстовый файл в один из 12 текстовых форматов. Одно из наиболее универсальных приложений для сканирования документа по фотографии.

1.5.3. Другие приложения

Наряду с вышеперечисленными приложениями были рассмотрены и другие программы, включая Document Scanner Pro[3], Genius Scan[4] и

Simple Scan[5]. Результаты работы разных сравниваемых систем на одном входном изображении приведены на рис. 1.1.

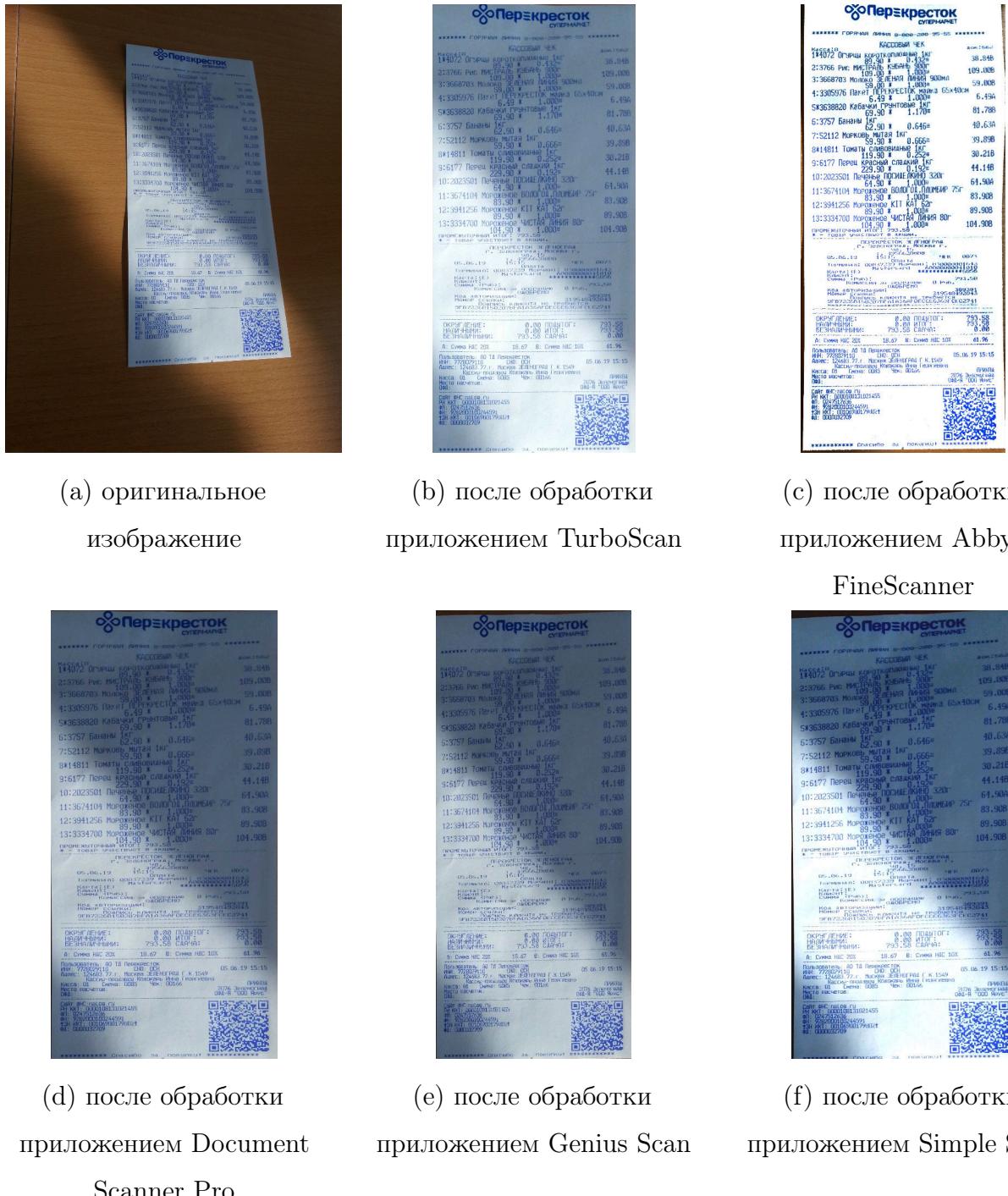


Рисунок 1.1 – Сравнение результатов работы различных приложений

Глава 2.

Исправление перспективных искажений

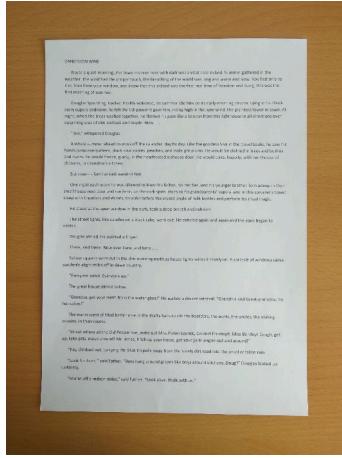
2.1. Описание искажения

Перспективные искажения – это явление, возникающее при перспективном проецировании объектов трехмерного пространства на двумерную поверхность, что имеет место в том числе и в процессе фотографирования. При съемке более удаленные от объектива фотокамеры объекты изображаются в меньших масштабах, а параллельные в пространстве прямые в общем случае оказываются на получающихся изображениях непараллельными. Зачастую такие изображения, называемые перспективными, кажутся наблюдателю естественными и правильными. Однако, в случае задачи сканирования документа, перспективные искажения становятся серьезной проблемой, мешающей распознаванию информации с изображения. Особенно существенными получаются перспективные искажения в случае, когда оптическая ось фотокамеры не перпендикулярна поверхности документа в процессе съемки.

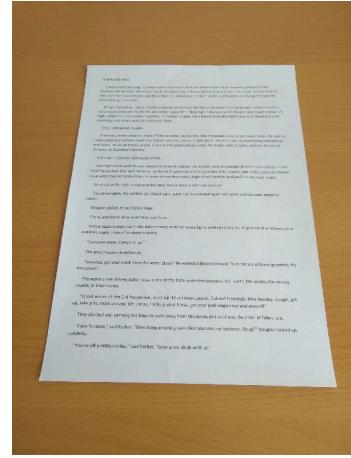
Примеры перспективных изображений приведены на рис. 2.1.

2.2. Идея метода и ограничения

Многие из существующих исследований, посвященных коррекции перспективных искажений на фотографиях документов, основаны на текстовой информации, содержащейся в документе. Например, в статьях [6] и [7] описывается метод, использующий идею восстановления



(a) изображение с несущественными
перспективными искажениями



(b) изображение с существенными
перспективными искажениями

Рисунок 2.1 – Примеры фотографий с перспективными искажениями

местонахождения точек схода (*vanishing points*). Возможность нахождения этих точек достигается за счет предположения, что текст на документе имеет определенную структуру, а именно строки и места начала и конца строк образуют четкие условно горизонтальные и вертикальные линии соответственно. Методы, предложенные в статьях [8] и [9] также решают задачу коррекции перспективы, основываясь на текстовой информации в документе. Но в отличие от предыдущих статей, они пытаются оценить вертикальную точку схода не за счет наличия четкой структуры параграфов, а за счет выделения вертикальных отрезков у отдельных символов (*vertical stroke boundary*). При этом в процессе работы их методы активно используют для достижения этой цели в том числе морфологические преобразования.

Произвольные документы могут иметь сложную структуру и содержать не только текст. Поэтому в этой главе будет предложен алгоритм коррекции перспективных искажений, работающий на любом документе, независимо от содержания и даже формы листа (единственное требование заключается в том, что документ должен быть прямоугольным, но при этом с любым соотношением сторон).

Идея данного алгоритма заключается в определении местоположения

границ документа на изображении, оценки соотношения длин сторон исходного документа и подборе перспективного преобразования, применение которого к начальной фотографии позволило бы получить изображение документа прямоугольной формы с заданным соотношением длин сторон. Алгоритм также, помимо коррекции перспективы, удаляет внешний по отношению к документу фон.

Данный алгоритм работает в следующих предположениях:

1. Документ имеет прямоугольную форму в выровненном состоянии.

Благодаря этому требованию, можно рассчитывать на то, что перспективная проекция документа будет иметь форму четырехугольника, и это обеспечит корректность дальнейшей оценки соотношения длин сторон документа. Кроме того, подавляющее большинство документов удовлетворяют этому условию.

2. На входном изображении присутствуют четкие границы документа.

В частности, данное предписание достигается за счет контрастного и однородного внешнего фона. Это значительно облегчает правильное определения сторон и углов документа. Алгоритм применим и в случае неоднородного и неконтрастного внешнего фона, но тогда может ухудшиться качество работы.

3. Геометрические искажения, вызванные деформациями и изгибами листа бумаги документа, отсутствуют. Этого можно добиться, если положить документ на ровную поверхность при фотографировании. Данное ограничение вызвано тем, что изогнутый документ на изображении может не иметь форму четырехугольника в общем случае.

2.3. Детали реализации

В данном разделе приводится подробное поэтапное описание реализации предложенного алгоритма по исправлению перспективных

искажений. Каждый следующий шаг алгоритма применяется к результату работы с предыдущего шага.

2.3.1. Уменьшение входного изображения

Будем считать, что исходная фотография удовлетворяет описанным выше ограничениям алгоритма. Входное изображение представляется из себя цветную картинку в цветовом пространстве RGB, обычно размером в несколько тысяч на несколько тысяч пикселей. Работать с таким большим изображением, имеющим объем несколько мегабайт, может быть достаточно ресурсоемким процессом. Так как данный алгоритм разрабатывался с намерением внедрения на мобильные устройства, то первым шагом алгоритма является уменьшение входного изображения с сохранением пропорций таким образом, чтобы большая сторона полученного изображения была порядка 300 пикселей¹. Данная процедура не слишком сильно влияет на результат применения алгоритма, так как в конечном итоге приведет к ошибке определения углов документа всего в несколько пикселей, что незаметно в масштабах изображения, существенно сокращая при этом количество требуемых вычислений.

2.3.2. Применение медианного фильтра

Текущая цель заключается в определении контуров документа на фотографии. Однако сразу пытаться найти границы на изображении бывает затруднительно из-за наличия большого количества мелких деталей, к коим относятся неоднородности на внешнем фоне и содержание документа, включая буквы, строки и столбцы таблиц, небольшие рисунки и другие структурные элементы документа. Поэтому на втором шаге алгоритма к текущему изображению 3 раза применяется медианный фильтр с размером ядра, равного 5. Данная операция часто успешно позволяет размыть все

¹Здесь и далее числовые параметры алгоритма были подобраны эмпирически в процессе тестирования



(a) исходное уменьшенное изображение

(b) результат одиночного применения медианного фильтра

(c) результат применения медианного фильтра три раза подряд

Рисунок 2.2 – Пример применения медианного фильтра для размытия содержания документа

относительно мелкие детали, оставив при этом достаточно четкую границу документа с внешним контрастным фоном. При этом медианный фильтр используется не единожды, а трижды, так как после одинарного применения на изображении зачастую остаются мелкие пятна на местах скопления символов. Пример результата данного шага продемонстрирован на рис. 2.2

2.3.3. Детектирование границ

На текущем шаге для обнаружения контуров документа на изображении применяется хорошо зарекомендовавший себя во многих задачах компьютерного зрения, где требуется выделение границ, оператор Кэнни[10]. Данный алгоритм состоит из нескольких этапов.

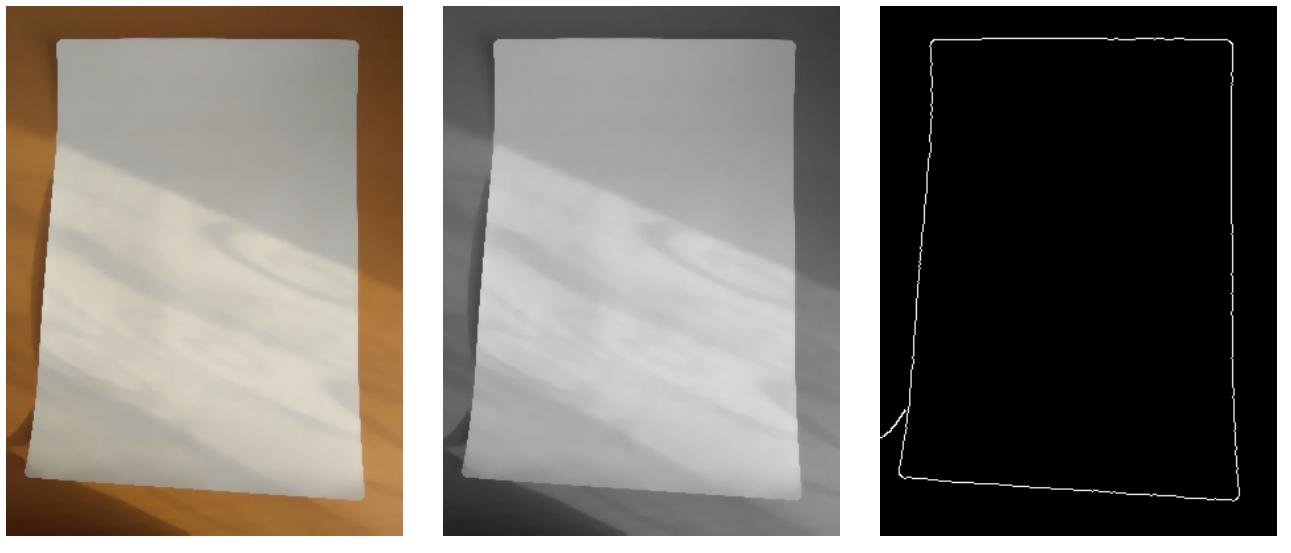
- Вначале входное изображение конвертируется в оттенки серого цвета. Этот шаг призван уменьшить последующие вычислительные затраты.
- Затем полученное одноканальное изображение сглаживается с

использованием фильтра Гаусса для подавления шума².

- Третий этап заключается в поиске вектора градиента для каждой точки с использованием операторов Собеля. Угол каждого вектора округляется до одного из 4 значений: 0° , 45° , 90° и 135° .
- На следующей стадии, называющейся подавлением немаксимумов, точками границ объявляются только те пиксели, в которых достигается локальный максимум градиента в направлении градиента. То есть, например, если угол градиента равен 0° , то такая точка объявляется граничной, если модуль градиента в этом пикселе больше модуля градиента слева и справа.
- После этого к выделенным на предыдущем этапе граничным точкам применяется двойная пороговая фильтрация. Выбираются два значения t_1 и t_2 , причем $t_1 < t_2$. Если значение градиента в пикселе превышает значение t_2 , то такая граница считается достоверной. Если значение меньше t_1 , то такая граница подавляется. Если значение лежит в промежутке от t_1 до t_2 , то про такие пиксели говорят, что они находятся в области неоднозначности.³
- Заключительная стадия призвана решить вопрос о том, какие из промежуточных точек также стоит отнести к граничным точкам. К итоговому результату, состоящему вначале только из достоверных границ, постепенно добавляются точки из области неоднозначности,

²В нашем случае данный шаг можно пропустить, так как к изображению уже применялся медианный фильтр, который тоже является отличным средством шумоподавления

³Примечание. Качество работы алгоритма нахождения границ сильно зависит от выбора параметров t_1 и t_2 . Например, если выбрать t_2 слишком большим, то из конечного результата пропадут многие точки границ, а если t_2 слишком мало, то в качестве границ часто будут выбираться шумовые точки. В нашей работе используется выбор параметров по правилу $t_2 = t$ и $t_1 = \frac{1}{2}t$, где t – пороговое значение бинаризации из метода Оцу[11]. Данный подход выбора параметров был предложен в статье [12], и отлично себя зарекомендовал во многих прикладных задачах (см., например, [13]).



(a) уменьшенное размытое изображение (b) в оттенках серого цвета (c) результат применения оператора Кэнни

Рисунок 2.3 – Пример применения оператора Кэнни для обнаружения границ

прилегающие вплотную к уже отобранным в качестве границ точкам. Данная операция прекращается, когда новое добавление невозможно.

Результатом работы данного алгоритма является бинарное изображение, белые пиксели которого соответствуют найденным граничным точкам. Пример работы алгоритма приведен на рис. 2.3. Отдельное внимание стоит уделить обнаруженной границе на рис. 2.3с внизу слева, которая не относится к контуру документа. Такие ложноположительные результаты определения границ возникают из-за неоднородного внешнего фона, в частности, как в данном случае, из-за теней.

2.3.4. Преобразование Хафа

Далее для выделения прямых линий, соответствующих контурам документа к бинарному изображению применяется преобразование Хафа[14].

Ненулевые элементы бинарного изображения можно рассматривать как точки на плоскости (см. рис. 2.5). Прямая линия на плоскости может

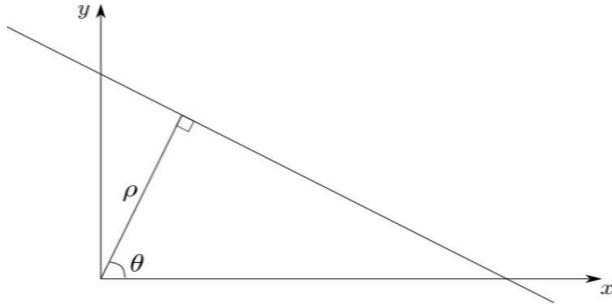


Рисунок 2.4 – Пояснение к методу заданию прямой на плоскости через параметры θ и ρ

быть параметрически задана в виде

$$\rho = x \cos \theta + y \sin \theta, \quad (2.1)$$

где θ – угол между осью абсцисс и перпендикуляром, проведенным к прямой из начала координат, а ρ – длина этого перпендикуляра (см. рис. 2.4). Заметим, что при $\sin \theta \neq 0$ (то есть, если прямая не вертикальная), уравнение (2.1) можно почленно разделить на $\sin \theta$ и получить привычное представление прямой на плоскости

$$y = -\operatorname{ctg} \theta \cdot x + \frac{\rho}{\sin \theta} \quad \text{или} \quad y = kx + b. \quad (2.2)$$

Таким образом, любая прямая однозначно характеризуется парой (θ, ρ) .

На этапе инициализации в преобразовании Хафа непрерывное двумерное пространство (θ, ρ) разбивается на равные дискретные ячейки некоторого заранее выбранного размера, и представляется в виде двумерной матрицы, называемой накопительным пространством или пространством Хафа. Каждая ячейка в таком пространстве соответствует некоторому множеству близких по углу и расстоянию до начала координат прямых, из которых выбирается лишь одна прямая в качестве представителя данной ячейки. В процессе работы для каждой такой ячейки считается количество точек плоскости, лежащие достаточно близко к прямой, соответствующей данной ячейке. Нетрудно видеть, что одна точка плоскости соответствует множеству прямых, задающих синусоиду в

пространстве (θ, ρ) . Таким образом, процедуру заполнения накопительной матрицы можно рассматривать как процесс голосования точек плоскости за множество ячеек, соответствующих такой синусоиде. Когда накопительная матрица заполнена, ячейки с наибольшим числом голосов соответствуют наиболее выразительным прямым на плоскости изображения.

В качестве прямых, являющихся аппроксимацией линий сторон документа, будем выбирать прямые по следующему правилу. Первая прямая будет соответствовать паре параметров (θ, ρ) , за которые были отданы наибольшее количество голосов. Остальные выбираемые линии соответствуют локальным максимумам среди параметров (θ, ρ) , которые набрали не менее 50% голосов от наибольшего максимума. В большинстве случаев, выбранных прямых оказывается около 4. Таким образом получаем наиболее вероятные прямые, на которых лежат границы документа.

На рис. 2.5 изображен пример нахождения прямых, на которых лежат стороны документа, при помощи преобразования Хафа. На рис. 2.5b продемонстрирована визуализация накопительного пространства. На ней черные точки соответствуют ячейкам матрицы с нулевым количеством голосов, и чем светлее точка, тем большее количество голосов было отдано за соответствующую ячейку.

2.3.5. Выделение углов документа

На текущем шаге требуется определить местонахождение углов документа. В оптимальном случае на предыдущем этапе нашлись ровно 4 прямые, содержащие стороны документа. Заметим, что углы тогда можно оценить как точки пересечения прямых, содержащих условно горизонтальные стороны документа, с прямыми, содержащими условно вертикальные стороны. Для этого необходимо разбить прямые на 2 группы на основании величины угла θ , который принимает значения в диапазоне от 0° до 180° . Это можно сделать, например, сопоставив углу θ точку

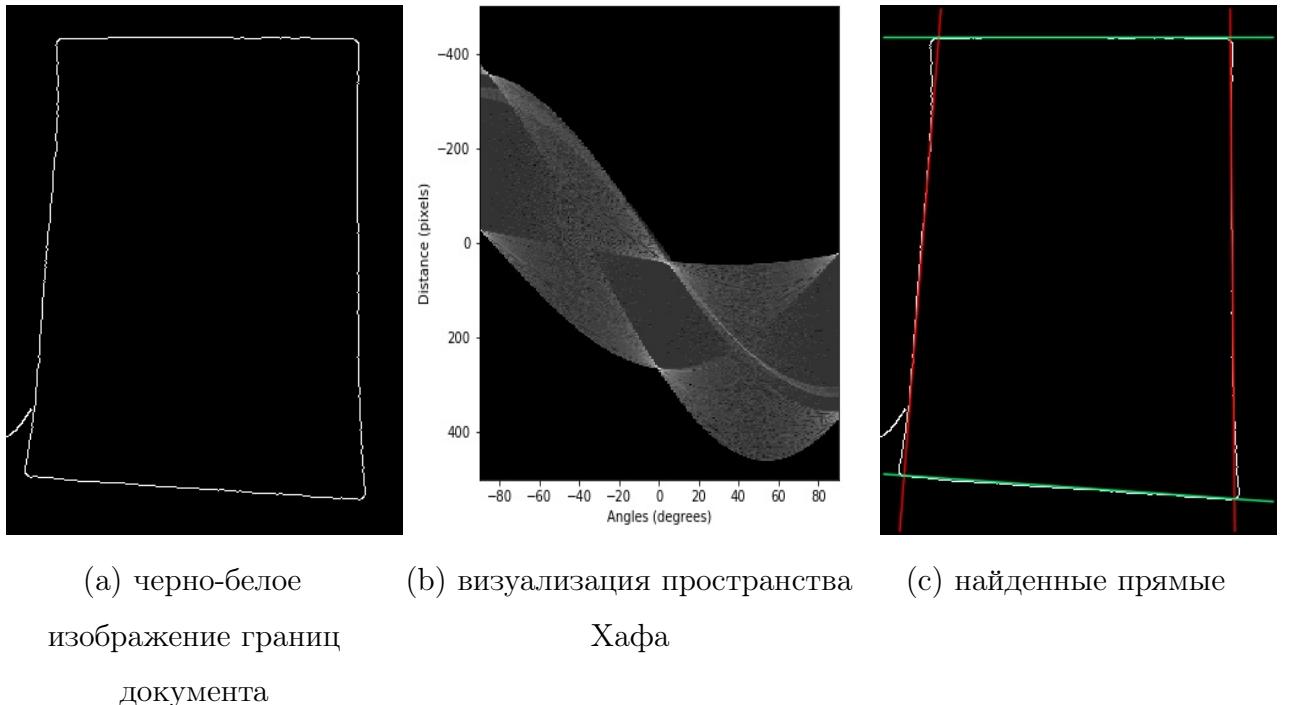


Рисунок 2.5 – Пример применения преобразования Хафа для выделения прямых

на единичной окружности в двумерном пространстве вида $(\cos 2\theta, \sin 2\theta)$. Домножение угла на коэффициент 2 нужно в этом случае для того, чтобы углы порядка 180° находились рядом с углами величины 0° , ведь они фактически задают прямые с одинаковым наклоном. Заметим, что при таком сопоставлении близкие по величине углы будут соответствовать близким точкам двумерного пространства в смысле евклидовой метрики, и с увеличением дуги между двумя углами будет расти расстояние между точками. Далее точки двумерного пространства можно кластеризовать на 2 класса, используя, например, метод *k-means*, и таким образом разбить прямые на 2 группы. Как уже говорилось, далее углы документа можно найти как точки пересечения прямых из первой группы с прямыми из второй. В случае, когда после преобразования Хафа нашлось больше четырех прямых, могут быть разные стратегии оценки местонахождения углов документа. Об одном из возможных подходов будет рассказано в разд. 2.4.

На этом же шаге можно вернуться от уменьшенного изображения

к оригинальному, увеличив координаты точек углов документа на уменьшенном изображении во столько раз, во сколько мы сжимали начальную фотографию до уменьшенной.

2.3.6. Оценка соотношения длин сторон документа

Итак, к этому шагу мы определили местонахождение углов документа на оригинальном изображении. Необходимо подобрать такое перспективное преобразование, которое бы отображало поверхность документа, имеющую форму произвольного четырехугольника в прямоугольник, соответствующий выровненному изображению документа. На данном этапе, однако, возникает вопрос о том, каковы параметры этого прямоугольника. В частности, нас интересуют длина стороны прямоугольника и соотношение длин его сторон. Рассмотрим сейчас метод оценки пропорции сторон документа, приняв во внимание геометрию получения фотографии. Данный метод был предложен в статье [15].

Без ограничения общности, будем считать, что прямоугольный документ шириной w и высотой h в процессе получения фотографии находился в плоскости $z = 0$ в глобальной системе координат (см. рис. 2.6). Пусть углы документа \mathbf{M}_i ($i = 1, \dots, 4$) имеют координаты $(0, 0)$, $(w, 0)$, $(0, h)$ и (w, h) в плоскости $z = 0$ соответственно. Получающаяся проекция прямоугольника на плоскость изображения имеет форму четырехугольника. Наблюдаемые углы документа обозначим через \mathbf{m}_i . Далее для удобства, если задан вектор $\mathbf{x} = [x_1, \dots, x_n]^T$, через $\tilde{\mathbf{x}}$ будем обозначать вектор $[x_1, \dots, x_n, 1]^T$ для упрощения работы с однородными координатами.

Будем использовать стандартную модель проективной камеры (*projective camera, pinhole camera model*). В такой модели произвольная точка пространства \mathbf{M} проецируется на плоскость изображения в точку \mathbf{m} по правилу:

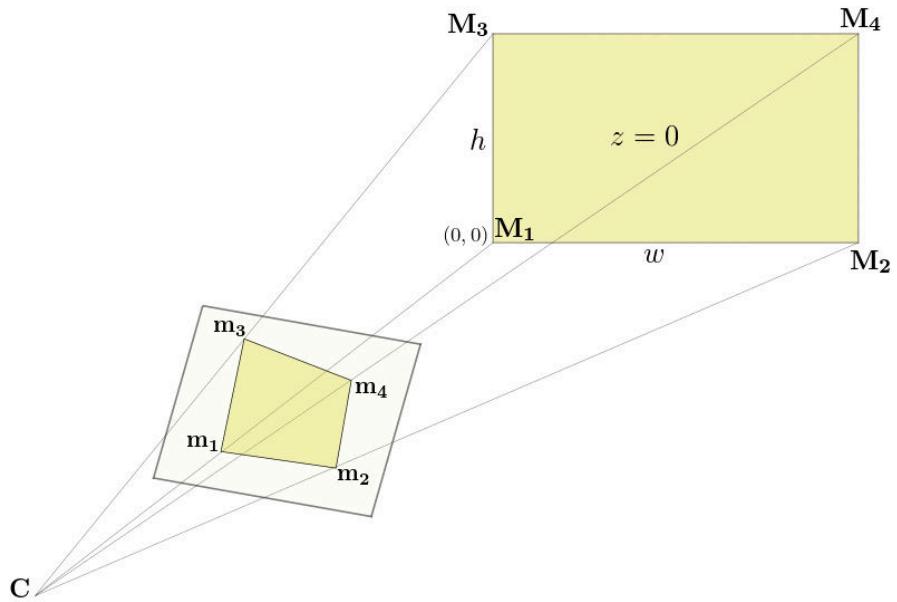


Рисунок 2.6 – Геометрия перспективного проецирования при получении изображения

$$\lambda \tilde{\mathbf{m}} = A[R|\mathbf{t}]\widetilde{\mathbf{M}}, \quad (2.3)$$

где $A = \begin{pmatrix} f & 0 & u_0 \\ 0 & sf & v_0 \\ 0 & 0 & 1 \end{pmatrix}$ – верхняя треугольная матрица внутренних параметров камеры, f - фокусное расстояние камеры, s - соотношение сторон одного пикселя, (u_0, v_0) – координаты точки проекции точек объекта в пространстве, находящихся на оптической оси камеры, $R = [\mathbf{r}_1|\mathbf{r}_2|\mathbf{r}_3]$ – ортогональная матрица размера 3×3 , определяющая поворот камеры относительно глобальной системы координат, \mathbf{t} – вектор сдвига (по-другому, $\mathbf{t} = R\mathbf{c}$, где \mathbf{c} – координаты центра камеры в глобальной системе координат). То есть пара (R, \mathbf{t}) описывает преобразование между глобальной системой координат и системой координат камеры.

Подставляя координаты углов в выражение (2.3), получаем:

$$\left\{ \begin{array}{l} \lambda_1 \tilde{\mathbf{m}}_1 = A\mathbf{t}, \\ \lambda_2 \tilde{\mathbf{m}}_2 = wA\mathbf{r}_1 + A\mathbf{t}, \\ \lambda_3 \tilde{\mathbf{m}}_3 = hA\mathbf{r}_2 + A\mathbf{t}, \\ \lambda_4 \tilde{\mathbf{m}}_4 = wA\mathbf{r}_1 + hA\mathbf{r}_2 + A\mathbf{t}. \end{array} \right. \quad (2.4a)$$

$$\left\{ \begin{array}{l} \lambda_2 \tilde{\mathbf{m}}_2 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1, \\ \lambda_3 \tilde{\mathbf{m}}_3 - \lambda_1 \tilde{\mathbf{m}}_1 = hA\mathbf{r}_2, \\ \lambda_4 \tilde{\mathbf{m}}_4 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1 + hA\mathbf{r}_2. \end{array} \right. \quad (2.4b)$$

$$\left\{ \begin{array}{l} \lambda_2 \tilde{\mathbf{m}}_2 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1, \\ \lambda_3 \tilde{\mathbf{m}}_3 - \lambda_1 \tilde{\mathbf{m}}_1 = hA\mathbf{r}_2, \\ \lambda_4 \tilde{\mathbf{m}}_4 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1 + hA\mathbf{r}_2. \end{array} \right. \quad (2.4c)$$

$$\left\{ \begin{array}{l} \lambda_2 \tilde{\mathbf{m}}_2 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1, \\ \lambda_3 \tilde{\mathbf{m}}_3 - \lambda_1 \tilde{\mathbf{m}}_1 = hA\mathbf{r}_2, \\ \lambda_4 \tilde{\mathbf{m}}_4 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1 + hA\mathbf{r}_2. \end{array} \right. \quad (2.4d)$$

Выполняя (2.4b) - (2.4a), (2.4c) - (2.4a) и (2.4d) - (2.4a), получим:

$$\left\{ \begin{array}{l} \lambda_2 \tilde{\mathbf{m}}_2 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1, \\ \lambda_3 \tilde{\mathbf{m}}_3 - \lambda_1 \tilde{\mathbf{m}}_1 = hA\mathbf{r}_2, \\ \lambda_4 \tilde{\mathbf{m}}_4 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1 + hA\mathbf{r}_2. \end{array} \right. \quad (2.5a)$$

$$\left\{ \begin{array}{l} \lambda_2 \tilde{\mathbf{m}}_2 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1, \\ \lambda_3 \tilde{\mathbf{m}}_3 - \lambda_1 \tilde{\mathbf{m}}_1 = hA\mathbf{r}_2, \\ \lambda_4 \tilde{\mathbf{m}}_4 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1 + hA\mathbf{r}_2. \end{array} \right. \quad (2.5b)$$

$$\left\{ \begin{array}{l} \lambda_2 \tilde{\mathbf{m}}_2 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1, \\ \lambda_3 \tilde{\mathbf{m}}_3 - \lambda_1 \tilde{\mathbf{m}}_1 = hA\mathbf{r}_2, \\ \lambda_4 \tilde{\mathbf{m}}_4 - \lambda_1 \tilde{\mathbf{m}}_1 = wA\mathbf{r}_1 + hA\mathbf{r}_2. \end{array} \right. \quad (2.5c)$$

Вычитание (2.5c) - (2.5a) - (2.5b) дает

$$\lambda_4 \tilde{\mathbf{m}}_4 = \lambda_2 \tilde{\mathbf{m}}_2 + \lambda_3 \tilde{\mathbf{m}}_3 - \lambda_1 \tilde{\mathbf{m}}_1. \quad (2.6)$$

Домножим векторно каждую часть равенства на $\tilde{\mathbf{m}}_4$:

$$0 = \lambda_2 \tilde{\mathbf{m}}_2 \times \tilde{\mathbf{m}}_4 + \lambda_3 \tilde{\mathbf{m}}_3 \times \tilde{\mathbf{m}}_4 - \lambda_1 \tilde{\mathbf{m}}_1 \times \tilde{\mathbf{m}}_4, \quad (2.7)$$

и теперь скалярно на $\tilde{\mathbf{m}}_3$:

$$\lambda_2 (\tilde{\mathbf{m}}_2 \times \tilde{\mathbf{m}}_4) \tilde{\mathbf{m}}_3 = \lambda_1 (\tilde{\mathbf{m}}_1 \times \tilde{\mathbf{m}}_4) \tilde{\mathbf{m}}_3. \quad (2.8)$$

Вводя обозначение $\lambda_2 = k_2 \lambda_1$, получаем выражение на k_2 :

$$k_2 = \frac{(\tilde{\mathbf{m}}_1 \times \tilde{\mathbf{m}}_4) \tilde{\mathbf{m}}_3}{(\tilde{\mathbf{m}}_2 \times \tilde{\mathbf{m}}_4) \tilde{\mathbf{m}}_3}. \quad (2.9)$$

Аналогичным образом можно ввести обозначение $\lambda_3 = k_3 \lambda_1$ и получить выражение на k_3 :

$$k_3 = \frac{(\tilde{\mathbf{m}}_1 \times \tilde{\mathbf{m}}_4) \tilde{\mathbf{m}}_2}{(\tilde{\mathbf{m}}_3 \times \tilde{\mathbf{m}}_4) \tilde{\mathbf{m}}_2}. \quad (2.10)$$

Подстановка (2.9) в (2.5a) дает

$$\mathbf{r}_1 = \lambda_1 w^{-1} A^{-1} \mathbf{n}_2, \quad (2.11)$$

где

$$\mathbf{n}_2 = k_2 \tilde{\mathbf{m}}_2 - \tilde{\mathbf{m}}_1. \quad (2.12)$$

Аналогично, подставляя (2.10) в (2.5b), получим

$$\mathbf{r}_2 = \lambda_1 h^{-1} A^{-1} \mathbf{n}_3, \quad (2.13)$$

где

$$\mathbf{n}_3 = k_3 \tilde{\mathbf{m}}_3 - \tilde{\mathbf{m}}_1. \quad (2.14)$$

Из свойств матрицы поворота R , мы имеем $\mathbf{r}_1 \cdot \mathbf{r}_2 = 0$. Поэтому из уравнений (2.11) и (2.13) следует:

$$\mathbf{n}_2^T A^{-T} A^{-1} \mathbf{n}_3 = 0. \quad (2.15)$$

Опять же из свойств матрицы поворота, $\mathbf{r}_1 \cdot \mathbf{r}_1 = 1$ и $\mathbf{r}_2 \cdot \mathbf{r}_2 = 1$. Из уравнений (2.11) и (2.13) получаем соответственно:

$$\left\{ \begin{array}{l} 1 = \lambda_1^2 w^{-2} \lambda_2 \mathbf{n}_2^T A^{-T} A^{-1} \mathbf{n}_2, \end{array} \right. \quad (2.16a)$$

$$\left\{ \begin{array}{l} 1 = \lambda_1^2 h^{-2} \lambda_2 \mathbf{n}_3^T A^{-T} A^{-1} \mathbf{n}_3. \end{array} \right. \quad (2.16b)$$

Деление (2.16b) на (2.16a) дает нам

$$\left(\frac{w}{h} \right)^2 = \frac{\mathbf{n}_2^T A^{-T} A^{-1} \mathbf{n}_2}{\mathbf{n}_3^T A^{-T} A^{-1} \mathbf{n}_3}. \quad (2.17)$$

Сейчас мы хотим извлечь всю возможную полезную информацию из расположения четырехугольника на изображении, при этом у нас нет какого-либо дополнительного знания о расположении этого четырехугольника в пространстве. Мы можем сделать несколько предположений о характеристиках камеры. Во-первых, разумно предположить, что пиксели камеры квадратные ($s = 1$), что справедливо в современных камерах, и, во-вторых, точки пространства, лежащие на оптической оси камеры, проецируются в центр изображения (u_0 и v_0 равны полуширине и полувысоте изображения соответственно). В этих предположениях можно вычислить

фокусное расстояние камеры f из уравнения (2.15)

$$f^2 = -\frac{1}{n_{23}n_{33}s^2} \left\{ \left[n_{21}n_{31} - (n_{21}n_{33} + n_{23}n_{31})u_0 + n_{23}n_{33}u_0^2 \right] s^2 + \left[n_{22}n_{32} - (n_{22}n_{33} + n_{23}n_{32})v_0 + n_{23}n_{33}v_0^2 \right] \right\}, \quad (2.18)$$

где n_{2i} (n_{3i}) – это i -я компонента вектора \mathbf{n}_2 (\mathbf{n}_3 соответственно). Получив такую оценку на фокусное расстояние, мы теперь знаем все параметры камеры, и искомое соотношение длин сторон прямоугольника может быть вычислено из формулы (2.17).

Оговоримся, что если $n_{23} = 0$ или $n_{33} = 0$, то оценка (2.18) на фокусное расстояние не может быть получена. Это происходит при $k_2 = 1$ или $k_3 = 1$ соответственно. Тем не менее, в статье [15] доказано, что это соответствует единственному случаю, когда на фотографии документ также имеет форму прямоугольника. При этих обстоятельствах, можно пренебречь влиянием перспективных искажений, и тогда оценка соотношения длин сторон документа тривиальна. В программной реализации считается, что наступает данный случай, если хотя бы одна из переменных k_2 или k_3 отличается от значения 1 меньше чем на 10^{-3} .

Итак, мы научились находить пропорции документа. Заметим, что абсолютный размер прямоугольника не может быть вычислен. Это очевидно, так как больший по размеру прямоугольник, находящийся дальше от камеры, давал бы абсолютно такое же изображение. Поэтому для восстановления документа в качестве длины его стороны можно взять некоторый характерный размер документа на изображении. Опишем здесь, как выбираются высота H и ширина W исправленного изображения. Пусть W_1 и W_2 – длины верхней и нижней сторон документа до исправления искажения, а H_1 и H_2 – длины боковых сторон документа до исправления искажения. Обозначим $\hat{H} = \max(H_1, H_2)$, $\hat{W} = \max(W_1, W_2)$ и $\hat{r} = \frac{\hat{W}}{\hat{H}}$. Пусть r – оценка на отношение ширины к высоте документа, полученная описанным в этом разделе методом (по нашим рассуждениям тогда должно

выполняться $\frac{W}{H} = r$). Определим тогда размеры выровненного изображения по правилу:

$$\begin{cases} W = \hat{W} \text{ и } H = \frac{W}{r}, & \text{если } \hat{r} \geq r, \\ H = \hat{H} \text{ и } W = rH, & \text{иначе.} \end{cases} \quad (2.19a)$$

$$(2.19b)$$

Заметим, что при таком задании высоты и ширины выровненного изображения каждый пиксель документа входного изображения будет соответствовать по крайней мере одному пикслю документа на выходном изображении, то есть не будет потери информации из-за сжатия.

2.3.7. Перспективное преобразование

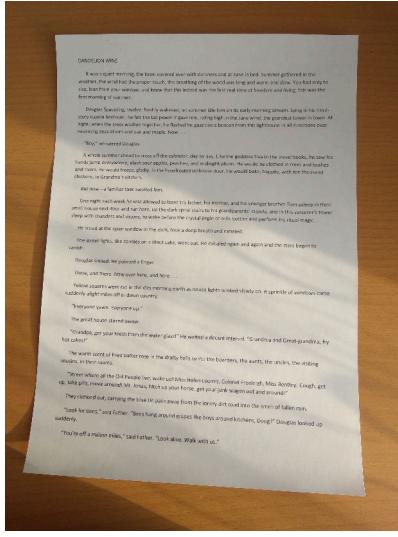
Теперь нам известно местоположение углов документа на изображении, а также параметры прямоугольника, в который необходимо его преобразовать. То есть у нас есть соответствие четырех точек на начальном изображении в четыре точки выходного изображения. По этому соответствуанию легко можно определить преобразование, которое бы перевело начальное изображение в выходное. В данной работе использовалась готовая реализация для нахождения требуемого перспективного преобразования и его применения из библиотеки OpenCV[16] (функции *getPerspectiveTransform* и *warpPerspective* соответственно).

Пример итогового результата на показан на рис. 2.7.

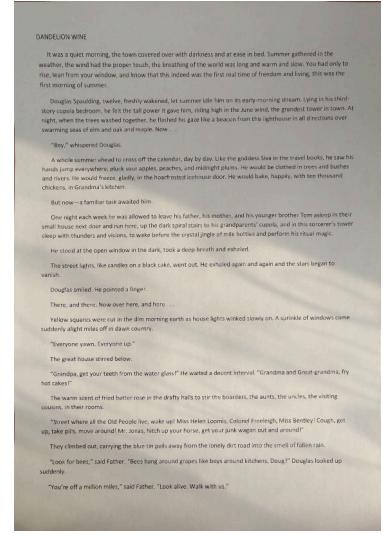
2.4. Примеры работы и предложения по улучшению

Успешные и неудачные примеры работы разработанного алгоритма на различных входных данных показаны в прилож. А. Прокомментируем полученные результаты.

На рис. A.1 и рис. A.2 показаны примеры в случае отсутствия заметных искажений и с явно выраженным перспективными искажениями соответственно. Алгоритм также справляется со своей задачей при наличии сложных структурных элементов на документе, включая таблицы (см.



(a) начальное изображение



(b) исправленное изображение

Рисунок 2.7 – Пример работы алгоритма по исправлению перспективных искажений

рис. A.3) и печати (см. рис. A.4). Предложенный метод устойчив к присутствию других видов искажений на изображении, таких как яркие тени (см. рис. A.5), блики (см. рис. A.10), и изогнутость документа (см. рис. A.9). Алгоритм исправляет даже сильно выраженные перспективные искажения, такие как на рис. A.7, достаточно точно при этом определяя верные пропорции документа. Предложенный способ определения углов документа как пересечения прямых, содержащих стороны документа, корректно работает даже в случае явного отсутствия этих углов на изображении, как в примере на рис. A.8.

Частая ситуация наличия файла, в котором лежит документ (см. рис. A.10) побудил к усовершенствованию алгоритма. Часто на таких изображениях после преобразования Хафа детектируются не одна прямая, соответствующая боковой стороне, а сразу две или три прямых (за счет присутствия белой полосы от файла, примыкающей к стороне документа). В таких случаях точек пересечения вертикальных прямых с горизонтальными больше четырех, поэтому появляется неоднозначность в определении местоположения углов документа. Однако за счет того, что некоторые прямые, соответствующие белой полосе от файла и прилегающей стороне

документа, расположены достаточно близко друг к другу, точки пересечения прямых отлично визуально разбиваются на 4 независимых группы вокруг углов документа. Поэтому в таких условиях было предложено в качестве углов документа выбирать центры кластеров при кластеризации точек пересечения прямых на 4 группы методом *k-means*. В таком случае обычно местоположение для двух углов документа удается установить абсолютно точно, а для двух других оценка получается немного смещенной, но близкой к действительности.

Отдельно акцентируем внимание на характерных случаях, в которых предложенный алгоритм работает неудачно. Часто это происходит из-за недостаточного внимания к ограничениям на входные данные, описанным в разд. 2.2. При недостаточно контрастном фоне не получается найти границы документа, как в примере рис. A.11. При неоднородности внешнего фона могут обнаружиться ложноположительные прямые, как на рис. A.12. Из-за наличия геометрических искажений, вызванных изгибами бумаги, изображение документа уже не может быть аппроксимировано четырехугольником, что также может быть причиной неудачи (см. рис. A.14). Наконец, возможны случаи, когда документ имеет сильно вытянутую прямоугольную форму, как на рис. A.13, что может явиться причиной того, что прямые, соответствующие коротким сторонам, не распознаются. Это наводит на мысль, что требуется доработка этапа поиска прямых как локальных максимумов в пространстве Хафа. Также возможны случаи, когда на документе присутствуют картинки с явно выраженными прямыми контрастными краями (см. рис. A.15). В этих случаях границы картинок могут быть восприняты алгоритмом за границы документа. В таких сложных ситуациях алгоритм может работать некорректно, поэтому изображения с документами подобного вида следует обрабатывать отдельно.

Глава 3.

Нормализация неравномерной освещенности

3.1. Описание искажения

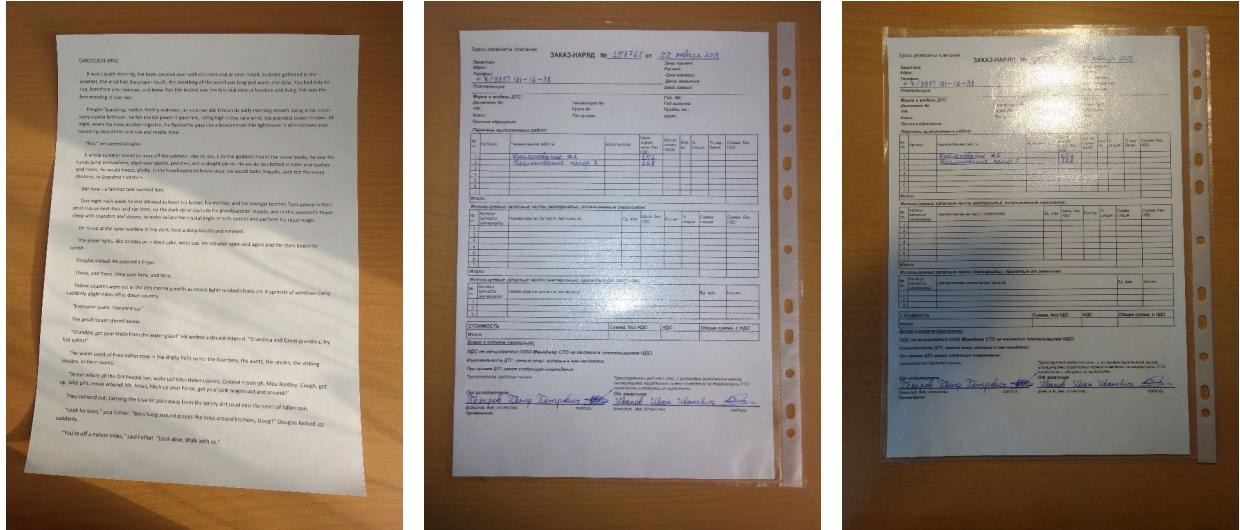
Неравномерная освещенность и блики на изображении – еще один из самых распространенных видов искажений. Неоднородное освещение чаще всего возникает из-за наличия сторонних объектов, которые препятствуют однаковому попаданию лучей от источника света на поверхность документа в процессе фотографирования. Если же документ в процессе съемки находился в пластиковом файле, то велика вероятность появления бликов на изображении, так как лучи света от источника отлично отражаются от поверхности файла и попадают в объектив фотокамеры. Последнее явление возникает особенно часто, если съемка производится со вспышкой.

Данные искажения могут существенно помешать успешному распознаванию информации с изображения документа. Примеры неравномерной освещенности и бликов изображены на рис. 3.1.

3.2. Идея метода и ограничения

Идея данного метода была предложена в работе [17].

Для входного изображения фоновым изображением будем называть такое изображение, которое могло быть получено в тех же условиях и тем же способом, что и входное изображение, если бы вместо оригинального документа был абсолютно такой же документ, но белый, без какого-либо содержания. Фактически фоновое изображение является оценкой



(a) изображение с наличием ярких теней (b) изображение с наличием бликов (c) изображение с яркими тенями и бликами

Рисунок 3.1 – Примеры изображений с неравномерным освещением

условий освещенности. Если для входного изображения известно фоновое изображение, то изображение с равномерным освещением можно вычислить, используя их разность.

В общем случае, у нас есть только входное изображение, требующее нормализации освещения. Главная идея предложенного в этой главе алгоритма заключается в оценке фонового изображения, и применения полученной оценки для нормализации освещения.

Из введенного определения фонового изображения возникают следующие ограничения на входные данные алгоритма:

1. Документ должен быть изготовлен на светлой бумаге. Данное требование выполнено для подавляющего большинства документов. Если это условие не выполняется, или документ содержит относительно большие по размеру сплошные темные участки, как например изображения, то алгоритм в процессе работы будет воспринимать их как тени и будет стремиться осветлить их, что может привести к неправильной передаче информации, содержащейся в документе. Пример такого входного изображения и соответствующего ему

результата работы предложенного метода будет рассмотрен в разд. 3.4.

2. Внешний фон на изображении, то есть область за пределами границы документа, должна отсутствовать, либо также быть светлого цвета. Поэтому рекомендуется использовать данный алгоритм после исправления перспективного искажения с помощью метода, предложенного в гл. 2.

3.3. Детали реализации

3.3.1. Представление данных

Для работы с освещенностью чаще всего бывает удобно работать в цветовой модели, содержащей в явном виде информацию о яркости или интенсивности. Поэтому, даже если входное изображение изначально представлено в цветовом пространстве RGB, сконвертируем его в цветовую модель HSL. Цвет пикселя в данной модели записывается с помощью трех координат: цветовой тон (*hue*), насыщенность (*saturation*) и интенсивность (*lightness*). При этом координата *hue* полностью отвечает за цвет, *saturation* характеризует чистоту цвета (минимальное значение этой координаты соответствует оттенку серого цвета, а максимальное – насыщенному цвету), а координата *lightness* описывает яркость или воспринимаемую интенсивность света. Такое представление цвета зачастую представляется удобным, и поэтому популярно среди дизайнеров и в системах машинного зрения. Предлагаемый в данной главе метод в процессе работы использует и преобразует только канал изображения, отвечающий за интенсивность света, оставляя два других канала неизменными.

3.3.2. Оценка фонового изображения

Будем рассматривать канал интенсивности как отдельное изображение в градациях серого цвета. Оценка фонового изображения получается

путем последовательного применения двух базовых морфологических операций к этому одноканальному изображению. Приведем описания данных преобразований.

Первое преобразование, называемое дилатацией или морфологическим расширением (*dilatation*), заключается в свертке изображения с некоторым ядром, вычисляющим максимальное значение среди пикселей, которые попали в окрестность, заданную формой и размером ядра. На практике, изображение после такого преобразования становится светлее и слегка размытым, а темные детали ослабляются или вообще исчезают. Применительно к оценке фонового изображения, дилатация позволяет пренебречь влиянием темного текста, содержащегося на документе.

Вторая операция, называемая эрозией или морфологическим сжатием (*erosion*), является двойственной к первой. Действие этого преобразования подобно дилатации, разница лишь в том, что во время свертки происходит поиск наименьшего значения, а не максимума. После действия эрозии можно наблюдать сужение светлых участков на изображении. Иллюстрация работы данных преобразований с использованием ядра, имеющего форму квадрата, представлено на рис. 3.2.

Для нахождения оценки фонового изображения последовательно применяются две данные операции с одинаковым квадратным ядром, причем размер этого ядра выбирается таким, чтобы соответствовать порядку характерного размера символа текста на изображении. Так как точный размер шрифта неизвестен, было решено использовать ядро размера 11×11 ¹. Такое допущение показало хороший результат во время тестирования алгоритма. Подробнее об этом будет ниже.

¹Примечание. Данная величина приводится для свертки с изображением, имеющим размеры как у входного. Заметим, что для уменьшения количества вычислений, как и в случае алгоритма исправления перспективных искажений, можно было предварительно уменьшить входное изображение, и тогда ядро понадобилось бы меньшего размера.

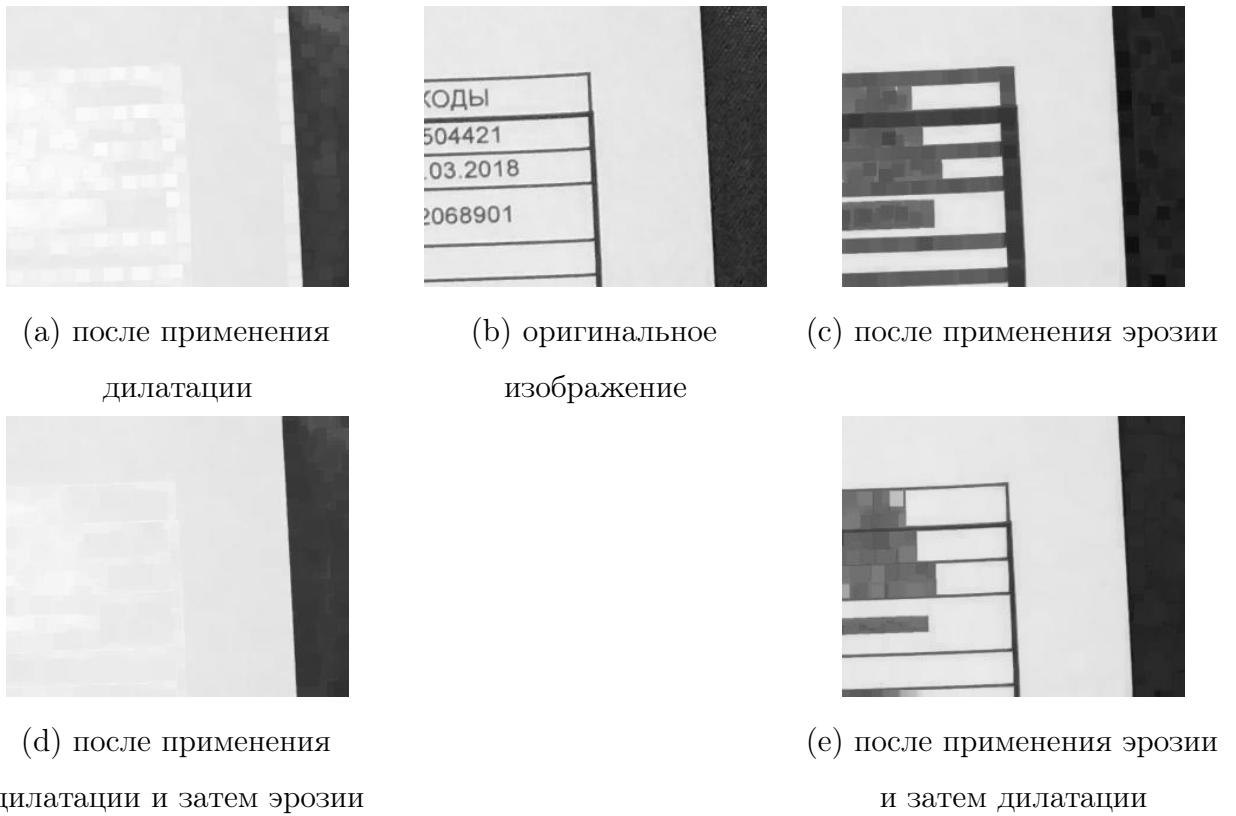


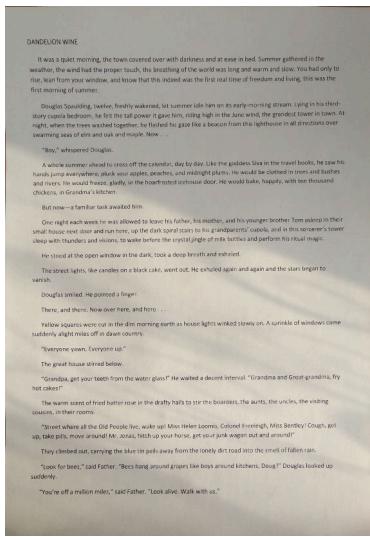
Рисунок 3.2 – Иллюстрация работы операторов морфологического расширения и сжатия

3.3.3. Вычисление исправленного изображения

После того как была найдена оценка на фоновое изображение, для нормализации освещенности канал яркости изначального изображения должен быть изменен по правилу:

$$g(x, y) = [f(x, y) - b(x, y)] + \overline{b(x, y)}, \quad (3.1)$$

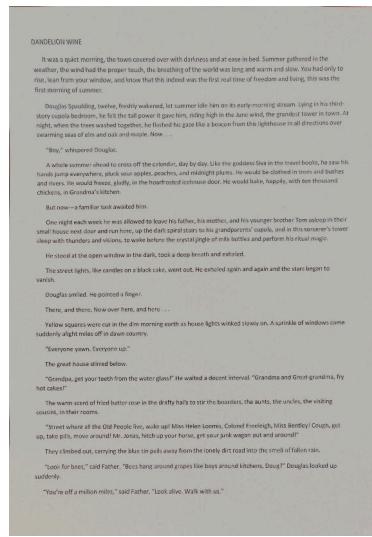
где $g(x, y)$ – это значение пикселя выходного изображения, соответствующего позиции (x, y) на изображении, $f(x, y)$ – значение пикселя входного изображения, $b(x, y)$ – значение пикселя фонового изображения, $\overline{b(x, y)}$ – среднее значение пикселей по всему фоновому изображению. Заметим, что просто разность интенсивностей входного изображения и оценки фонового искажения без прибавления среднего значения не является корректным решением. Несложно видеть, что среднее значение яркости входного и фонового изображений – это величины одного порядка, поэтому при таком



(а) начальное изображение



(б) оценка освещения



(c) результат коррекции

Рисунок 3.3 – Пример работы алгоритма коррекции освещения

подходе выходное изображение было бы практически черное.

Демонстрация работы предложенного алгоритма с визуализацией оценки фонового изображения приведена на рис. 3.3.

3.4. Примеры работы

Примеры работы описанного метода нормализации освещенности приведены в прилож. Б. На рис. Б.1, рис. Б.2, рис. Б.3, рис. Б.4, рис. Б.5 и рис. Б.6 показано, как алгоритм работает в условиях выполнения ограничений, описанных в разд. 3.2, в случаях наличия ярких теней на входном изображении. Можно заметить, что на исправленных изображениях освещение документа стало равномерным, при этом содержание документа не сильно изменилось. Особого внимания заслуживают примеры рис. Б.2, рис. Б.3 и рис. Б.4, так как они содержат цветные элементы, такие как заполненные синей ручкой поля документа и печать, цвет которых при преобразованиях поменялся несущественно. То есть предложенный метод, в отличие от, например, бинаризации, способен сохранять цвета.

В случае наличия бликов на поверхности файла, в котором содержится

документ, как на рис. Б.7, алгоритм не смог устраниТЬ искажение. Более того, информация, которая была рядом с областью отблеска, стала восприниматься чуть хуже.

Наличие внешнего фона светлого цвета (см. рис. Б.3) не ухудшает работу описанного метода. Однако следует отметить, что в этом случае граница документа на итоговом изображении видна не так хорошо, как до исправления искажения. В случае, когда присутствует контрастный внешний фон, как на рис. Б.8 и рис. Б.9, на выходном изображении наблюдается затемненность изначально белого листа бумаги документа (за счет влияния среднего значения фонового изображения по формуле (3.1)), и иногда неестественный цвет внешнего фона. Это в очередной раз подчеркивает целесообразность использования описанного метода нормализации освещенности только после исправления перспективных искажений и удаления внешнего по отношению к документу фона.

Наконец, если документ содержит темные рисунки, как в примере на рис. Б.10, то после нормализации освещения эти рисунки становятся заметно светлее, что может негативно сказаться на чтении или распознавании данных элементов.

Глава 4.

Измерение качества работы преобразований

4.1. Обоснование необходимости метрики

До сих пор при оценке качества методов, предложенных в гл. 2 и гл. 3, мы давали оценку их работе, основываясь только лишь на визуальном восприятии документа на фотографии до обработки изображения и после. Данный подход обладает двумя явными недостатками. Во-первых, не всегда бывает очевиден вывод о качестве результата работы алгоритма, особенно в случае наличия существенных искажений, сильно затрудняющих понимание информации на документе. В таких случаях, восприятие качества работы проделанных преобразований может быть очень индивидуальным. Во-вторых, хотелось бы иметь не только качественную оценку, чтобы делать вывод об успешности работы алгоритмов, но и количественную. Численная оценка особенно важна при сравнении разных методов исправления искажений, чтобы судить о том, какой метод работает в той или иной ситуации лучше, и насколько.

В общем случае документов произвольного формата и содержания сравнение способности изображений передавать без изменений информацию, которая присутствует на документе, является сложной задачей, предполагающей сравнение изображений друг с другом. В таком случае придумать метрику для сравнения качества работы различных алгоритмов исправления искажений затруднительно. Поэтому ограничимся случаем, когда документ не включает в себя каких-либо сложных структурных

элементов, таких как рисунки, графики, формулы, таблицы, а содержит лишь простой печатный текст, причем расположенный в одной колонке.

4.2. Описание метрики

В случае простых текстовых документов можно построить метрику, основываясь на том, насколько точно изображение документа передает его текстовое содержание. Это можно сделать с помощью оптического распознавания текста. В нашей работе в качестве системы распознавания текста был сделан выбор в пользу сводного ПО Tesseract 3.0[18], так как оно имеет одно принципиальное отличие от многих других современных систем OCR. Другие системы зачастую используют двухэтапный подход к распознаванию текста: на первом этапе на изображении распознаются отдельные символы, а на втором этапе символы, в которых система не уверена после первого этапа, заменяются на символы, которые скорее всего соответствуют данному слову или контексту предложения. Программа Tesseract 3.0 выполняет только первый из перечисленных этапов, что делает эту систему особенно подходящей к нашему случаю, когда конечной задачей является оценка качества распознавания, а не угадывание содержания документа.

Итак, можно оценить насколько точно изображение передает текстовое содержание документа. Пусть t_o – это последовательность символов, отражающая то, что в действительности изображено на документе. Заметим, что t_o можно получить, используя содержание электронной версии документа, которая использовалась при печати, или с помощью того же распознавания с некоторого эталонного изображения, относительно точно передающего содержание (например, это может быть отсканированное изображение). Пусть t_r – последовательность символов, которая была распознана с изображения, которое мы анализируем, с помощью выбранной системы OCR. Тогда оценкой качества передачи текстовой информации,

содержащейся на данном изображении, можно определить величину q :

$$q = 1 - \frac{d(t_o, t_r)}{\max(|t_o|, |t_r|)}, \quad (4.1)$$

где $d(t_o, t_r)$ – это расстояние Левенштейна между последовательностями символов t_o и t_r , то есть наименьшее число операций вставки, удаления или замены символа, которые необходимы для получения из одной последовательности символов другой. За $|t|$ здесь обозначено число символов в тексте t . Подчеркнем, что данная метрика q может принимать значения от 0 до 1, причем 1 соответствует случаю, когда текст t_r был распознан верно (то есть полностью совпадает с текстом t_o), а 0 соответствует случаю, когда текст был распознан полностью неверно (например, в случае, когда t_r – пустая последовательность). Таким образом, данная величина q для рассматриваемого изображения может рассматриваться как точность передачи текстовой информации.

Если имеется несколько изображений одного и того же текстового документа, полученных с помощью некоторой последовательности преобразований, то об успешности преобразований можно судить, анализируя изменения соответствующих значений величины q .

4.3. Примеры измерения качества

Примеры применения введенной метрики для измерения качества работы алгоритмов приведены в прилож. В. На основании данной метрики можно оценить, насколько хорошо работают предложенные в этой работе методы.

Примеры на рис. B.1, рис. B.2, рис. B.3, рис. B.4 и рис. B.5 демонстрируют, что алгоритм коррекции перспективы дает значительный прирост к успешности распознавания текста на документах. Особенно резкое улучшение наблюдается на изображениях, где перспективные искажения существенны, и лист документа повернут таким образом, что

линии текста немного наклонены. Однако в отдельных случаях наблюдается уменьшение значения метрики после коррекции перспективы, как на рис. В.6. В этом конкретном случае после применения алгоритма для исправления перспективных искажений текст, оказавшийся в тени, не был распознан системой OCR Tesseract, хотя до обработки этот текст читался системой относительно успешно. По-видимому, это произошло из-за встроенного в эту систему метода предварительной бинаризации изображения.

Алгоритм нормализации освещенности также в большинстве случаев улучшает качество распознавания. Если документ изначально находился в хороших условиях освещения, как в примерах рис. В.1 и рис. В.2, нормализация освещения лишь немногого увеличивает значение метрики. В случаях явно выраженных ярких теней (см. рис. В.6), закрывающих собой текст документа, нормализация освещенности значительно способствует успешности распознавания. Лишь в редких случаях, как на рис. В.7, нормализация освещенности совсем немного уменьшает точность распознавания.

При работе с документами в файлах оба преобразования оказываются бесполезными. Пример на рис. В.8 демонстрирует, что после исправления искажений каждым из методов значение метрики существенно уменьшается. После коррекции перспективы это происходит за счет белой полосы справа от файла, которая устраняется не полностью, в результате чего система распознает на ее месте символы, которых на самом деле на документе нет. Алгоритм нормализации освещенности в свою очередь не справляется с бликом, ухудшая распознавание символов вокруг светового пятна.

Глава 5.

Итоги и результаты работы

По итогам работы были разработаны и реализованы в программном коде методы исправления перспективных искажений и нормализации освещенности. Тестирование предложенных алгоритмов на изображениях из собранного набора данных показали высокую эффективность при решении поставленной задачи коррекции искажений, что подтверждается как при визуальной проверке результатов их работы (см. прилож. Г), так и с использованием предложенной метрики (см. прилож. В). Эксперименты подтвердили улучшение качества распознавания текста на изображениях документов после автоматической обработки, произведенной разработанным приложением.

Применение данных методов при сканировании печатных документов с использованием фотокамер минимизирует негативное влияние наиболее часто встречающихся искажений, что позволяет создавать цифровые копии документов, которые более точно передают их содержание. В частности, предложенные алгоритмы рекомендованы к использованию на стадии предварительной обработки фотографий перед оптическим распознаванием информации с изображения документа.

Список использованных источников

- [1] TurboScan. <https://play.google.com/store/apps/details?id=com.piksoft.turboscan>. accessed June 2019.
- [2] ABBYY FineScanner AI. <https://www.finescanner.com>. accessed June 2019.
- [3] Document Scanner Pro – Free Scan to PDF & JPG. <https://play.google.com/store/apps/details?id=com.docscannerpro.pdfscan>. accessed June 2019.
- [4] Genius Scan – PDF Scanner. <https://play.google.com/store/apps/details?id=com.thegrizzlylabs.geniusscan.free>. accessed June 2019.
- [5] Simple Scan – Free PDF Scanner App. <https://play.google.com/store/apps/details?id=comsimplescan.scanner>. accessed June 2019.
- [6] Paul Clark and Majid Mirmehdi. Rectifying perspective views of text in 3d scenes using vanishing points. *Pattern Recognition*, 36(11):2673 – 2686, 2003.
- [7] Yusuke Takezawa, Makoto Hasegawa, and Salvatore Tabbone. Camera-captured document image perspective distortion correction using vanishing point detection based on radon transform. pages 3968–3974, 12 2016.

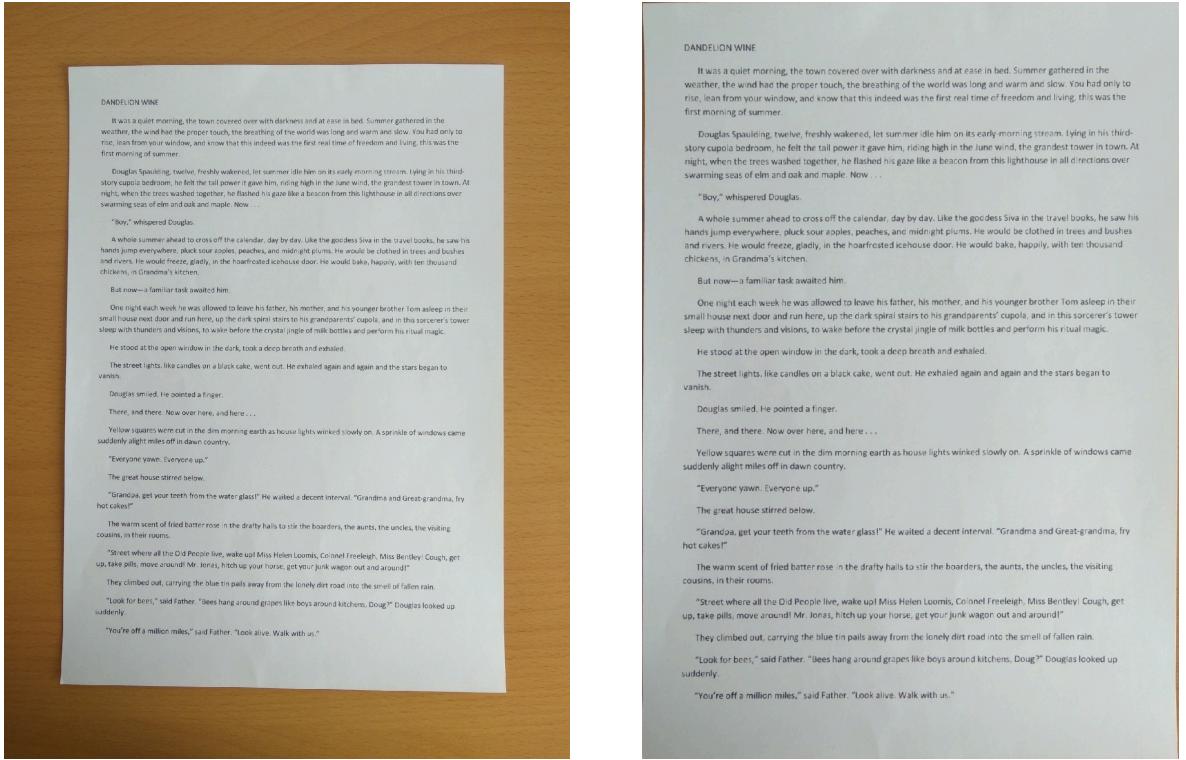
- [8] Shijian Lu, Ben M. Chen, and C.C. Ko. Perspective rectification of document images using fuzzy set and morphological operations. *Image and Vision Computing*, 23(5):541 – 553, 2005.
- [9] L. Miao and S. Peng. Perspective rectification of document images based on morphology. In *2006 International Conference on Computational Intelligence and Security*, volume 2, pages 1805–1808, Nov 2006.
- [10] J Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, June 1986.
- [11] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, Jan 1979.
- [12] Mei Fang, GuangXue Yue, and QingCang Yu. The study on an application of otsu method in canny operator. 01 2009.
- [13] B. D. Setiawan, A. N. Rusydi, and K. Pradityo. Lake edge detection using canny algorithm and otsu thresholding. In *2017 International Symposium on Geoinformatics (ISyG)*, pages 72–76, Nov 2017.
- [14] Paul C. Hough V. Method and means for recognizing complex patterns, December 1962.
- [15] Zhengyou Zhang and Li-Wei He. Whiteboard scanning and image enhancement. *Digital Signal Processing*, 17(2):414 – 432, 2007.
- [16] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- [17] Pandore: A library of image processing operators (Version 6.6). [Software]. Greyc Laboratory. <https://clouard.users.greyc.fr/Pandore>, 2018. [accessed August 2018].

- [18] Anthony Kay. Tesseract: An open-source optical character recognition engine. *Linux J.*, 2007(159):2–, July 2007.

Приложение А.

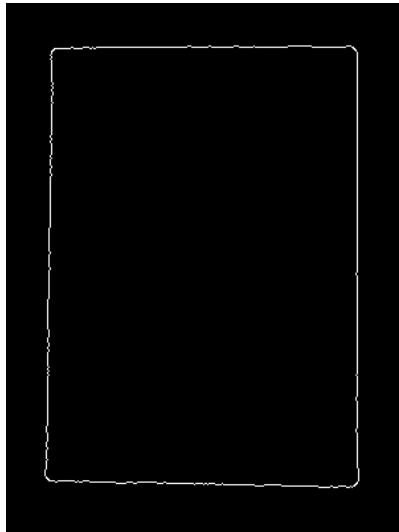
Примеры работы алгоритма исправления перспективных искажений

В данном приложении приводятся различные характерные примеры работы предложенного алгоритма по исправлению перспективных искажений. На изображениях можно встретить начальное изображение, результат работы алгоритма и некоторые промежуточные шаги, такие как результат применения оператора Кэнни, визуализация пространства Хафа с выделенными зеленым и красным цветами максимумами и выделенные теми же цветами соответствующие им прямые. Желтым цветом выделены точки пересечения найденных прямых. Голубым цветом выделены предполагаемые углы документа на изображении.

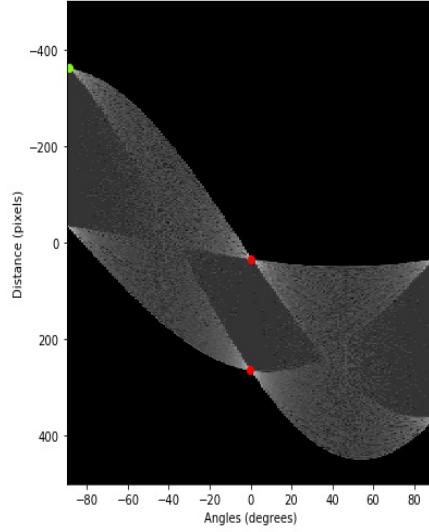


(a) начальное изображение

(b) исправленное изображение

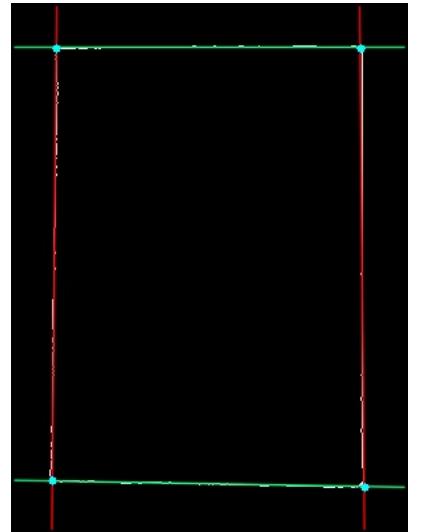


(c) результат
детектирования границ



(d) визуализация пространства

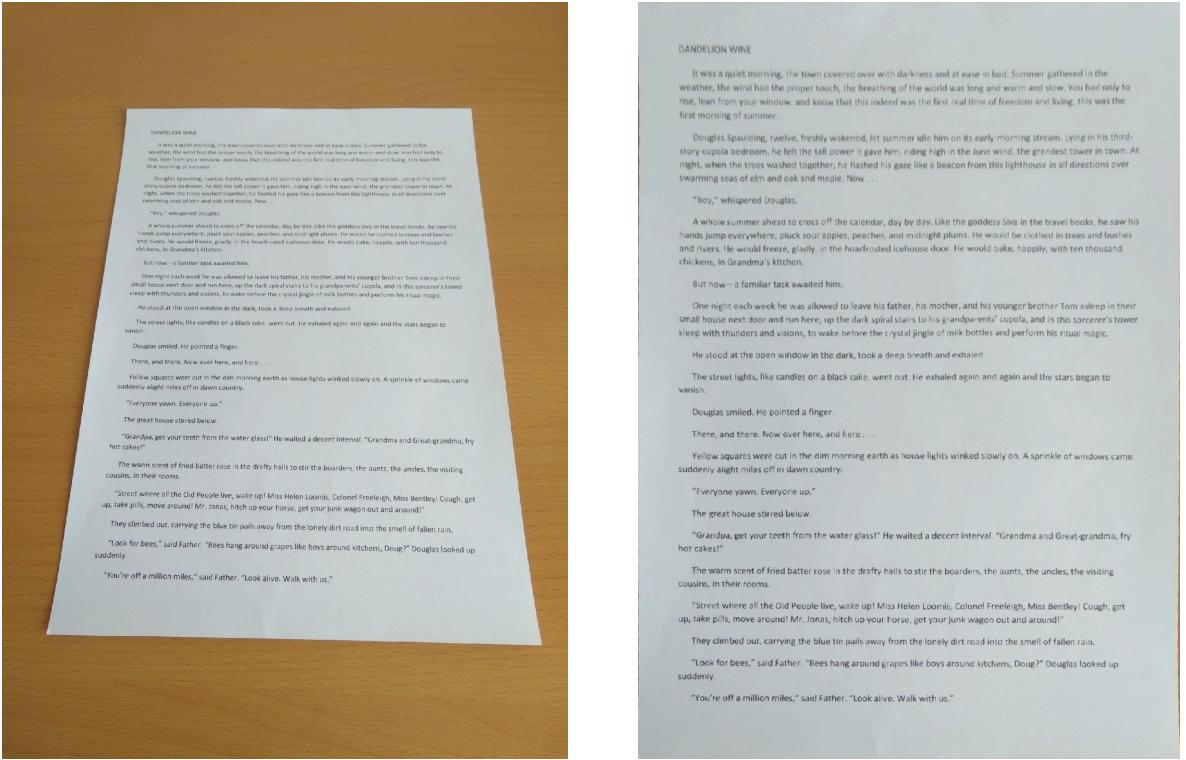
Хафа



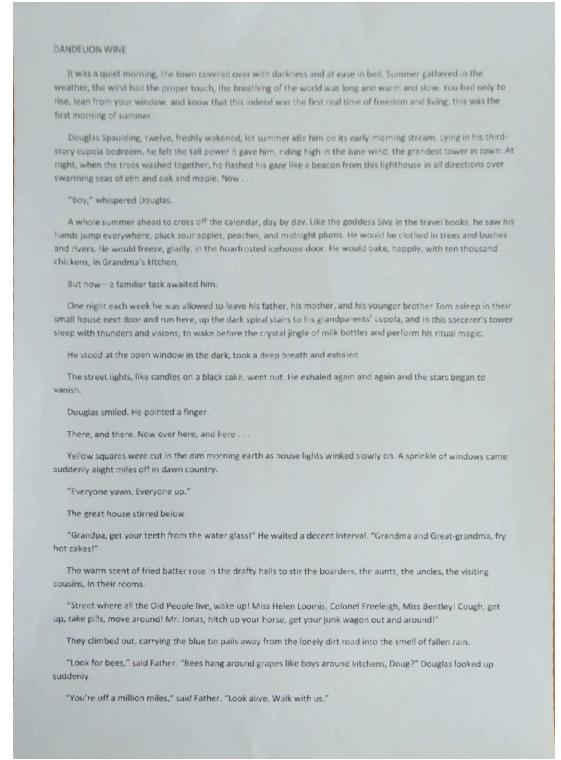
(e) найденные прямые и

углы

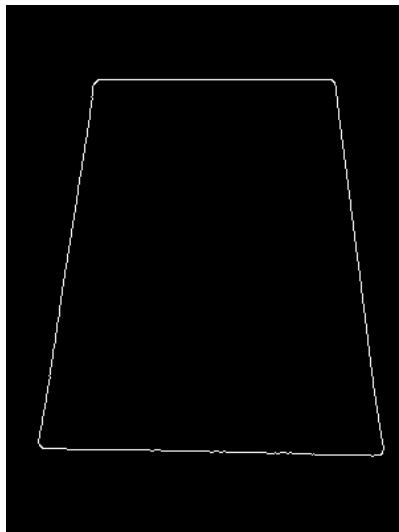
Рисунок А.1 – Пример работы алгоритма коррекции перспективы в случае незначительных искажений



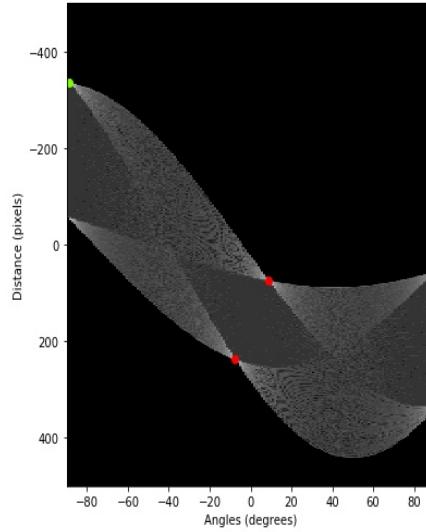
(a) начальное изображение



(b) исправленное изображение

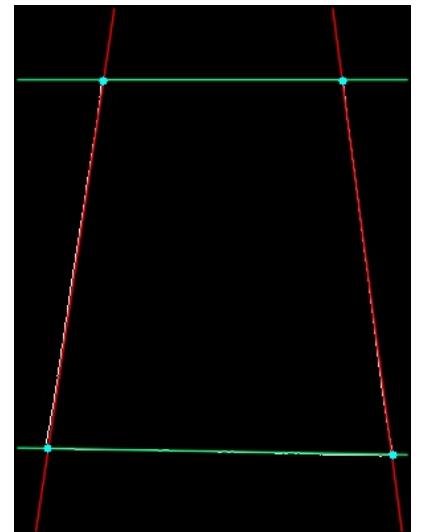


(c) результат
детектирования границ



(d) визуализация пространства

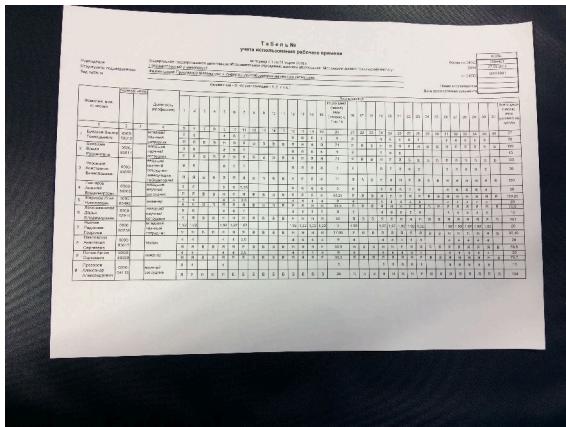
Хафа



(e) найденные прямые и

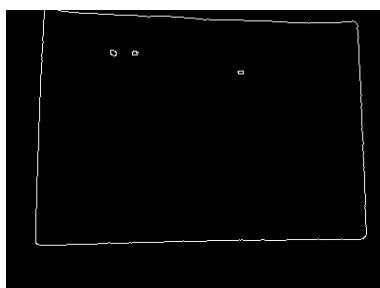
углы

Рисунок A.2 – Пример работы алгоритма коррекции перспективы в случае искаженного вертикального листа формата А4

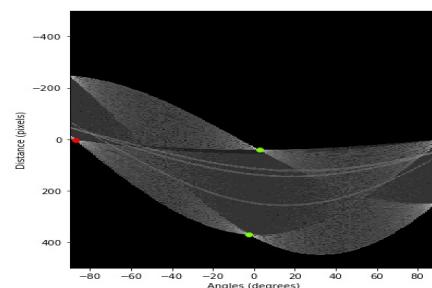


(а) начальное изображение

(b) исправленное изображение

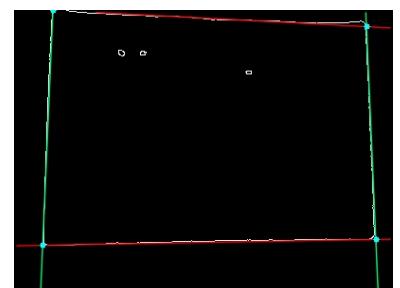


(c) результат
детектирования границ



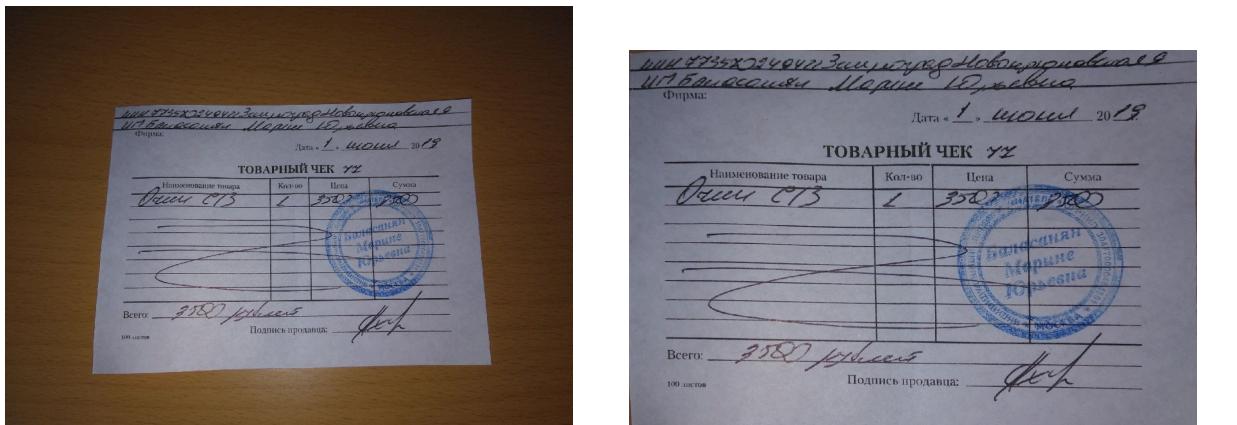
(d) визуализация пространства

Xada



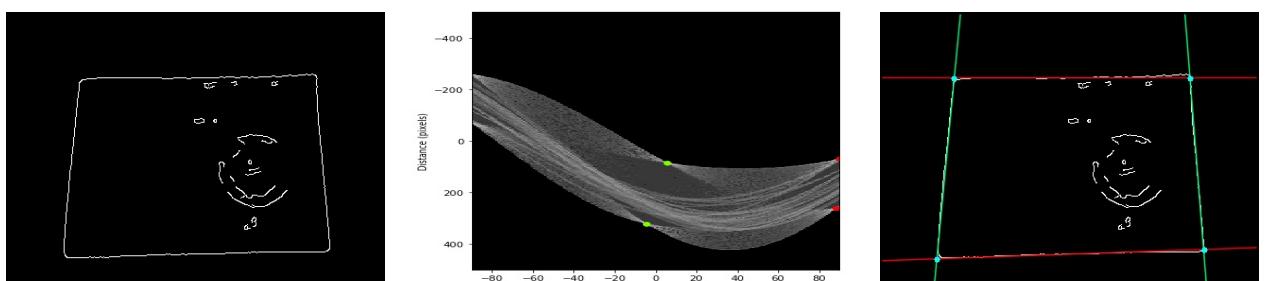
(e) найденные прямые и
углы

Рисунок А.3 – Пример работы алгоритма коррекции перспективы в случае наличия таблиц на документе



(a) начальное изображение

(b) исправленное изображение

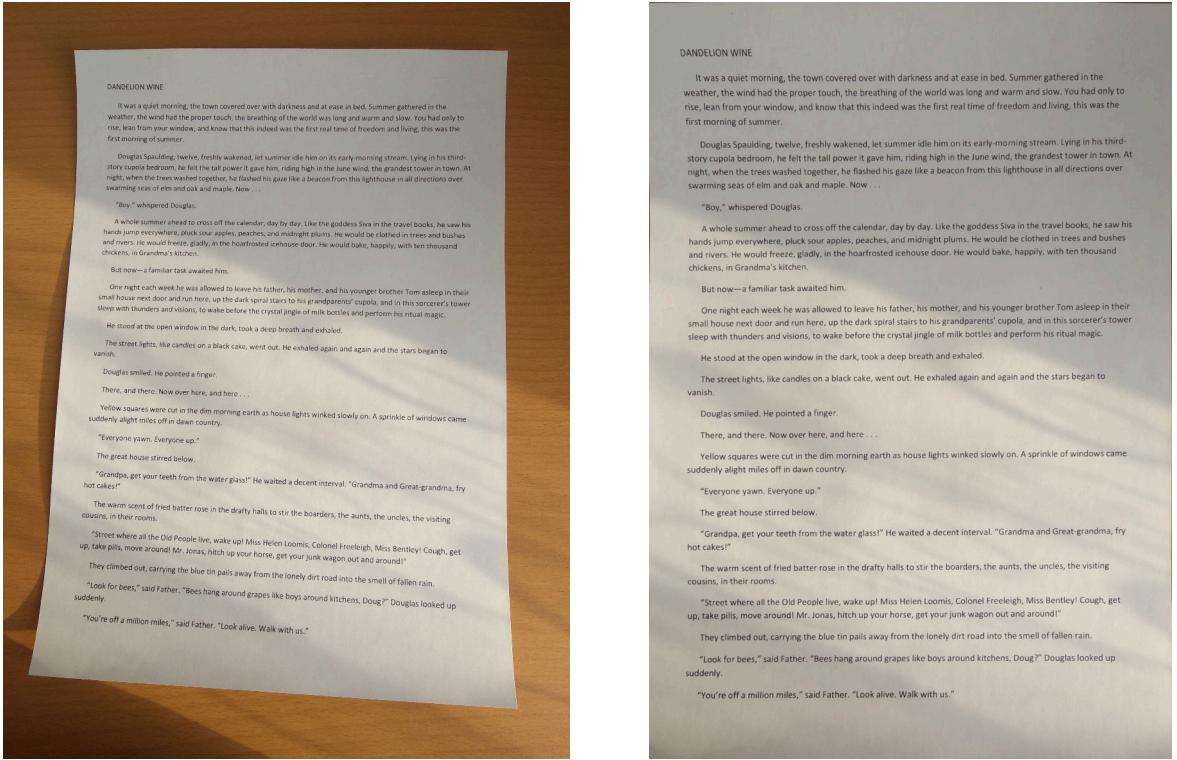


(c) результат
детектирования границ

(d) визуализация пространства
Хафа

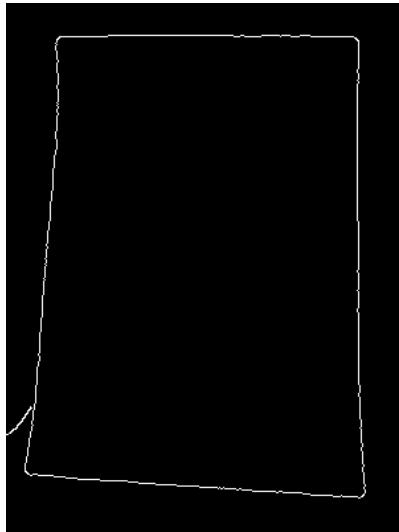
(e) найденные прямые и
углы

Рисунок А.4 – Пример работы алгоритма коррекции перспективы в случае документа с
печатью

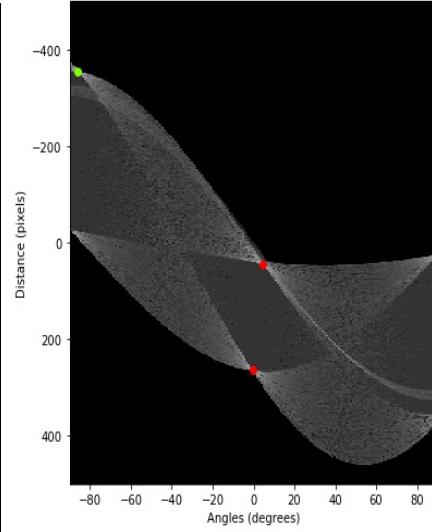


(a) начальное изображение

(b) исправленное изображение

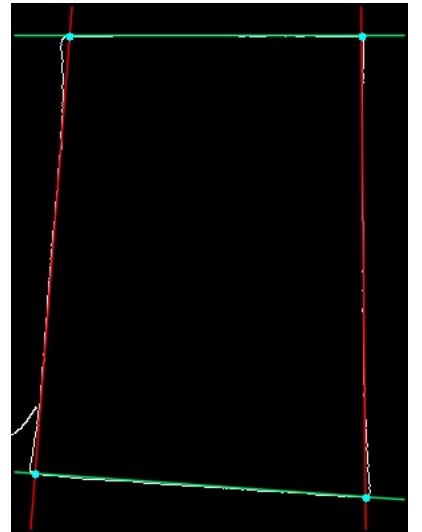


(c) результат
детектирования границ



(d) визуализация пространства

Хафа



(e) найденные прямые и

углы

Рисунок А.5 – Пример работы алгоритма коррекции перспективы в случае наличия неравномерного освещения



(a) начальное изображение

(b) исправленное изображение

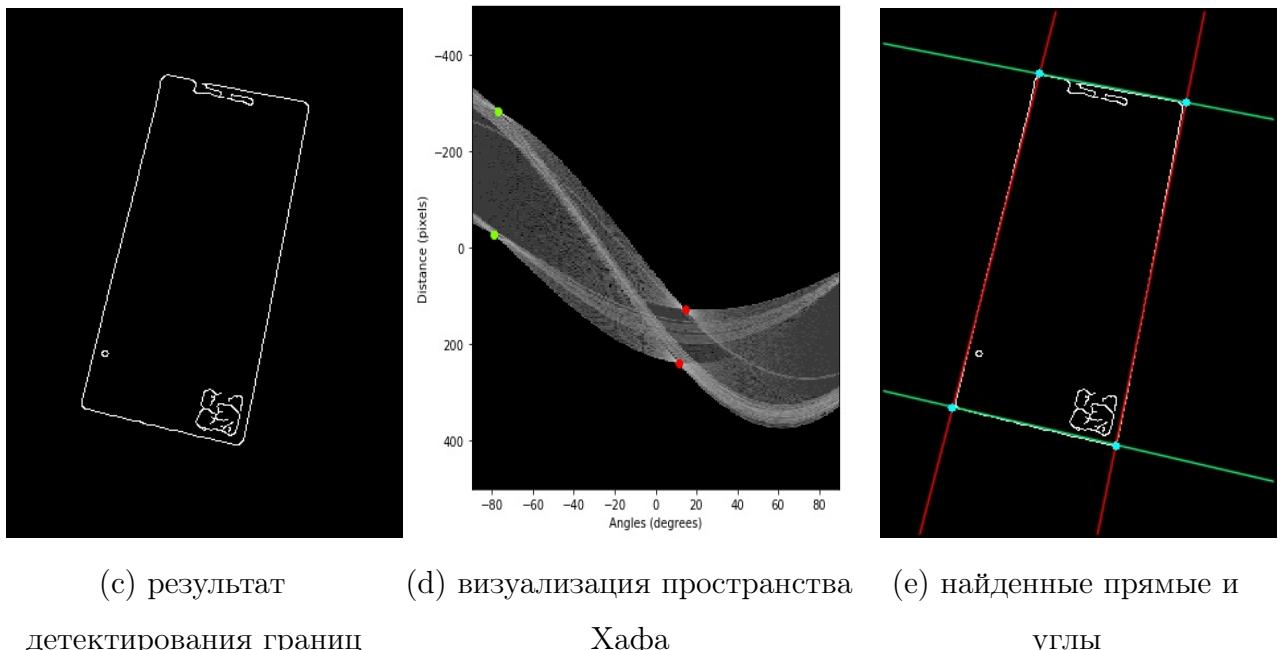
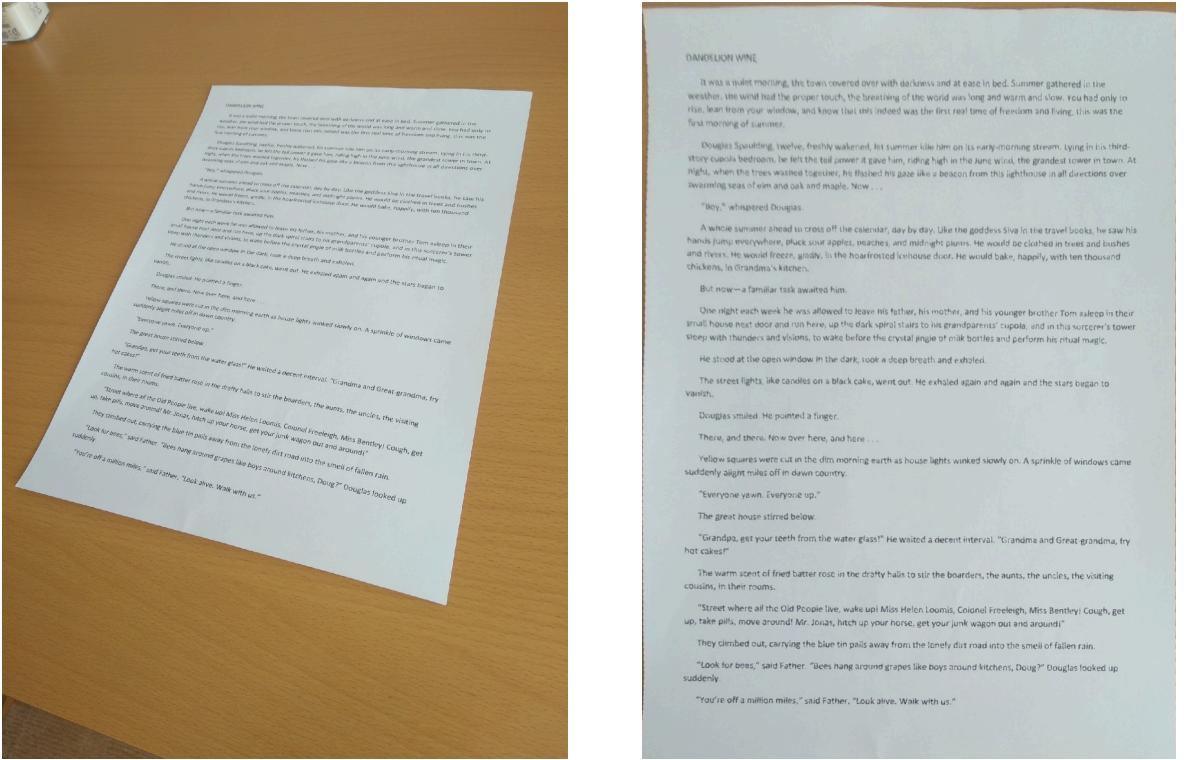
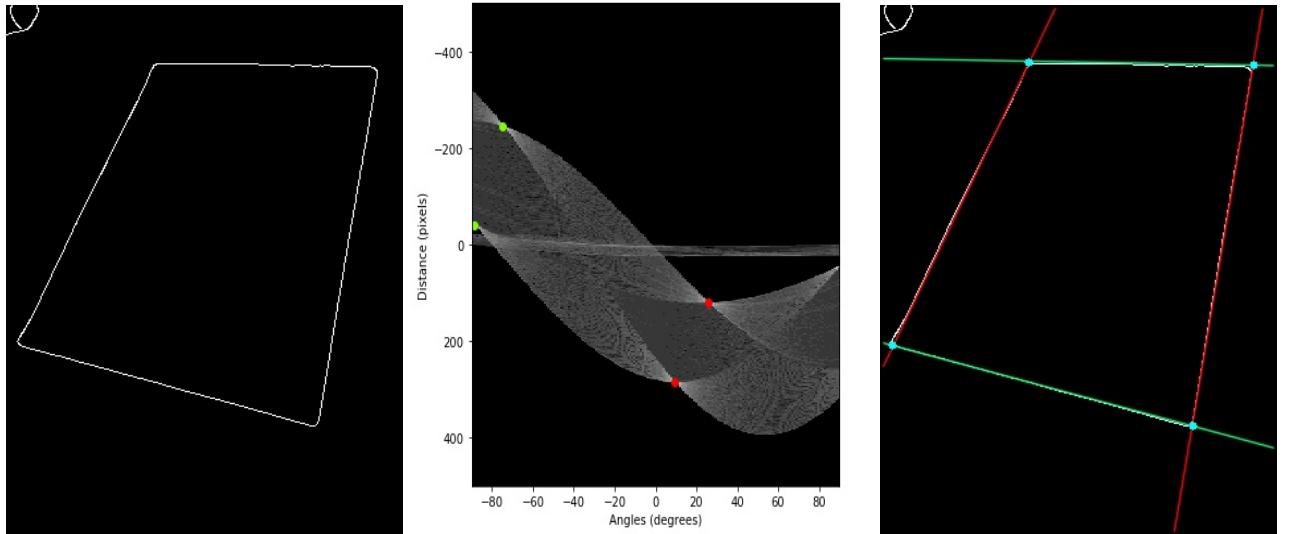


Рисунок А.6 – Пример работы алгоритма коррекции перспективы для сильно вытянутого по вертикали документа



(a) начальное изображение

(b) исправленное изображение



(c) результат

(d) визуализация пространства

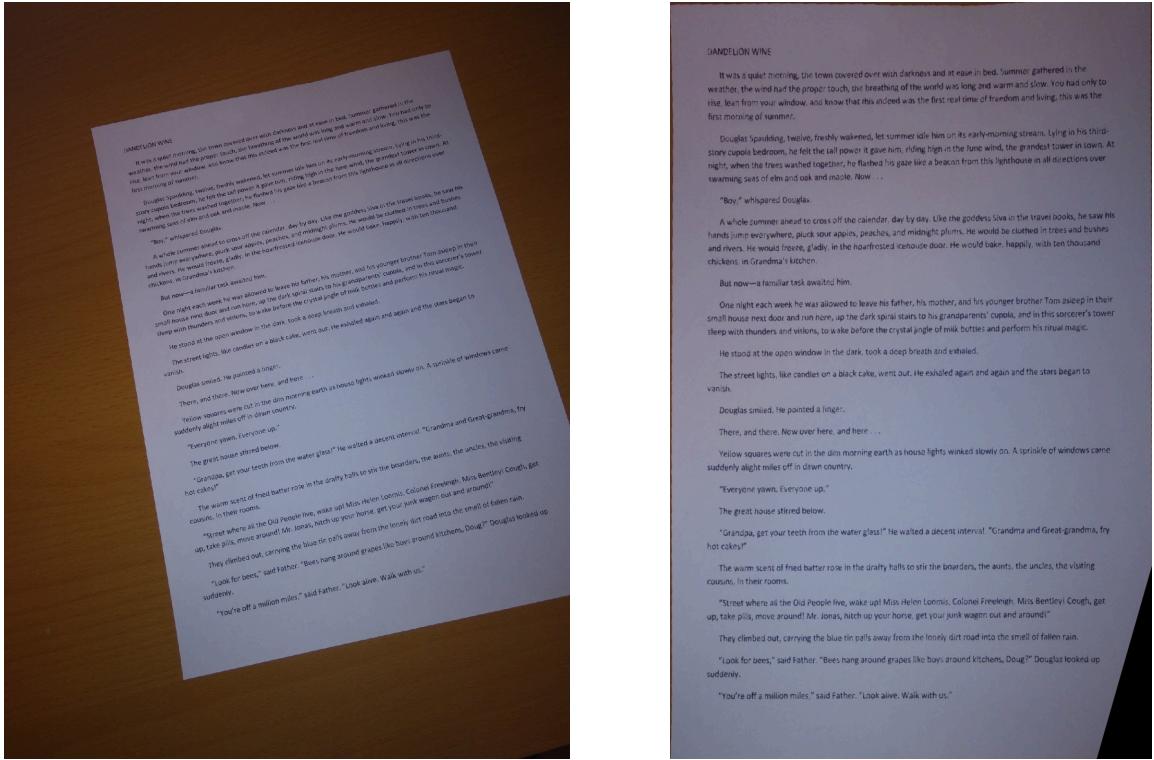
(e) найденные прямые и

детектирования границ

Хафа

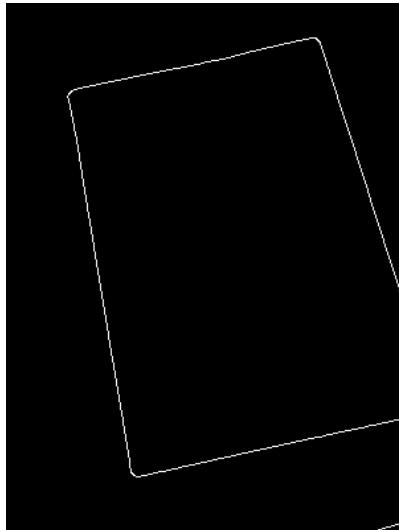
углы

Рисунок А.7 – Пример работы алгоритма коррекции перспективы в случае сильно выраженного искажения

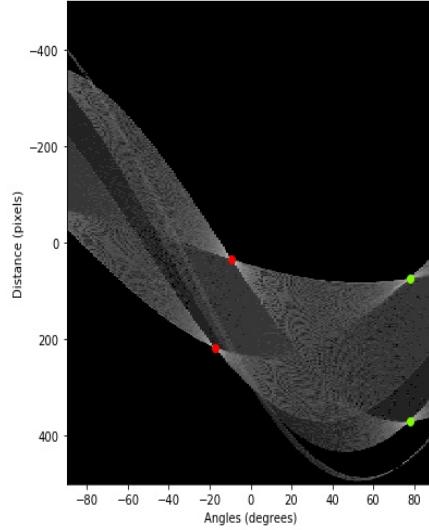


(a) начальное изображение

(b) исправленное изображение

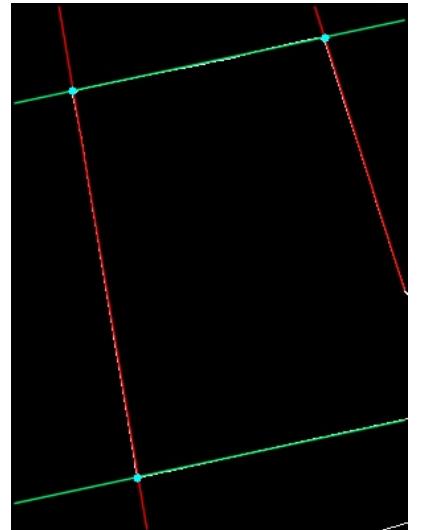


(c) результат
детектирования границ



(d) визуализация пространства

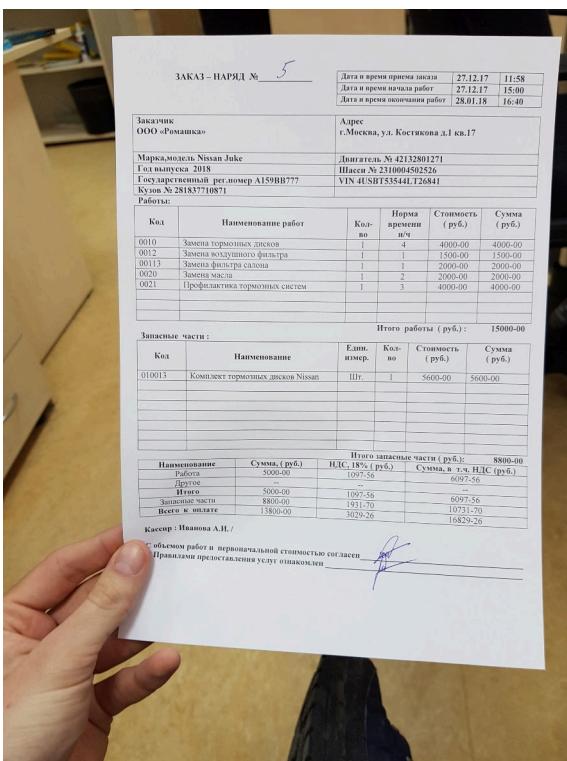
Хафа



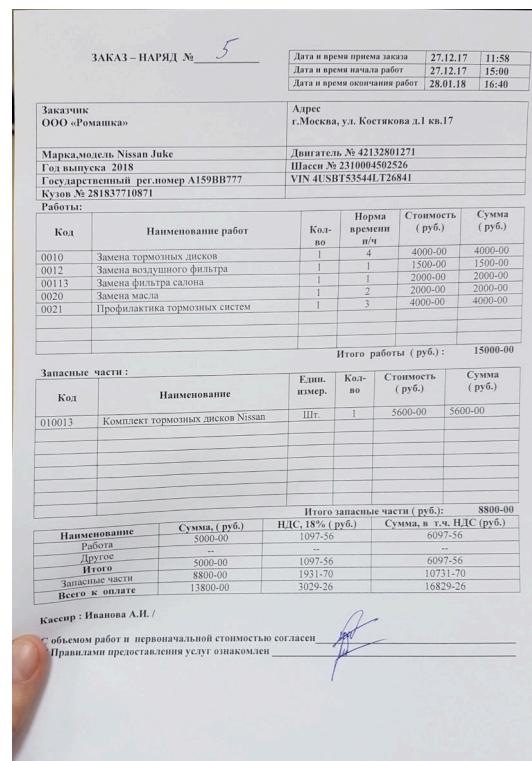
(e) найденные прямые и

углы

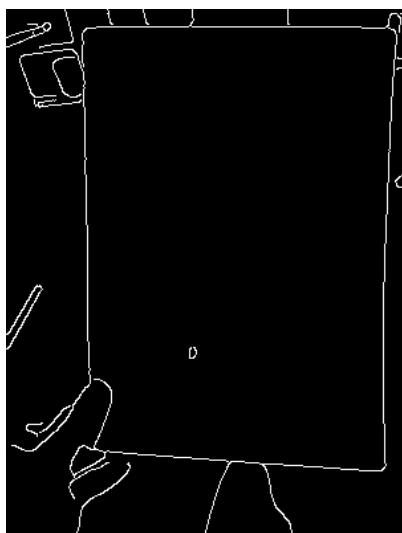
Рисунок A.8 – Пример работы алгоритма коррекции перспективы в случае отсутствующего на изображении угла документа



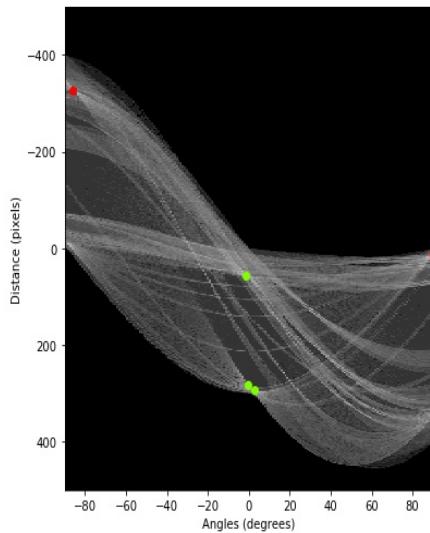
(a) начальное изображение



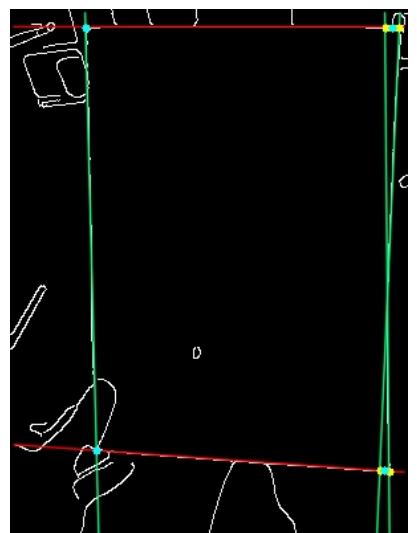
(b) исправленное изображение



(c) результат детектирования границ

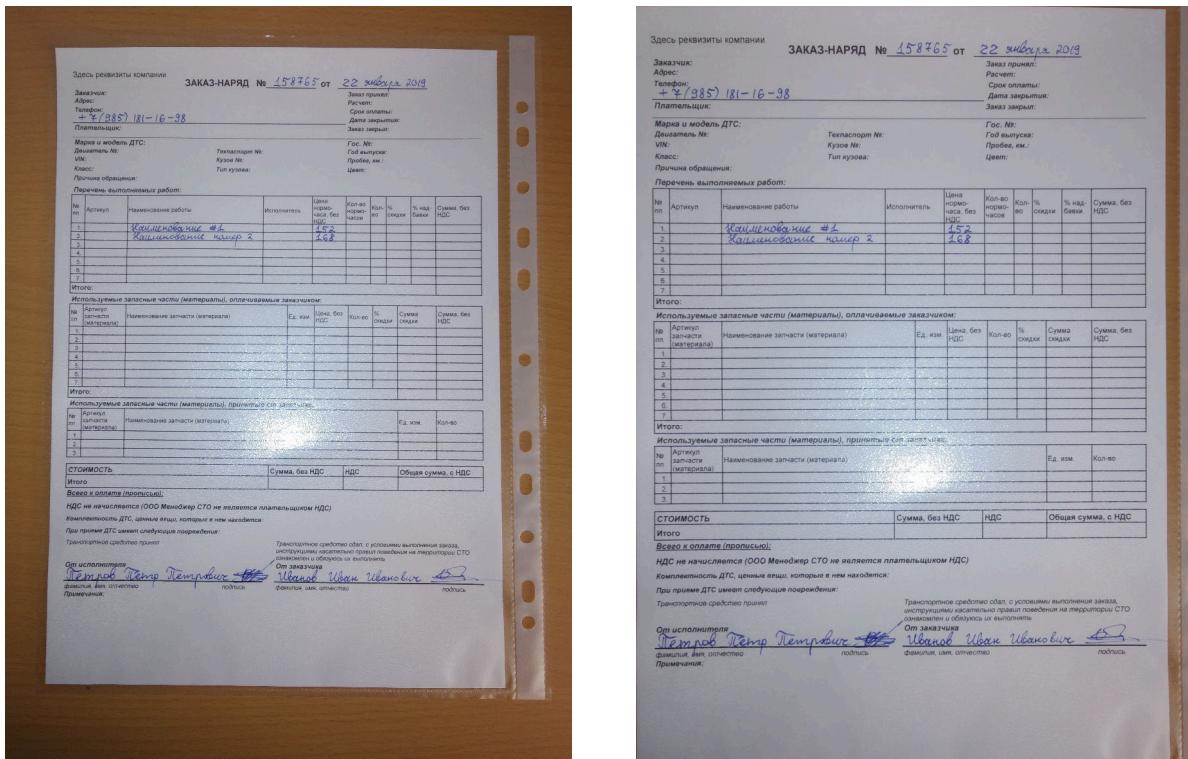


(d) визуализация пространства Хафа



(e) найденные прямые и углы

Рисунок A.9 – Пример работы алгоритма коррекции перспективы в случае неоднородного фона и наличия геометрических искажений, вызванных держанием документа на весу



(a) начальное изображение

(b) исправленное изображение

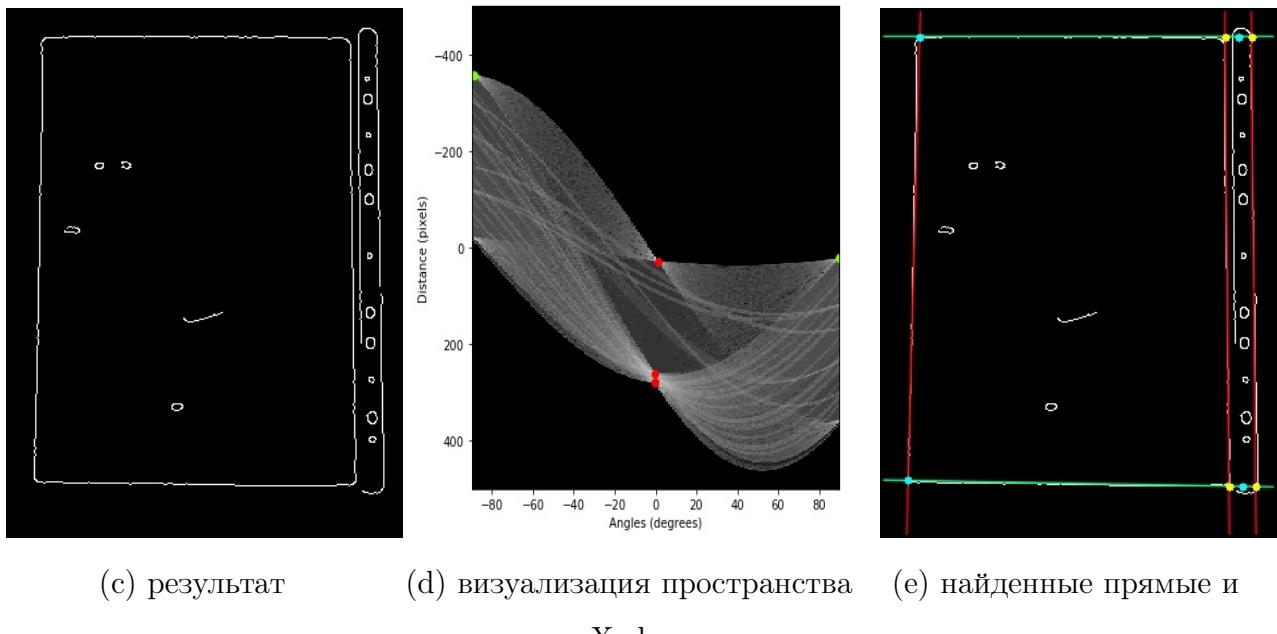
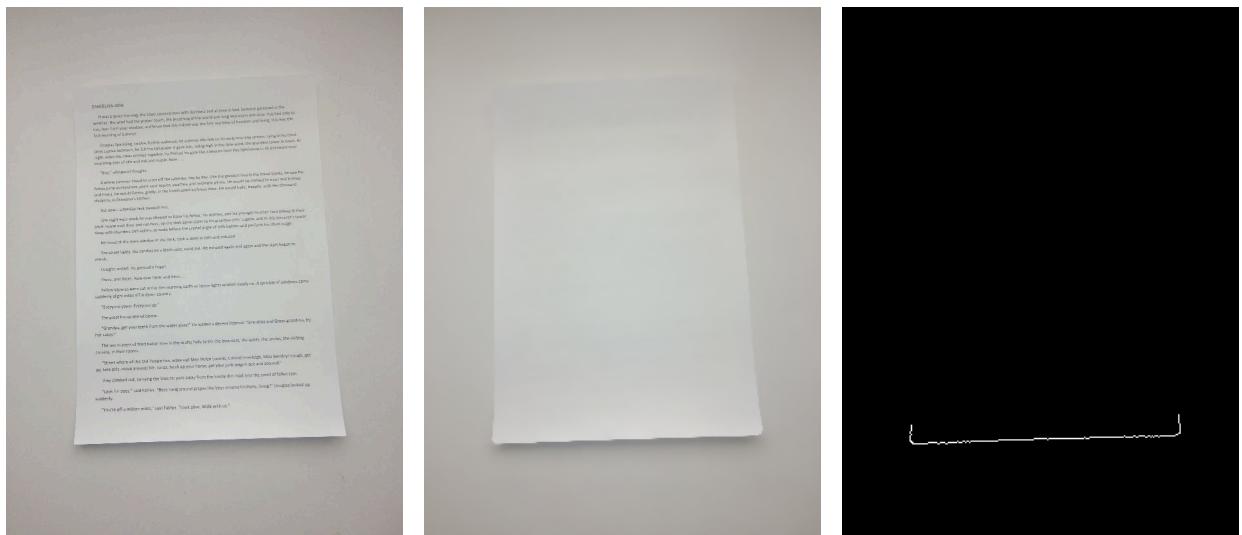
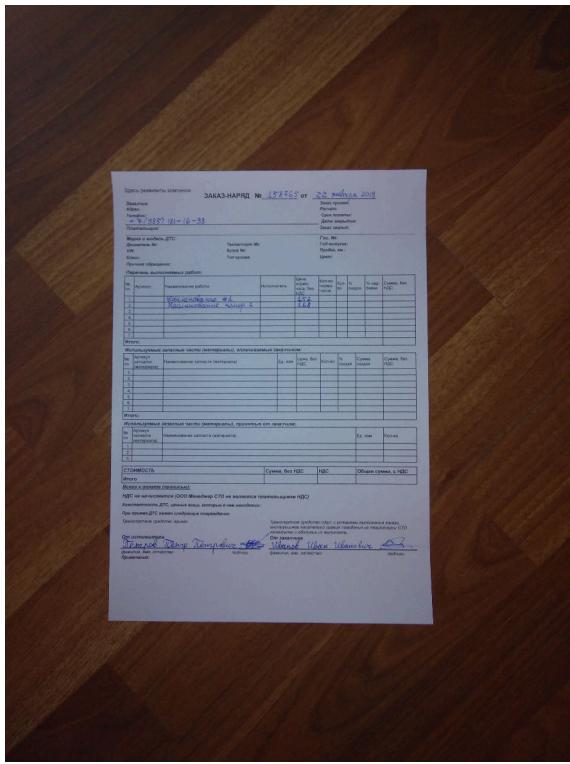


Рисунок А.10 – Пример работы алгоритма коррекции перспективы для документа в файле

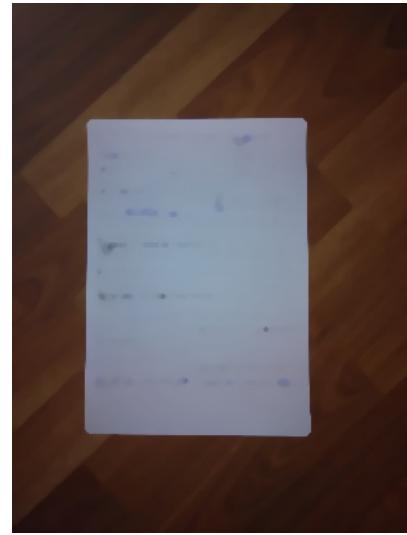


(a) начальное изображение (b) после применения
медианного фильтра (c) результат детектирования
границ

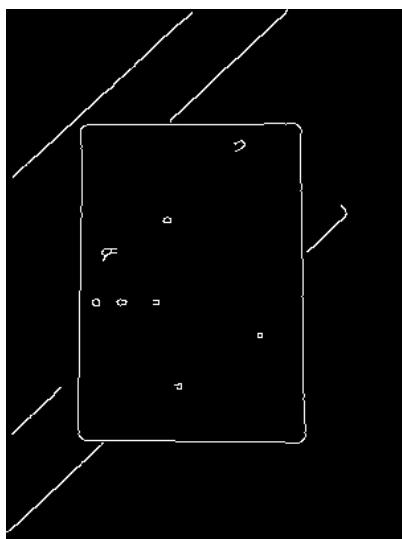
Рисунок А.11 – Неудачный пример работы алгоритма коррекции перспективы из-за неконтрастного фона



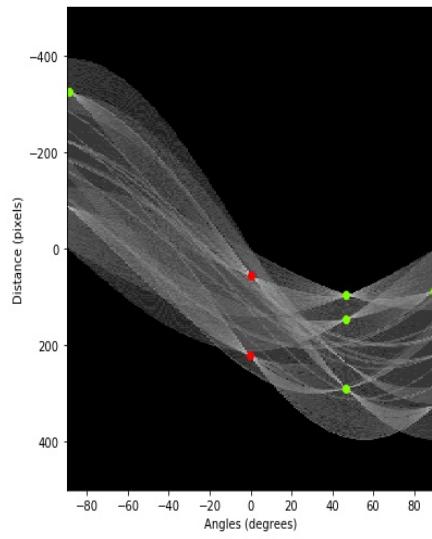
(a) начальное изображение



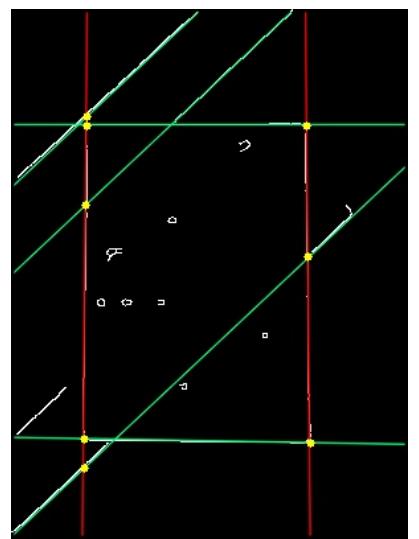
(b) результат применения медианного фильтра



(c) результат детектирования границ



(d) визуализация пространства Хафа



(e) найденные прямые и их точки пересечения

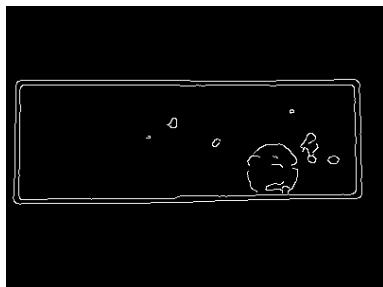
Рисунок А.12 – Неудачный пример работы алгоритма коррекции перспективы из-за наличия явно выраженных прямых в структуре фона



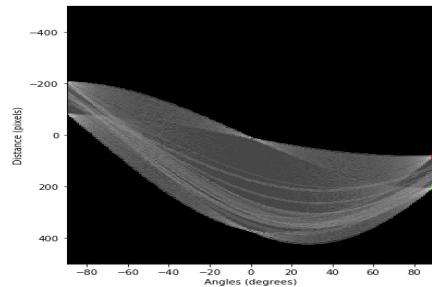
(a) начальное изображение



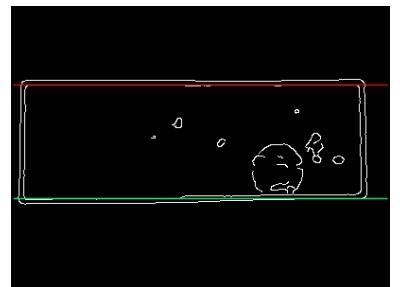
(b) результат применения медианного фильтра



(c) результат детектирования границ

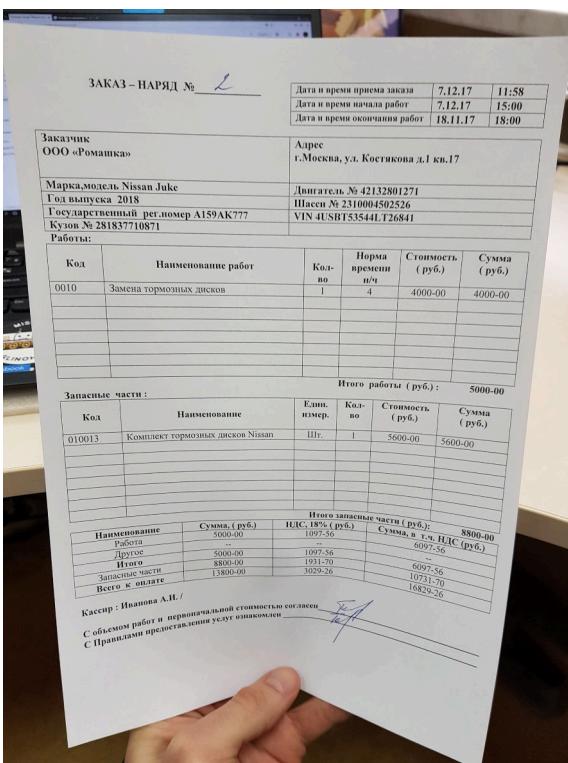


(d) визуализация пространства Хафа

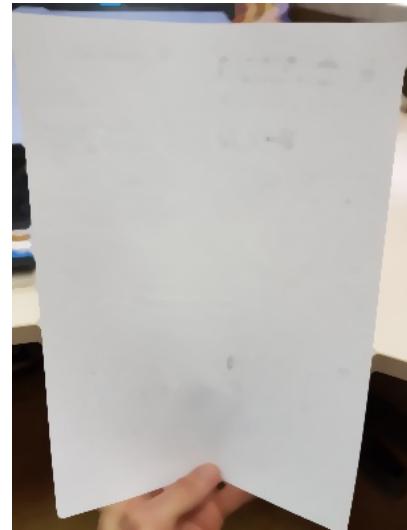


(e) найденные прямые

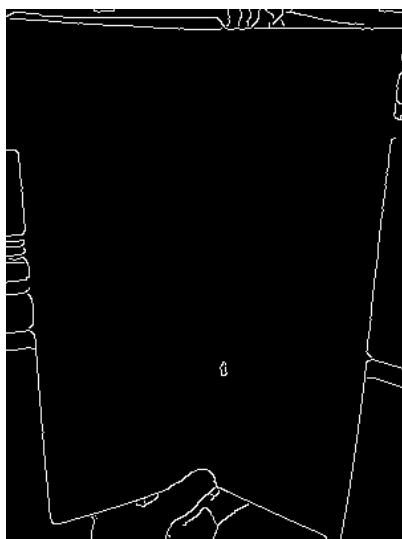
Рисунок А.13 – Неудачный пример работы алгоритма коррекции перспективы из-за сильно вытянутого по горизонтали документа



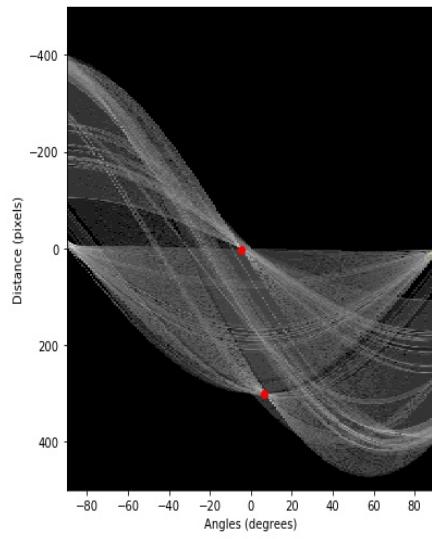
(a) начальное изображение



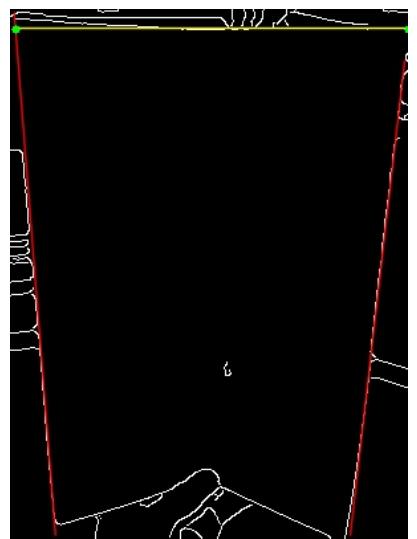
(b) результат применения медианного фильтра



(c) результат детектирования границ

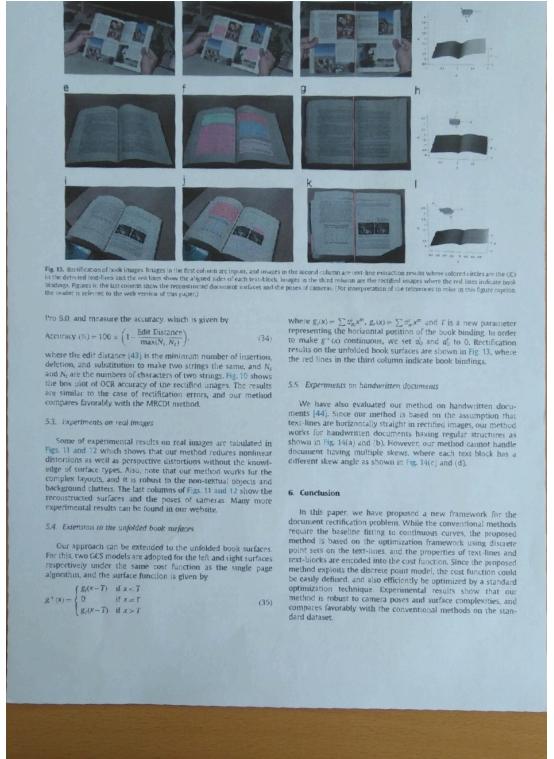
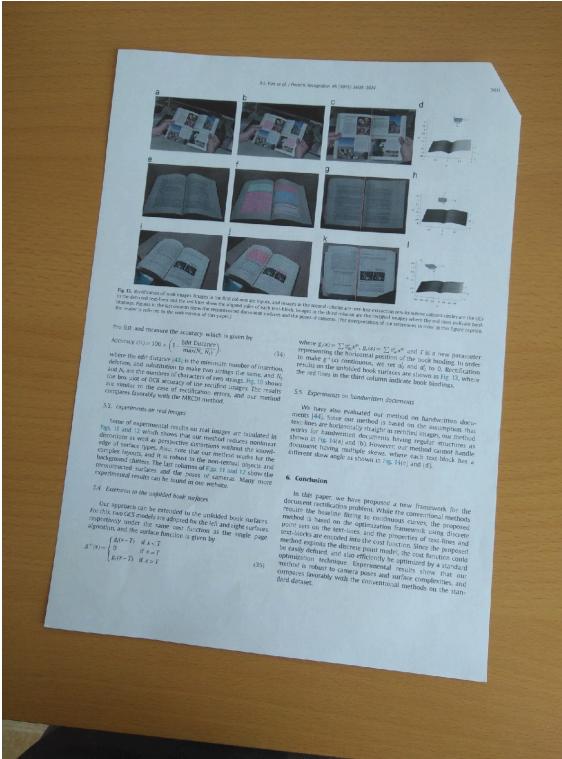


(d) визуализация пространства Хафа



(e) найденные прямые и их точки пересечения

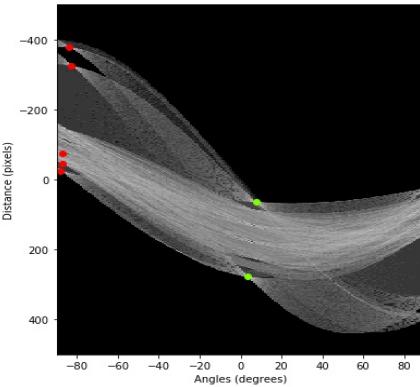
Рисунок А.14 – Неудачный пример работы алгоритма коррекции перспективы из-за сильно выраженного изгиба документа



(a) начальное изображение

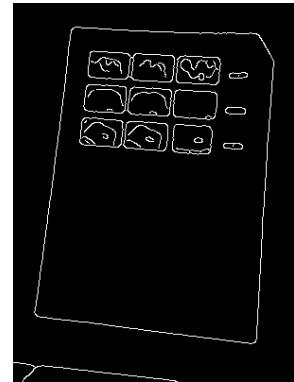


(c) результат применения медианного фильтра

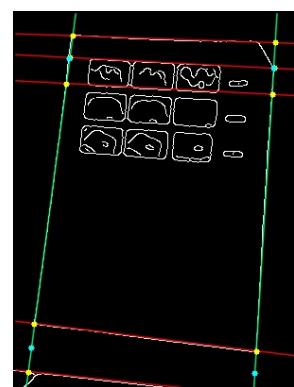


(e) визуализация пространства Хафа

(b) исправленное изображение



(d) результат детектирования границ



(f) найденные прямые и углы

Рисунок А.15 – Неудачный пример работы алгоритма коррекции перспективы из-за наличия картинок с ярко выраженной прямой границей

Приложение Б.

Примеры работы алгоритма нормализации освещенности

В данном приложении содержатся примеры работы предложенного алгоритма по исправлению освещенности. В примерах приводятся оригинальное изображение документа (до или после коррекции перспективного искажения), предполагаемое фоновое изображение, несущее в себе информацию об освещении, и результат исправления освещенности.

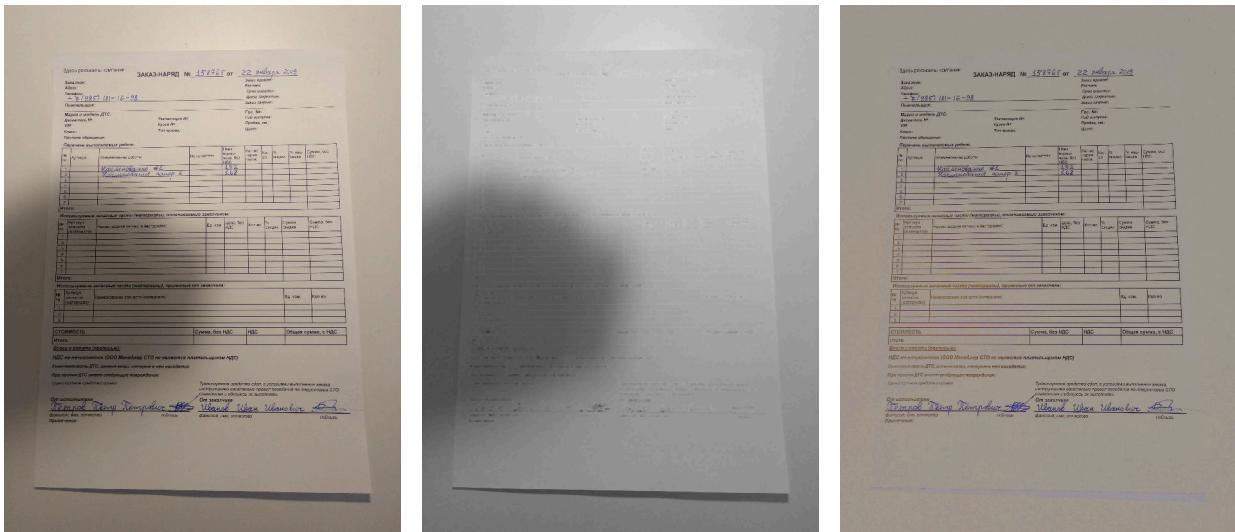


(a) начальное изображение (b) оценка освещения (c) результат коррекции

Рисунок Б.1 – Пример работы алгоритма коррекции освещения в случае наличия ярких теней



Рисунок Б.2 – Пример работы алгоритма коррекции освещения в случае наличия ярких теней и цветной информации на документе



(a) начальное изображение (b) оценка освещения (c) результат коррекции

(б) оценка освещения

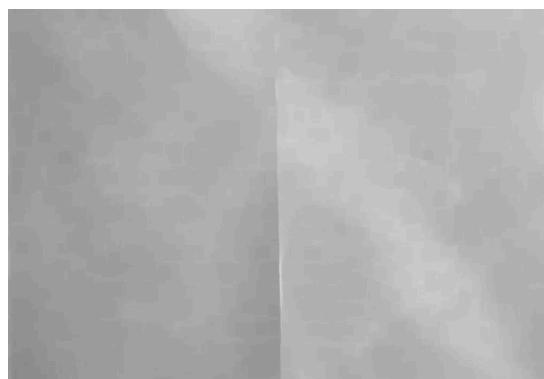
(c) результат коррекции

Рисунок Б.3 – Пример работы алгоритма коррекции освещения в случае наличия ярких теней, цветной информации на документе и светлого внешнего фона



(а) начальное изображение

(b) результат коррекции



(с) оценка освещения

Рисунок Б.4 – Пример работы алгоритма коррекции освещения в случае наличия ярких теней и цветных элементов

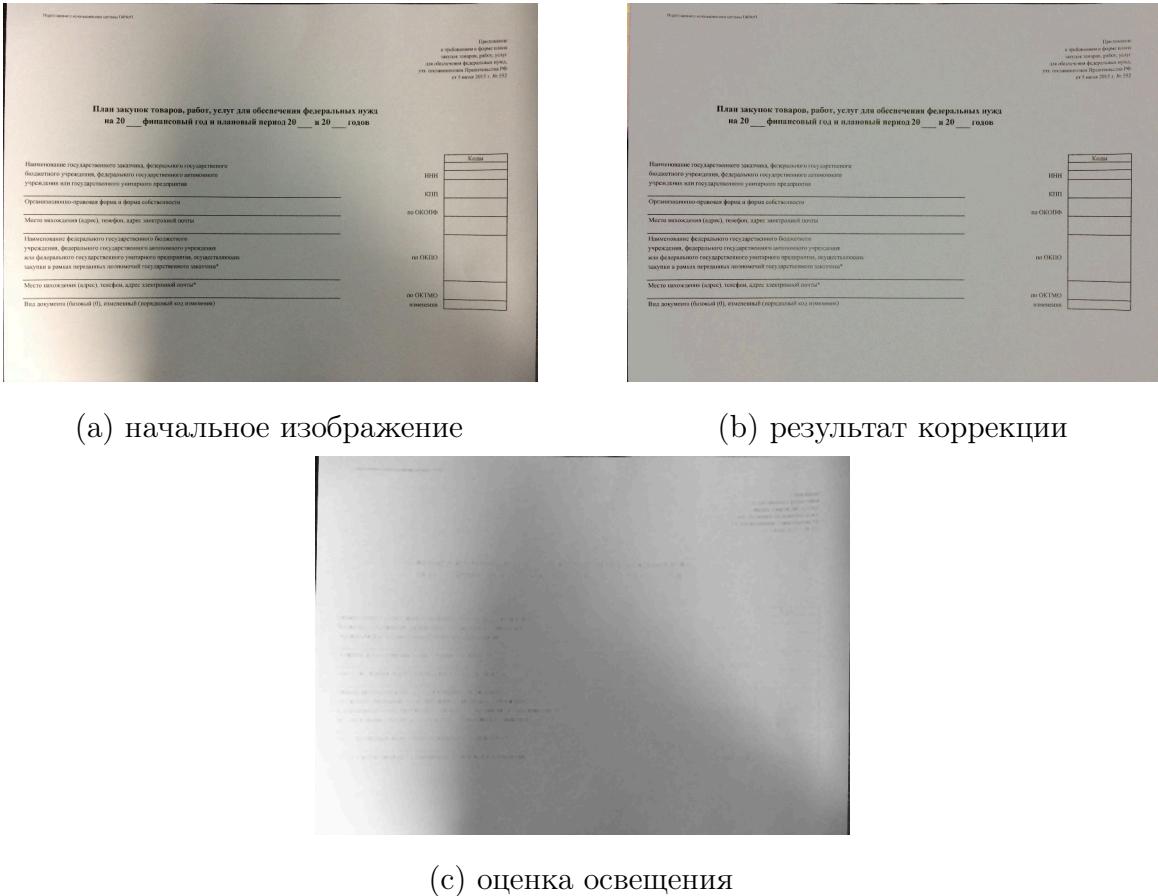


Рисунок Б.5 – Пример работы алгоритма коррекции освещения в случае наличия ярких теней

(а) начальное изображение

(b) результат коррекции

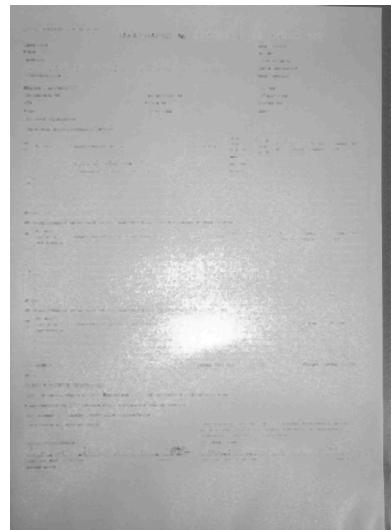


(c) оценка освещения

Рисунок Б.6 – Пример работы алгоритма коррекции освещения в случае наличия ярких теней

Заказчик		ЗАКАЗ-НАРЯД № 157765 от 22.02.2019				
Фамилия:	Имя:	Заказчик:	Год приемки:			
Адрес:		Приемка:	Бригада:			
Телефон:	+7 (385) 181-16-98	Срок выполнения:	До конца рабочей недели			
Планерщик:		Завершена:				
Марка и модель ДТС:		Год:	Год приемки:			
Базисная №:		Километр:	Продолж. на:			
Время:		Тип приемки:	Часы:			
Причина обращения:						
Первичные вспомогательные работы № Адресуемая работа Исполнитель Срок приемки Код приемки % выполн. % предп. Сроки без НДС						
1.	<i>Капитальный ремонт № 1</i>		<i>102</i>			
2.	<i>Капитальный ремонт № 2</i>		<i>102</i>			
3.						
4.						
5.						
6.						
7.						
8.						
9.						
10.						
Вид приемки: приемка частей (изделий), сплошной приемки						
№	Краткое описание приемки (изделия) или приемка по позиции	Ед. изм	Приемка	% выполн.	Сроки	Сроки без
1.	Проверка качества	штук	Приемка	100	предп.	НДС
2.	Проверка качества	штук	Приемка	100	предп.	НДС
3.	Проверка качества	штук	Приемка	100	предп.	НДС
4.	Проверка качества	штук	Приемка	100	предп.	НДС
5.	Проверка качества	штук	Приемка	100	предп.	НДС
6.	Проверка качества	штук	Приемка	100	предп.	НДС
7.	Проверка качества	штук	Приемка	100	предп.	НДС
8.	Проверка качества	штук	Приемка	100	предп.	НДС
9.	Проверка качества	штук	Приемка	100	предп.	НДС
10.	Проверка качества	штук	Приемка	100	предп.	НДС
Итого:						
Документы, поданные заказчиком (материала), проверены СДН 20.02.2019						
№	Наименование документа (изделия)	Ед. изм	Приемка	% выполн.	Сроки	Сроки без
1.	Проверка качества	штук	Приемка	100	предп.	НДС
2.	Проверка качества	штук	Приемка	100	предп.	НДС
3.	Проверка качества	штук	Приемка	100	предп.	НДС
4.	Проверка качества	штук	Приемка	100	предп.	НДС
5.	Проверка качества	штук	Приемка	100	предп.	НДС
6.	Проверка качества	штук	Приемка	100	предп.	НДС
7.	Проверка качества	штук	Приемка	100	предп.	НДС
8.	Проверка качества	штук	Приемка	100	предп.	НДС
9.	Проверка качества	штук	Приемка	100	предп.	НДС
10.	Проверка качества	штук	Приемка	100	предп.	НДС
СТОИМОСТЬ				Сумма без НДС НДС Общая сумма с НДС		
Всего в смете (расходах):						
СДН № 14-01-000 Мартовский СДН не является платежным документом (НДС)						
При приемке СДН не выставляется, подпись в смете недостаточна						
При приемке СДН не выставляется, подпись недостаточна						
Транспортные расходы оплачиваются с документом выставляемым заказчиком, в соответствии с условиями договора о перевозке грузов СДН						
Организаторы: <i>Петров Константин Петрович</i> <i>Шевченко Иван Иванович</i> <i>Лебедев Геннадий Геннадьевич</i>						
Уполномоченный для подписания						
Примечание:						

(а) начальное изображение

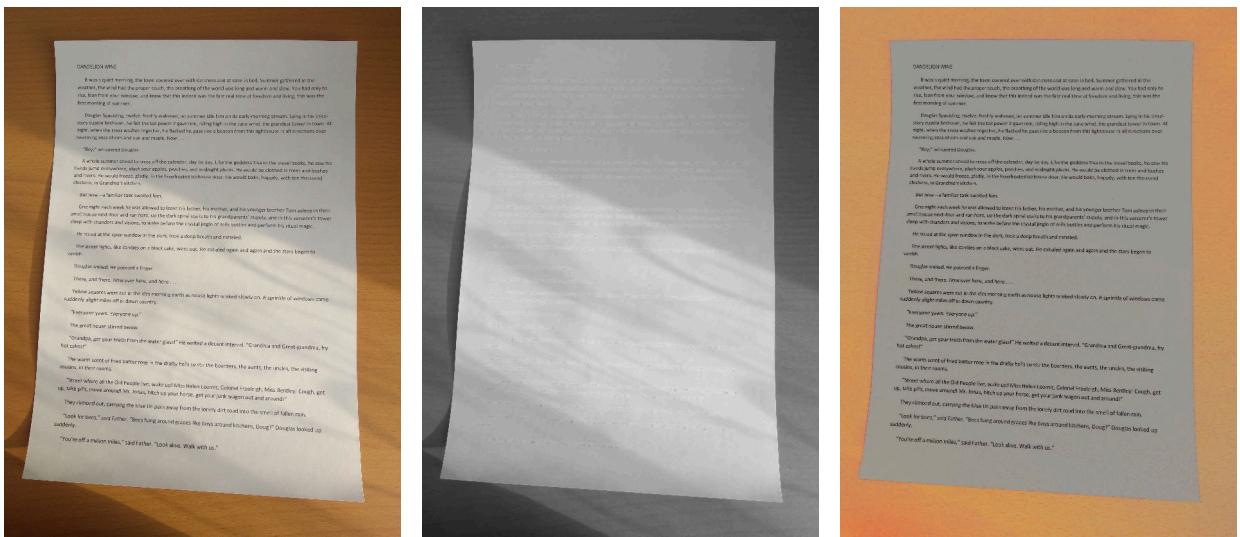


(б) оценка освещения

Заказчик		ЗАКАЗ-НАРД № 158765 от 22 января 2019						
Заказчик Юридическое лицо: Телефон: +7(985) 181-10-98 Платежные реквизиты:		Заказ-наряд Форма приемки Форма отгрузки Бланк приемки Бланк отгрузки						
Марка и модель АТС: Имя: Фамилия: Место работы: Контактный телефон: Прочее обозначение:		Год: Приемка Приемка Приемка Приемка						
Перечень выполненных работ:								
№ столбца	Наименование работы	Исполнитель	Дата выполнения	Время затраченное на выполнение	Норма затрат времени	% нормы	% затраченного времени	Сумма, без НДС
1.	Комплексное тестирование № 1			152				
2.	Комплексное тестирование № 2			162				
3.								
4.								
5.								
6.								
7.								
8.								
9.								
10.								
Итого:								
Проверка, выполнение работ (маркировка, выявление нарушений)								
1.	Гарантийное	Приемлено исполнителем	Код, подпись исполнителя	Код, подпись приемщика	Код, подпись	Сумма, без НДС	Сумма, без НДС	
2.								
3.								
4.								
5.								
6.								
7.								
8.								
9.								
10.								
Метод:								
Подпись/печать ответственного за исполнение (исполнителя, руководителя подразделения)								
1.	Исполнитель	Приемщик	Код, подпись исполнителя	Код, подпись приемщика	Код, подпись	Сумма, без НДС	Сумма, без НДС	
2.								
3.								
4.								
5.								
6.								
7.								
8.								
9.								
10.								
Сторонность		Сумма, без НДС НДС Сумма суммы с НДС						
Итого:								
Всего в сопроводительном документе:								
НДС не начисляется (НДС-Импорт СТО не является физическим лицом)								
НДС начисляется (НДС-Импорт СТО является физическим лицом)								
НДС не начисляется (НДС-Импорт СТО не является физическим лицом)								
НДС начисляется (НДС-Импорт СТО является физическим лицом)								
Примечание: в случае отсутствия работ, в результате выполнения которых СТО не получила определенного количества продукции, или если количество продукции, фактически полученной СТО, отличается от заявленного в настоящем документе, то в нем должны быть отражены факты, обстоятельства, причины и меры по устранению выявленных недостатков.								
Ознакомление с условиями приемки								
Приемщиком								
Исполнитель: <u>Петр Петрович Петров</u> Должность: <u>директор</u> Фамилия, имя, отчество: <u>Петров Петр Иванович</u> Подпись: <u>Петров Петр Иванович</u>								
Приемщиком								
Исполнитель: <u>Александр Иванов</u> Должность: <u>заместитель директора</u> Фамилия, имя, отчество: <u>Иванов Александр Иванович</u> Подпись: <u>Иванов Александр Иванович</u>								

(c) результат коррекции

Рисунок Б.7 – Пример работы алгоритма коррекции освещения в случае наличия
бликов на поверхности файла

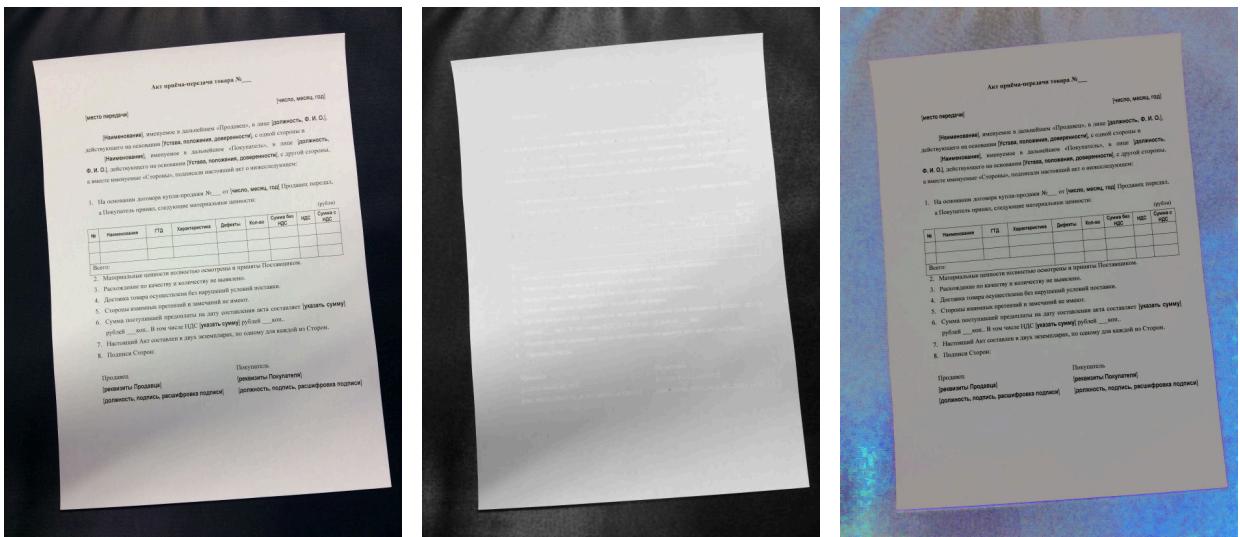


(а) начальное изображение

(б) оценка освещения

(c) результат коррекции

Рисунок Б.8 – Пример работы алгоритма коррекции освещения в случае наличия контрастного внешнего фона



(а) начальное изображение

(б) оценка освещения

(c) результат коррекции

Рисунок Б.9 – Пример работы алгоритма коррекции освещения в случае наличия контрастного внешнего фона темного цвета

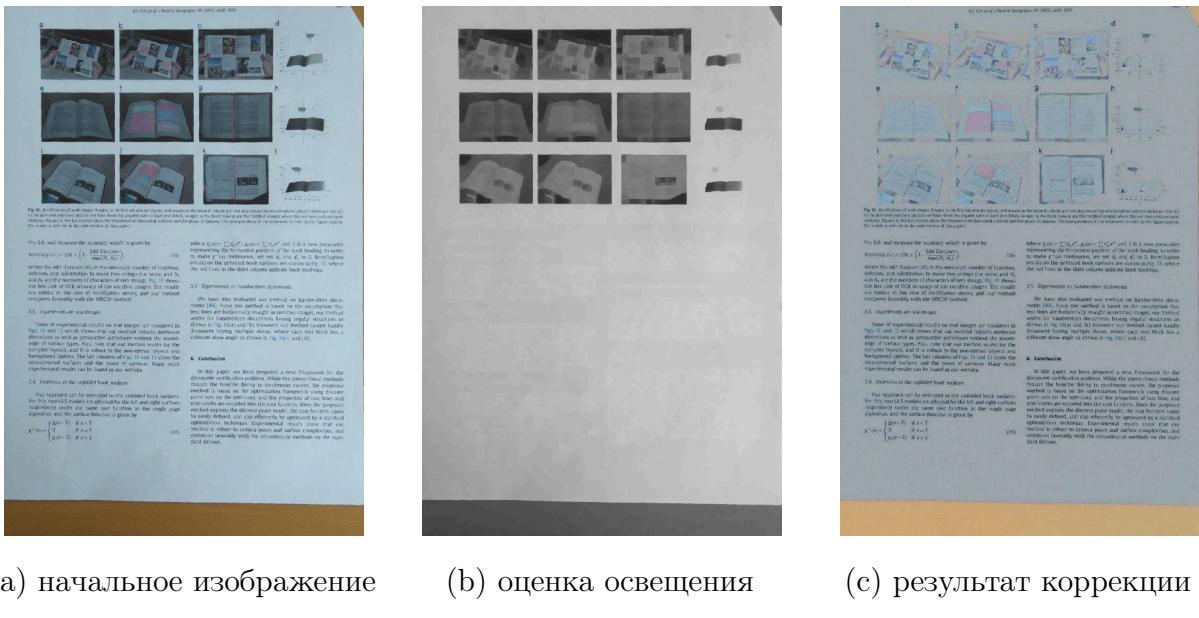
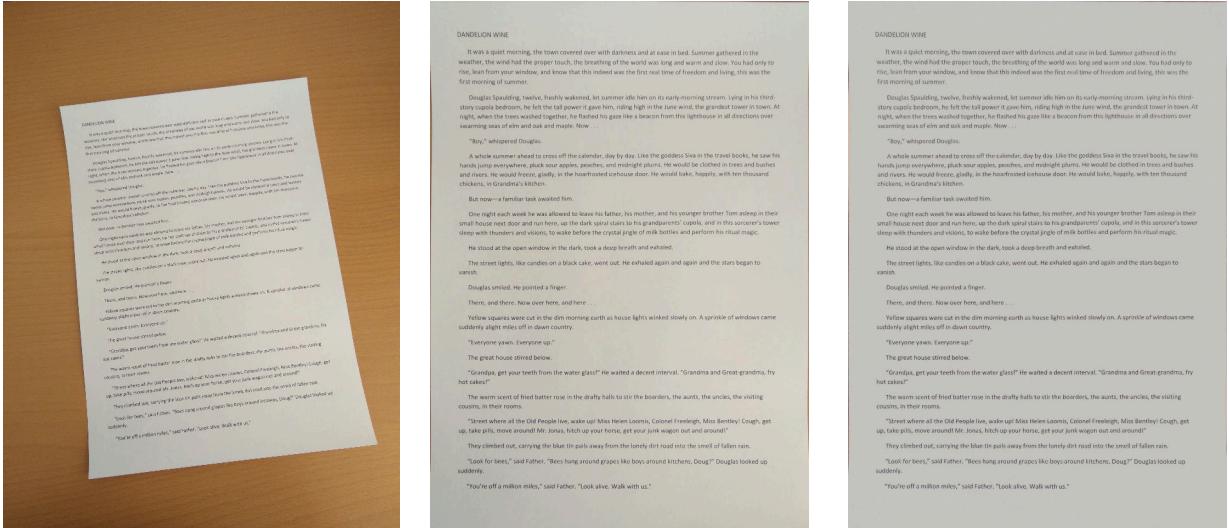


Рисунок Б.10 – Пример работы алгоритма коррекции освещения в случае документа, содержащего темные изображения

Приложение В.

Примеры измерения качества работы предложенных алгоритмов

В данном приложении приводятся несколько различных примеров, демонстрирующих качество работы преобразований изображений по исправлению искажений, описания которых содержатся в гл. 2 и гл. 3. Каждый пример представляет из себя входную фотографию, после коррекции перспективы и после исправлений перспективных искажений и нормализации освещенности. Для каждого изображения приведено значение метрики q , описание которой приводится в гл. 4.



(a) начальное изображение

$q = 0.367$

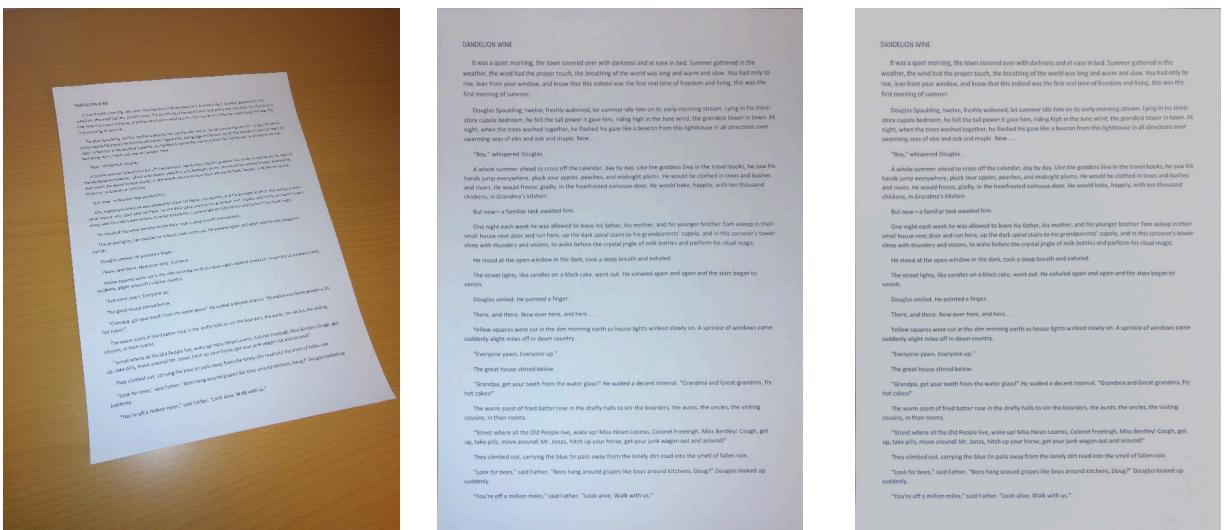
(b) после коррекции

$q = 0.998$

(c) после дополнительной
нормализации освещенности

$q = 0.999$

Рисунок В.1 – Оценка качества предложенных преобразований



(a) начальное изображение

$q = 0.282$

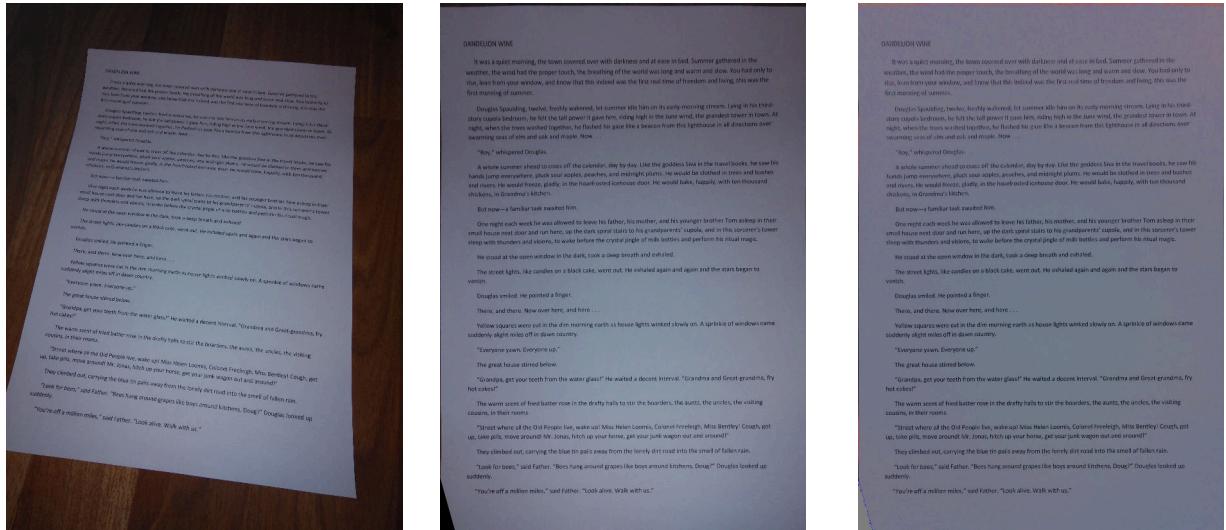
(b) после коррекции

$q = 0.996$

(c) после дополнительной
нормализации освещенности

$q = 0.999$

Рисунок В.2 – Оценка качества предложенных преобразований



(a) начальное изображение

$$q = 0.46$$

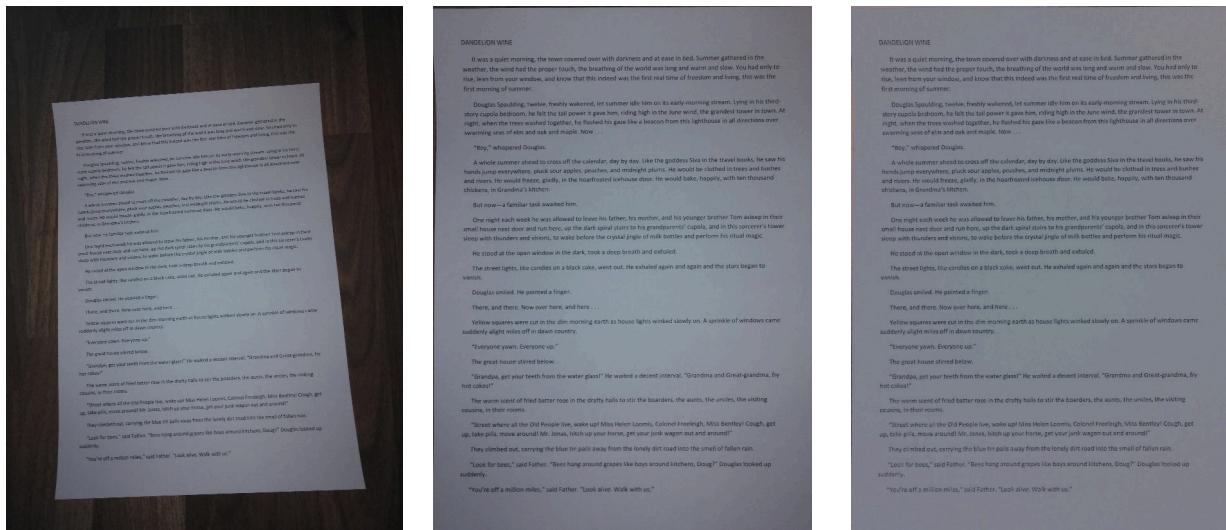
(b) после коррекции перспективы

$$q = 0.49$$

(c) после дополнительной нормализации освещенности

$$q = 0.97$$

Рисунок В.3 – Оценка качества предложенных преобразований



(a) начальное изображение

$$q = 0.30$$

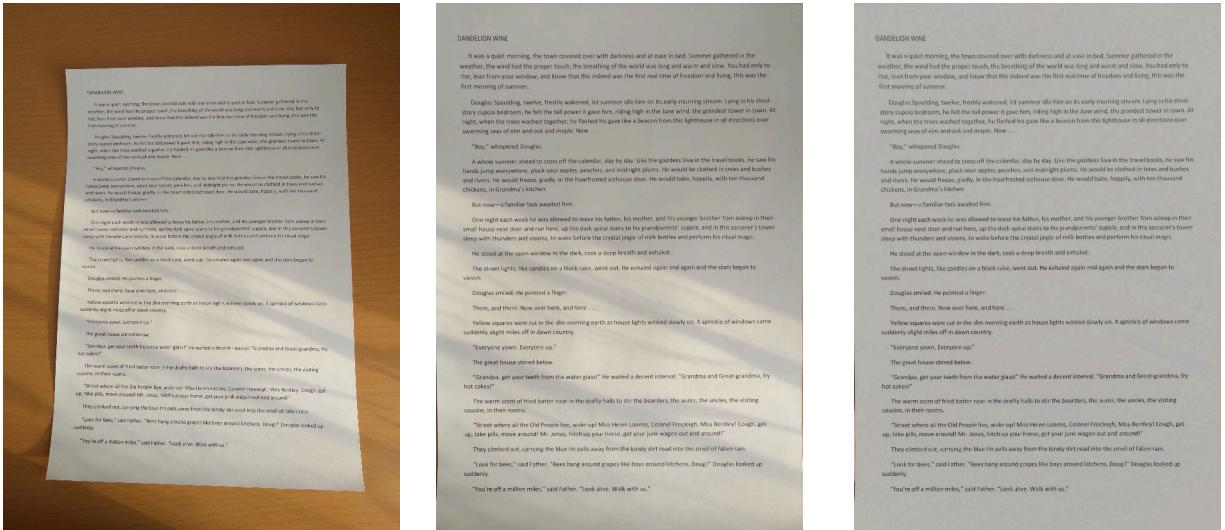
(b) после коррекции перспективы

$$q = 0.92$$

(c) после дополнительной нормализации освещенности

$$q = 0.99$$

Рисунок В.4 – Оценка качества предложенных преобразований



(a) начальное изображение

$q = 0.938$

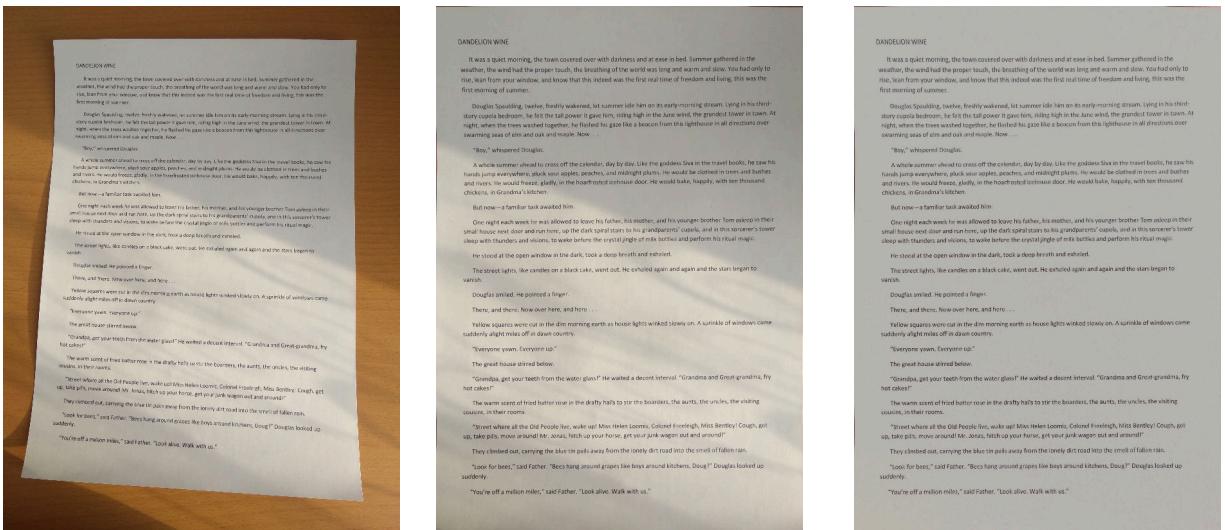
(b) после коррекции

$q = 0.987$

(c) после дополнительной
нормализации освещенности

$q = 0.988$

Рисунок B.5 – Оценка качества предложенных преобразований



(a) начальное изображение

$q = 0.784$

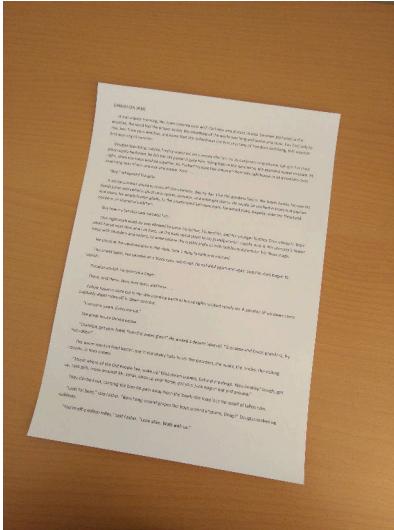
(b) после коррекции

$q = 0.345$

(c) после дополнительной
нормализации освещенности

$q = 0.999$

Рисунок B.6 – Оценка качества предложенных преобразований



(a) начальное изображение

$$q = 0.0479$$

DANDY WINE

It was a quiet morning, the town covered over with darkness and areas in light. Summer gathered in the window, the wind had the power to catch, the breathing of the world was long and warm and slow. You had only to rise, lean from your window, and know that this indeed was the first real time of freedom and living. This was the first morning of summer.

Douglas Spaulding, twelve, freshly wakened, let summer sit him on its early-morning stream. Lying in his third-story cupola bedroom, he felt the tall power it gave him, riding high in the June wind, the grandest tower in town. At night he would sleep, giddily, in the heart-achmed icemouse door. He would baffle, Tappily, with ten thousand swarming roes of elm and oak and maple. Now...

"Boy," whispered Douglas.

A whole summer ahead to cross off the calendar, day by day. Like the goddesses like in the travel books, he has hands jump everywhere, pluck your apples, peaches, and midnight plums. He would be clothed in trees and bushes and flowers. He would sleep, giddily, in the heart-achmed icemouse door. He would baffle, Tappily, with ten thousand swarming roes of elm and oak and maple. Now...

But now—a familiar look assailed him.

One right each week he was allowed to leave his father, his mother, and his younger brother Tom asleep in their small house next door and run home, as the dark spirit starts to his grandparents' cupola, and in this son's tower sleep with shudders and visions, to wake before the crystal jugs of milk bottles and perform his usual magic.

He stood at the open window, in the dark, took a deep breath and exhaled.

The street lights, like candles on a black cane, went out. He exhaled again and again and the stars began to vanish.

Douglas smiled. There, and there. Now, over here, and here...

Yellow squares were cut in the dim morning earth as house lights worked slowly on. A sprinkle of windows came suddenly alight miles off in down country.

"Everyone yawn. Everyone up."

The great house shuddered below.

"Grandma, get your teeth from the water glass!" he waled a decent interval. "Grandma and Great-grandma, try not to fall."

The warm scent of fried batter rose in the drafty halls to stir the boarders, the aunts, the uncles, the visiting cousins, in their rooms.

"Street where all the Old People live, wake up! Miss Helen Loonies, Colored Freewheel, Miss Bentley Cough, get up, take pills, move around Mr. Jones, hitch up your horse, get your junk wagon out and around!"

They climbed out, carrying the blue tin pails away from the lonely dirt road into the smell of fallen roses.

"Look for bees," said Father. "Bees hang around grapes like boys around kitchens, Doug?" Douglas looked up suddenly.

"You're off a million miles," said Father. "Look alive. Walk with us."

(b) после коррекции

$$\begin{matrix} & \text{перспективы} \\ q & = 0.9992 \end{matrix}$$

DANDY WINE

It was a quiet morning, the town covered over with darkness and areas in light. Summer gathered in the window, the wind had the power to catch, the breathing of the world was long and warm and slow. You had only to rise, lean from your window, and know that this indeed was the first real time of freedom and living. This was the first morning of summer.

Douglas Spaulding, twelve, freshly wakened, let summer sit him on its early-morning stream. Lying in his third-story cupola bedroom, he felt the tall power it gave him, riding high in the June wind, the grandest tower in town. At night he would sleep, giddily, in the heart-achmed icemouse door. He would baffle, Tappily, with ten thousand swarming roes of elm and oak and maple. Now...

"Boy," whispered Douglas.

A whole summer ahead to cross off the calendar, day by day. Like the goddesses like in the travel books, he saw his hands jump everywhere, pluck your apples, peaches, and midnight plums. He would be clothed in trees and bushes and flowers. He would sleep, giddily, in the heart-achmed icemouse door. He would baffle, Tappily, with ten thousand swarming roes of elm and oak and maple. Now...

But now—a familiar look assailed him.

One right each week he was allowed to leave his father, his mother, and his younger brother Tom asleep in their small house next door and run home, as the dark spirit starts to his grandparents' cupola, and in this son's tower sleep with shudders and visions, to wake before the crystal jugs of milk bottles and perform his usual magic.

He stood at the open window in the dark, took a deep breath and exhaled.

The street lights, like candles on a black cane, went out. He exhaled again and again and the stars began to vanish.

Douglas smiled. There, and there. Now, over here, and here...

Yellow squares were cut in the dim morning earth as house lights worked slowly on. A sprinkle of windows came suddenly alight miles off in down country.

"Everyone yawn. Everyone up."

The great house shuddered below.

"Grandma, get your teeth from the water glass!" he waled a decent interval. "Grandma and Great-grandma, try not to fall."

The warm scent of fried batter rose in the drafty halls to stir the boarders, the aunts, the uncles, the visiting cousins, in their rooms.

"Street where all the Old People live, wake up! Miss Helen Loonies, Colored Freewheel, Miss Bentley Cough, get up, take pills, move around Mr. Jones, hitch up your horse, get your junk wagon out and around!"

They climbed out, carrying the blue tin pails away from the lonely dirt road into the smell of fallen roses.

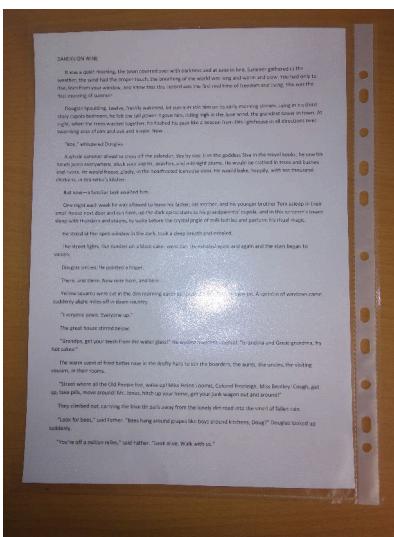
"Look for bees," said Father. "Bees hang around grapes like boys around kitchens, Doug?" Douglas looked up suddenly.

"You're off a million miles," said Father. "Look alive. Walk with us."

(c) после дополнительной нормализации освещенности

$$q = 0.9989$$

Рисунок В.7 – Оценка качества предложенных преобразований



(a) начальное изображение

$$q = 0.95$$

DANDY WINE

It was a quiet morning, the town covered over with darkness and areas in light. Summer gathered in the window, the wind had the power to catch, the breathing of the world was long and warm and slow. You had only to rise, lean from your window, and know that this indeed was the first real time of freedom and living. This was the first morning of summer.

Douglas Spaulding, twelve, freshly wakened, let summer sit him on its early-morning stream. Lying in his third-story cupola bedroom, he felt the tall power it gave him, riding high in the June wind, the grandest tower in town. At night he would sleep, giddily, in the heart-achmed icemouse door. He would baffle, Tappily, with ten thousand swarming roes of elm and oak and maple. Now...

"Boy," whispered Douglas.

A whole summer ahead to cross off the calendar, day by day. Like the goddesses like in the travel books, he saw his hands jump everywhere, pluck your apples, peaches, and midnight plums. He would be clothed in trees and bushes and flowers. He would sleep, giddily, in the heart-achmed icemouse door. He would baffle, Tappily, with ten thousand swarming roes of elm and oak and maple. Now...

But now—a familiar look assailed him.

One right each week he was allowed to leave his father, his mother, and his younger brother Tom asleep in their small house next door and run home, as the dark spirit starts to his grandparents' cupola, and in this son's tower sleep with shudders and visions, to wake before the crystal jugs of milk bottles and perform his usual magic.

He stood at the open window, in the dark, took a deep breath and exhaled.

The street lights, like candles on a black cane, went out. He exhaled again and again and the stars began to vanish.

Douglas smiled. There, and there. Now, over here, and here...

Yellow squares were cut in the dim morning earth as house lights worked slowly on. A sprinkle of windows came suddenly alight miles off in down country.

"Everyone yawn. Everyone up."

The great house shuddered below.

"Grandma, get your teeth from the water glass!" he waled a decent interval. "Grandma and Great-grandma, try not to fall."

The warm scent of fried batter rose in the drafty halls to stir the boarders, the aunts, the uncles, the visiting cousins, in their rooms.

"Street where all the Old People live, wake up! Miss Helen Loonies, Colored Freewheel, Miss Bentley Cough, get up, take pills, move around Mr. Jones, hitch up your horse, get your junk wagon out and around!"

They climbed out, carrying the blue tin pails away from the lonely dirt road into the smell of fallen roses.

"Look for bees," said Father. "Bees hang around grapes like boys around kitchens, Doug?" Douglas looked up suddenly.

"You're off a million miles," said Father. "Look alive. Walk with us."

(b) после коррекции

$$\begin{matrix} & \text{перспективы} \\ q & = 0.88 \end{matrix}$$

DANDY WINE

It was a quiet morning, the town covered over with darkness and areas in light. Summer gathered in the window, the wind had the power to catch, the breathing of the world was long and warm and slow. You had only to rise, lean from your window, and know that this indeed was the first real time of freedom and living. This was the first morning of summer.

Douglas Spaulding, twelve, freshly wakened, let summer sit him on its early-morning stream. Lying in his third-story cupola bedroom, he felt the tall power it gave him, riding high in the June wind, the grandest tower in town. At night he would sleep, giddily, in the heart-achmed icemouse door. He would baffle, Tappily, with ten thousand swarming roes of elm and oak and maple. Now...

"Boy," whispered Douglas.

A whole summer ahead to cross off the calendar, day by day. Like the goddesses like in the travel books, he saw his hands jump everywhere, pluck your apples, peaches, and midnight plums. He would be clothed in trees and bushes and flowers. He would sleep, giddily, in the heart-achmed icemouse door. He would baffle, Tappily, with ten thousand swarming roes of elm and oak and maple. Now...

But now—a familiar look assailed him.

One right each week he was allowed to leave his father, his mother, and his younger brother Tom asleep in their small house next door and run home, as the dark spirit starts to his grandparents' cupola, and in this son's tower sleep with shudders and visions, to wake before the crystal jugs of milk bottles and perform his usual magic.

He stood at the open window in the dark, took a deep breath and exhaled.

The street lights, like candles on a black cane, went out. He exhaled again and again and the stars began to vanish.

Douglas smiled. There, and there. Now, over here, and here...

Yellow squares were cut in the dim morning earth as house lights worked slowly on. A sprinkle of windows came suddenly alight miles off in down country.

"Everyone yawn. Everyone up."

The great house shuddered below.

"Grandma, get your teeth from the water glass!" he waled a decent interval. "Grandma and Great-grandma, try not to fall."

The warm scent of fried batter rose in the drafty halls to stir the boarders, the aunts, the uncles, the visiting cousins, in their rooms.

"Street where all the Old People live, wake up! Miss Helen Loonies, Colored Freewheel, Miss Bentley Cough, get up, take pills, move around Mr. Jones, hitch up your horse, get your junk wagon out and around!"

They climbed out, carrying the blue tin pails away from the lonely dirt road into the smell of fallen roses.

"Look for bees," said Father. "Bees hang around grapes like boys around kitchens, Doug?" Douglas looked up suddenly.

"You're off a million miles," said Father. "Look alive. Walk with us."

(c) после дополнительной нормализации освещенности

$$q = 0.82$$

Рисунок В.8 – Оценка качества предложенных преобразований в случае документа в файле

Приложение Г.

Примеры работы предложенных алгоритмов

В данном приложении приводятся успешные примеры работы предложенных методов для устранения искажений. Каждый пример состоит из изначального изображения документа, после применения преобразования коррекции перспективы и затем после нормализации освещения.

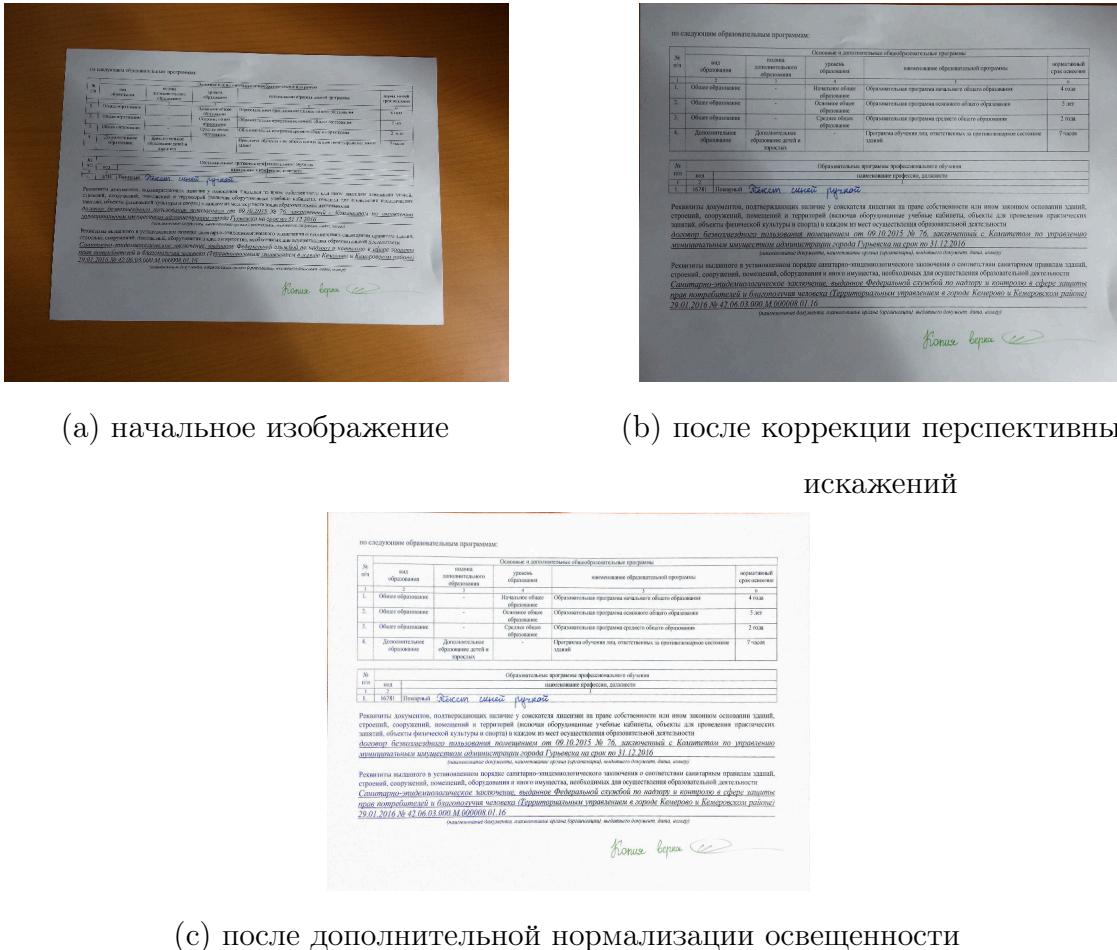


Рисунок Г.1 – Пример работы предложенных методов исправления искажений

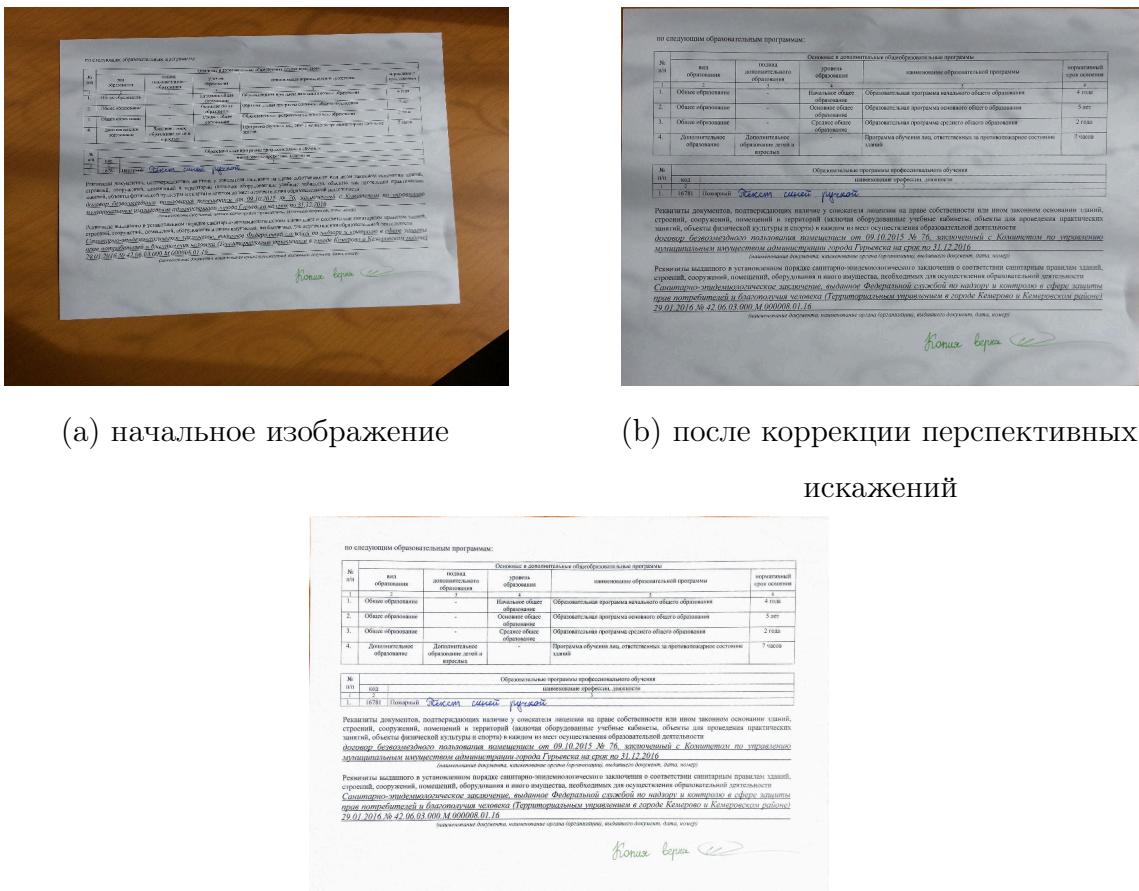
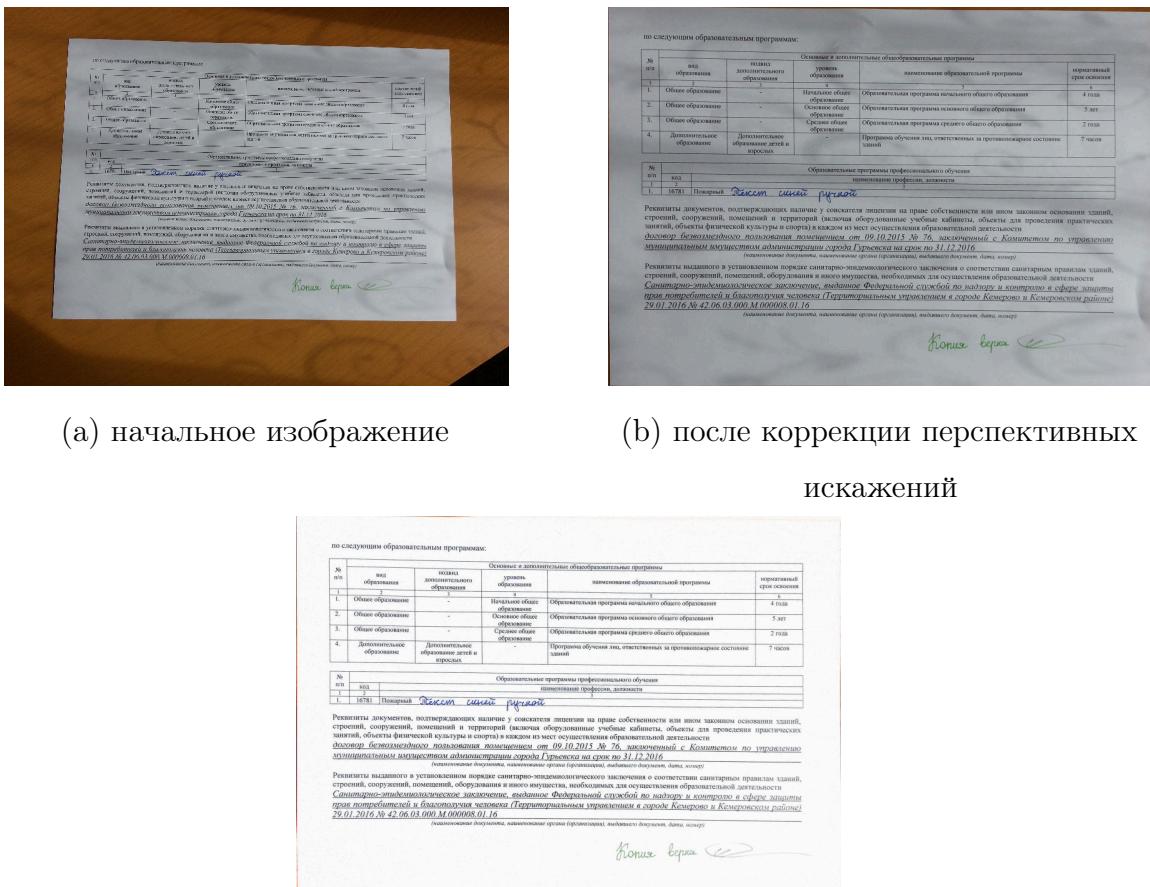


Рисунок Г.2 – Пример работы предложенных методов исправления искажений



(a) начальное изображение

(b) после коррекции перспективных искажений

(c) после дополнительной нормализации освещенности

Рисунок Г.3 – Пример работы предложенных методов исправления искажений

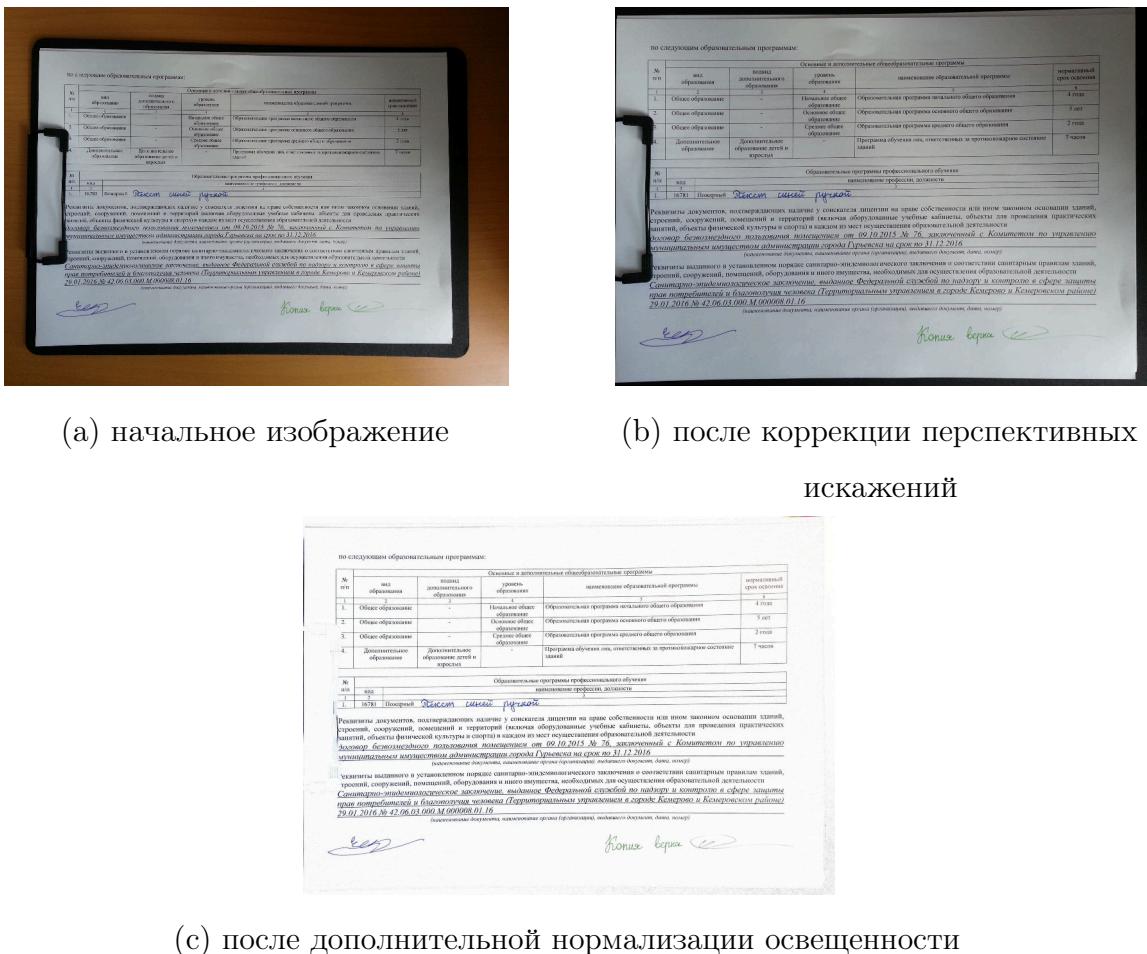
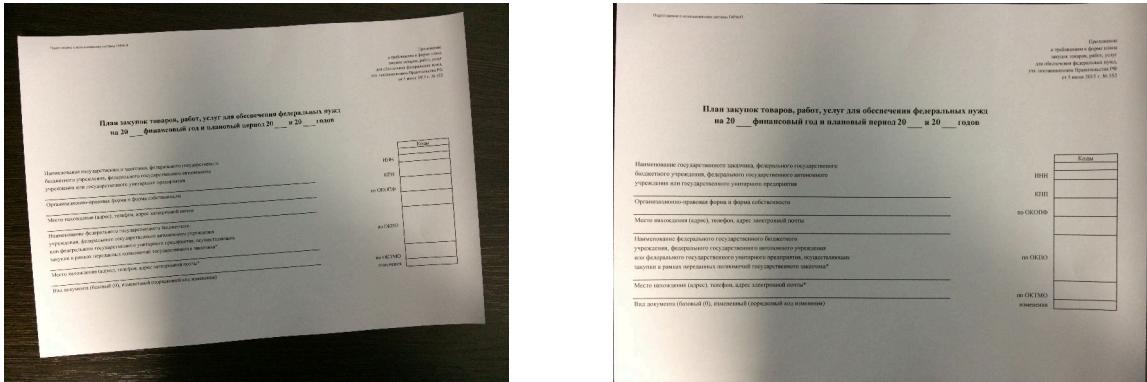
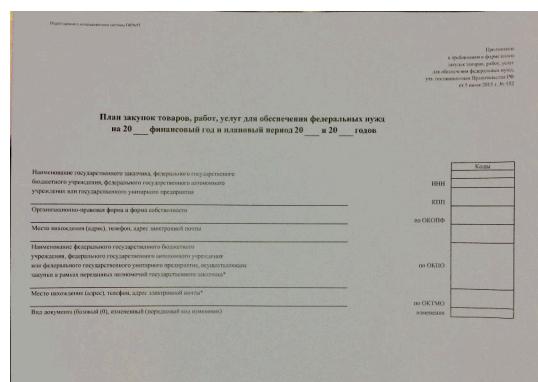


Рисунок Г.4 – Пример работы предложенных методов исправления искажений



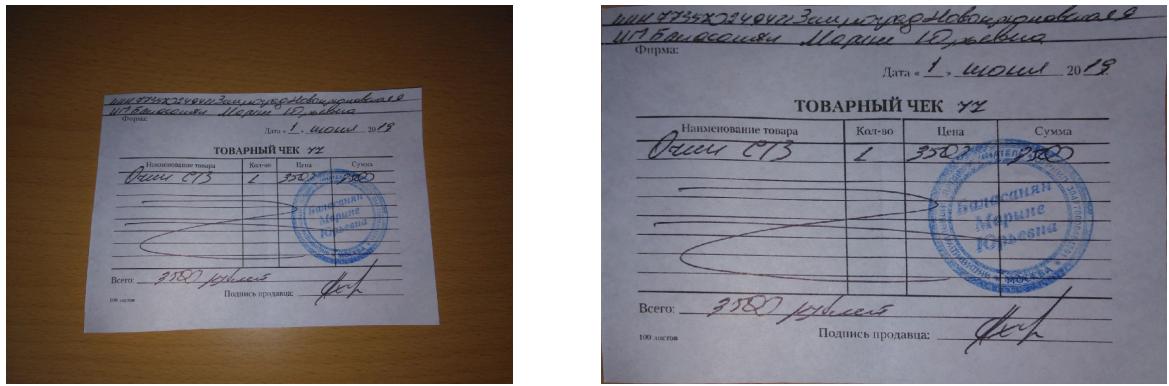
(a) начальное изображение

(b) после коррекции перспективных искажений



(c) после дополнительной нормализации освещенности

Рисунок Г.5 – Пример работы предложенных методов исправления искажений



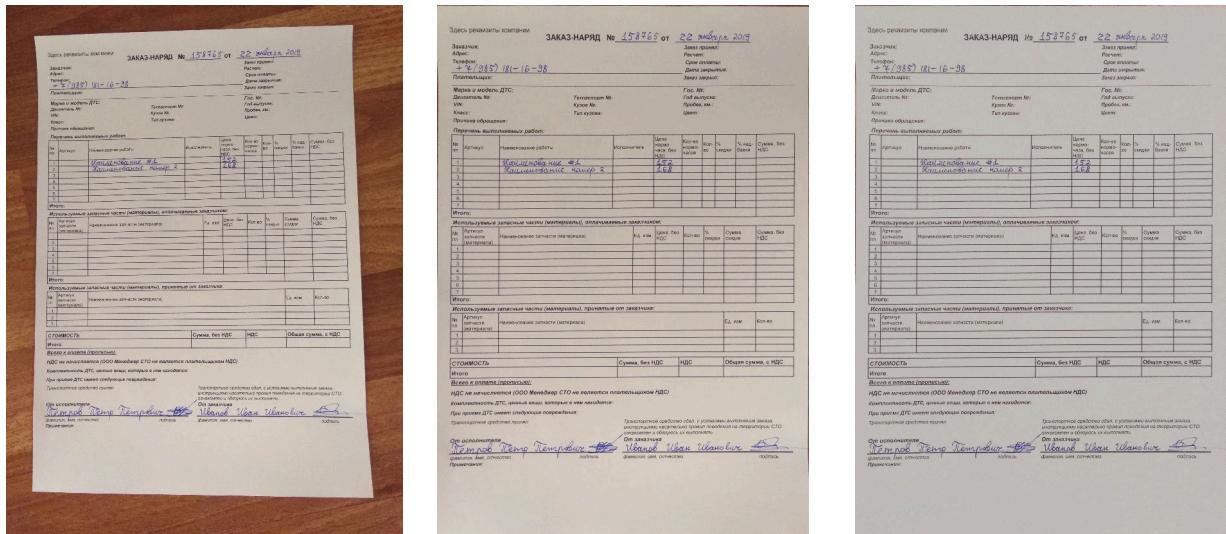
(a) начальное изображение

(b) после коррекции перспективных искажений



(c) после дополнительной нормализации освещенности

Рисунок Г.6 – Пример работы предложенных методов исправления искажений



(a) начальное изображение

(b) после коррекции

перспективных искажений

(c) после дополнительной

нормализации освещенности

Рисунок Г.7 – Пример работы предложенных методов исправления искажений

БУЛГАЛЬСКИЙ БАЛАНС
на 30 Июня 2009 г.

		4 023 000
		ДОБРО
Форма №1 по ОКУД		2000036
Организация - производство стеклопакетов		00004996
Идентификационный номер налогоплательщика		31015021450252010001
Вид деятельности - производство стеклопакетов		по СНКЭД 26.14
Признак собственности		Частная собственность
Открытие лицензии		по АКОБиС: счет 47
Единая система учёта и отчётов		по НДСИ: счет 39
Местонахождение (адрес)		37291, Нижегородская обл., Великий Новгород, Волчанская, д. 15
Дата утверждения		
Дата открытия / приемки		
АКТИВ		
	код	На начало отчетного года
	поквартально	На конец отчетного года
1.		
Нематериальные активы		
Библиотечные активы		110 692 598
Основные средства		120 125453 120811
Нематериальное имущество		130 11538 6530
Долгосрочные финансовые активы		140 50115 50112
Отложенные налоговые активы		200 25454 25451
Итого по разделу 1		190 172931 169891
2. Оборотные активы		
Запасы		210 141693 155260
в том числе:		
сырье, материалы и др. аналогичные ценности		211 97474 97387
изделия, полуфабрикаты, товары для продажи		212 273 273
головные проекции и т.д. для инженерных		214 399510 406346
работ спортивных сооружений		215 1796 1796
изделия для сельского хозяйства		216 1000 1000
Налог на добавленную стоимость по приобретенным ценностям		240 68910 44671
в том числе:		
запасы и аналогичные ценности		241 50204 36200
финансовые активы		250 290 2290
недвижимое имущество		251 113 113
другие запасы		270 182 182
Денежные средства		280 36645 36642
Прочие оборотные активы		290 371921 363811
БАЛАНС		

(а) начальное изображение

(b) после коррекции

БУХГАЛТЕРСКИЙ БАЛАНС
на 30 Июня 2009 г.

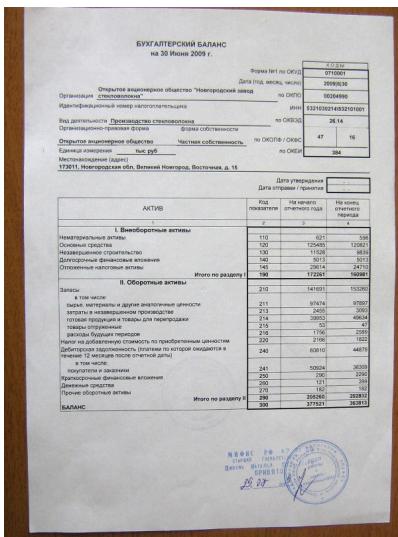
Форма №1 на ОГРН	Х-02
Дат (год, месяц, число)	2009/06/30
по ОГРН	50004990
по СКРН	26.14
по ОКОНХ	47
по ОГЕИМ	384

Организация Открытое акционерное общество "Новгородская земля строительство"
Юридический адрес: Промышленный проспект, дом 100
Организационно-правовая форма: Акционерное общество
Отрасль экономики: Неспециализированная деятельность по управлению
Начальник собственности: Единый измеритель - тыс руб
Индивидуальный идентификатор: 172011, Новгородская обл., Великий Новгород, Восточная, д. 15

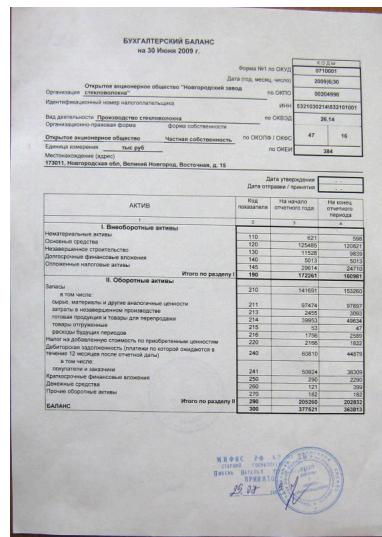
Дата утверждения			
Дата открытия / приемки			
АКТИВ	Код пояснения	На начало отчетного периода	На конец отчетного периода
I. Финансовые активы			
Нематериальные активы			
Оценочные активы	110	621	595
Незавершенное строительство	120	1254560	1260611
Долгосрочные финансовые вложения	130	1250000	1250000
Стоимость участия в уставном капитале	140	5013	5013
Итого по разделу I	190	172261	169081
II. Оборотные активы			
Запасы			
в том числе:			
бум. материалы и друг. нематериальные ценности	211	87474	97937
запасы производственных мощностей	212	2730	3500
птицы племенные и телята для откорма	214	99653	48624
товары, работы, услуги	215	1796	457
расходы будущих периодов	216	2166	2002
иные запасы	220	100	100
на конец отчетного периода. По приобретенным ценностям (дебиторская задолженность, остаток кредиторской задолженности в конце отчетного периода (дата))	240	10815	44879
Итого по разделу II	290	205260	202832
САДАНО		377521	363613

(b) после коррекции перспективных искажений (c) после дополнительной нормализации освещенности

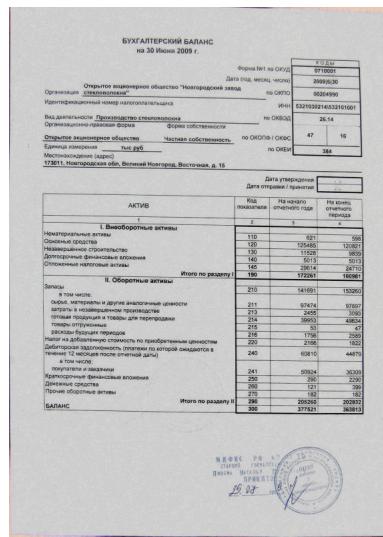
Рисунок Г.8 – Пример работы предложенных методов исправления искажений



(а) начальное изображение



(b) после коррекции



(b) после коррекции перспективных искажений (c) после дополнительной нормализации освещенности

Рисунок Г.9 – Пример работы предложенных методов исправления искажений

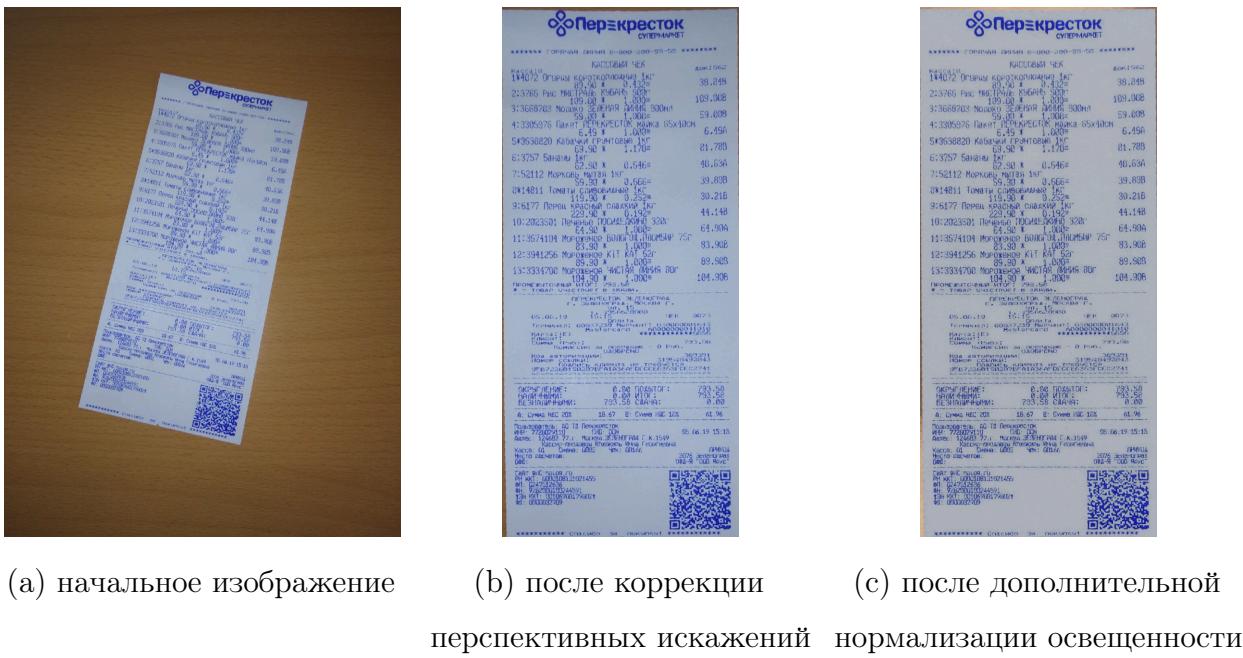


Рисунок Г.10 – Пример работы предложенных методов исправления искажений

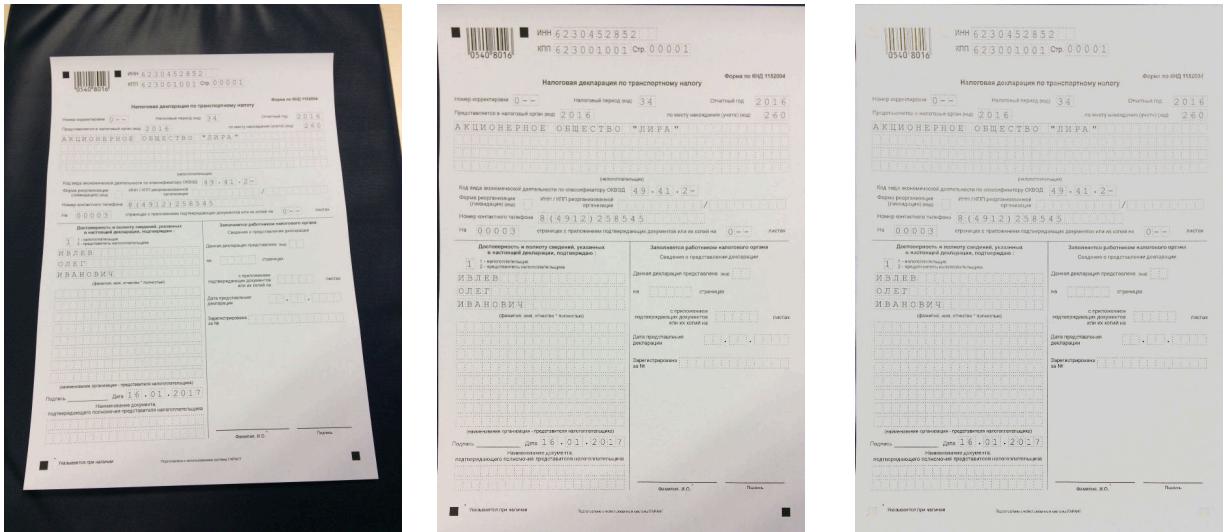
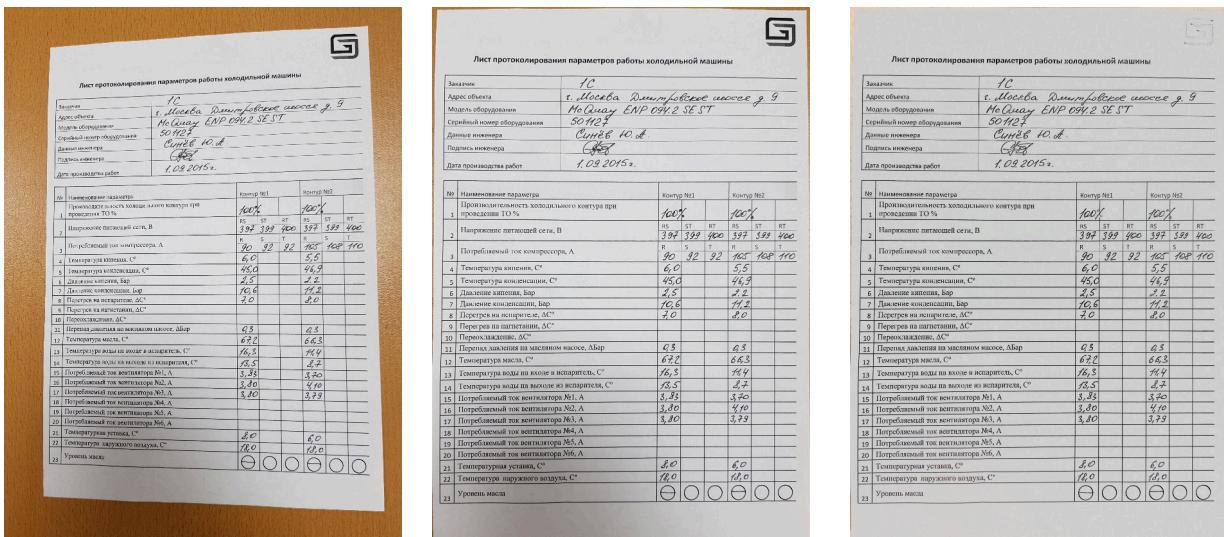
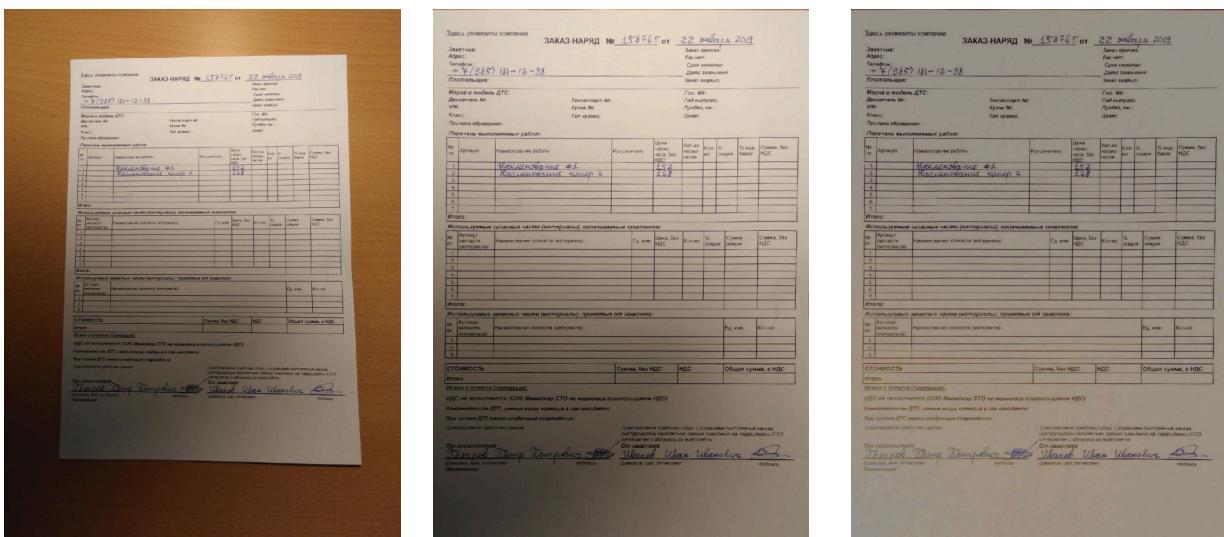


Рисунок Г.11 – Пример работы предложенных методов исправления искажений



(a) начальное изображение (b) после коррекции (c) после дополнительной
перспективных искажений нормализации освещенности

Рисунок Г.12 – Пример работы предложенных методов исправления искажений



(a) начальное изображение (b) после коррекции (c) после дополнительной
перспективных искажений нормализации освещенности

Рисунок Г.13 – Пример работы предложенных методов исправления искажений

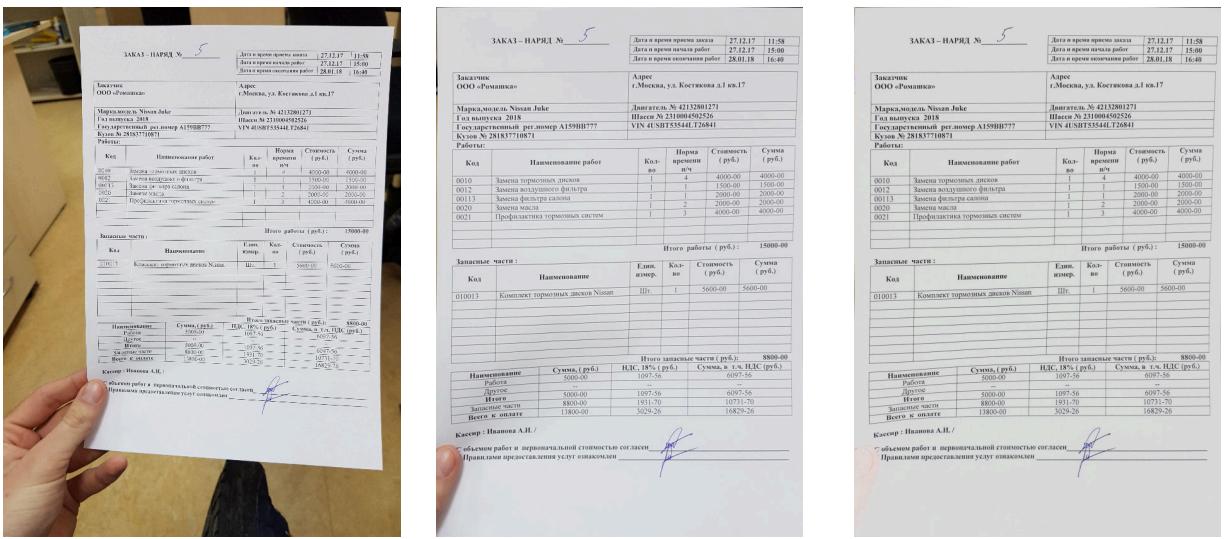


Рисунок Г.14 – Пример работы предложенных методов исправления искажений

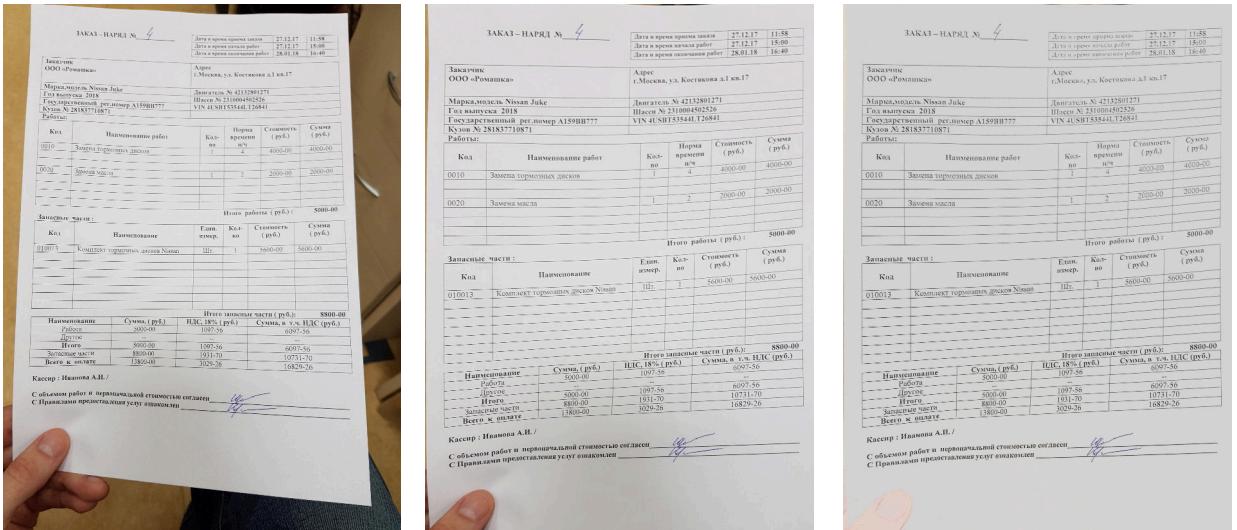
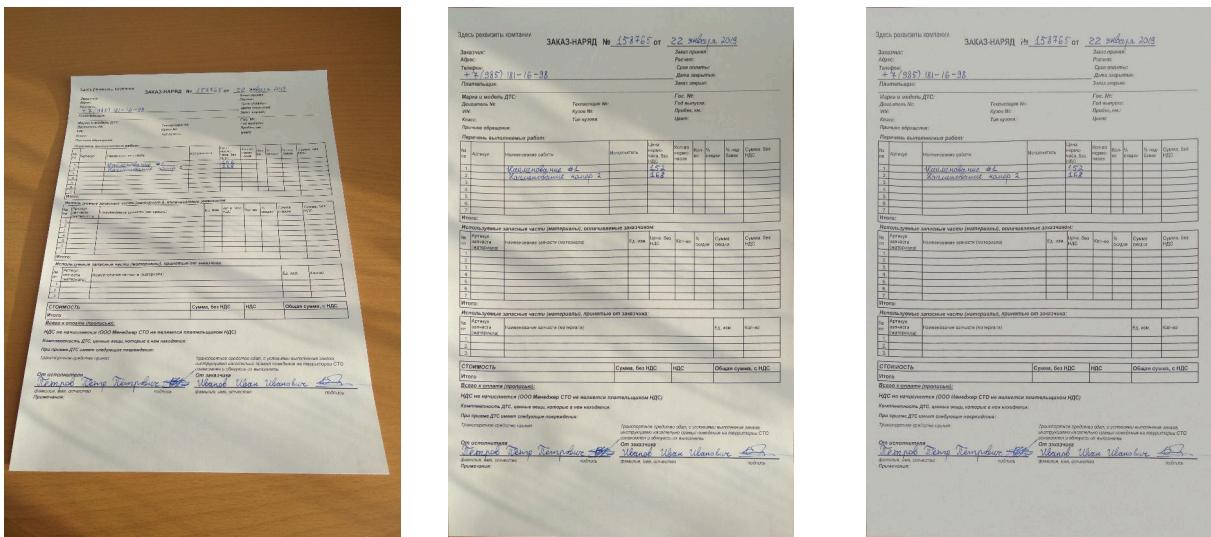
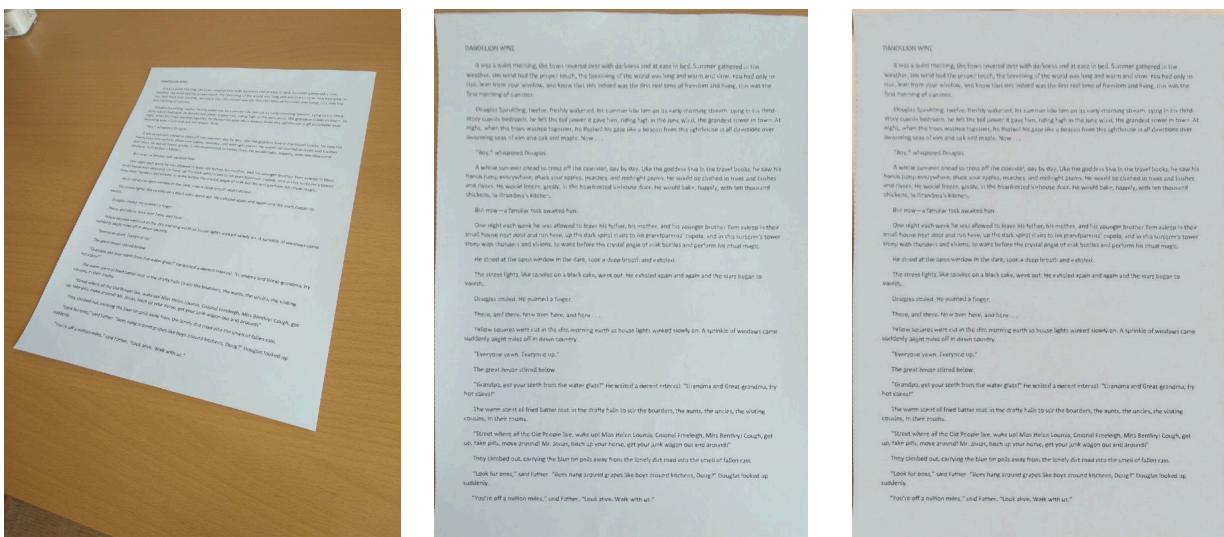


Рисунок Г.15 – Пример работы предложенных методов исправления искажений



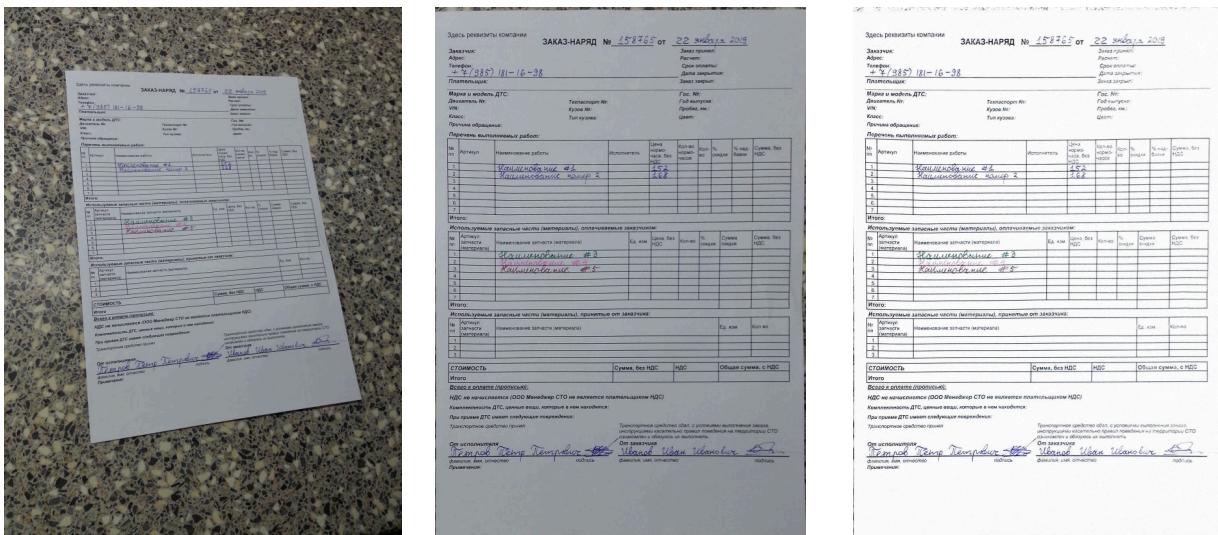
(a) начальное изображение (b) после коррекции (c) после дополнительной
перспективных искажений нормализации освещенности

Рисунок Г.16 – Пример работы предложенных методов исправления искажений



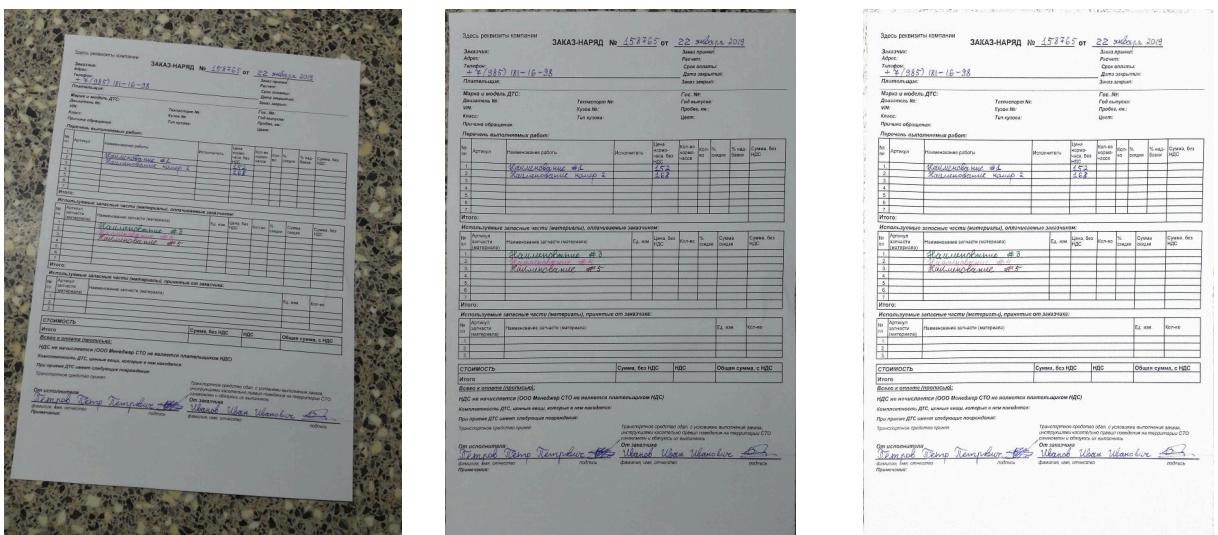
(a) начальное изображение (b) после коррекции (c) после дополнительной
перспективных искажений нормализации освещенности

Рисунок Г.17 – Пример работы предложенных методов исправления искажений



(a) начальное изображение (b) после коррекции (c) после дополнительной
перспективных искажений нормализации освещенности

Рисунок Г.18 – Пример работы предложенных методов исправления искажений



(a) начальное изображение (b) после коррекции (c) после дополнительной
перспективных искажений нормализации освещенности

Рисунок Г.19 – Пример работы предложенных методов исправления искажений