

На правах рукописи

Хомяков Олег Олегович

**МОДЕЛЬ, МЕТОД, АЛГОРИТМ И ВЫЧИСЛИТЕЛЬНЫЙ МОДУЛЬ
ОБРАБОТКИ ИЗОБРАЖЕНИЙ СИМВОЛЬНОЙ МАРКИРОВКИ**

Специальность 2.3.2. Вычислительные системы и их элементы

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Курск–2024

Работа выполнена в Юго-Западном государственном университете на кафедре вычислительной техники

Научный руководитель: кандидат технических наук, доцент, старший научный сотрудник кафедры вычислительной техники, федеральное государственное бюджетное образовательное учреждение высшего образования "Юго-Западный государственный университет"
Панищев Владимир Славиевич

Официальные оппоненты: доктор технических наук, доцент заведующий кафедрой физики и прикладной математики
Муромский институт (филиал) федерального государственного бюджетного образовательного учреждения высшего образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»
Орлов Алексей Александрович

кандидат технических наук, доцент заведующий кафедрой «Программирование» федеральное государственное бюджетное образовательное учреждение высшего образования «Пензенский государственный технологический университет»
Мартышкин Алексей Иванович

Ведущая организация: федеральное государственное бюджетное образовательное учреждение высшего образования «Волгоградский государственный технический университет»

Защита диссертации состоится 24 декабря 2024 года в 10:00 на заседании диссертационного совета 24.2.435.05, созданного на базе Юго-Западного государственного университета, по адресу: 305040, г. Курск, ул. 50 лет Октября, 94.

С диссертацией можно ознакомиться в библиотеке Юго-Западного государственного университета и на сайте <https://swsu.ru/dis/>

Автореферат разослан « » 2024г.

Ученый секретарь
диссертационного совета 24.2.435.05

Титенко Евгений Анатольевич

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Научно-технический прогресс определил применение специализированных вычислительных систем (ВС) и их элементов для решения различных расчетно-инженерных, научно-исследовательских, проблемно-ориентированных задач, задач управления промышленными объектами, имеющих в своем составе модули-обработчики по различным видам информации. Символьная информация и вычислительные элементы (средства) ее получения и обработки служат основой для работы высокопроизводительных ВС, способных расширять их функциональные возможности за счет формализации интеллектуальных действий, свойственных человеку в части распознавания и интерпретации текстов и текстовых образов. В связи с этим в архитектуру высокопроизводительных ВС при решении данного класса задач целесообразно вводить вычислительные модули (ВМ) для подготовки видеоданных, предобработки изображений в части устранения искажений, шумов, избыточности и распознавания символьных данных по их изображениям.

Специализированные системы компьютерного зрения (ССКЗ), отвечающие за распознавание символьной информации (СИ) на этапе ее подготовки и ввода, являются частью специализированных ВС промышленного применения в гибких производственных системах (ГПС), ведущих потоковую (на конвейере) обработку символьной маркировки продукции. При этом вычислительные процессы распознавания и обработки символьной информации (ОСИ) в условиях априорных искажений символов, смешанных алфавитов, неопределенности стилей и характеристик записи символов являются недостаточно проработанными как с алгоритмической, так и с аппаратной точек зрения.

Используемые ССКЗ создаются преимущественно с использованием программных решений или на базе стандартных архитектур ВС последовательного типа, что ограничивает их скоростные характеристики. Известные аппаратные решения для распознавания и классификации символьных данных по их изображениям отличаются повышенной аппаратной сложностью, зависимостью качества распознавания от исходных (тестовых) наборов. Как следствие, ВМ распознавания и классификации символьной маркировки продукции (характеристик объекта) по своей производительности становятся узким местом для параллельных архитектур, специализированных ВС. Исходя из этого, существует **объективное противоречие** между точностью и скоростью распознавания символьной маркировки продукции на **вычислительных модулях (ВМ)** для специализированных ВС.

Научно-техническая задача разработки методов и аппаратных средств распознавания и классификации маркировки символьной продукции в условиях неопределенности (несколько алфавитов) и искажений/зашумлений изображений символов является актуальной.

Степень разработанности темы. В настоящее время по вопросам разработки ССКЗ и обработки СИ опубликовано большое количество научно-исследовательских работ. В частности, существует широкий выбор методов для распознавания текста и классификации символов. В научной литературе известно большое количество работ, посвященных этой теме, как зарубежных, так и

отечественных авторов Ю.И. Журавлев, И.Б. Гуревич, В.А. Сойфер, А.А. Орлов, В.В. Корнеев, В.Г. Хорошевский, И.А. Каляев, W. Pratt, R. Gonzalez, R. Woods, R.O.Duda, P.E. Hart. Тем не менее в известных трудах недостаточно проработаны вопросы и технические решения, обеспечивающие в комплексе потоковое распознавание и классификацию символьной маркировки продукции по их изображениям. Большинство реализованных подходов к построению ССКЗ ограничены использованием моноалфавитов (латиница) и не учитывают различные виды искажений, что снижает точность распознавания. Таким образом, в данной области существует необходимость в разработке новых моделей, методов, алгоритмов и ВМ для вычислительных систем.

Соответствие паспорту специальности. Согласно паспорту специальности 2.3.2. Вычислительные системы и их элементы, проблематика, рассмотренная в диссертации, соответствует пункту 5 и 6 паспорта специальности (5. Разработка научных методов и алгоритмов организации арифметической, логической, символьной и специальной обработки данных, хранения и ввода-вывода информации, в части разработки метода и алгоритма обработки изображений, содержащих СИ. 6. Разработка научных подходов и методов, архитектурных и структурных решений, обеспечивающих эффективную техническую реализацию аппаратно-программных систем и комплексов за счет оптимизации применяемой электронной компонентной базы, элементов вычислительных систем и встраиваемого программного обеспечения, в части разработки архитектурного решения для вычислительного модуля обработки изображений, обеспечивающего повышение комплексной характеристики «точность × пропускная способность»).

Объектом исследования являются элементы вычислительной системы для обработки изображений и распознавания символьных данных (маркировки) на них.

Предметом исследования являются модели, методы, алгоритмы и вычислительные модули обработки символьной информации на изображениях маркировки продукции.

Цель диссертационной работы. Повышение комплексной характеристики «точность × пропускная способность» обработки символьной информации на маркировке продукции.

Задачи.

1. Провести сравнительный анализ существующих моделей, методов, алгоритмов и вычислительных модулей обработки СИ, постановка задачи обработки изображений с маркировкой продукции.

2. Разработать математическую модель и метод распознавания символьной маркировки с возможностью настройки на конкретные виды маркировок.

3. На основе математической модели и метода создать конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, реализующий распознавание символьной маркировки продукции.

4. Разработать структурную организацию вычислительного модуля обработки символьной информации.

5. Провести экспериментальные исследования разработанного алгоритма и вычислительного модуля.

Научная новизна и положения, выносимые на защиту.

1. Математическая модель обработки изображения, отличающаяся использованием последовательности разработанных частных математических моделей для выделения объекта, предобработки изображения объекта, распознавания текста и классификации характеристик объекта, в совокупности позволяющая получить характеристики объекта, содержащего символьную маркировку.

2. Метод обработки изображений, содержащих символьную информацию, отличающийся применением двумерных дескрипторов и выделения объекта на основе комбинации методов Канни и Рамера-Дугласа-Пекера, а также применением рекуррентной сверточной нейронной сети, позволяющий вести комплексную обработку областей текста в условиях геометрической, информационной, цветовой неопределенностей для последующей классификации объектов.

3. Конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, осуществляющий распознавание символьной маркировки продукции, основанный на разработанной модели и методе обработки изображений, отличающийся распараллеливанием процесса обработки, настройкой к различным видам характеристик и устойчивостью к искажениям, что обеспечивает повышенную точность распознавания.

4. Структурная организация вычислительного модуля обработки символьной информации, которая учитывает особенности созданного алгоритма обработки изображений, содержащих СИ, отличающаяся применением конвейерно-параллельной обработки и позволяющая повысить комплексный показатель «точность × пропускная способность».

Практическая значимость. Разработанные модель, метод, алгоритм и вычислительный модуль обработки символьной информации на маркировке продукции может быть применен в составе ССКЗ для решения широкого спектра задач, таких как обнаружение и распознавание символьной информации в условиях конвейеризированных производств, а также автоматизации учета и инвентаризации складских помещений. Разработанное решение обеспечивает повышение комплексного показателя «точность × пропускная способность», в сравнении с аналогами до 39%-57% в зависимости от уровня шума на изображении.

Результаты диссертационной работы внедрены в ООО «Ценозавр» (г. Курск), а также используются в образовательном процессе кафедры вычислительной техники Юго-Западного государственного университета (г. Курск) при проведении занятий по дисциплинам «Цифровая обработка и анализ изображений» и «Архитектура систем обработки, анализа и интерпретации данных» направления подготовки 09.04.01 Информатика и вычислительная техника.

Методология и методы исследования. Для решения поставленных задач используются: методы проектирования вычислительных систем и их элементов, методы цифровой обработки изображений, метод компьютерного моделирования, теории алгоритмов, проектирования ЭВМ, математического моделирования

Обоснованность и достоверность научных положений и выводов основывается на использовании общепринятых методов компьютерного зрения,

машинного обучения и математического моделирования с экспериментальной проверкой отдельных гипотез на тестовой выборке, содержащей различные виды исследуемых изображений, и сопоставлением полученных данных с результатами теоретических исследований.

Публикации.

По теме диссертационного исследования опубликовано 11 работ, из которых 3 публикации входят в перечень изданий, утвержденных ВАК Минобрнауки России, получено положительное решение о выдаче патента.

Апробация

Результаты диссертационного исследования докладывались, обсуждались и получили положительную оценку на 8 Международных, Всероссийских и Региональных научно-технических конференциях: Труды Международной научно-технической конференции «Информационные технологии и математическое моделирование систем 2019» (Одинцово, 2019), Международный научно-исследовательский конкурс «Лучший исследовательский проект 2020» (Петрозаводск, 2020), национальный суперкомпьютерный форум «НСКФ-2021» (Переславль-Залесский, 2021), XIV Всероссийская межвузовская научная конференция «Наука и образование в развитии промышленной, социальной и экономической сфер регионов России» (Муром, 2022), Международная научная конференция молодых ученых «Исторические, философские, методологические проблемы современной науки» (Курск, 2022), VI Всероссийская научно-практическая конференция «Актуальные вопросы современной науки: теория и практика научных исследований» (Пенза, 2022), XVII Международная научно-техническая конференция «Оптико-электронные приборы и устройства в системах распознавания образов и обработки изображений» (Курск, 2023), Международный научный форум «Наука и инновации – современные концепции» (Москва, 2024). Работа представлялась на конкурсе инновационных проектов «Инновация и изобретение года» (номинация «Инновационная идея»), (Курск, 2021).

Личный вклад автора. В работах, опубликованных в соавторстве, лично соискателем предложены: [1,3,4] – математическая модель обработки изображений, [2,5] – метод обработки изображений, содержащих символьную информацию, [3,6,7,10] – конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, [9,12] – структурная организация вычислительного модуля обработки символьной информации.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована значимость рассматриваемой в диссертации темы, определены общая цель и конкретные задачи исследования, выделена научная новизна, практическая значимость исследования, положения, выносимые на защиту, и информация о структуре диссертации.

В первом разделе представлен обзор существующих решений и подходов к реализации ССКЗ и ОСИ, методов обработки изображений и видеопотоков, рассмотрена организация вычислительных модулей. Приведены описания и представлены основные преимущества и недостатки архитектур ВМ. Описаны основные шаги процесса обработки символьной информации, такие как: поиск и

выделение объекта для распознавания, поиск текста, распознавание текста, определение характеристик объекта.

Проведенный анализ показал, что вычислительная сложность средств распознавания и классификации символьных данных в условиях структурной неопределённости и искажений является критичной для ССКЗ, что негативно сказывается на производительности ВС в целом.

Также показано, что использование нейросетевых решений в совокупности с гибридными архитектурами ВМ обеспечивает распределение частных задач (подготовка изображений, распознавание) между различными блоками ВМ. Этот подход позволяет эффективно использовать ресурсы аппаратного обеспечения и ускорить процесс обработки данных.

Сущность предлагаемого в диссертации подхода к реализации ССКЗ заключается в объединении множества частных моделей обработки изображений в единую систему и аппаратной поддержке вычислительно трудоемких функций предобработки, распознавания изображений, классификации текстовой маркировки с применением двумерных дескрипторов изображений и сверточной рекуррентной нейронной сети (НС) и нейросети с архитектурой EAST. Будучи объединенными едиными формами и форматами входных и выходных данных предлагаемые, единая модель, метод, алгоритм и вычислительный модуль с конвейерно-параллельной структурой позволяют извлекать и обрабатывать сведения геометрического, топологического, цветового и символьного характера из исходных изображений, что обеспечивает повышение комплексной характеристики «точность × пропускная способность» обработки СИ.

Во втором разделе разработана математическая модель обработки изображения маркировки продукции и разработаны оригинальные частные модели, обеспечивающие повышение точности распознавания и классификации характеристик объекта на основе маркировочных данных в потоке.

Входными данными модели является кадр видеопотока (RGB изображение), выходными – классифицированная текстовая информация о характеристиках объекта.

Математическую модель обработки изображения M_{result} составляют следующие частные модели:

- математическая модель кадра видеопотока I_{src} ;
- математическая модель M_{roi} получения области изображения, содержащего объект I_{roi} :
 - функция $F_{\text{move}}(I_{\text{src}}, \vec{I}_{\text{src}})$ определения области изображения, содержащего объект;
 - функция $F_{\text{contours}}(F_{\text{move}})$ поиска контуров объекта;
 - функция $F_{\text{con}}(F_{\text{contours}})$ выделения объекта;
 - функция $F_{\text{desc}}(F_{\text{con}})$ формирования дескрипторов объекта;
 - функция $F_{\text{comp}}(F_{\text{desc}}, E_{\text{desc}})$ сравнения дескрипторов объекта с эталонными дескрипторами;
- математическая модель M_{obj} предобработки изображения объекта:

- функция F_{hsv} (I_{roi}) трансформации RGB изображения в HSV изображение;
- функция F_{bin} (F_{hsv} , L_b , L_t) бинаризации по нижней и верхней границам;
- функция F_{filter} (F_{bin}) удаления шумов;
- функция F_{dilate} (F_{filter}) дилатации изображения с уменьшенным уровнем шума;
- функция F_{bord} (F_{dilate}) поиска границ объекта;
- функция $F_{\text{transform}}$ (F_{bord} , I_{roi}) исправления аффинных и проекционных искажений;
- математическая модель M_{text} поиска и распознавания текста:
 - функция F_{find} (I_{obj}) поиска области текста I_{text} ;
 - функция F_{base} (I_{text}) поиска базовой линии;
 - функция F_{split} (F_{base} , I_{text}) разбиения на символы;
 - функция F_{rec} (F_{split}) распознавания текста;
- математическая модель M_{class} получения характеристик:
 - функция F_{info} (T , I_{text}) получения дополнительной информации геометрического и цветового характера, описывающей область текста;
 - функция F_e (T , E , F_{info}) сравнения признаков области текста и эталонных признаков;
 - функция F_{out} (F_e , T) формирования результата.

Математическая модель кадра изображения представлена в виде матрицы I_{src} интенсивности цвета точек изображения в RGB представлении:

$$I_{\text{src}} = \begin{pmatrix} i_{0,0} & \cdots & i_{x_{\text{src}},0} \\ \vdots & \ddots & \vdots \\ i_{0,y_{\text{src}}} & \cdots & i_{x_{\text{src}},y_{\text{src}}} \end{pmatrix}, \quad (1)$$

где $i_{x_{\text{src}},y_{\text{src}}}$ – вектор, содержащий интенсивность цвета точки (ИЦТ) изображения в RGB представлении, используемый диапазон интенсивности находится в пределах между 0 и 255 (стандартное 8-битное изображение), x_{src} и y_{src} – размеры изображения в пикселях по вертикали и горизонтали.

Математическая модель M_{roi} получения области изображения, содержащего объект формирует двумерную матрицу I_{roi} размерности x_{roi} и y_{roi} , элементы которой представляют собой ИЦТ и в общем виде представляется как:

$$M_{\text{roi}} = F_{\text{comp}}(F_{\text{desc}}(F_{\text{con}}(F_{\text{contours}}(F_{\text{move}}(I_{\text{src}}, \vec{I}_{\text{src}}(t_{\text{start}}, t_{\text{end}}))))), E_{\text{desc}}(I_e)), (2)$$

где \vec{I}_{src} – вектор, содержащий матрицы изображений I_{src} , полученные в моменты времени от t_{start} до t_{end} , E_{desc} – функция получения дескрипторов шаблонного изображения, I_e – матрица шаблонного изображения.

Функция F_{move} (I_{src} , \vec{I}_{src}) определения области изображения, содержащего объект формирует вектор размерности n_{move} , элементы которого представляют собой координаты точек, изменивших значения яркости относительно $\vec{I}_{\text{src}}(t_{\text{start}}, t_{\text{end}})$.

Функция F_{contours} (F_{move}) поиска контуров объекта формирует вектор C_{roi} , представляющий собой последовательность координат точек, описывающих границы объекта, для которых выполняется одно из следующих условий:

$$\begin{aligned} i_{\text{move}}(x_{\text{move}}, y_{\text{move}}) &= \{x_{\text{move}} = \max(x_0, x_1, \dots, x_n)\} \\ \text{или } i_{\text{move}}(x_{\text{move}}, y_{\text{move}}) &= \{x_{\text{move}} = \min(x_0, x_1, \dots, x_n)\} \end{aligned} \quad (3)$$

при $x_i \in F_{\text{move}}(I_{\text{src}}, \vec{I}_{\text{src}}(t_{\text{start}}, t_{\text{end}}))$ и $y_{\text{move}} = \text{const}$,

где i_{move} – элемента вектора C_{roi} , x_{move} , y_{move} – координаты точки по горизонтали и вертикали

Функция F_{con} (F_{contours}) выделения объекта формирует двумерную матрицу, элементы которой представляют собой ИЦТ, описывающие область, в котором находится объект и рассчитывается по выражению:

$$F_{\text{con}}(F_{\text{contours}}) = \begin{pmatrix} i_{\min_x(C_{\text{roi}}), \min_y(C_{\text{roi}})} & \cdots & i_{\max_x(C_{\text{roi}}), \min_y(C_{\text{roi}})} \\ \vdots & \ddots & \vdots \\ i_{\min_x(C_{\text{roi}}), \max_y(C_{\text{roi}})} & \cdots & i_{\max_x(C_{\text{roi}}), \max_y(C_{\text{roi}})} \end{pmatrix}, \quad (4)$$

где C_{roi} – вектор точек, в которых происходило движение, \min и \max – функции поиска точки с минимальным и максимальным значением координат по горизонтали и вертикали.

Функция F_{desc} (F_{con}) формирования дескрипторов объекта создает дескрипторы, представляющие собой двумерную матрицу, содержащую координаты 500 особых точек дескрипторов:

$$F_{\text{desc}}(F_{\text{con}}) = \begin{pmatrix} d_{x_1}, d_{x_2}, \dots, d_{x_{500}} \\ d_{y_1}, d_{y_2}, \dots, d_{y_{500}} \end{pmatrix}, \quad (5)$$

где d_x и d_y – координаты особых точек изображения, определяемых распределением по Гауссу относительно центра изображения.

Функция F_{comp} (F_{desc} , E_{desc}) сравнения дескриптора объекта с эталонным дескриптором формирует двумерную матрицу I_{roi} в случае, если удовлетворяется следующее условие:

$$\frac{E_{11} + E_{00}}{E_{00} + E_{01} + E_{10} + E_{11}} > T_{\text{desc}}, \quad (6)$$

где T_{desc} – пороговое значение соответствия дескрипторов; E_{01} – количество точек, для которых результат бинарного теста дескриптора области с объектом равен 0, а эталонного 1; E_{10} – количество точек, для которых результат бинарного теста дескриптора области с объектом равен 1, а эталонного 0; E_{11} – количество точек, для которых результат бинарного теста двух дескрипторов равен 1; E_{00} – количество точек, для которых результат бинарного теста двух дескрипторов равен 0; при этом бинарный тест τ определяется как:

$$\tau(d_x, d_x d'_x d'_y) = \begin{cases} 1, & \text{при } i_{d_x d_y} < i_{e_{d'_x d'_y}}; \\ 0, & \text{при } i_{d_x d_y} \geq i_{e_{d'_x d'_y}} \end{cases} \quad (7)$$

где d'_x и d'_y – координаты особых точек эталонного изображения, $i_{e_{d'_x d'_y}}$ – ИЦТ эталонного изображения.

Математическая модель M_{obj} предобработки изображения формирует изображение I_{obj} , представляющее собой двумерную матрицу, элементами которой являются ИЦТ и в общем виде может быть записана как

$$M_{obj} = F_{transform} \left(F_{bord} \left(F_{dilate} \left(F_{filter} \left(F_{bin} (F_{hsv}(I_{roi}), L_b, L_t) \right) \right) \right), I_{roi} \right). \quad (8)$$

Функция $F_{hsv}(I_{roi})$ трансформации RGB изображения в HSV изображение формирует двумерную матрицу изображения I_{hsv} :

$$I_{hsv} = \begin{pmatrix} i_{hsv_{0,0}} & \cdots & i_{hsv_{x_{hsv},0}} \\ \vdots & \ddots & \vdots \\ i_{hsv_{0,y_{hsv}}} & \cdots & i_{hsv_{x_{hsv},y_{hsv}}} \end{pmatrix}, \quad (9)$$

где $i_{hsv_{x_{hsv},y_{hsv}}}$ – вектор содержащий цветовой тон, насыщенность и яркость изображения в цветовой модели HSV, x_{hsv} и y_{hsv} – размерности изображения по горизонтали и вертикали.

Функция $F_{bin}(F_{hsv}, L_b, L_t)$ бинаризации по нижней и верхней границам преобразует I_{hsv} в двумерную матрицу I_{bin} , с элементами i_{bin} формируемыми по следующему условию:

$$i_{bin} = \begin{cases} 1, & \text{при } L_b \leq i_{hsv} \leq L_t \\ 0, & \text{в ином случае} \end{cases}. \quad (10)$$

Функция $F_{filter}(F_{bin})$ удаления шумов преобразует элементы I_{hsv} , формируя двумерную матрицу I_{filter} размерности I_{hsv} , элементы которой формируются как $g(x_g, y_g)$

$$g(x_g, y_g) = \sum_{u=-k}^k \sum_{v=-k}^k \frac{1}{2\pi\sigma^2} \exp\left(-\frac{u^2 + v^2}{2\sigma^2}\right) i_{bin_{x_g-u, y_g-v}}. \quad (11)$$

Функция $F_{dilate}(F_{filter})$ дилатации изображения с уменьшенным уровнем шума применяет к I_{filter} операцию дилатации с использованием матрицы T_{dilate}

$$T_{dilate} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}. \quad (12)$$

Функция $F_{bord}(F_{dilate})$ поиска границ объекта формирует вектор C_{obj} размерности n_c , содержащий последовательность точек C_{obj} , описывающих границы объекта на изображении, полученном в результате F_{dilate} :

$$C_{obj} = (C_{obj_0}, C_{obj_1}, \dots, C_{obj_{n_c}}). \quad (13)$$

Функция $F_{transform}(F_{bord}, I_{roi})$ исправления аффинных и проекционных искажений формирует двумерную матрицу I_{obj} , элементами которой являются вектора i_{obj} , содержащие ИЦТ таким образом, что:

$$I_{obj}(x_{obj}, y_{obj}) = I_{roi} \left(\frac{P_{00}x_{obj} + P_{01}y_{obj} + P_{02}}{P_{20}x_{obj} + P_{21}y_{obj} + P_{22}}, \frac{P_{10}x_{obj} + P_{11}y_{obj} + P_{12}}{P_{20}x_{obj} + P_{21}y_{obj} + P_{22}} \right), \quad (14)$$

где P – матрица трансформации размерности 3×3 , получаемая с использованием вектора границ объекта C_{obj} , x_{obj} и y_{obj} координаты точек по вертикали и горизонтали.

Математическая модель M_{text} поиска и распознавания текста формирует вектор, содержащий распознанный текст T и в общем виде может быть записана как:

$$M_{text} = F_{rec}(F_{split}(F_{base}(F_{find}(I_{obj})), F_{find}(I_{obj}))). \quad (15)$$

Функция $F_{\text{find}}(I_{\text{obj}})$ поиска области текста I_{text} формирует вектор, содержащий координаты области текста

$$\begin{aligned} I_{\text{text}} &= (T_{bb0}, T_{bb1}, \dots, T_{bbn}) \\ T_{bbi} &= (x_{\text{left}}, y_{\text{top}}, x_{\text{right}}, y_{\text{bottom}}), \end{aligned} \quad (16)$$

где x_{left} и y_{top} , x_{right} и y_{bottom} – координаты, описывающие верхние левые, и правые нижние координаты области текста.

Функция $F_{\text{base}}(I_{\text{text}})$ поиска базовой линии формирует координаты линии, по которой будет происходить разбиение на символы.

Функция $F_{\text{split}}(F_{\text{base}}, I_{\text{text}})$ разбиения на символы использует базовую линию для поиска координат, описывающих область с символом I_{char} .

$$I_{\text{char}_n} = \begin{pmatrix} x_{\text{charL}} = na_{mx} - m_w/2 \\ y_{\text{charT}} = na_{my} - m_h/2 \\ x_{\text{charR}} = na_{mx} - m_w/2 \\ y_{\text{charB}} = na_{my} - m_w/2 \end{pmatrix}, \quad (17)$$

где x_{charL} , y_{charT} , x_{charR} , y_{charB} – координаты левого верхнего и нижнего правого углов прямоугольника, описывающего символ текста, n – номер символа в строке; a_{mx} , a_{my} – координаты середины символа по горизонтали и вертикали, рассчитываемые как среднее от верхней, нижней и базовой линий текста; m_w , m_h – максимальный размер символов текста по горизонтали и вертикали.

Функция $F_{\text{rec}}(F_{\text{split}})$ распознавания текста формирует вектор T , содержащий распознанные символы текста.

Математическая модель M_{class} получения характеристик в общем виде представлена как

$$M_{\text{class}} = F_{\text{out}}(F_e(T, E, F_{\text{info}}(T, I_{\text{text}})), T). \quad (18)$$

Функция $F_{\text{info}}(T, I_{\text{text}})$ получения дополнительной информации геометрического и цветового характера, описывающего область текста формирует вектор:

$$F_{\text{info}}(T, I_{\text{text}}) = \begin{pmatrix} I_{\text{color}} \\ I_{\text{size}} \\ I_{\text{relations}} \end{pmatrix}, \quad (19)$$

где I_{color} – цветовая характеристика области текста, I_{size} – информация о размере области текста, $I_{\text{relations}}$ – информация о положении области текста относительно других.

Функция $F_e(T, E, F_{\text{info}})$ сравнения признаков области текста и эталонных признаков. Для расчета вводится матрица эталонных признаков характеристик (включающих цветовую и геометрическую информацию об области текста) E :

$$E = \begin{pmatrix} e_{0,0} & \dots & e_{n,0} \\ \vdots & \ddots & \vdots \\ e_{0,q} & \dots & e_{n,q} \end{pmatrix}, \quad (20)$$

где n – количество признаков, q – количество типов характеристик, e – эталонное значение признака.

Тогда значение F_e выражается как:

$$F_e(T, E, F_{\text{info}}(T, I_{\text{text}})) = \min \left(\text{dist}(T, F_{\text{info}}(T, I_{\text{text}}), E) \right), \quad (21)$$

где dist – функция расчета расстояния между эталонными признаками и признаками текста.

Функция $F_{out}(F_e, T)$ формирования результата создает вектор результата R , представляемый в виде:

$$R = (d, x, k), \quad (22)$$

где d – тип характеристики, x – текст, описывающий характеристику объекта, k – вектор критериев оценки, содержащий вычисленные расстояния между эталонными признаками и признаками текста.

Таким образом, обобщенная математическая модель обработки изображения M_{result} представлена как:

$$M_{result} = M_{class}(M_{text}(M_{obj}(M_{roi}(I_{src}))))). \quad (23)$$

Новизна математической модели заключается в интеграции моделей поиска объектов, предобработки, распознавания текста и классификации характеристик в единую структуру, позволяя получить характеристики объекта, содержащего символьную маркировку.

На основе разработанной математической модели разработан метод обработки изображений, содержащих символьную информацию для распознавания символьной маркировки (характеристик объекта), состоящий из следующих этапов:

1. Прием кадра(ов) изображения и поиск двумерного объекта-контейнера (продукция, товар) (основан на выражениях и функциях 1-7).
2. Предобработка области изображения с найденным объектом-контейнером (основан на выражениях и функциях 8-14).
3. Поиск и распознавание символьной маркировки внутри объекта-контейнера (основан на выражениях и функциях 15-17).
4. Классификация характеристик объектов и формирование их геометрических, топологических, цветовых и текстовых признаков (основан на выражениях и функциях 18-22).

Первый этап является проверочным, его цель – поиск объекта-контейнера и проверка его геометрической формы на соответствие заданным шаблонам видов продукции (товара). В результате принимается решение о дальнейшей обработке кадра(ов) изображения при условии, что он содержит объект для распознавания. Этап отличает применение двумерных дескрипторов исходного объекта и заданных шаблонов, основанных на редукции количества проверяемых пикселей из матриц изображений. Формирование и сравнение двумерных дескрипторов реализовано по методу ORB (Oriented FAST and Rotated BRIEF), что позволяет определить наличие объекта для распознавания, содержащего символьную маркировку.

Второй этап является подготовительным, он осуществляет предобработку изображений, включая компенсацию аффинных искажений и искажений перспективы на изображениях области объекта. В результате определяются прямоугольные блоки изображения, содержащие текст для последующего распознавания. Этап отличает применение метода Гаусса для удаления шумов и выделение границ прямоугольных блоков изображений на основе комбинации методов Канни и Рамера-Дугласа-Пекера.

Третий этап является основным, он состоит в поиске и выделении из матрицы пикселей требуемых областей текста и их последующем посимвольном

распознавании. Области текста выделяются с помощью нейронной сети с архитектурой EAST, что позволяет сегментировать изображение, содержащие области текста произвольной структуры. Разделение текста на символы реализовано методом поиска базовой линии строк, что сокращает ошибки распознавания первого рода за счет уменьшения влияния межсимвольной дистанции и перекрытия символов на точность разбиения для изображений, содержащих различных шрифт и стиль текста, а также курсивное начертание. Посимвольное распознавание символов реализовано нейронной сетью с архитектурой сверточной рекуррентной сети CRNN. Особенность этапа – потенциальный параллелизм за счет независимой обработки нескольких кадров. Выходные данные третьего этапа – символьная последовательность.

Четвертый этап является завершающим. Он необходим для классификации характеристик исходных объектов по их геометрическим, топологическим и цветовым характеристикам, что позволяет получить текстовую информацию о характеристиках объекта. Конструктивный вид функции классификации представлен количественно-разностной функцией невязки между эталонными и выделенными признаками.

Новизна метода заключается в применении двумерных дескрипторов и выделения объекта на основе комбинации методов Канни и Рамера-Дугласа-Пекера, а также применением рекуррентной сверточной нейронной сети, позволяющей вести комплексную обработку пространственно-цветовых характеристик областей текста, таким образом формируя завершённый цикл обработки изображений объектов с символьной маркировкой.

В третьем разделе разработан конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, реализующий разработанный метод. Блок-схема алгоритма представлена на рисунках 1 и 2.

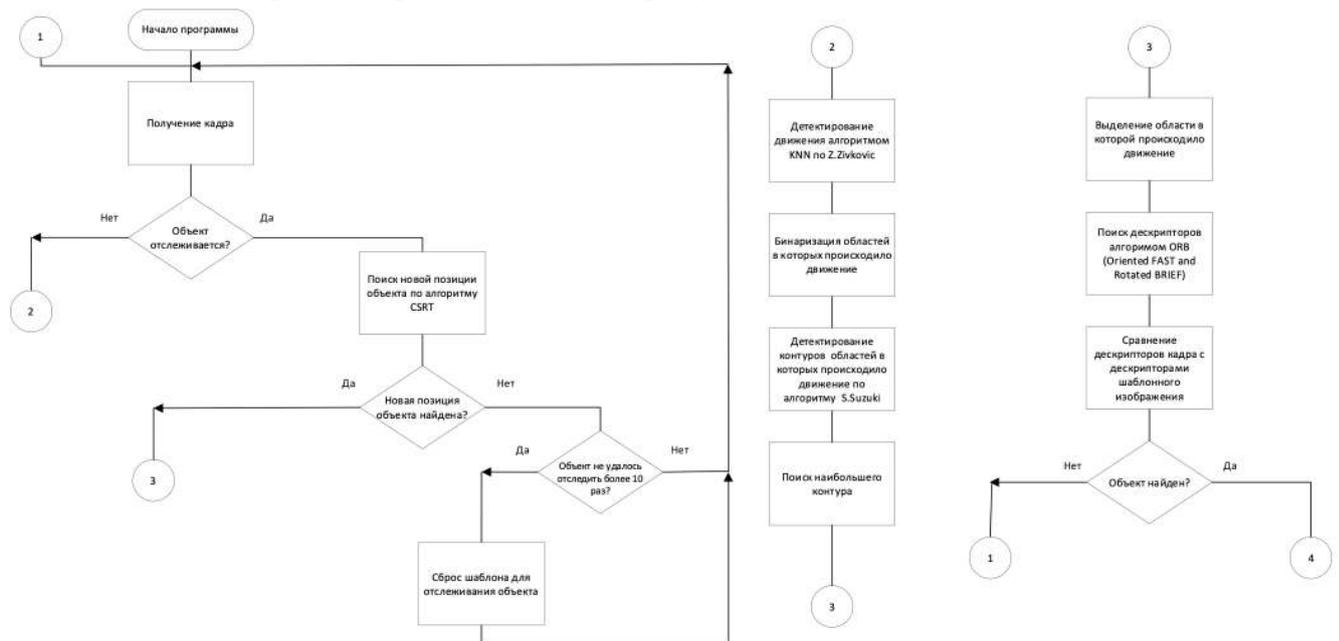


Рисунок 1 – Блок-схема алгоритма обработки изображений, содержащих СИ

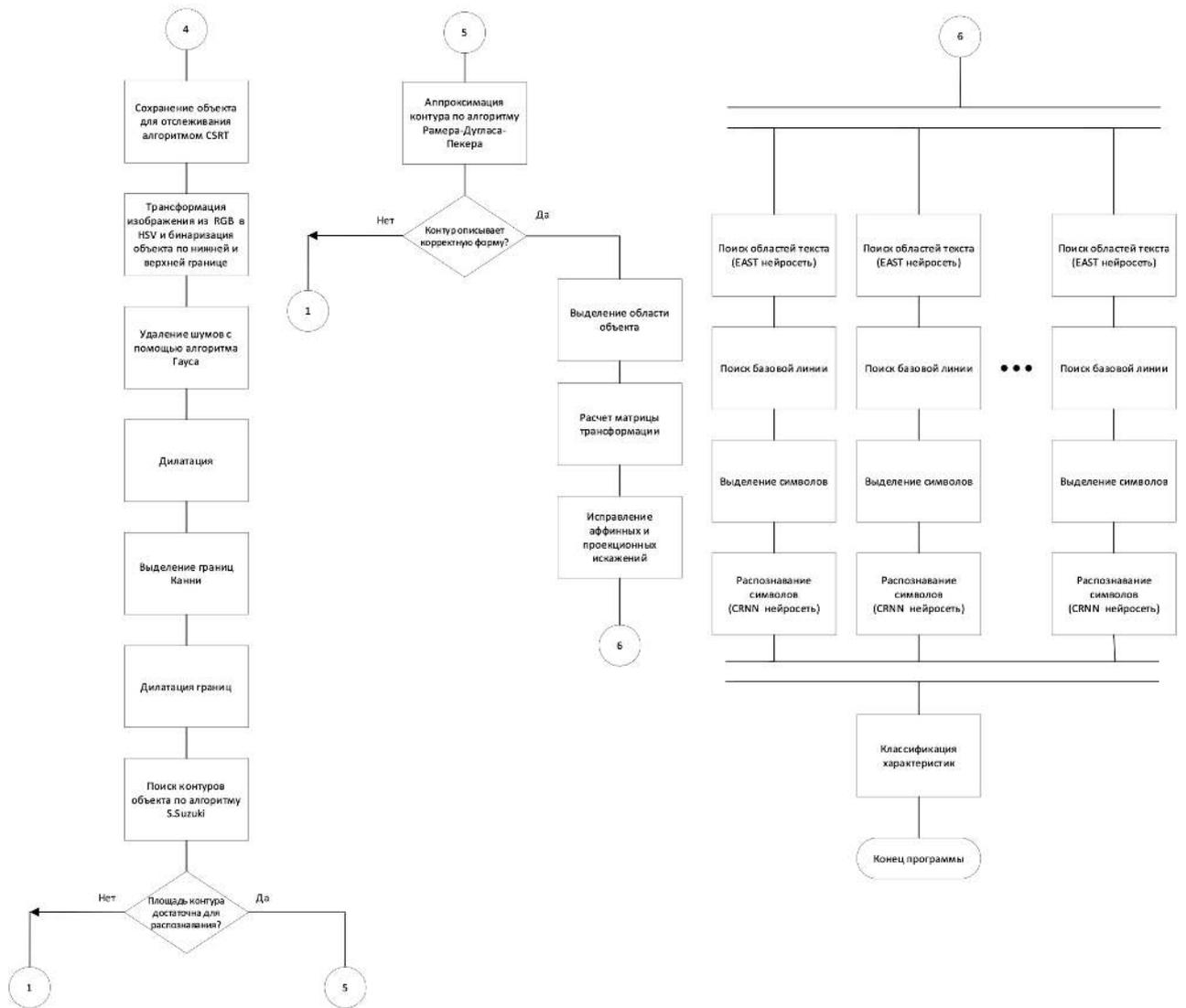


Рисунок 2 – Блок-схема алгоритма обработки изображений, содержащих СИ (продолжение)

Входными данными алгоритма являются кадры видеопотока. В основе алгоритма лежит последовательность действий, включающая обнаружение объекта на кадре, предварительную обработку изображения, распознавание символьной информации и определение характеристик объекта. При этом данные, полученные на этапе предобработки и поиска областей текста, сохраняются для дальнейшего использования при распознавании текста и определения характеристик объекта. Результатом работы алгоритма является вектор, содержащий символьную информацию, тип характеристик объекта и набор критериев оценки.

Произведено обучение нейросетей для распознавания текста. Для обучения использовались публичные выборки: COCO-text для нейросети с архитектурой EAST, SynthText для русского языка и MJsynth для английского языка для нейросети с архитектурой CRNN. Данные выборки были дополнены реальными изображениями объектов, содержащих маркировку. Для увеличения точности работы сети и увеличения устойчивости к различным искажениям первоначальная выборка была дополнена изображениями, содержащие искажения.

Проведено тестирование решения, статистические данные о точности распознавания при различном уровне шума, распределяемому по нормальному

закону, с доверительными интервалами для среднего (достоверность 0.99) представлены в таблицах 1 и 2. Пример работы алгоритма представлен на рисунке 3.

Таблица 1 – Статистические данные по точности обнаружения объекта в кадре при различных искажениях

Уровень шума (PSNR)	Площадь перекрытия объекта	Точность обнаружения объекта в кадре
От 20 dB	До 5%	98% ± 0.3%
15-20 dB	До 15%	81% ± 0.7%
10-15 dB	До 30%	21% ± 0.9%

Таблица 2 – Точность распознавания символов разработанного конвейерно-параллельного алгоритма обработки изображений, содержащих СИ

Уровень шума (PSNR)	Точность распознавания	Ошибки первого рода (распознанные символы, не присутствовавшие на изображении)	Ошибки второго рода (некорректно распознанные символы, присутствовавшие на изображении)
От 20 dB	97.7% ± 0.5%	2%	0.4%
15-20 dB	91.8% ± 0.7%	6.6%	1.7%
10-15 dB	80% ± 0.8%	17.2%	2.9%

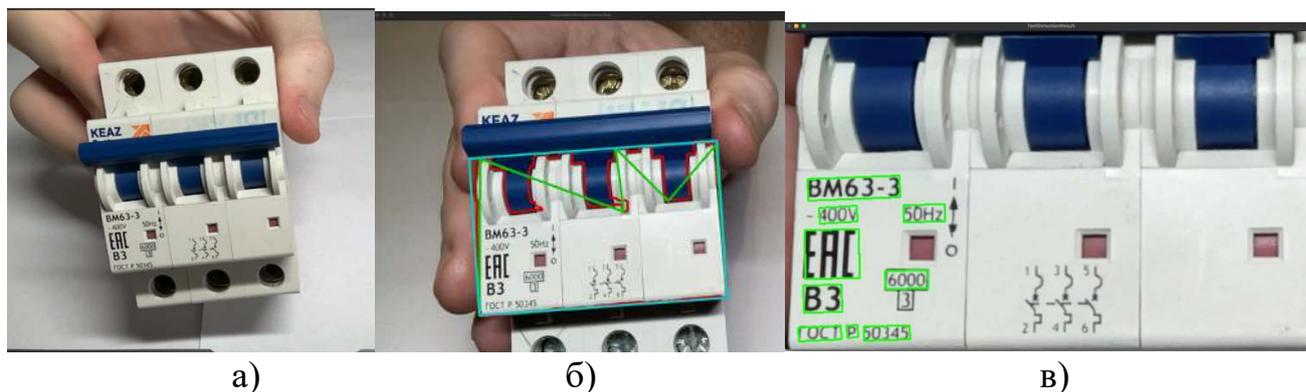


Рисунок 3 – Пример работы конвейерно-параллельного алгоритма обработки изображений, содержащих СИ, а) – исходное изображение, б) – поиск объекта для распознавания, в) – поиск слов, содержащих характеристики объекта

Предлагаемое решение позволило осуществлять распознавание символьной маркировки, обладающее способностью адаптации к разнообразным видам маркировок и искажений, что обеспечивает точность распознавания от 97% до 80% в зависимости от уровня шума на изображении. При этом основной проблемой при распознавании изображений с высоким уровнем шума является некорректное определение участка изображения как символа, что видно из увеличения процента ошибок первого рода до 17,15% при увеличении уровня шума на изображении.

В четвертой главе рассматривается структурная организация вычислительного модуля обработки символьной информации. Вычислительный модуль реализует разработанный в третьей главе алгоритм и включает в себя следующие блоки: получения изображения (1), предобработки изображения (2), классификации символьной информации (3), формирования выходного значения.

Из блока получения изображения данные направляются в блоки анализа цветовой информации, удаления шумов и поиска области объекта в блоке предобработки. Затем информация о области объекта и цветовых данных передается в блоки поиска слов и символов в блоке классификации символов. Результаты классификации символов поступают на вход блока классификации характеристик, где определяется текстовая информация, являющаяся маркировкой продукции. Завершает работу блок формирования результата распознавания. Структурная организация вычислительного модуля представлена на рисунке 4.

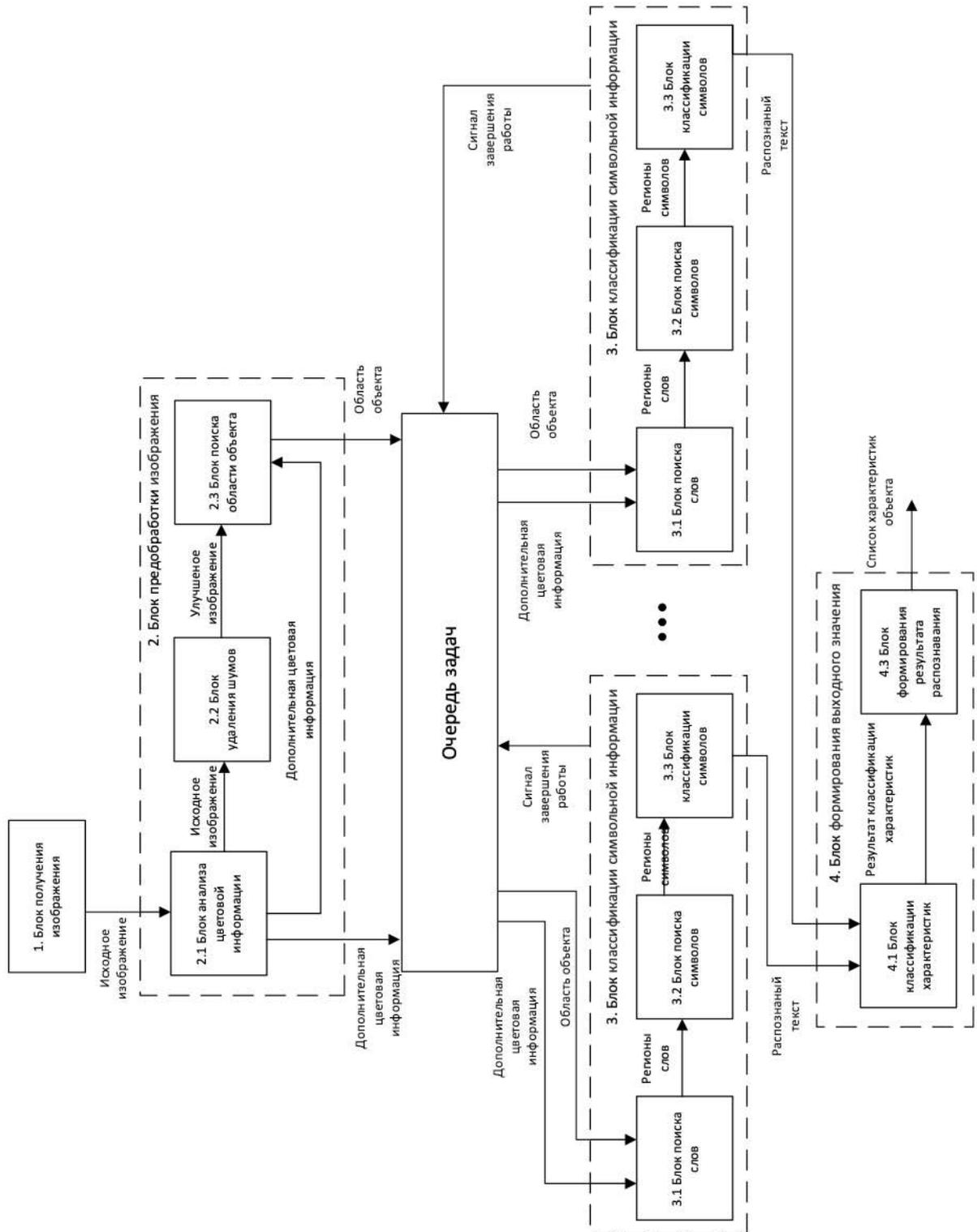


Рисунок 4 – Структурная организация вычислительного модуля

Разработанное решение было протестировано для определения комплексной характеристики «точность × пропускная способность» оцениваемой в количестве корректно распознанных кадров в секунду (крк/с) вычисляемых как $t = a \times b$, где a – точность распознавания символов в диапазоне $[0..1]$, b – пропускная способность. Сравнение с аналогами приведено в таблице 3. Результаты тестирования при различном уровне шума, распределяемому по нормальному закону, приведены в таблице 4.

Таблица 3 – Сравнение с аналогами

Критерий	Поддержка программной реализации	Поддержка аппаратной реализации	Поддержка ускорения на графических процессорах	Возможность параллельной работы	Сохранение порядка обработки кадров
Разработанный вычислительный модуль	Да	Да	Да	Да	Да
Keras (CRNN)	Да	Нет	Да	Да	Нет
TesseractOCR	Да	Нет	Нет	Нет	Нет
Устройство распознавания символов (на основе ПЛИС с процессором ARM – Cortex A9)	Нет	Да	Нет	Да	Нет
Устройство распознавания символов (программная реализация на процессоре AMD Phenom II 925)	Да	Нет	Нет	Да	Нет
Устройство распознавания символов с графическим процессором (AMD Phenom II 925 и Nvidia GTX 1060)	Да	Да	Да	Да	Нет

Компьютерное моделирование для проведения тестирования проводилось с использованием процессора с частотой 2.3GHz, GPU с частотой 1290MHz и объемом памяти 4GB. Ограничения и допущения модуля: входные изображения имеют разрешение 1920x1080 пикселей, освещенность сцены находится в диапазоне от 150 до 1500 люкс, пиковое соотношение сигнала к шуму (PSNR) не ниже 10 dB, объект для распознавания в кадре занимает не менее 8.5% от площади кадра.

При этом пропускная способность обусловлена следующими показателями:

1. Блоки 1 и 2 (получения изображения и предобработки изображения) обрабатывают 1 кадр за 35мс, что позволяет достичь пропускной способности до 33.3 кадров в секунду
2. Блок 3 (расознавания символьной информации) обрабатывает один кадр за 400мс, что при использовании 14 параллельных блоков позволяет достичь пропускной способности до 35 кадров в секунду

3. Блок 4 (формирования выходного значения) обрабатывает 1 кадр за 5мс.

Таблица 4 – Результаты тестирования по комплексной характеристике «точность × пропускная способность»

Решение	Точность × пропускная способность при PSNR от 20 dB	Точность × пропускная способность при PSNR 15-20 dB	Точность × пропускная способность при PSNR 10-15 dB
Разработанный вычислительный модуль	$0.97 \times 33.3 = 32.3$ крк/с	$0.91 \times 33.3 = 30.3$ крк/с	$0.80 \times 33.3 = 26.6$ крк/с
Keras (CRNN)	$0.93 \times 5.7 = 5.3$ крк/с	$0.84 \times 5.7 = 4.8$ крк/с	$0.6 \times 5.7 = 3.4$ крк/с
TesseractOCR	$0.89 \times 22.2 = 19.7$ крк/с	$0.75 \times 22.2 = 16.7$ крк/с	$0.51 \times 22.2 = 11.3$ крк/с
Устройство распознавания символов (на основе ПЛИС с процессором ARM – Cortex A9)	$0.98 \times 7.3 = 7.1$ крк/с	-	-
Устройство распознавания символов (программная реализация на процессоре AMD Phenom II 925)	$0.98 \times 7.9 = 7.7$ крк/с	-	-
Устройство распознавания символов с графическим процессором (AMD Phenom II 925 и Nvidia GTX 1060)	$0.98 \times 14 = 13.7$ крк/с	-	-

В отличие аналогов, разработанный вычислительный модуль использует подход, который сохраняет порядок обработки данных, что позволяет обеспечивать последовательность распознавания при работе с видеопотоками, где каждый кадр содержит информацию, логически связанную с предыдущими и последующими изображениями. Таким образом, разработанная структурная организация вычислительного модуля обработки символьной информации позволяет обрабатывать символьную информацию на маркировке продукции и превосходит аналоги по показателю «точность × пропускная способность», на величину от 39% до 57% в зависимости от уровня шума.

В заключении сформулированы основные результаты диссертационной работы, рекомендации и перспективы дальнейшей разработки темы.

К перспективам дальнейших исследований можно отнести возможность оптимизации вычислений, с целью повышения скорости работы алгоритма и вычислительного модуля, его реализующего. В частности, могут быть рассмотрены перспективы использования альтернативных архитектур нейронных сетей для распознавания символов и особенности применения вычислительного модуля в составе комплексных ССКЗ.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

Диссертационная работа посвящена решению актуально научно-технической задачи, связанной с разработкой методов и средств конвейерно-параллельной обработки для распознавания и классификации маркировки продукции в условиях неопределённости и искажений/зашумлений изображений символов. В ходе исследования получены следующие основные результаты:

1. Проведен анализ существующих моделей, методов, алгоритмов и вычислительных модулей для обработки изображений и распознавания символов. Выделена необходимость создания моделей, методов, алгоритмов и вычислительных модулей, направленных на ликвидацию проблем выполнения распознавания маркировки продукции, возникающих при наличии шумов на изображении и работе с несколькими алфавитами.

2. Разработана математическая модель обработки изображений, отличающаяся использованием последовательности разработанных частных математических моделей для выделения объекта, предобработки изображения объекта, распознавания текста и классификации характеристик объекта, в совокупности позволяющая получить характеристики объекта, содержащего символьную маркировку.

3. Метод обработки изображений, содержащих символьную информацию, включающий последовательность проверочного, подготовительного, основного (распознавание) и классификационного этапов и отличающийся применением двумерных дескрипторов и выделения объекта на основе комбинации методов Канни и Рамера-Дугласа-Пекера, а также применением рекуррентной сверточной нейронной сети, позволяющий вести комплексную обработку областей текста в условиях геометрической, информационной, цветовой неопределенностей для последующей классификации характеристик объектов.

4. Разработан конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, который реализует распознавание символьной маркировки продукции в соответствии с созданной математической моделью и методом обработки изображений, содержащих СИ, и отличающийся совмещением предварительной обработки изображений и параллельной классификации символьной информации.

5. Разработана структурная организация вычислительного модуля обработки символьной информации, основанная на конвейерно-параллельном алгоритме обработки изображений, содержащих СИ. Спроектированная структурная организация вычислительного модуля позволяет повысить комплексный показатель «точность × пропускная способность», в сравнении с аналогами до 39% (для изображений с уровнем шума (PSNR) от 20dB), до 44% (PSNR – 15-20dB) и до 57% (PSNR – 10-15dB).

СПИСОК НАУЧНЫХ РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях, рекомендованных перечнем ВАК РФ

1. **Хомяков О.О.** Распознавание символьной информации для автоматизации производственных процессов / В.С. Панищев, М.И. Труфанов, О.Г. Добросердов, О.О. Хомяков // Известия Юго-Западного государственного университета. Курск – 2021. – № 25 – С. 122-137.

2. **Хомяков О.О.** Метод распознавания текстовых данных на изображениях / В.С. Панищев, О.О. Хомяков, Д.В. Титов, С.И. Егоров // Известия ЮФУ. Технические науки. Ростов-на-Дону – 2023. – № 4 (234) – С. 57-65.

3. **Хомяков О.О.** Математическая модель и параллельный алгоритм обработки изображений, содержащих символьную информацию / О.О. Хомяков, В.С. Панищев, В.С. Титов, Э.И. Ватутин. // Труды МАИ – 2024 – № 137 – С. 25.

Публикации в других изданиях РИНЦ:

4. **Хомяков О.О.** Распознавание символьной информации на маркировке автоматических выключателей / В.С. Панищев, О.О. Хомяков // Информационные технологии и математическое моделирование систем. Одинцово – 2019. – С. 106-109.

5. **Хомяков О.О.** Улучшение качества распознавания маркировки автоматических выключателей для частных случаев / В.С. Панищев, О.О. Хомяков // Лучший исследовательский проект. Петрозаводск – 2020. – С. 169-174.

6. **Хомяков О.О.** Сравнение алгоритмов обработки данных в задаче распознавания символьной маркировки / В.С. Панищев, О.О. Хомяков, М.Р.Л. Мбуйи, Н.К. Мухамадсадилов // Национальный суперкомпьютерный форум. Переславль-Залесский – 2021.

7. **Хомяков О.О.** Определение производителя продукции с помощью технологии оптического распознавания символов на примере молочной продукции / О.О. Хомяков, В.С. Панищев, Н.К. Мухамадсадилов, К.С. Калугин // Всероссийская научная конференция Регионы России. Муром – 2022. – С. 334-335.

8. **Хомяков О.О.** Развитие искусственного интеллекта в цифровом мире: философские аспекты. / О.О. Хомяков // Исторические, философские, методологические проблемы современной науки. Курск – 2022. – С. 221-227

9. **Хомяков О.О.** Система распознавания характеристик продукции с использованием текстовой маркировки / О.О. Хомяков, А.И. Безмельцев, А.А. Савоськин // Актуальные вопросы современной науки: теория и практика научных исследований. Пенза – 2022. – С. 62-64

10. **Хомяков О.О.** Увеличение точности классификации характеристик продукции с использованием алгоритмов анализа структуры документа / В.С. Панищев, О.О. Хомяков // Оптико-электронные приборы и устройства в системах распознавания образов и обработки изображений. сборник материалов XVII Международной научно-технической конференции. Курск – 2023. – С. 181-182.

11. **Хомяков О.О.** Увеличение пропускной способности алгоритма распознавания символов в видеопотоке при условии наличия искажений // Наука и инновации – современные концепции. Москва. – 2024. – С. 129-132.

Патенты, свидетельства об интеллектуальной собственности:

12. **Хомяков О.О.** Устройство для распознавания символьной информации маркировки продукции / Хомяков О.О., Панищев В.С., Савоськин А. А., Безмельцев А. И. / заявка № 2024113679 от 21.05.2024, решение о выдаче патента от 26.09.2024.