Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное

образовательное учреждение высшего образования

«Комсомольский-на-Амуре государственный университет»

Кафедра «Прикладная математика»

(наименование кафедры прикрепления)

**РЕФЕРАТ по НАУЧНЫМ ИССЛЕДОВАНИЯМ**

за 1 полугодие 4 года обучения

Направление подготовки 09.06.01 Информатика и вычислительная техника

(код и наименование направления подготовки)

Направленность подготовки 05.13.18 Математическое моделирование, численные методы и комплексы программ

(код и наименование направленности подготовки)

Аспирант Столяров Виталий Яварович

(ФИО аспиранта)

Оценка за научно-исследовательскую деятельность \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(оценка дифференцированного зачета)

Оценка за подготовку научно-квалификационной работы (диссертации) на соискание ученой степени кандидата наук \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(оценка дифференцированного зачета)

Окончательная проверка реферата проведена \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(дата проверки реферата и подпись научного руководителя)

Комсомольск-на-Амуре, 2021

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Научные исследования** | Перечень работ блока НИ | Перечень работ  каждого блока НИ | | Размещение материала, подтверждающего выполнение работ НИ |
| Научно исследовательская деятельность | Участие в конференциях, симпозиумах, семинарах, и т.д., 2 | | Приложение А |
|  |  | |
| Подготовка научно - квалификационной  работы (диссертации) на соискание  ученой степени  кандидата наук | Подготовка и оформление рукописи научно-квалификационной работы (диссертации), 50% | | Приложение Б |

**Приложение А**

(обязательное)

Опубликовано статей (РИНЦ, ВАК)

**УНИВЕРСАЛЬНЫЙ ПРОГРАММНЫЙ КОМПЛЕКС ПОДГОТОВКИ ИЗОБРАЖЕНИЙ ИДЕНТИФИКАТОРОВ ПЕРЕД РАСПОЗНАВАНИЕМ**

**Universal software package for preparing images of identifiers before recognition.**

**Stolyarov Vitalij Yavarovich – postgraduate of Komsomolsk on Amure State University (Russia, Komsomolsk-on-Amure); phone 8(984)1758172, e-mail: stolyarov.vitalij@list.ru.**

**Столяров Виталий Яварович** – аспирант Комсомольского-на-Амуре государственного университета (Россия, Комсомольск-на-Амуре); тел. 8(984)1758172. E-mail: stolyarov.vitalij@list.ru.

**Аннотация.** В настоящее время цифровая обработка изображений является одним из важнейших направлений применения современной вычислительной техники. На сегодняшний день известны следующие задачи обработки изображений: фильтрация и улучшение визуального восприятия, восстановление отсутствующих участков, обнаружение объектов и их идентификация, оценка геометрических трансформаций и совмещение изображений, оценка параметров изображений (задачи измерения различных характеристик изображения или их отдельных элементов: вероятностные характеристики изображения, положение и размеры объектов и т.д.), сжатие изображений. Проблемы распознавания изображений кроме классической задачи распознавания фигур заданной формы на изображении ставят новые задачи распознавания линий и углов на изображении, распознавания края изображения. В данной исследовательской работе автор решал проблему локализации текста на изображении при различных углах поворота. Для распознавания текста на изображении необходимо на начальном этапе локализовать область изображения, содержащую текстовые символы, затем определить угол отклонения. Задача распознавания актуальна в самых разных областях: при анализе телевизионных трансляций, поиске информации в медиабазах, административном обслуживании в банках, страховых компаниях и т.д. В статье описывается оптимизированная автором методика обнаружения текстовых областей и выделения символов на изображениях, снятых с металлических поверхностей, на которых лазерной гравировкой нанесён некоторый текст. Разработан специализированный алгоритм обработки (очистки) изображения, получаемого с USB-камеры. Подход основан на предварительном повороте изображения при помощи преобразования Хафа и применении методов математической морфологии. Цель работы заключалась в создании программного компонента, пригодного для распознавания идентификационных номеров, выполненных лазерной гравировкой на металлических деталях. В результате исследования найдено применение преобразованию Хафа для вращения изображения и получен оптимизированный алгоритм предварительной обработки текстовых изображений, снятых с металлических поверхностей с помощью USB-устройства. Практическая значимость данного алгоритма обуславливается использованием в программном обеспечении, разрабатываемом для Комсомольского-на-Амуре авиастроительного завода, в качестве программного компонента предварительной обработки изображений перед распознаванием.

**Annotation**. Currently, digital image processing is one of the most important areas of application of modern computer technology. The following image processing tasks are known today: filtering and improving visual perception, restoring missing areas, detecting objects and identifying them, evaluating geometric transformations and combining images, evaluating image parameters (tasks of measuring various image characteristics or their individual elements: probabilistic characteristics of the image, position and size of objects, etc.), image compression. The problems of image recognition, in addition to the classical problem of recognizing figures of a given shape in an image, pose new problems of recognizing lines and angles in an image, recognizing the edges of an image. In this research work, the author solved the problem of localizing text in an image at various angles of rotation. To recognize text in the image, it is necessary at the initial stage to localize the image area containing text characters, then determine the angle of deviation. The recognition problem is relevant in a variety of areas: in the analysis of television broadcasts, the search for information in media outlets, administrative services in banks, insurance companies, etc. The article describes a technique optimized by the author for detecting text areas and highlighting characters in images shot from metal surfaces on which some text is laser-engraved. A specialized algorithm for processing (cleaning) an image received from a USB camera has been developed. The approach is based on the preliminary rotation of the image using the Hough transform and the application of mathematical morphology methods. The purpose of the work was to create a software component suitable for recognition of identification numbers made by laser engraving on metal parts. As a result of the study, the application of the Hough transform for image rotation has been found and an optimized algorithm for the preliminary processing of text images taken from metal surfaces using a USB device has been obtained. The practical significance of this algorithm is determined by the use in the software developed for the Komsomolsk-on-Amur aircraft factory as a software component for pre-processing images before recognition.

**Ключевые слова:** обработка изображений перед распознаванием, математическая морфология, лазерная гравировка, преобразование Хафа, динамическое определения угла поворота изображения.

**Keywords**: image processing before recognition, mathematical morphology, laser engraving, Hough transform, dynamic determination of the angle of rotation of the image.

УДК 004.421

**Введение**

На предприятиях с каждым годом возрастает потребность в прикладных программах, способных считывать и распознавать идентификаторы на стандартных текстовых бирках, надписях, маркировках и шильдиках деталей, инструмента и блоков используемого при производстве оборудования. В силу того, что эти объекты подвергаются активной эксплуатации, качество изображения надписей на них, снимаемых для распознавания в компьютерных системах, имеет далеко не превосходное качество. Среди зарубежных учёных, которые внесли неоценимый вклад в развитие анализа изображений, стоит отметить Р. Дериша, Б. Хорна, К. Шапиро, Г. Финлейсона, а среди отечественных – В. Л. Арлазарова, М. М. Бонгарда, Ю. В. Визильтера, С. Ю. Желтова, Ю. И. Журавлева, Д. С. Лебедева, Б. М. Миллера, В. А. Сойфера и П. А. Чочи.

В реальном производстве большая доля надписей и маркировок выполняется лазерной гравировкой на металле. Угол поворота, тон, цвет, наличие шероховатости металлической поверхности затрудняют выделение области текстовых маркировок при их компьютерном считывании. Отсюда возникает потребность в разработке (усовершенствовании) методик предварительной обработки изображения перед передачей его в стандартный OСR (модуль распознавания текста) в автоматизированных системах машинного зрения. Одним из главных шагов предварительной обработки изображения является локализация области изображения, содержащей текст.

В связи с этим с каждым днём всё больше и больше возрастает интерес к подходам, используемым при распознавании рукописных, печатных и выполненных лазерной гравировкой текстам [1–8], существующие методы и алгоритмы регулярно пересматриваются и улучшаются с точки зрения решения задач поиска, распознавания и выявления смысла разнообразной текстовой информации на изображениях и видеозаписях [9–14; 15–17; 1]. В качестве решения к задаче детектирования местоположения текста используется большое число различных алгоритмов, которые базируются на использовании локальных особенностей изображений. Такими алгоритмами являются алгоритмы SWT [18] и MSER [19]. Методика, предложенная автором, отличается от подходов, используемых в алгоритмах SWT и MSER. Общая схема предложенной методики представлена на рис. 1.

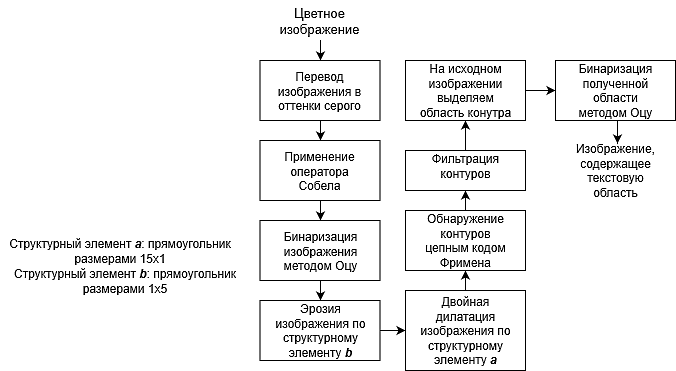
****

Рис. 1. Схема предложенной методики

В качестве входных параметров метода рассматривается цифровое дискретное изображение размером *M* на *N* точек, а элементы такого изображения будем называть пикселями. Важно заметить, что алгоритм работает исключительно с полутоновыми изображениями, поэтому необходимо на первом шаге преобразовать исходное цветное изображение в полутоновое.

Операции распознавания на изображениях определённых объектов, как правило, предваряются обработкой изображений для создания условий, повышающих эффективность и качество выделения и распознавания искомых или изучаемых объектов. При выделении текстовой области на изображении вышеописанным подходом было выяснено, что при повороте детали на угол более 20-25º возникают проблемы выделения текстовой области, обусловленные тем, что предложенный автором алгоритм, основанный на методах математической морфологии, не является изотропным, из-за того, что дифференцирование проводится только в горизонтальном направлении (см. рис. 2).



Рис. 2. Иллюстрация недостатка выделения текстовой области методами   
математической морфологии при повороте бирки на угол более 20º

|  |
| --- |
| рис 2 |
| Рис. 3. Задание прямой на плоскости  параметрами ρ и ϴ |

Целью данной работы является построение программного модуля для предварительной обработки изображений перед распознаванием, основанного на предложенном автором методе математической морфологии, и в случае отклонения детали на изображении свыше 20º автор предлагает методику, которая основывается на предварительном анализе областей изображений, построении пространства Хафа, на основе которого определяется угол отклонения бирки с лазерной гравировкой от горизонтального положения на изображении.

**Определение угла поворота**

***Общая идея преобразования Хафа для обнаружения прямых линий на изображении бирки, выполненной лазерной гравировкой на металле***. Основная идея преобразования Хафа для поиска прямых состоит в том, что любая точка бинарного изображения может принадлежать некоторому множеству прямых линий [9]. Возьмём точку  из заданного множества *n* точек и рассмотрим общее уравнение прямой на плоскости *xy* в форме с угловым коэффициентом . Однако удобнее представить прямую с помощью двух других параметров ρи ϴ*.* Параметр ρ – это длина перпендикуляра, опущенного на прямую из начала координат, а ϴ *–* это угол между данным перпендикуляром и осью *x* (см. рис. 3)*.*

Плоскость (ρ, ϴ)иногда называют пространством Хафа для набора прямых в двумерном случае, или фазовым пространством [20].

Очевидно, что через точку *A* с координатами  проходит бесконечно много прямых, которые удовлетворяют уравнению  при различных значениях *a* и *b.* Все прямые, проходящие через точку *A*, соответствуют следующему уравнению [21]:

|  |  |
| --- | --- |
| . | (1) |

Привлекательность преобразования Хафа с точки зрения вычислений проистекает из возможности разбиения пространства параметров ρϴна так называемые «ячейки накопления». В ячейке с координатами (*p*, *q*)накапливается значение *A*(*p*, *q*)для прямоугольника в пространстве параметров, соответствующего точке . Исходные значения во всех ячейках накопления равны нулю. Далее для каждой точки из заданного множества точек в плоскости *xy* принимаем параметр ϴ равным поочередно каждому разрешённому дискретному значению на оси ϴ и находим соответствующее ему значение ρ путём решения уравнения (1). После этого найденное значение ρ округляется до ближайшего разрешённого дискретного значения на оси ρ. Если выбор значения приводит к решению , увеличиваем накопленное значение в соответствующей ячейке (происходит процедура голосования): *A*(*p*, *q*)*= A*(*p*, *q*)*+*1*.* После выполнения описанной процедуры для всех исследуемых точекзаписанное в ячейке (*p*, *q*)значение *A*(*p*, *q*)*= P* говорит о том, что в плоскости *xy* имеется *P* точек, которые лежат на прямой . Точность попадания точек на эту прямую определяется размерами ячеек накопления на плоскости (ρ, ϴ).Вычислительная сложность данного метода линейна по *n* числу точек в заданном множестве на плоскости *xy.*

***Описание методики оптимизации применения преобразования Хафа к изображениям бирки, выполненной лазерной гравировкой на металле.*** Для того чтобы снизить время выполнения алгоритма, автор предлагает предварительно анализировать область на количество содержащихся в ней пикселей объекта. Предполагая, что цвет пикселя фона равен нулю, а цвет пикселя объекта равен единице, выведем следующую методику анализа области.

Введём фильтр-маску *w*(*x*, *y*)размерами *m*×*n*,все значения данной маски равны единице. Далее необходимо «просканировать» введённой маской исходное изображение. Сканирование заключается в поэлементном умножении области изображения и маски, если сумма произведений элементов менее 60 % от площади маски, то данная область не является интересующей и к ней не применяется преобразование Хафа. Автор считает оптимальным параметры размера маски принимать не более 7×7 за счёт узких областей вокруг текстовой области (см. рис. 2 слева). Выведем формулу для описанного сканирования областей:

|  |  |
| --- | --- |
|  | (2) |

где *a =* (*m –* 1)/2, *b =* (*m* – 1)/2.

Используем приведённую выше методику сканирования областей для построения пространства Хафа изображения, показанного на рис. 2 слева. Результат проиллюстрирован на рис. 4, где по горизонтали параметр ϴ, по вертикали – ρ. Области изображения с ярко выраженными всплесками говорят о том, что в соответствующих полярных точках находятся самые длинные прямые исходного изображения.

ϴ

ρ

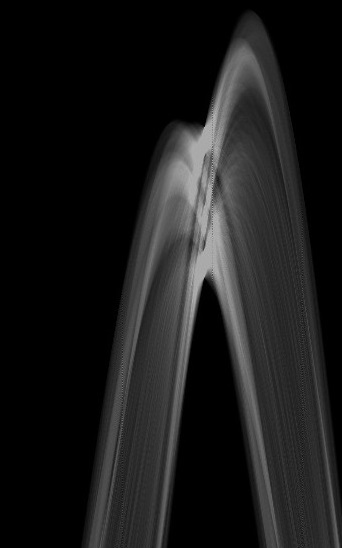


Рис. 4. Пространство Хафа   
для исходного изображения

**Описание методики получения угла поворота с использованием преобразования Хафа к изображениям бирки, выполненной лазерной гравировкой на металле**

Применительно к изображениям бирок с лазерной гравировкой, автором предлагается следующий прием предварительной обработки изображений для снижения угла отклонения изображения бирки, выполненной лазерной гравировкой на металле от горизонтальной прямой:

-все изображения бирок данного типа обладают широкой окантовкой вокруг текстовой области (рис. 5 слева). Пользуясь этим свойством, можно определить, что самая длинная прямая (в пространстве Хафа имеет самое большое число голосов), содержащая пиксели объекта будет проходить по или внутри линии окантовки. В качестве примера на рисунке 4 справа отображена самая длинная прямая обнаруженная алгоритмом Хафа (белая линия проходящая сквозь окантовку). Зная параметр *ϴ* этой прямой в дальнейшем можно применить аффинное преобразование «поворот».

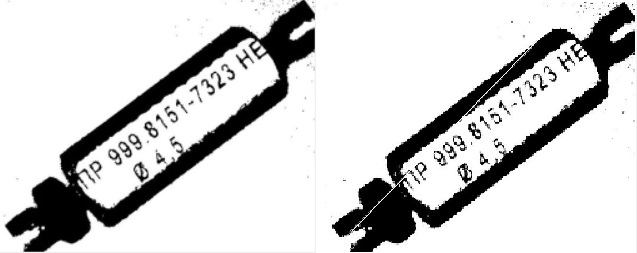


Рис.5. Слева – исходное изображение, справа – найденная прямая с помощью преобразования Хафа

Полный алгоритм работы программного модуля, реализующий предлагаемую методику поиска угла отклонения изображения бирки, выполненной лазерной гравировкой на металле с предварительным анализом областей, состоит из следующих шагов:

1. Съем изображения;
2. Выполнить процесс бинаризации с адаптивным порогом;
3. Произвести анализ областей используя формулу (2)
4. Выполнить цикл по всем пикселям области, в котором:
   1. Если пиксель черного цвета, тогда выполнить процедуру голосования, для  *0º ≤ ϴ ≤180º* согласно уравнению (1);
5. Определить в пространстве Хафа точку с максимальным числом голосов;
6. По параметру *ϴ* найденной точке на четвертом шаге, выполнить аффинное преобразование «поворот» на исходном изображении.

Результат решения проблемы, показанной на рис.1 представлен на рис. 6.

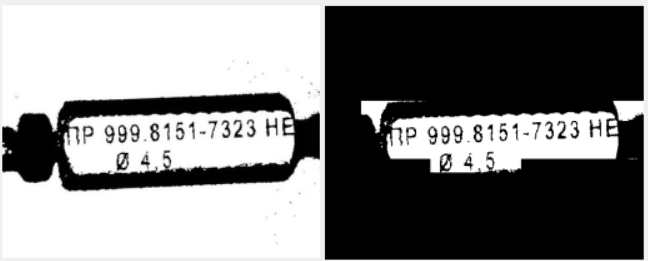


Рис. 6. Результат аффиного преобразования «поворот» на угол ϴ.

После поворота изображения можно приступать к предложенной автором методики локализации текста способами математической морфологии.

**Локализация текста**

В контексте машинного зрения морфология относится к описанию свойств формы областей на изображении. Операции математической морфологии изначально были определены как операции над множествами, но скоро выяснилось, что они также полезны в задачах обработки множества точек в двумерном пространстве.

Множествами в математической морфологии представляются объекты на изображении. Легко заметить, что множество всех фоновых пикселов бинарного изображения является одним из вариантов его полного описания [14].

Входными данными для аппарата математической морфологии являются два изображения: обрабатываемое и специальное, зависящее от вида операции и решаемой задачи. Такое специальное изображения принято называть примитивом или структурным элементом. В нашем случае мы задаем два структурных элемента:

1. Структурный элемент b для эрозии. В результате применения операции эрозии все объекты, меньшие чем структурный элемент, стираются, объекты, соединённые тонкими линиями становятся разъединёнными и размеры всех объектов уменьшаются. В нашем случае эрозия осуществляется в вертикальном по не высокому шаблону, что позволяет очистить изображение от импульсного шума, и более точно локализовать область по высоте символов идентификатора.

2. Структурный элемент a для дилатации. Дилатация увеличивает область изображения, расширяя его пиксели и тем самым способствуя объединению областей изображения, которые были разделены шумом и др. В нашем случае область текста размывается, это пригодится для выделения контура текстовой области.

Разработанная морфологическая операция является модификацией операции размыкания, которая в общем случае описывается следующим образом [6]:

,

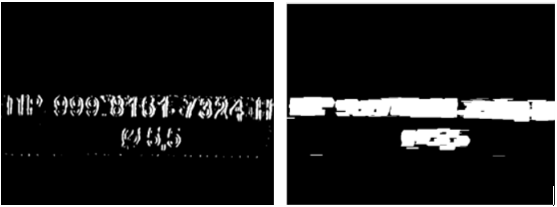
где *А* – исходное изображение, b – структурный элемент, ⊝ - операция эрозия, ⨁ - операция дилатация.

Предложенная автором формула имеет следующий вид:

, (3)

где *А* – исходное изображение, b – структурный элемент, применяемый при операции «эрозия», a – структурный элемент применяемый для операции «дилатация».

То есть – автор предлагает применять двойную дилатацию для улучшения качества размытия, в результате чего получается область, которая является фоном для текста (рисунок 7). На рисунке 7 слева показано изображение, полученное в результате применения дифференциального оператора Собеля, и последующей его бинаризацией, справа – результат применения выведенной автором формулы к изображению слева.

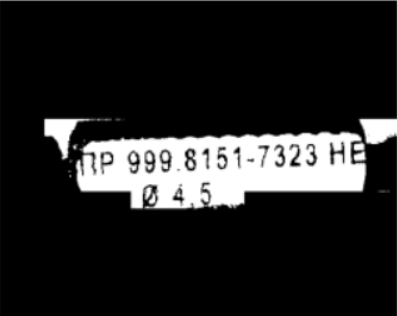
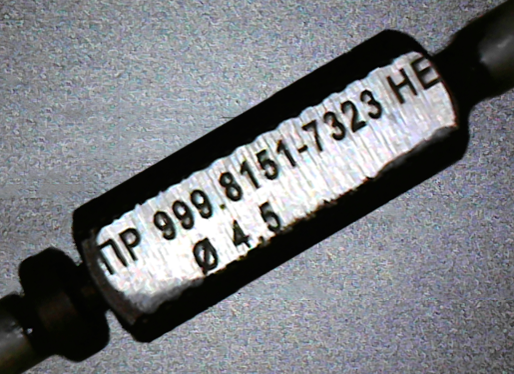


исходное обработанное

Рис. 7– Результат применения формулы (5) к предварительно обработанному изображению оператором Собеля и методом Оцу.

Далее над изображением, полученным в результате применения формулы (3), необходимо провести контурный анализ. Контурный анализ производим цепным кодом Фримена [12] с аппроксимацией Teh-Chin [12]. Обнаруженные контуры соединяются, накладываются на исходное изображение, и выполняется бинаризация полученной тестовой области.

Результат работы полного программного компонента представлен на рисунке 8.



исходное обработанное

**Рис. 8 – Финальный результат работы программы**

На рисунке 8 справа представлено исходное изображение, слева – то же изображение, пропущенное через разработанный автором программный компонент, в котором сначала было применено преобразование Хафа с предварительным анализом области, а далее полученный результат прошел через фильтр основанный на морфологической обработке изображений.

**РЕЗУЛЬТАТЫ**

Работа, предложенной автором методики по обнаружению текстовых областей с предварительным поворотом изображения детали (если трубуется) проверялась на реальных изображения бирки с лазерной гравировкой (подобные рис. 8 слева). Автор предлагает сравнить предложенную методику, основанную на использованиии методов математической морфологии с известными на сегодняшний день алгоритмами SWT и MSER, основная идея работы которых заключается в том, что буквы и цифры на изображении, как правило, имеют постоянную толщину штриха [10].

Для сравнения работы алгоритмов для локализации текста используются следующие метрики [10]:

1. Точность локализации. Эта метрика отражает вероятность того, что выделенные текстовые области являются корректными;
2. Полнота локализации. Эта метрика отражает вероятность того, что корректныеmтекстовые области были выделены;
3. Время работы алгоритмов.

Результаты сравнения работы алгоритмов представлены в таблице 1.

**Табл. 1. Результаты сравнения работы алгоритмов**

|  |  |  |  |
| --- | --- | --- | --- |
| Метод | Точность локализации, % | Полнота локализации, % | Время работы с одним изображением, с. |
| SWT | 91 | 83 | 0.23 |
| MSER | 93 | 88 | 0.21 |
| **Методика, предложенная автором** | **96** | **85** | **0.17** |

Как видно из табл. 1, предложенная автором методика показывает высокую скорость в сравнении с другими известными алгоритмами. Не самый лучший показатель полноты локализации, обусловлен тем, что предложенная автором методика не изотропна, т.к. применение дифференциального оператора к изображению осуществляется только в горизонтальном направлении, но этот недостаток устраняется вышеописанным предварительным анализом областей с последующим применением алгоритма Хафа.

**Заключение**

Основная цель работы была достигнута. На практике был реализован программный компонент с предложенной автором методикой, оптимизированный алгоритм предварительной обработки текстовых изображений, снятых с металлических поверхностей с помощью USB-устройств. Программный компонент продемонстрировал высокое качество выделения текстовых групп из изображений, что позволило, в дальнейшем, при передаче этих изображений в стандартные OCR системы получить 100% распознавание снятых текстовых надписей.

**ЛИТЕРАТУРА**

1. Кузьмицкий, Н. Н. Обнаружение фрагментов текста на изображениях реальных сцен на базе свёрточной нейросетевой модели / Н. Н. Кузьмицкий // Информатика. – 2015. – № 2(46). – С. 12-21.

2. Yin, X.-C. Multi-orientation scene text detection with adaptive clustering / X.-C. Yin, W.-Y. Pei, J. Zhang,   
H.-W. Hao // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2015. – Vol. 37, Issue 9. – P. 1930-1937. – DOI: 10.1109/TPAMI.2014.2388210.

3. Zuo, Z.-Y. Multi-strategy tracking based text detection in scene videos / Z.-Y. Zuo, S. Tian, X.-C. Yin // 13th International Conference on Document Analysis and Recognition (ICDAR). – 2015. – P. 66-70. – DOI: 10.1109/ICDAR.2015.7333727.

4. Koo, H. I. Scene text detection via connected component clustering and nontext filtering / H. I. Koo, D. H. Kim // IEEE Transactions on Image Processing. – 2013. – Vol. 22, Issue 6. – P. 2296-2305. – DOI: 10.1109/TIP.2013.2249082.

5. Болотова, Ю. А. Распознавание автомобильных номеров на основе метода связных компонент и иерархической временной сети / Ю. А. Болотова, В. Г. Спицын, М. Н. Рудомёткина // Компьютерная оптика. – 2015. – Т. 39, № 2. – С. 275-280. – DOI: 10.18287/0134-2452-2015-39-2-275-280.

6. Столяров, В. Я. Оценка эффективности применения контурной обработки к изображениям с лазерной гравировкой / В. Я. Столяров, В. А. Тихомиров // Актуальные научные исследования в современном мире: сборник Междунар. науч.-практ. конф., София, Болгария, 2018 (World of Sciens). – С. 84-88.

7. Epshtein B., Ofek E., Wexler Y. Detecting text in natural scenes with stroke witdh transform // Proceedings of International Conference on Computer Vision and Pattern Recognition, 2012. – P. 2963-2970.

8. Neumann, L. A method for text localization and recognition in real-world images / L. Neumann, J. Matas // 10th Asian conference on Computer vision, 2012. – P. 770-783.

9. Гонсалес, Р Цифровая обработка изображений: учеб. пособие / Р. Гонсалес, Р. Вудс; пер. с англ.:   
Л. И. Рубанова, П. А. Чочиа; под ред. П. А. Чочиа. – М.: Техносфера, 2012. – 1104 с.

10. Otsu, N. A Threshold Selection Method from Gray-Level Histograms / N. Otsu // IEEE Transactions on Systems, Man, and Cybernetics. – 1979. – Vol. 9, No. 1. – Pp. 62-66.

11. Визильтер, Ю. А. Обработка и анализ цифровых изображений с примерами на LabView и Imaq Vision: учеб. пособие для вузов / Ю. А. Визильтер, В. А. Князь, С. Ю. Желтов; под ред. В. А. Мовчан. – М.: ДМК-Пресс, 2016. – 464 с.

12. Фурман, Я. А. Введение в контурный анализ; приложение к обработке изображений и сигналов: учеб. пособие для вузов / Я. А. Фурман, А. К. Передреев, А. В. Кревецкий; под ред. Я. А. Фурман. – М.: Физматлит, 2003. – 592 с.

13. Красильников, Н. А. Цифровая обработка 2D- и 3D-изображений: учеб. пособие для вузов / Н. А. Красильников. – СПб.: BHV, 2011. – 608 с.

14. Ghosh P., Deguchi K. Mathematics of shape description: a morphological approach to image processing and computer graphics [Wiley, 2008].

15. Jung, K. Text information extraction in images and video: a survey / K. Jung, K. Kim, A. K. Jain // Pattern Recognition. – 2004. – Vol 5. – P. 977-997.

16. Шапиро, Л. Компьютерное зрение / Л. Шапиро, Дж. Стокман; пер. с англ. – М.: Бином, 2006. – 752 с.

17. Андрианов, А. И. Локализация текста на изображениях сложных графических сцен [Электронный ресурс] / А. И. Андрианов // Современные проблемы науки и образования. – 2013. – № 3. – Режим доступа: https://www.science-education.ru/ru/article/view?id=9311 (дата обращения: 12.05.2020).

18. Epshtein, B. Detecting text in natural scenes with stroke witdh transform / B. Epshtein, E. Ofek, Y. Wexler // Proceedings of International Conference on Computer Vision and Pattern Recognition. – 2012. – P. 2963-2970.

19 Neumann, L. A method for text localization and recognition in real-world images / L. Neumann, J. Matas // 10th Asian conference on Computer vision. – 2012. – P. 770-783.

20. Кэлер, А. Изучаем OpenCv 3: учеб. пособие / А. Кэлер, Г. Брэдски; пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2017. – 826 с.

21. Дегтярева, А. Преобразование Хафа [Электронный ресурс] / А. Дегтярева, В. Вежневец // Компьютерная графика и мультимедиа. – 2003. – Вып. № 1(1). – Режим доступа: https://cgm.computergraphics.ru/content/view/36 (дата обращения: 12.05.2020).

**Приложение Б**

(обязательное)

Подготовка и оформление рукописи научно-квалификационной работы

**Актуальность темы.** В настоящее время цифровая обработка изображений является одним из важнейших направлений применения современной вычислительной техники. На сегодняшний день известны следующие задачи обработки изображений: фильтрация и улучшение визуального восприятия, восстановление отсутствующих участков, обнаружение объектов и их идентификация, оценка геометрических трансформаций и совмещение изображений, оценка параметров изображений (задачи измерения различных характеристик изображения или их отдельных элементов: вероятностные характеристики изображения, положение и размеры объектов и т. д.), сжатие изображений. Проблемы распознавания изображений кроме классической задачи распознавания фигур заданной формы на изображении ставят новые задачи распознавания линий и углов на изображении, распознавания края изображения.

Типичными задачами компьютерного зрения являются обнаружение, отслеживание и классификация объектов. Цель обнаружения – выявить на изображении объект. Для достижения поставленной цели мы можем использовать границы (контуры), особые точки (например, углы), информацию о цвете и так далее. Отслеживание применяется в работе с камерами наружного наблюдения. При этом также можно воспользоваться контурами, особыми точками и информацией о цвете, а можно вычесть фон из текущего кадра (при условии, что камера статична). Чтобы выполнить классификацию, нужно распознать обнаруженный объект. Для распознавания можно выполнить попиксельное сравнение с шаблоном, сравнить контуры или особые точки, осуществить поиск по обученному классификатору (например, с помощью каскадов Хаара) и др. В последнее время для распознавания объектов все чаще используются глубокие сверточные нейронные сети. С компьютерным зрением тесно связана еще одна технология – обработка изображения. В действительности, прежде чем пытаться обнаружить объекты на изображении, необходимо выполнить предварительную обработку. В частности, в подавляющем количестве случаев нужно изменить размеры изображения и выполнить выравнивание гистограммы. Кроме того, важным моментом является процесс сглаживания изображения, чтобы избавиться от цифрового шума, который возникает в случае использования дешевых моделей камер, а также при различных специфических режимах съемки. Многие алгоритмы очень чувствительны к шумам, и их наличие может привести, например, к неверно обнаруженным границам объекта или ложным особым точкам.

Ученные достаточно усердно, в течении многих лет, занимаются исследованиями в области компьютерного зрения, но до сих пор так и не приблизились к способностям человека распознавать образы. Причиной этому можно выделить ряд проблем.

Первая проблема скрывается в том, что цвет пикселя в цифровом изображении задается числовыми значениями. Причем этих чисел в полноцветных изображениях три: первое число определяет количество красного цвета, второе – зеленого и третье – синего (это есть цветовая модель RGB). Вторая проблема заключается в том, что пиксель имеет квадратную форму, а в видеорядах встречается прямоугольная форма элементов изображения. При масштабировании таких элементов, на определенном этапе, мы не сможем определить, что за фигура перед нами. Третья проблема возникает при сжатии изображения, которое используется для уменьшения размера файла. Например, в случае сохранения в формате JPG появляются артефакты сжатия, с которыми нужно как-то справляться.

Таким образом, в основе анализа и обработки изображений как прикладной технико-математической дисциплины лежат задачи математической формализации яркостно-геометрических моделей изображений (объектов) и на основе этих формализованных моделей строятся процедуры (методы) анализа наблюдаемых изображений. Примером такого рода моделей наиболее общего характера является достаточно изученный класс разложений (ряды Фурье, обобщенные ряды Фурье и другие). Данные методы позволяют выявлять внутреннюю организацию математических объектов, исследовать критические свойства, регулярным образом порождать различные наборы характеристик. В области анализа изображений стремление к созданию в достаточной мере общих и в то же время предметно адекватных моделей и процедур для различных прикладных задач привело к появлению целого ряда, на первый взгляд, абсолютно различных методов современного компьютерного зрения, таких как согласованная фильтрация и корреляционное обнаружение, пространственно-частотные и частотные методы на основе двумерного преобразования Фурье [5],[61] и вейвлет-анализа [38], морфологических подход Ю.П. Пытьева, математическая морфология Серра, метод «нормализации фона», преобразование Хафа [182] и обобщенное преобразование Хафа, структурно-лингвистический подход и большое количество других.

Значительный вклад в разработку методов и алгоритмов обработки изображений и машинного зрения применительно к обсуждаемым задачам внесли, как отечественные, так и зарубежные ученные: Ю.П.Пытьев [86,87], Л.П. Ярославский [110-111], Ю.И. Журавлев [60], Ж. Серра [150, 151], Д. Марра [76], Е. Дэвис [135-142], Д. Баллард [117, 118], А. Демпстер [145], Г.Шаффер [152], У. Гренадер [42], М.Павель [153], Дж. Ту [102], К. Фу [104] и многие другие. В течении последних десятилетий было создано множество хороших систем машинного зрения, в которых тех или иных сочетания используются вышеупомянутые парадигмы и подходы. Однако отсутствие единого математического формализма для описания яркостно-геометрической структуры изображений и соответствующей единой методики разработки алгоритмов анализа изображений до сих пор является серьезной проблемой, замедляющей и затрудняющей проектирование новых приложений и практических систем машинного зрения. Рассматриваемые в данной диссертационной работе морфологический подход к анализу изображений, математический аппарат проективных морфологических разложений позволяют раскрыть единую математическую природу большинства перечисленных подходов и таким образом обеспечивается максимально-возможная гибкость и вариативность предметно-ориентированного структурного анализа изображений.

Другой не менее важной проблемой современного анализа изображений есть необходимость учета вероятностной природы реальных изображений, учета априорной информации, обеспечение устойчивости методов анализа изображений относительно различного рода шумов и искажений. Предлагаемый в данной диссертационной работе вычислительный алгоритм, применяемых для параметрической идентификации геометрических элементов растрового изображения с предварительным анализом областей предназначен для описания геометрических аспектов (угла наклона и др.) структурного анализа изображения.

**Цель работы** состоит в разработке методики подготовки изображения с лазерной гравировкой перед распознаванием. Данная методика будет основываться на теории и методах морфологического анализа изображения с предварительным устранением шумов на изображении и проведением яркостно-геометрических процедур. Практическая направленность работы заключается в разработке структуры и методологии программно-алгоритмического обеспечения ряда прикладных систем анализа изображений.

Для достижения цели исследования решаются следующие **задачи**:

1. анализ существующих методов обработки и анализа изображений, исследование их критических свойств, включая проектирование новых модификаций и обобщение данных методов, повышающих их эффективность и расширяющих область их действия;
2. разработка математической модели системы снятия изображений, учитывающую описание штатных и нештатных ситуаций
3. разработка вычислительного метода, описывающего яркостно-геометрические аспекты анализа изображений;
4. разработка вычислительного метода, описывающего вероятностные аспекты анализа изображений;
5. разработка вычислительного метода, описывающего морфологические аспекты анализа изображений;
6. проектирование структуры программного обеспечения подготовки изображения перед распознаванием.

Объект исследования – система снятия изображений.

Предмет исследования – алгоритмы математической морфологии, вычислительные алгоритмы, применяемый для параметрической идентификации геометрических элементов растрового изображения, кратно масштабные иерархические структуры данных, дискретные дифференциальные операторы, вычисляющий приближённое значение градиента яркости изображения.

Методы исследования основаны на общей методологии  
математического моделирования исследуемых процессов, объектно-ориентированном программировании. При решении задач были  
использованы теория математической морфологии, теория множеств и универсальной алгебры, теория вероятности и теория графов. Применялись интегрированные среды для языков программирования: C# (Visual Studio), Python (PyCharm, Spyder) и пакеты прикладного программного обеспечения. Проведены экспериментальные апробации разработанных алгоритмов в различных предметных областях.

**Научная новизна работы**:

1. Предложена математическая модель локализации текста на основе методов морфологических разложений, опирающихся на однородные структурные модели, представляющие изображение в виде объединения независимых линейных проекций на образующие (структурные примитивы) из некоторого набора, причем коэффициенты линейной связи образующих и анализируемого образа составляют регулярный вектор морфологического разложения, характеризующий данный образ.
2. Предложен вычислительный метод применяемый для параметрической идентификации геометрических элементов растрового изображения, основанный на модификации преобразования Хафа.
3. Предложен вычислительный метод, работающий на основе модифицированного дискретного дифференциального оператора Собеля, вычисляющего приближённое значение градиента яркости изображения.
4. Предложен подход выяснения угла отклонения детали, основанный на разбиении изображения на пирамиды, с разработанным автором анализом каждой пирамиды.

Практическая значимость. Предложенные в работе алгоритмы могут  
быть использованы: для подготовки изображения перед распознаванием идентификаторов, выполненных лазерной гравировкой на металле, номерных знаков транспортных средств (ТС).

Апробация работы. Результаты работы докладывались и получили одобрение:

* на пятой научно – практической конференции молодых ученных и специалистов «Исследования и перспективные разработки в машиностроении», Комсомольск-на-Амуре, 2018г.
* на фестивале научно-технического творчества молодежи «Инноватор года 2018», Комсомольск-на-Амуре, 2018 г.
* на международной научно – практической конференции «Актуальные научные исследования в современном мире», София, 2018 г.
* на научно-технической конференции студентов и аспирантов ФГБОУ ВО «КнАГУ», Комсомольск-на-Амуре, 2019
* The 35th IBIMA conference, Seville, 2020

**Основные положения, выносимые на защиту**:

1. Математическая модель локализации текста на основе методов морфологических разложений;
2. Вычислительный метод применяемый для параметрической идентификации геометрических элементов растрового изображения;
3. Вычислительный метод, вычисляющий приближённое значение градиента яркости изображения;
4. Вычислительный метод, вычисляющий угол отклонения детали на изображении;
5. Программный модуль подготовки изображения перед распознаванием.

Достоверность полученных результатов, научных положений, выводов, изложенных в диссертации, подтверждается результатами экспериментов c обработанными изображениями, а также использованием информационных данных для натурных экспериментов, полученных из открытых источников и путем экспериментов в ФГБОУ ВО «КнАГУ».

Публикации.Результаты диссертационного исследования опубликованы  
в восьми научных работах, в числе которых две статьи, опубликованы в ведущих рецензируемых журналах, включенных в перечень ВАК, одна статья индексируется в базах Scopus и WebOfScience. Получено одно  
свидетельство о регистрации программ для ЭВМ.

Структура и объем диссертационной работы. Диссертация состоит  
из введения, четырех глав, заключения, списка сокращений и условных  
обозначений, списка литературы из ??? наименования, приложения. Полный  
объем составляет ??? страницы, ??? рисунков, ??? таблицы.

Первая глава носит обзорный характер и содержит анализ современного состояния исследований в области цифровой обработки изображений. Приведена классификация изображений. Определена специфика изображений с лазерной гравировкой как объекта исследований. Анализ обобщенной схемы этапов цифровой обработки изображений позволяет сделать вывод о большой значимости этапа предварительной обработки изображения при решении широкого круга задач, связанных с анализом изображений и распознаванием образов, возникающих при построении систем технического зрения, роботизированных комплексов, систем автоматизированного анализа и т.п.

Во второй главе описывается математическая модель разработанной методики предварительной обработки изображений, выполненных лазерной гравировкой

В третье главе…

В четвертой главе…

В заключении…

**Глава 1. Анализ методов и особенностей предварительной цифровой обработки полутоновых изображений**

# Классификация изображений. Особенности изображений с лазерной гравировкой

Несмотря на большое число способов регистрации изображений, задача всегда одинакова: сформировать цифровое изображение, основанное на данных, которые воспринимаются чувствительными компонентами. От большинства сенсоров поступает аналоговый сигнал в виде непрерывно меняющегося напряжения, амплитуда и форма которого связанны с регистрируемым физическим явлением. Чтобы получить цифровое изображение, необходимо преобразовать непрерывно поступающий сигнал в цифровую форму. Данная операция состоит из двух процессов – дискретизации и квантования.

Первоочередной принцип, который находится в основе дискретизации и квантования изображений показан на рисунке 1.1. В левом верхнем углу показано изображение-источник *f(x, y)*, которое мы хотим трансформировать в цифровую форму. В правом верхнем углу рисунка 1.1 – профиль вдоль линии сканирования между точками *А* и *В.* Изображение является непрерывным по координатам *х* и *у*, а также по амплитуде. Представление координат в виде конечного множества отсчетов называется дискретизацией, а представление амплитуды значениями из конечного множества – квантованием [15]. На рисунке 1.1 слева снизу представлена процедура дискретизации и квантования изображения.

Таким образом изображение определяется как двумерная функция *f(x, y)*, где *x* и *y* — координаты в пространстве (конкретно на плоскости) и значение *f* которой в любой точке, задаваемой парой координат *(x, y)*, называется интенсивностью или уровнем серого (яркостью) изображения в этой точке [15]. В случае, если величины *x*, *y* и *f* принимают конечное число дискретных значений, то идет речь о цифровом изображении. Цифровой обработкой изображений называется обработка цифровых изображений с помощью цифровых вычислительных машин (компьютеров). Важно отметить, то что цифровое изображение состоит из конечного числа элементов, каждый из которых расположен в конкретном месте и принимает определенное значение. Эти элементы называются элементами изображения или пикселями.

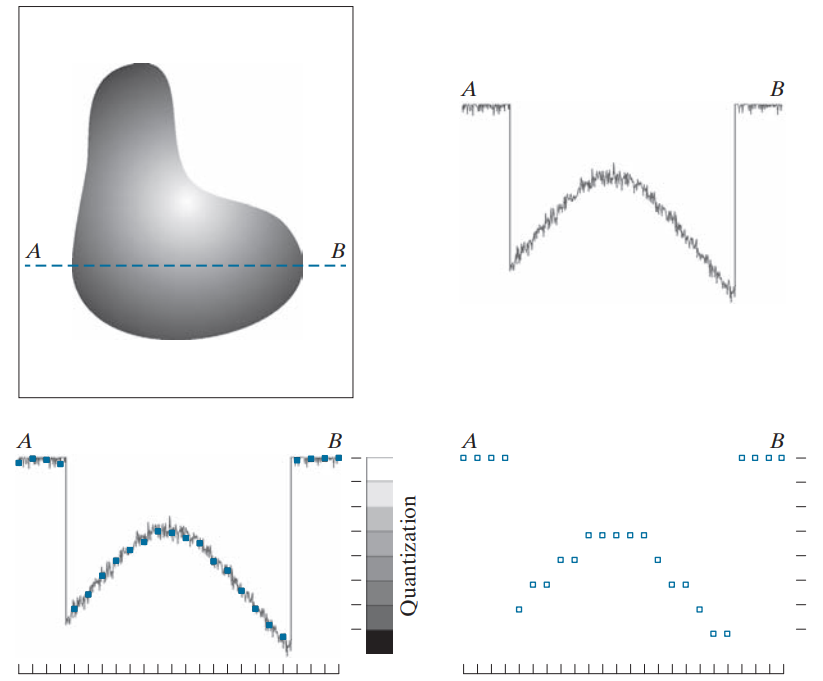


Рисунок 1.1 – Формирование цифрового изображения

Компьютерная графика – это область информатики, занимающаяся созданием, хранением и обработкой различных изображений (рисунков, чертежей, мультипликации) на компьютере [16]. На сегодняшний день существует четыре вида компьютерной графики: векторная, фрактальная, трехмерная. Векторная графика представляет изображение в виде совокупности крайне простых геометрических объектов. Такие объекты являются основой для построения изображения и называются примитивами. Примитивами могут быть отрезки, маленькие дуги, окружности, сплайны и т.д. Графика называется векторной потому, что набор примитивов, формирующие заданный графический объект, имеет название – вектор. Фрактальная графика основана на разделе математики – фрактальной геометрии. Термин фрактал ввел французский математик Бенуа Мандельброт. Этим термином он назвал геометрическую фигуру, которая состоит из частей, подобных целой фигуре. Трехмерная графика оперирует с объектами в трехмерном пространстве. Для построения изображения, которое выглядит как объемное, используется так называемое полигональное моделирование. Растровое изображение, представляет собой двумерный массив, элементы которого (пиксели) содержат информацию об уровне серого или цвете. Далее речь будет идти только о растровых изображениях, которые, в свою очередь, делятся на: бинарные, полутоновые, цветные индексированные, полноцветные [40].

Бинарные изображения, характеризуются только двумя значениями уровня яркости - 0 или 1, могут быть получены в результате обработки полутоновых, цветных индексированных или полноцветных изображений методами бинаризации с фиксированным или адаптивным порогом [15].

Черно-белые (бинарные) изображения намного легче получать, хранить и обрабатывать, чем изображения, в которых имеется много уровней яркости. Однако, в силу того, что в бинарных изображениях кодируется информация лишь о силуэте объекта, область их применения достаточно ограничена.

Одним из наиболее распространенных типов изображений, применяемых при исследованиях различного рода является полутоновое изображение, которое состоит из элементов, принимающих одно из значений интенсивности какого-либо цвета. В большинстве случаев используется глубина одного из основных цветов или уровня серого 8 бит на элемент изображения [15, 80].

Цветное индексированное изображение — такое изображение, цвет каждого элемента которого задаётся в специальной таблице — палитре. Таким образом, каждый элемент изображения (пиксел) представлен в памяти не реальным цветом, а неким условным индексом, который расшифровывается по таблице цветов (палитре) в реальные компоненты цвета, что позволяет для каждого такого элемента построить «истинный» цвет в заданной цветовой модели [58].

В отличие от цветных индексированных, элементы полноцветных изображений непосредственно хранят информацию о яркостях цветовых составляющих, однако их использование требует больших вычислительных затрат. Выбор типа изображения зависит от решаемой задачи, от того, насколько полно и без потерь нужная информация может быть представлена с заданной глубиной цвета и производится на этапе формулирования технических требований при создании любой системы обработки изображений.

Полутоновые изображения нашли наибольшее применение в биомедицинской диагностике [6,7,63,75], в основном, из-за специфики методов интроскопии, используемых для их получения. Далее будут рассматриваться только полутоновые изображения (шкала серого цвета). Методы, применяемые при обработке цветных и полутоновых изображений аналогичны. Отличие состоит в том, что для первых ощутимо возрастают затраты на вычисления - обработка производится по каждой из цветовых компонент отдельно с их последующим микшированием.

Специфика изображений с лазерной гравировкой в первую очередь связана со способами их получения. Гравировка – это нанесение рельефного изображения на изделие. Обычно для переноса эскиза на поверхности используются ручные или электрические граверы. Они позволяют добиться достаточно красивых и оригинальных результатов, но требуют больших затрат времени и огромного мастерства. Кроме того, в качестве основы могут использоваться далеко не все материалы, а лишь достаточно твердые, вроде дерева, металла, пластика или стекла, так как гравер снимает часть материала в месте соприкосновения и оставляет углубление.

При длительной эксплуатации детали с идентификатором, выполненным с помощью лазерной гравировки, возникают дополнительные артефакты в области нанесения идентификатора (поскольку эта область априори становится более чувствительной к воздействию в следствии применения лазерного гравера). Для зрительной системы человека такие артефакты не являются большой проблемой, и мы можем однозначно определить последовательность символов на идентификаторе. В компьютерной зрительной системе дела обстоят иначе. Алгоритмам распознавания достаточно тяжело чем является элемент на изображении или это символ «П» или «1 1». Помимо артефактов непосредственно на самом идентификаторе, дополнительные артефакты (в виде помех) возникают при регистрации изображения в вычислительном устройстве. Таким образом цифровое изображение состоит из самой детали и совокупности артефактов: первые артефакты – возникают в силу эксплуатации детали, минимизировать которые практически невозможно; вторые – в результате переноса изображения из реального мира в виртуальный, минимизировать которые возможно, используя более дорогие системы снятия изображений и реализацией дополнительных фильтров обработки и реконструкции изображений. Сущность технологического процесса работы с изображениями с лазерной гравировкой изображен на рисунке 1.2.

Система снятия и регистрации изображений

Деталь с идентификатором

**+**

Эксплуатационные артефакты

**+**

помехи

Фильтр обработки и реконструкции изображение

Зашумленное (с дефектами и артефактами) цифровое изображение детали с идентификатором

Очищенное изображение

Рисунок 1.2 – Технологический процесс обработки изображения с лазерной гравировкой

# Анализ групп основных методов предварительной обработки изображений

Обработка изображений не является одношаговым процессом [83], следовательно, основные этапы обработки изображения, проводимой с целью улучшения его восприятия модулем, отвечающим за распознавание, можно изобразить в виде структурной схемы, приведенной на рисунке 1.3 [79].

визуальная интерпретация, визуализация

Регистрация изображения

оцифровка, кодирование

предварительная обработка

восстановление изображения

сегментация, выделение признаков, классификация, распознавание, интерпретация

обработка изображения

несемантическая обработка

семантическая обработка

результат

повышение качества изображения

Рисунок 1.3 – Основные этапы обработки изображений для последующей визуализации и анализа

Процесс обработка изображений начинается с момента получения (регистрации) изобра­жений при помощи системы формирования изображений, за исключением тех ситуаций, при которых изображение было представлено в цифровой форме заранее. Наиболее важным на этом шаге является верный выбор системы формирования изображе­ния из светового потока и крайне точный подбор параметров фиксиро­вания интересующего признака объекта на изображении (освещение, диапа­зон длин волн и т.п). Оцифровка необходима для преобразования изображе­ния в цифровую форму, кодирование - требуется в случае передачи изобра­жений по каналам связи.

Блок предварительной обработки изображения чрезвычайно важен для выделения сложно раз­личимых деталей, или подчеркивания интересующих характеристик на ис­ходном изображении. Как правило, именно на этом этапе производятся самые важные работы с изображением: локальная фильтрация, геометрические и координатные преобразования для устранения искажений, внесенных при формировании изображений.

Восстановление изображений в своей основе опирается на математиче­ские или вероятностные модели искажений изображения. Возмущения в изо­бражении, вызванные расфокусировкой оптики, размытостью изображений, движением объекта, погрешностями в датчике, либо при передаче сигналов изображения могут корректироваться с помощью данных моделей, однако, требуют априорной информации о природе возмущения [31, 38].

Сегментация делит изображение на области постоянных признаков и неоднородностей. Полученные знания о геометрической форме объекта позволяют применять морфологические операторы для ее анализа и модификации, а дополнительные данные, как например, о среднем уровне яркости, площади, периметре объекта могут быть использованы при классификации объектов [16]. На базе различения классов объектов на изображении и выбора при­знаков выполнятся присвоение объекту уникального числового или тек­стового идентификатора (распознавание).

На выходе любого из показанных на рисунке 1.3 блоков, может производится процедура отображения результатов обработки.

Количество модулей представленной схемы, задействованных в системе обработки изображений, зависит от цели конкретной прикладной задачи. На­пример, в задаче распознавания текстовых идентификаторов, выполненных лазерной гравировкой. Все операции на этапе семантической обработки (рисунок 1.3) являются обязательными. В этом случае, главная задача обработки изображения состоит не в повышении его ка­чества, оцениваемого визуально, а в повышении качества для алгоритмов распознавания. Сохра­нение свойств исходного изображения при этом, как правило, желатель­но, но не всегда обязательно.

На рисунке 1.4 приведена классификация методов предварительной обработки изображений в за­висимости от области действия, используемых алгоритмов и размера обраба­тываемой области изображения. Переход в частотную область при реализации алгоритмов обработки, как правило, вызван необходимо­стью обрабатывать изображения больших размерностей [14, 65]. При таких ситуация оправдывается использование быстрых дискретных ортогональных преобра­зования (например, быстрое преобразование Фурье (БПФ)). Относительно размера обрабатываемой области изображения, в данной работе будут рас­сматриваться «окна» размером 3x3 и 5x5 пикселей, поскольку являются наиболее оптималь­ными с точки зрения вычислительных затрат [58, 82].

Методы предварительной обработки изображений

Область действия

Алгоритмы

Размер обрабатываемой области изображения

Координатная

Частотная

Линейный

Нелинейные

Поэлементные операторы

локальные операторы

глобальные операторы

Рисунок 1.4 – Методы предварительной обработки изображений

Линейная фильтрация нашла свое применение в ситуациях, где необходимо подавить шумы на изображении, для компенсации неравномерности чувствительности, для создания эффектов размытия изображений. Такая фильтрация также широко применяется в задачах выделения контуров на изображении, подчеркивания верхних пространст­венных частот [56,58]. Как и линейные фильтры с конечной импульсной характеристикой, нелинейные фильтры работают по скользящей маске (или скользящее окно). Но, при линейной фильтрации производится расчет ли­нейной комбинации отсчетов сигнала, а при нелинейной фильтрации производятся нелинейные преобразования отсчетов сигнала в определяемой маской фильтра окрестности элементов.

Предварительная обработка изображения как этап процедуры повыше­ния качества изображения состоит из большого набора элементарных алгорит­мов. Наиболее распространенные из них приведены на рисунке. 1.5 [15,82]:

Основные алгоритмы предварительной обработки изображений

Нелинейные преобразования

Фильтрация в частотной области (ДОП)

Линейные преобразования

Свертка со скользящими масками

Интерполяция

временное суммирование

Рисунок 1.5 – Основные алгоритмы предварительной обработки изображений

Одна из важнейших характеристик любого метода обработки изображений – это является ли он линейным или нелинейным. Линейный операторы крайне важны для обработки изображений, так как они опираются на значительную совокупность хорошо изученных теоретически и практических результатов. Нелинейные операторы исследованы в большей степени хуже, поэтому область их применения более узкая [15].

Интерполяция является базовым инструментом, который широко используется при решении задач такого рода, как увеличение и уменьшение изображений, их поворот и коррекция геометрических искажений [31]. Сущность интерполяции заключается в том, что это процесс, при котором уже имеющиеся данные используются для оценки значений в неизвестных точках.

Чтобы понять свертку со скользящими масками, важно прояснить близкородственную к ней концепцию – корреляция. Корреляция есть процесс движения маски фильтра по изображению и вычисление суммы произведений значений элементов маски и значений пикселей, на которые попадают соответствующие элементы маски, для всех точек изображения. Механизмы свертки такие же, но за исключением того, что предварительно маска фильтра поворачивается на 180º [15].

Геометрические преобразования [40,43,58], используемые для устранения геометрических искажений, вносимых, как правило, на этапе регистрации изображения (рисунок 1.6) должны обязательно присутствовать при реализации большинства процедур обработки изображений.

Геометрические преобразования изображений

пространственный сдвиг

отражение

масштабные преобразования

поворот

преобразование системы координат

Рисунок 1.6 – Геометрические преобразования изображений

Обычно, этап предварительной подготовки изображения состоит из нескольких операций, формирующих своеобразный «набор», причем выбор тех или иных процедур обусловлен как спецификой изображений, так и технически­ми требованиями к системе обработки изображений.

Обзор существующих специализированных программных средств для обработки изображений с текстовыми идентификаторами (приложение 1) показывает, что почти любой (за редким исключением) программный комплекс включает блок предварительной обработки, необходимый для улучшения изображе­ний, предназначенных для анализа системами распознавания [7,8,10]. Это подтверждает актуальность выбранного направле­ния исследования, связанного с совершенствованием алгоритмов предвари­тельной обработки изображений.

# Линейные дискретные преобразования полутоновых изображений

Если рассуждать с точки зрения реализации, методы обработки изображений подразделяются на два основных класса [38, 82]: локальные и глобальные. Главное отличие – в размере участка изображения, подвергаемого обработке. Выбор подхода зависит от конкретной задачи, которую решает исследователь. Преимущество методов при глобальной реализации заключается в простоте их исполнения и быстродействии. Локальные методы владеют более обширными функциональными возможностями, а конкретно, они умеют учитывать характеристики локальных областей, т.е. являются более гибкими в практической реализации [31,79].

Методы глобальной обработки изображений могут быть реализованы либо непосредственно по полю изображения, либо с помощью спектров [58].

Существует три основных области применения глобальных двумерных ортогональных преобразований для обработки изображений [82,84]:

* для выделения характеристик признаков изображения;
* для кодирования изображений, в том случае, когда ширина спектра уменьшается за счет отбрасывания или грубого квантования малых по величине коэффициентов преобразования;
* для сокращения размерности при выполнении вычислений;

К преобразованиям такого вида можно отнести преобразования Фурье, синусные, косинусные, вейвлет-преобразования, а также преобразования Карунена-Лоева, Уолша, Хаара и Адамара [14, 53, 54]. К сожалению, указанные преобразования не имеют в своем роде необходимой универсальности при решении широкого спектра задач, суть которых заключается в цифровой фильтрации.

Для того чтобы перенести исходное изображение в частотную область и его восстановления применяются дискретные прямое (ПДП) и обратное преобразования Фурье (ОДП) или другие дискретные ортогональные преобразования. Фильтрация в частотной области состоит в умножении Фурье-образа изображения на частотную характеристику фильтра [79].

В качестве альтернативы дискретным ортогональным преобразованиям Фурье для решения задач предварительной обработки изображений, устранения шумов и помех, оптимальной винеровской фильтрации, повышения качества восприятия изображений системами распознавания, выделения границ объектов и их сегментации в целях классификации и распознавания образов, анализа спектров изображений, вычисления дискретной свертки, корреляционной функции и других операциях над спектрами наибольшее применение нашли дискретные ортогональные преобразования Хартли, Хаара, Уолша-Адамара [40, 41, 64, 84].

ДПФ является базой для реализации большого количества методов цифровой фильтрации [18, 24]. К достоинствам данного метода можно отнести возможность применения быстрых алгоритмов ДПФ, что позволяет проектировать системы с обработкой данных в реальном масштабе времени [38]. Недостатком является необходимость ограничения бесконечного спектра при аналогово-цифровом преобразовании [58, 79]. Далее по тексту, в представленной работе, пойдет речь преимущественно о пространственных методах обработки изображений как более предпочтительных с точки зрения легкости технической реализации.

# Анализ возможностей применения дискретной свертки для предварительной обработки полутоновых изображений

Дискретную свертку можно рассматривать как инструмент, который необходим для построения любого линейного фильтра, инвариантного к трансляциями [64,65]. Прежде чем дать формальное описание дискретной свертки, нам необходимо дать точный ответ на вопрос «Что такое фильтр?».

Основное различие между фильтрами и точечными (попиксельными) операциями состоит в том, что фильтры, как правило, используют более одного пикселя исходного изображения для вычисления каждого нового значения пикселя результирующего изображения. В качестве пояснительного примера рассмотрим задачу сглаживания изображения. Изображения выглядят резкими в первую очередь в местах, где локальная интенсивность резко возрастает или падает (т.е. там, где разница между соседними пикселями велика). С другой стороны, мы воспринимаем изображение как размытое или нечеткое, если функция локальной интенсивности является плавной.

Таким образом, первое что приходит на ум – процедура сглаживания изображения может заключаться в простой замене каждого пикселя на среднее значение соседних пикселей. Чтобы определить новое значение пикселя в сглаженном изображении , мы используем исходный пиксель в той же позиции плюс восемь соседних пикселей для вычисления среднего арифметического этих девяти значений:

**(1.1)**

В относительных координатах изображения формула 1.1 приобретает вид:

**(1.2)**

В более компактном виде, формула 1.2:

**(1.3)**

Пример с локальным усреднением демонстрирует все важные элементы типичного фильтра (так называемого линейного фильтра).

Размер области фильтра является важным параметром фильтра, поскольку он определяет, сколько исходных пикселей вносит вклад в каждое результирующее значение пикселя, и, таким образом, определяет пространственную протяженность (поддержку) фильтра. Например, сглаживающий фильтр в уравнении 1.2 использует область размером 3 × 3, которая центрирована в текущей координате *(u, v)*. Подобные фильтры с большим размером, такие как 5×5, 7×7 или даже 21×21 пиксель, очевидно, будут иметь более сильные эффекты сглаживания.

Форма области фильтра не обязательно должна быть квадратной или прямоугольной. На самом деле, круглая (дискообразная) форма области была бы предпочтительнее для получения эффекта изотропного размытия (то есть того, который является одинаковым во всех направлениях изображения). Другой вариант - присвоить пикселям в сканируемой области разные веса, например, чтобы сильнее выделить пиксели, расположенные ближе к центру области. Кроме того, сканирующая область фильтра не обязательно должна быть смежной и может даже не содержать самого исходного пикселя (представьте, например, кольцевую область фильтра). Теоретически область фильтра может иметь даже бесконечный размер.

Традиционное и хорошо исследованное разбиение на линейные и нелинейные фильтры основано на математических свойствах функции фильтра; то есть, вычисляется ли результат из исходных пикселей с помощью линейного или нелинейного выражения [65]. Таким образом уравнение для свертки *g(x)* от последовательности *f(x)* с ядром свертки *h(x)*:

**(1.4)**

где *h(n-i,m-j)* также имеет название импульсная характеристика (пропагатором точки), ядро окна, скользящая маска, окно, фильтр или шаблон. В зависимости от способа определения данных используется или линейная свертка, или циклическая (круговая) [54].

На рисунке 1.7 и рисунке 1.8 отображены иллюстрации того, как последовательности представлены в обоих случаях:



Рисунок 1.7 – Здесь: (а) – Исходное изображение f; (б) – ядро свертки h;

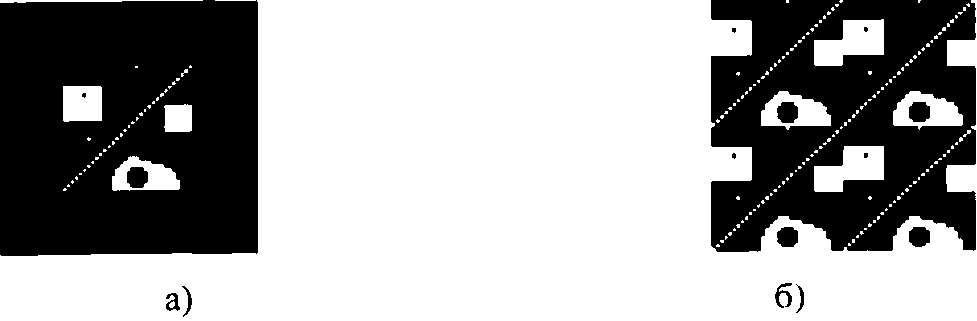
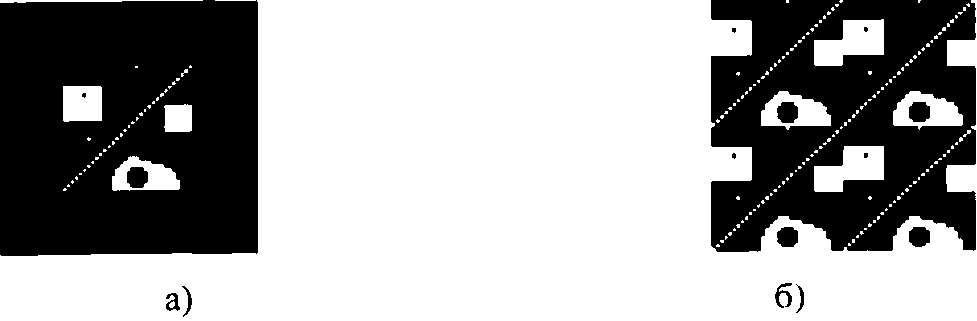


Рисунок 1.8 – Здесь: (а) – Исходное изображение f; (б) – ядро свертки h;

При определении ядра свертки начало координат маски не всегда размещается в левом верхнем углу области действия. Как правило, в результирующем изображении имеется смещение (рисунок 1.9).



Рисунок 1.8 – Здесь: (а) – 3х3 ядро линейно свертки; (б) – 3х3 ядро круговой свертки (периодическое с тем же периодом, как изображение)

Аналоговый прототип изображения, как функцию двух переменных в пространстве финитных бесконечно дифференцируемых действительных функций на прямоугольной области G плоскости (*x, y*)может быть формально определен как:

Матрица дискретной модели изображения на равномерной периодической сетке с шагом ∆*x*=∆*y* имеет вид:

,

где функция Дирака,

здесь – целая часть α.

Скользящие маски (операторы дискретной свертки) для обработки изображений представляют собой таблицы действительных чисел (весовых коэффициентов) размером *PQ*. Их действие на локальную область изображения с центром в элементе состоит в усреднении с весами – коэффициентами маски – значений яркости элементов изображения в окне *PQ*.

Таким образом, значение яркости элемента изображения – отклика для маски размерности 33:

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

определятся из соотношения

**(1.5)**

Выражение (1.5) определяет двумерную дискретную свертку скользящей маски М размерности 33 с изображение:

Применение фильтра к изображению – с технической точки зрения является довольно простым процессом, визуально продемонстрированный на рисунке 1.7. Для каждой точки изображения *(i, j)* выполняются следующие шаги, которые составляют алгоритм вычисления элемента отклика:

1. Ядро (маска) фильтра *M* перемещается по исходному изображению *f*  так, чтобы его начало координат *M(0, 0)* совпадало с пикселем (*i, j*) изображения*.*
2. Все коэффициенты фильтра *M(p,q)* умножаются на соответствующий элемент изображения *f(p + i, q + j)*, и результаты складываются.
3. Итоговая сумма присваивается элементу отклика.
4. Маска смещается на один пиксель исходного изображения, например, вправо вдоль строки и вычисляется следующий пиксель отклика.
5. После прохода текущей строки фильтр смещается к началу следующей и так далее, до тех пор, пока все изображение не подвергнется обработке.

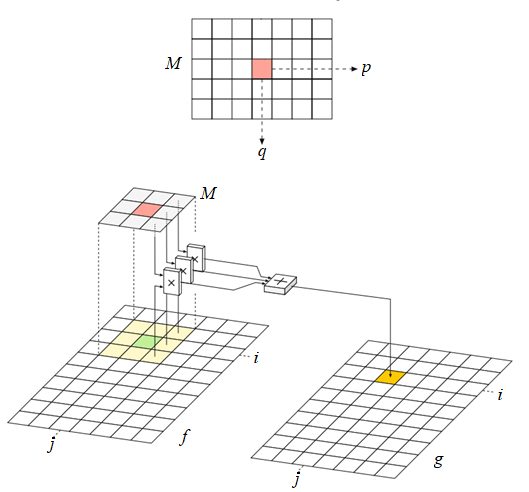


Рисунок 1.7 – Механизм двумерной дискретной свертки

# Анализ возможностей применения методов математической морфологии для предварительной обработки полутоновых изображений

## **1.5.1. Базовые понятия математической морфологии**

Применение аппарата математической морфологии является одним из сравнительно новых направлений в области анализа изображений. Начало математической морфологии, использующей представления теории множеств и интегральной геометрии, было положено работами французских исследователей Ж.Матерона [81] и Дж.Серра [85], которые занимались изучением и решением проблемам минералогии и петрографии. Цель исследований, указанных ученных состояла в количественном описании физических и механических свойств материалов с помощью анализа их геометрической структуры. За последующее время математическая морфология достигла состояния серьезного инструмента обработки изображений с основным применением в материаловедении, исследовании цитологических препаратов, анализе медицинских изображений.

Введем некоторые основные понятия из теории множеств, которые будут необходимы нам в дальнейшем. Пусть  - *n*-мерное пространство. Ниже обычно предполагается, что  или , где - *n*-мерное евклидово пространство, а - *n*-мерное дискретное пространство (*n*-мерная решетка). Применительно к изображениям, как правило, рассматриваются двумерные пространства. Если  и  - множества в , то *объединением множеств*  и  называется множество , (т.е. множество, состоящее из таких элементов , которые принадлежат  или ), а *пересечением множеств*  и  называется множество . Множество  называется *дополнением множества* . *Разностью множеств*  и  называется множество . Множество  называется пустым, если оно не содержит ни одного элемента. Обозначается такое множество как . Справедливы следующие соотношения:

;

; (1.6)

.

Определим на  *индикаторную функцию* множества  следующим образом:

.

Определим также меру множества :

 - для непрерывного пространства  и

- для дискретного пространства .

Для изображений эти определения означают, что мерой множества  является его площадь в непрерывном случае, а в дискретном – количество узлов решетки, входящих в множество.

## **1.5.2. Операции математической морфологии**

Двухградационное изображение можно рассматривать в качестве индикаторной функции набора множеств в (как, например, индикаторную функцию множества  на рис.1.8). Для заданного множества можно зафиксировать некоторый элемент (не обязательно принадлежащий этому множеству), который назовем центром (или началом) множества . Обозначим через  множество , центр которого помещен в точку . Одним из базовых понятий математической морфологии является понятие *структурного элемента*. Структурный элемент - это множество, состоящее из двух непересекающихся подмножеств  и , для которых определено общее начало.

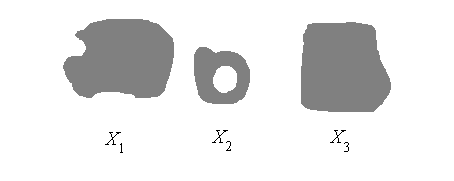


Рисунок 1.8 – Двухградационное изображение

*HM-преобразование.*

Согласно [85], базовым преобразованием, позволяющим строить набор различных операций математической морфологии, является преобразование Hit or Miss. Назовем это преобразование «HM-преобразование». Для данного множества  и данного структурного элемента  результат HM-преобразования определяется как:

 (1.7)

(Здесь через  обозначено дополнение множества .)

Как видно на рис.1.9, в результате HM-преобразования на исходном изображении выделяются элементы, окрестность которых совпадает со структурным элементом (форма окрестности определяется формой структурного элемента). Условие (1. 7) выполняется для элементов, лежащих на нижней границе  (например, 1- 4 позиции структурного элемента). В позиции 5 , но , в позиции 6 , наоборот, , но , а в позиции 7 не выполняются оба условия.

Применяя HM-преобразование с различными структурными элементами можно выделять специфические геометрические особенности изображений.

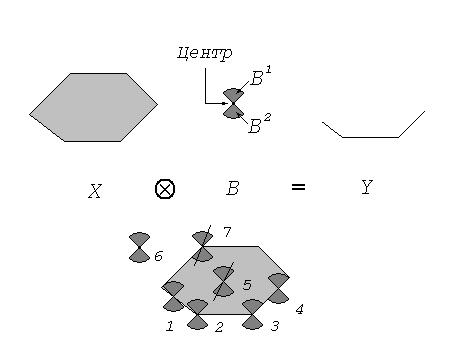


Рисунок 1.9 – НМ–преобразование

*Эрозия.*

Частным случаем HM-преобразования является операция *эрозии* *(erosion)*. Положим, что в структурном элементе  подмножество  - пусто (). При этом условие  всегда выполняется, и в множество  включаются только те элементы исходного множества , для которых выполняется условие  (рис. 1.10) [37]:

.  (1.8)

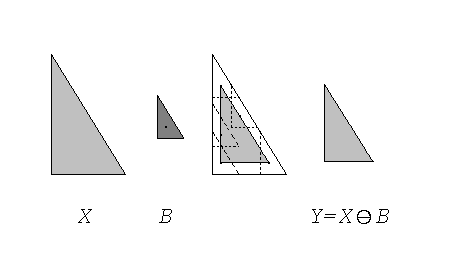


Рисунок 1.10 - Эрозия

*Дилатация.*

Операция, которая является двойственной к эрозии, называется *дилатация*, которая определяется следующим образом (рис.1.11) [37]:

.  (1.9)

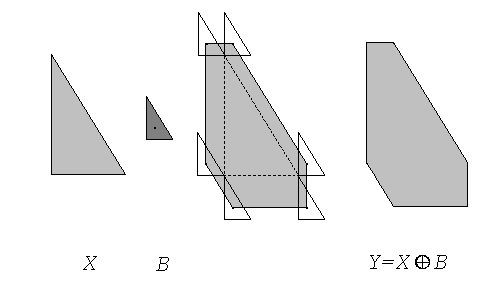


Рисунок 1.11 - Дилатация

## **1.5.2. Применение эрозии и дилатация**

Эрозия и дилатация – операции, предназначенные в первую очередь для выявления морфологических особенностей изображений, причем для выявления различных особенностей используются различные структурные элементы. Например, эрозия посредством круга с радиусом  позволяет найти в изображении объекты, минимальный поперечный размер которых превышает . Если же в качестве структурного элемента взять две точки, смещение между которыми определяется вектором , эрозия позволит выделить объекты, имеющие соседей в направлении и на расстоянии, заданных эти вектором (рис.10.9). (Под объектами здесь подразумеваются односвязные множества).