

На правах рукописи



Воскресенский Евгений Михайлович

**Параметрическая и структурная адаптация систем распознавания  
текстовых меток на видеоизображениях**

Специальность 05.13.01 – Системный анализ, управление и обработка  
информации (промышленность)

Автореферат диссертации на соискание ученой степени  
кандидата технических наук

Рыбинск – 2010

Работа выполнена в Институте менеджмента и информационных технологий (филиале) Санкт-Петербургского государственного политехнического университета в г. Череповце

Научный руководитель кандидат технических наук  
Царев Владимир Александрович

Официальные оппоненты: доктор технических наук, профессор  
Юдин Виктор Васильевич

кандидат технических наук  
Юдина Ольга Вадимовна

Ведущая организация Институт систем обработки  
изображений РАН, г. Самара

Защита состоится 7 июля 2010 г. в 12 часов на заседании диссертационного совета Д 212.210.04 в Рыбинской государственной авиационной технологической академии им. П. А. Соловьева по адресу: 152934, г. Рыбинск, Ярославская область, ул. Пушкина, 53.

С диссертацией можно ознакомиться в библиотеке Рыбинской государственной авиационной технологической академии им. П. А. Соловьева

Автореферат разослан « 4 » июня 2010 г.

Ученый секретарь  
диссертационного совета

Конюхов Б. М.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

### Актуальность темы

Системы распознавания текстовых меток на видеоизображениях (СРТМ), используемые в составе различных информационно-управляющих систем, применяются для оптической идентификации (как правило, движущихся) объектов контроля, имеющих регистрационные надписи, например, промышленных изделий, упакованных продуктов, наземных транспортных средств, грузовых контейнеров, денежных купюр и т.п. Целью применения таких систем является избавление человека от рутинных операций, а также решение задач, с которыми он физически не способен справиться в режиме реального времени (конвейерная обработка изделий, поиск в базе данных и т.п.).

СРТМ являются частным случаем OCR-систем, представленных системами распознавания текста различного происхождения (печатного, рукопечатного, рукописного). Методы оптического распознавания текста описаны в работах таких российских и зарубежных ученых как Горский Н.Д., Арлазаров В.Л., Ян Д.Е., Nishida H., Impedovo S. и др. Достаточно большое количество научных публикаций посвящено описанию различных прикладных СРТМ. На основе этих публикаций можно сделать вывод, что в настоящее время сложился общепринятый подход к представлению алгоритмической части СРТМ в виде многоуровневой системы эвристических алгоритмов анализа изображений и распознавания графических образов.

Как правило, алгоритмы СРТМ, используемых в промышленности и на транспорте, обладают множеством параметров, настройка которых позволяет адаптировать систему к заданным условиям эксплуатации при первичном внедрении СРТМ. Адаптация СРТМ также необходима при существенных изменениях условий ее эксплуатации (оптической схемы, освещения, фона и пр.) и при изменении требований пользователя к показателям эффективности системы. Эффективность СРТМ принято характеризовать, с одной стороны, «качеством» распознавания (на входных видеоизображениях как содержащих образ текстовой метки, так и не содержащих такого образа). С другой стороны – длительностью рабочего цикла по обработке одного видеокдра, которая часто ограничена требованием распознавания в реальном времени. В настоящее время эффективность большинства эксплуатируемых СРТМ зачастую не соответствует современным требованиям, что обусловлено не только несовершенством используемых алгоритмов, но и нередко неудачно настроенными параметрами.

Количество параметров СРТМ, как правило, довольно велико. Например, распространенные российские системы распознавания автомобильных номеров обладают более чем тремя десятками только открытых для пользователя параметров используемых алгоритмов. Зависимость значений критериев эффективности СРТМ от значений настраиваемых параметров носит сложный алгоритмический характер, что исключает применение аналитических методов вычисления оптимальных параметров. Кроме того, однократное вычисление значений критериев эффективности СРТМ при заданных значениях параметров

алгоритмов нередко требует существенных вычислительных затрат времени, что затрудняет применение известных методов оптимизации. В результате в настоящее время настройка СРТМ производится, в основном, вручную и приблизительно. Такой подход требует от настройщика системы высокой квалификации, сравнимой с квалификацией разработчика, и не гарантирует нахождение оптимальных параметров.

С учетом того факта, что количество ежегодных внедрений отдельных видов СРТМ в России уже измеряется тысячами и непрерывно растет, проблема обеспечения эффективности функционирования таких систем и автоматизации процесса их внедрения является весьма актуальной.

### **Цель работы**

Целью диссертационной работы является повышение эффективности функционирования систем распознавания текстовых меток на видеоизображениях путем совершенствования процессов разработки и адаптации таких систем с учетом условий эксплуатации.

### **Основные задачи**

Для достижения указанной цели в работе поставлены и решены следующие основные задачи:

- Разработка математической модели СРТМ, представляющей СРТМ в виде множественной композиции эвристических алгоритмов.
- Разработка на базе предложенной математической модели методов параметрической адаптации СРТМ к заданным условиям эксплуатации.
- Разработка способа повышения эффективности СРТМ за счет ее структурной адаптации, не требующей модификации используемых алгоритмов.
- Апробация предложенных модели и методов на примере оптоэлектронной системы идентификации объектов подвижного состава железнодорожного транспорта.

**Предмет исследования** – методы и средства параметрической и структурной адаптации алгоритмического обеспечения СРТМ.

### **Методы исследования**

Для решения поставленных задач использованы методы теории вероятностей и математической статистики, основы теории оптимизации и принятия решений, методы обработки изображений, математический аппарат теории распознавания образов.

### **Научная новизна**

В работе выделен и исследован класс иерархических эвристических алгоритмов анализа изображений и распознавания графических образов, лежащих в основе большинства современных СРТМ. Предложен набор моделей, методов и программных средств, обеспечивающих повышение эффективности функционирования СРТМ за счет совершенствования процессов разработки и внедрения таких систем. Научной новизной обладают:

1. Математическая модель СРТМ, обеспечивающая описание и исследование СРТМ в виде множественной композиции базовых специализированных эвристических алгоритмов анализа видеоизображений. Модель включает в себя вероятностные критерии эффективности всей СРТМ и критерии

эффективности отдельных составляющих алгоритмов и устанавливает зависимость между такими критериями.

2. Метод параметрической адаптации СРТМ. Метод основан на модели из п. 1 и заключается в декомпозиции СРТМ на алгоритмы-компоненты эвристического анализа изображений, статистической оценке значений вероятностных критериев эффективности системы и оценке затрат времени на анализ системой единицы входных данных. Он позволяет обеспечить практическую решаемость задачи вычисления квазиоптимальных (оптимальных в рамках заданных экзаменационной последовательности изображений и дискретизации значений параметров) значений параметров большинства существующих СРТМ по сравнению с известными способами их настройки.
3. Метод параметрической адаптации подсистемы принятия решений СРТМ. Метод основан на модели из п. 1 и заключается в декомпозиции СРТМ на алгоритмы-компоненты, формирующих списки альтернативных решений на промежуточных этапах анализа входного видеоизображения. Метод позволяет обеспечить практическую решаемость задачи вычисления квазиоптимальных (оптимальных в рамках заданных экзаменационной последовательности изображений и дискретизации значений параметров) значений параметров подсистемы принятия решений, управляющей размерами списков альтернативных решений.
4. Способ структурной адаптации СРТМ, обеспечивающий без сопутствующей модификации образующих СРТМ алгоритмов уменьшение среднего времени анализа последовательности входных видеоизображений как содержащих, так и не содержащих образы текстовых меток движущихся объектов контроля.

#### **Практическая ценность**

1. Предложена методика оценки затрат времени на подготовку экзаменационных данных и на вычисление оптимальных параметров алгоритмов СРТМ, на основе которых производится целенаправленный выбор способа декомпозиции СРТМ.
2. Предложена методика кластеризации экзаменационной выборки видеоизображений, позволяющая сопоставить подмножествам видеоизображений подходящие квазиоптимальные наборы параметров.
3. Разработаны алгоритмы и произведена параметрическая адаптация модуля распознавания системы идентификации объектов подвижного состава железнодорожного транспорта «ARSCIS», разработка которой выполнялась на базе ИМИТ СПбГПУ и научно-производственной компании «Малленом» (г. Череповец) по заказам администрации станции Череповец Северной Железной Дороги, ОАО «Лукойл-Ухтанефтепереработка», ОАО «Лукойл-Волгограднефтепереработка».
4. Разработаны программные инструментальные средства параметрической адаптации алгоритмов СРТМ, используемые в научно-производственной компании «Малленом» при разработке оптоэлектронных систем контроля, среди которых: семейство систем идентификации автотранспортных

средств «Автомаршал», системы контроля печатной продукции, системы распознавания текстовой маркировки на стальных слябах, трубах и др.

5. Результаты исследований используются в учебном процессе ИМИТ СПбГПУ в рамках дисциплины «Основы теории распознавания образов» для специальности 230105 – Программное обеспечение вычислительной техники и автоматизированных систем.

### **Апробация работы**

Основные положения диссертации докладывались и обсуждались на всероссийской научно-практической конференции «Образование. Наука. Бизнес. Особенности регионального развития и интеграции» (Череповец, 2004-2008 гг.), Межвузовском конкурсе-конференции «Технологии Microsoft в теории и практике программирования» (Санкт-Петербург, 2005 г., 2007 г., 2008 г.), IX международной научной конференции «Распознавание образов и анализ изображений: новые информационные технологии» (Йошкар-Ола, 2007), международной научной конференции «Распознавание образов и обработка информации 2007» (Минск, 2007), всероссийской научной конференции студентов и аспирантов «Молодые исследователи – регионам» (Вологда, 2008 г.), Вторых ежегодных смотрах-сессиях аспирантов и молодых ученых Вологодской области по отраслям наук (Вологда, 2008 г.), IX международной научной конференции «Распознавание образов и анализ изображений: новые информационные технологии» (Нижний Новгород, 2008), а также на научных семинарах научно-технической лаборатории систем технического зрения и экспертных систем ИМИТ СПбГПУ, кафедры ИУС ФТК СПбГПУ и кафедры МПО ЭВС РГАТА им. П.А. Соловьева.

### **Публикации**

По материалам диссертации опубликовано 16 печатных работ (из них 4 в научных изданиях, рекомендованных ВАК, одна монография).

### **На защиту выносятся:**

1. Математическая модель СРТМ.
2. Метод параметрической адаптации СРТМ.
3. Метод параметрической адаптации подсистемы принятия решений СРТМ.
4. Способ структурной адаптации СРТМ.

### **Структура и объем работы**

Диссертационная работа состоит из введения, четырех глав, заключения и приложения. Объем работы составляет 181 страницу. Работа содержит 32 рисунка, 9 таблиц, список использованной литературы из 86 наименований, 5 приложений на 18 страницах.

## **СОДЕРЖАНИЕ РАБОТЫ**

*Во введении* обоснована актуальность работы, сформулирована цель и поставлены задачи проводимых исследований, определены научная новизна и практическая значимость результатов, приведены сведения о публикациях и апробации полученных результатов.

**В первой главе** дается общая характеристика СРТМ как особого подкласса OCR-систем распознавания текста, а также определяется круг проблем, связанный с разработкой и адаптацией СРТМ.

СРТМ включает в себя ЭВМ, средства ввода видеоданных, а также алгоритмическую часть, обычно реализуемую в виде единого программного модуля распознавания. В большинстве случаев он представляется в виде композиции алгоритмов, решающих отдельные логические подзадачи распознавания текстовых меток. Алгоритм локализации (АЛ) выделяет прямоугольную зону видеоизображения, потенциально содержащую образ текстовой метки. Алгоритм сегментации (АС) выделяет на зоне множество сегментов, предположительно являющихся образами символов метки. Алгоритм распознавания (АР) сопоставляет каждому сегменту код символа либо отклоняет сегмент. Алгоритм формирования решений (АФР) составляет из распознанных сегментов с учетом их взаимного расположения коды текстовых меток. Часто, чтобы не потерять истинное решение, между промежуточными этапами передаются списки вариантов решений с некоторыми эвристическими оценками их правдоподобия, а итоговое решение по изображению выбирает алгоритм принятия решений (АПР).

В ходе своей работы СРТМ приходится, как правило, анализировать видеоизображения как содержащие образ текстовой метки (информативные), так и не содержащие (неинформативные). При этом с учетом плохой формализации задачи анализа изображений, решаемой СРТМ, разработчики таких систем вынуждены использовать эвристические алгоритмы локализации, сегментации и распознавания. В связи с этим эффективность анализа информативных изображений характеризуется вероятностью правильного распознавания текстовой метки  $P_{right}$  и вероятностью ошибки  $P_{err}$ , а неинформативных – вероятностью ложного решения  $P_{false}$ . Четвертым критерием является среднее время анализа отдельного изображения  $T$ , влияющее на эффективность функционирования модуля в режиме реального времени. Таким образом, СРТМ характеризуется критериями:

$$Q = (P_{right}, P_{err}, P_{false}, T). \quad (1)$$

Статистические оценки значений критериев  $Q$  характеризуют эффективность алгоритма  $A$ , осуществляющего полный цикл обработки входного видеоизображения в СРТМ с параметрами  $u = \{u_1, u_2, \dots, u_k\}$  в заданных условиях эксплуатации. Каждое место внедрения характеризуется аппаратным обеспечением СРТМ  $h$  и полученной по месту внедрения экзаменационной выборкой видеоизображений  $B$ . Таким образом, основным средством обеспечения заданной эффективности  $\hat{Q}(A, B, u, h) = (\hat{P}_{right}, \hat{P}_{err}, \hat{P}_{false}, \hat{T})$  является настройка параметров СРТМ.

Обычно множество настраиваемых параметров довольно велико (до нескольких десятков). Зависимость  $\hat{Q}(u)$  носит сложный алгоритмический характер, что исключает применение аналитических методов вычисления оптимальной конфигурации  $u^*$ . Кроме того, однократное вычисление оценки

$\hat{Q}$  обычно связано с существенными вычислительными затратами. Все это приводит к тому, что на современном этапе настройка параметров модуля распознавания, как правило, выполняется вручную самими разработчиками. Это не гарантирует нахождение оптимальной конфигурации и требует существенного объема трудовых затрат при адаптации СРТМ, поэтому актуальной задачей является создание методов, позволяющих рассчитывать оптимальные в заданных условиях эксплуатации значения параметров СРТМ без привлечения разработчиков системы.

Кроме того, эффективность современных СРТМ далека от идеальной по причине несовершенства современных методов анализа изображений. Например, для современных российских систем распознавания государственных регистрационных знаков автомобилей вероятность правильного распознавания в благоприятных условиях, как правило, составляет 0.8-0.9. Это обуславливает необходимость применения в СРТМ так называемой «мягкой схемы» принятия решения, в которой на промежуточных этапах анализа данных формируется не единственное решение, а списки альтернативных решений с оценками их правдоподобия. Такой подход дает возможность «не потерять» истинное решение на промежуточных этапах, но при этом приводит к значительному увеличению времени полного цикла анализа входного изображения. Таким образом, актуальной является также задача совершенствования подсистемы принятия решений в составе СРТМ.

**Вторая глава** посвящена моделированию алгоритмической части СРТМ. Современные методы оптического распознавания текстовой информации достаточно подробно описаны в литературе, однако мало исследованной остается задача синтеза модуля распознавания с заданными характеристиками из базовых алгоритмов с учетом условий эксплуатации. В связи с этим выделен класс  $A^C$  иерархических эвристических алгоритмов распознавания текстовых меток, которому соответствует большинство описанных в литературе СРТМ:

$$\begin{aligned} & 1) b \in \{Inf, Empty\}, \\ A(b) \in A^C & \Leftrightarrow 2) \exists Correct(A(b)) \in \{true, false\}, \\ & 3) A = A_1 \circ A_2 \circ \dots \circ A_n, \exists Correct_i(A_i(b_i)) \in \{true, false\}, \end{aligned} \quad (2)$$

где  $b$  – вход алгоритма  $A$ ;  $Inf$  – класс информативных входов, по которым алгоритм  $A$  должен возвращать некоторое непустое решение;  $Empty$  – класс неинформативных входов, по которым алгоритм  $A$  должен возвращать пустое решение;  $Correct$  – функция, определяющая, является ли решение алгоритма  $A$  по входу  $b$  правильным;  $A_i$  – алгоритм-компонент алгоритма  $A$ ,  $i=1..n$ ;  $Correct_i$  – функция, определяющая, является ли решение алгоритма-компонента  $A_i$  по входу  $b_i$  правильным. Важно, что значения функций  $Correct$  и  $Correct_i$  могут быть определены экспертом практически на любом наборе данных. При этом в качестве эксперта при распознавании текстовых меток на видеоизображениях может выступать любой взрослый человек.

Разработана математическая модель СРТМ, отождествляемой с некоторым  $A \in A^C$ . Модель позволяет установить функциональную зависимость значений



критериев эффективности  $Q$  от значений критериев эффективности алгоритмов-компонентов. Кроме того, она задает способы композиции алгоритмов-компонентов, образующих  $A$ , из базовых (неделимых) алгоритмов-компонентов, в число которых входят АЛ, АС, АР, АФР, АПР. Два базовых алгоритма являются смежными, если выход одного из них (вышележащего) является входом другого (нижележащего). В работе предложены спецификации входных и выходных данных для базовых алгоритмов СРТМ в целях обеспечения возможности создания их композиций. Композиция смежных базовых алгоритмов образует алгоритм-компонент  $A_i$ , алгоритмы-компоненты (далее алгоритмы) в свою очередь образуют  $A$ :

$$A_i = A_k^b \circ A_{k+1}^b \circ \dots \circ A_l^b, \quad A = A_1 \circ A_2 \circ \dots \circ A_n, \quad (3)$$

где  $A_k^b, A_{k+1}^b, \dots, A_l^b$  – смежные базовые алгоритмы,  $1 \leq k \leq l \leq m$ . Таким образом, существуют различные варианты декомпозиции алгоритма  $A$ . Алгоритм  $A_i$  характеризуется критериями  $Q^i = (P(C_i), P(F_i), T_i)$ , где  $C_i$  – «выход алгоритма  $A_i$  правильный»,  $T_i$  – затраты времени на анализ единицы входных данных. Если алгоритм способен давать пустой ответ, то дополнительным критерием может выступать вероятность события  $F_i$  – «выход алгоритма  $A_i$  не пустой». Статистические оценки  $\hat{P}(C_i)$ ,  $\hat{P}(F_i)$ ,  $\hat{T}_i$  рассчитываются по массиву прецедентов, созданному экспертом, в котором экзаменационные примеры представляют собой решения композиции вышележащих относительно  $A_i$  алгоритмов по входам экзаменационной последовательности  $B$  алгоритма  $A$ .

Каждому исходу анализа видеоизображения соответствует множество последовательностей событий  $C_i, \bar{C}_i, F_i, \bar{F}_i$ , определяющих связь  $Q$  и  $Q^i$ :

$$P_{right} = P(C_1 | I)P(C_2 | IC_1) \dots P(C_n | IC_1 C_2 \dots C_{n-1}), \quad (4)$$

$$P_{err} = 1 - P_{right} - P_{empty}, \quad P_{empty} = P(\bar{F}_1 | I) + P(\bar{F}_2 | IF_1) + \dots + P(\bar{F}_n | IF_1 F_2 \dots F_{n-1}), \quad (5)$$

$$P_{false} = P(F_1 | N)P(F_2 | NF_1) \dots P(F_n | NF_1 F_2 \dots F_{n-1}). \quad (6)$$

Здесь события  $I$  и  $N$  указывают на различные (информативные и неинформативные) входы. Среднее время анализа входных данных  $T$  рассчитывается с учетом вероятности пустого выхода каждого алгоритма-компонента:

$$T = T_1 + T_2 P(F_1) + T_3 P(F_2 | F_1) + \dots + T_n P(F_n | F_1 F_2 \dots F_{n-1}). \quad (7)$$

**В третьей главе** предложены методы параметрической и структурной адаптации СРТМ. Обычно разработчик СРТМ способен задать ограниченное множество допустимых значений каждого параметра  $u_j \in \{v_{j,1}, \dots, v_{j,m}\}$  используемых алгоритмов-компонентов, и все комбинации значений параметров образуют конечное множество конфигураций  $Conf$ . Поскольку критерии  $\hat{Q}$  носят статистический характер, то оптимальные в рамках  $Conf$  и  $B$  параметры в общем случае являются квазиоптимальными. Задача вычисления оптимальных параметров  $u^*$  при заданных  $A, B, h, Conf$

формулируется как задача максимизации  $\hat{P}_{right}$  с учетом ограничений на значения всех критериев:

$$u^* = \arg \max_{u \in Conf} (\hat{P}_{right}(u)), \hat{P}_{right}(u^*) \geq P_{right}^{tr}, \hat{P}_{err}(u^*) \leq P_{err}^{tr}, \quad (8)$$

$$\hat{P}_{false}(u^*) \leq P_{false}^{tr}, \hat{T}(u^*) \leq T^{tr}.$$

Иерархическая структура СРТМ позволяет производить отдельную настройку алгоритмов-компонентов. Способ композиции из базовых алгоритмов алгоритмов-компонентов, подлежащих отдельной настройке, предлагается выбирать с учетом затрат времени на составление массива прецедентов и на длительность процесса вычисления  $u^*$ . При ограниченном количестве значений каждого параметра и заданной конфигурации вышележащих алгоритмов  $\{u^1, u^2, \dots, u^{i-1}\}$  задачу адаптации алгоритма  $A_i$  можно свести к формированию подмножества  $Conf^{i,filter}$ :

$$Conf^{i,filter} = \{u^i \mid P(C_i \mid u^i) \geq C_i^{tr}, T_i(u_i) \leq T_i^{tr}\}. \quad (9)$$

где пороги  $C_i^{tr}$ ,  $T_i^{tr}$  задаются с учетом (8). Множество  $Conf^{i,filter}$  может быть довольно большим. Его сокращение до множества оптимальных по Парето конфигураций сопряжено с возможностью потерять оптимальную конфигурацию  $u^*$  СРТМ. Это обусловлено тем, что выходы  $A_i$  образуют экзаменационную последовательность для  $A_{i+1}$ , состав которой влияет на эффективность  $A_{i+1}, A_{i+2}, \dots, A_n$ . Поэтому оптимальная по  $\hat{P}(C_i)$  и  $\hat{T}_i$  конфигурация  $A_i$  не обязательно является составляющей оптимальной конфигурации  $u^*$  алгоритма  $A$  (глобально оптимальной). Для выявления всех потенциально глобально оптимальных конфигураций алгоритма  $A_i$  предлагается использовать прецедентный критерий – вектор  $V = (V_1, V_2, \dots, V_{sz})$ ,  $V_j \in \{0,1\}$ ,  $V_j = 1 \Leftrightarrow A_i(Inp_j) \sim E_j$ , иначе  $V_j = 0$ . Здесь  $E_j$  – эталонный результат анализа экзаменационного входа  $Inp_j$ , « $\sim$ » – знак эквивалентности,  $sz$  – количество экзаменационных примеров. Необходимое условие глобальной оптимальности (НУГО) конфигурации  $u^i$ , которой соответствует прецедентный критерий  $V^i$ , следующее: конфигурация  $u^i$  должна быть оптимальна по Парето по векторному критерию  $q^i = (V_1^i, V_2^i, \dots, V_{sz}^i, \hat{T}_i)$  на множестве  $Conf^i$ . Тогда результатом настройки алгоритма  $A_i$  является сформированное из  $Conf^{i,filter}$  множество удовлетворяющих НУГО конфигураций  $Conf^{i,opt}$ . Если  $Conf^{i,filter}$  не пусто, то формируется множество оптимальных по Парето конфигураций:

$$Conf^{i,opt} = \{u^i \mid u^i \in Conf^{i,filter}, Pareto(q^i(u^i)) = true\}, \quad (10)$$

где  $Pareto$  – функция, определяющая, является ли конфигурация  $u^i$ , характеризующаяся векторным критерием  $q^i(u^i)$ , оптимальной по Парето на множестве  $Conf^{i,filter}$ .

Метод вычисления  $u^*$  заключается в последовательном, то есть от вышележащих алгоритмов к нижележащим, вычислении множеств  $Conf^{i,opt}$ , образующих в результате множество потенциально оптимальных конфигураций  $Conf^{opt}$  (если  $\forall i \in 1..n \Rightarrow Conf^{i,opt} \neq \emptyset$ ). Для элементов  $Conf^{opt}$  вычисляются критерии эффективности, исходя из которых выбирается  $u^*$ .

Настройка параметров представленного в работе модуля распознавания идентификационных номеров объектов подвижного состава железнодорожного транспорта предложенным методом заняла менее 16 часов. При тех же условиях настройка методом перебора потребовала бы несколько лет. Настройка модуля распознавания, входящего в состав оптоэлектронной системы «ARSCIS», позволила повысить значение статистической оценки вероятности правильной идентификации ж/д вагонов по месту внедрения системы на станции Череповец на 3.4% по сравнению с ручной настройкой, произведенной разработчиками системы, без ухудшения значений остальных статистических критериев.

Как было отмечено выше, в целях повышения эффективности СРТМ применяется «мягкая схема» принятия решений. В работе предложен алгоритм усечения списков (АУС) вариантов решений, принимающий решения об усечении списков до приемлемого размера либо их отклонении. Параметр АУС  $K_i$  ограничивает длину СВР алгоритма  $A_i$ . Решение о принятии или отклонении СВР выполняется по результатам сравнения осредненной оценки правдоподобия элементов усеченного списка с порогом  $E_i$ .

Задачей настройки подсистемы принятия решений СРТМ является поиск оптимальных при заданных  $A$ ,  $B$ ,  $h$  и конечном  $Conf^{AVC}$  значений порогов  $u^{AVC} = \{K, E\}$ . Часто требуется производить настройку АУС отдельно от параметров алгоритмов-компонентов (при фиксированной конфигурации  $u$ ), для решения этой задачи разработан следующий метод. Пусть  $A^{AVC} \in A^C$ , при этом  $A^{AVC} = \{A_1, A_2, \dots, A_n\}$ , где каждый  $A_i$  возвращает СВР. Для расчета оптимальных значений порогов  $u^{AVC}$  параметры АУС устанавливаются в такие значения  $K_i^{\max}$  и  $E_i^{\min}$ , при которых СВР пропускаются и не усекаются. В итоге выходом  $A$  является полный список вариантов решений, сформированных по результатам анализа видеоизображения. Далее по множеству  $B$  статистически оцениваются  $\hat{P}_{right}(u^{AVC})$ ,  $\hat{P}_{err}(u^{AVC})$ ,  $\hat{P}_{false}(u^{AVC})$ .

Каждое решение  $d_j$  из полного списка решений  $d = \{d_1, d_2, \dots, d_k\}$  алгоритма  $A$  образовано последовательностью промежуточных решений алгоритмов-компонентов:  $d_j^1, d_j^2, \dots, d_j^{n-1}$ ,  $d_j = d_j^n$ . Каждое промежуточное решение  $d_j^i$  занимает определенную позицию  $p(d_j^i)$  в соответствующем СВР, характеризующемся оценкой правдоподобия  $e(d_j^i)$ , а его корректность устанавливается с использованием массива прецедентов. Характеристики решений  $d_j$ :

$$Pos(d_j) = \{p(d_j^1), p(d_j^2), \dots, p(d_j^n)\}, \quad (11)$$

$$Eval(d_j) = \{e(d_j^1), e(d_j^2), \dots, e(d_j^n)\}, \quad (12)$$

позволяют установить исход анализа примера  $B_j$  при любой конфигурации  $u^{AVC}$  и сформировать множество:

$$Conf^{AVC, filter} = \left\{ u^{AVC} \mid u^{AVC} \in Conf^{AVC}, \hat{P}_{right}(u^{AVC}) \geq P_{right}^{tr}, \hat{P}_{err}(u^{AVC}) \leq P_{err}^{tr}, \hat{P}_{false}(u^{AVC}) \leq P_{false}^{tr} \right\}. \quad (13)$$

Из полученного множества удаляются конфигурации, не удовлетворяющие ограничению  $T^{tr}$ :

$$Conf^{AVC, opt} = \{u^{AVC} \mid u^{AVC} \in Conf^{AVC, filter}, T(u^{AVC}) \leq T^{tr}\}, \quad (14)$$

где  $T(u^{AVC})$  либо измеряется экспериментально, либо оценивается предложенным в работе способом, который основан на статистическом расчете среднего количества вызовов каждого алгоритма-компонента на одно видеоизображение при заданном  $u^{AVC}$ . Экспериментальной базой для исследования метода послужил модуль распознавания идентификационных номеров объектов подвижного состава железнодорожного транспорта, описанный в гл. 4.

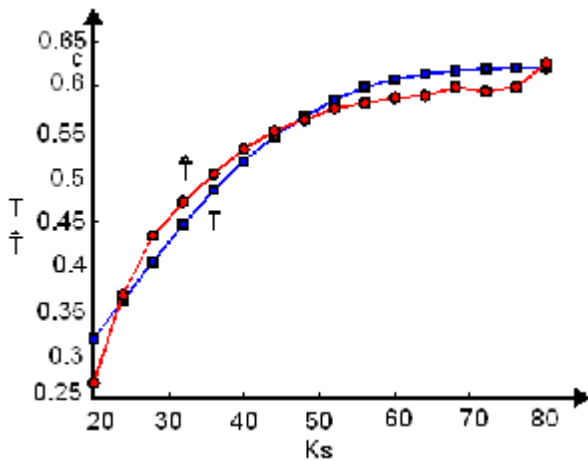


Рис. 1. Оцененные и измеренные средние затраты времени на анализ видеоизображения

На рисунке 1 представлены графики зависимости времени анализа видеоизображения от порога на размер СВР АС  $K_s = 20..45$ . График с квадратными маркерами соответствует зависимости  $\hat{T}$ , вычисленной предложенным методом, второй график отображает измеренное время  $T$ . Рисунок показывает, что оценка  $\hat{T}$  достаточно близка к  $T$ , что говорит о возможности использования предложенного способа в процессе настройки алгоритмов  $A^{AVC}$ .

В ходе эксперимента настройка предложенным методом параметров АУС указанной системы при  $|Conf^{AVC}| = 40000$  заняла около 20 минут. При аналогичных условиях непосредственное «переборное» вычисление значений критериев эффективности каждой конфигурации займет 31 сутки.

В главе 3 также предложен способ структурной адаптации СРТМ, функционирующих в режиме реального времени, направленный на снижение затрат времени  $T$ . Для этого в состав СРТМ включается алгоритм, синтезированный из АС и АР, который производит классификацию локализованных зон на содержащие и не содержащие текстовую метку

посредством анализа их фрагментов. Обосновано, что отсеивание неинформативных зон позволяет сократить средние затраты времени на анализ последовательности видеоизображений, существенная доля которых не содержит текстовых меток.

**В четвертой главе** дано описание разработанных алгоритмов модуля распознавания системы идентификации объектов подвижного состава железнодорожного транспорта и результатов экспериментов по их адаптации с использованием предложенных в гл. 3 методов.

Алгоритм локализации на первом этапе производит наложение фильтра на исходное изображение  $I_{m \times n}$ , что дает матрицу  $E_{m \times n}$ , характеризующую уровень изменения яркости в окрестности каждого пикселя  $I_{i,j}$ . На втором этапе производится анализ  $E_{m \times n}$  с целью локализации зон  $Z_{h \times w}$ , характеризующихся повышенным средним значением элементов  $E_{i,j}$  относительно окружающей зону элементов  $E_{i,j}$ . Оценкой правдоподобия зоны является ее среднее значение  $E_{i,j}$ . Зона анализируется алгоритмом сегментации, основанном на применении гистограммы яркости. Алгоритм распознавания реализован в виде комплекса многослойных перцептронов, производящих двухэтапную (класс символа, подтверждение) классификацию сегментов. Алгоритм формирования решений, основанный на поиске в глубину, формирует из распознанных сегментов с учетом их координат варианты текстовых меток. В качестве итогового выбирается решение с максимальной суммарной оценкой правдоподобия сегментов, образующих решение.

В работе приведен пример требований технического задания к разработанной СРТМ и дано подробное описание процесса параметрической адаптации системы, а также представлены результаты экспериментов по ее структурной адаптации. В состав разработанного модуля распознавания был включен алгоритм классификации зон (АКЗ), производящий сегментацию фрагментов зон, распознавание сегментов и принятие решения о классификации на основании количества распознанных сегментов (табл. 1). Экзаменационная последовательность состояла из 550 кадров с изображениями железнодорожных цистерн и вагонов с видимым идентификационным номером на борту, а также 688 изображений цистерн и вагонов без номера в кадре.

Таблица 1

Результаты экспериментов по сравнению эффективности типовой СРТМ и ее версии с алгоритмом классификации зон

Критерии эффективности	Без АКЗ	АКЗ
Время анализа инф. кадра, сек.	0.60	0.63
Время анализа неинф. кадра, сек.	0.71	0.25
Среднее время анализа кадра (23% информативных кадров), сек.	0.68	0.34

Эксперименты показали, что статистические показатели эффективности СРТМ с АКЗ при подобной структурной адаптации идентичны показателям

исходной СРТМ. В то же время использование АКЗ позволило сократить среднее время анализа видеоизображения примерно на 50%, с учетом того, что в заданных условиях около 23% кадров содержат образ номера. Доверительный интервал для  $\hat{P}_{right} = 0.65$  при надежности оценки  $g = 0.95$  равен  $(0.62, 0.68)$ , оценка  $\hat{t} = 0.60$  характеризуется  $P(|t - \hat{t}| \leq 0.02) = 0.923$ .

**В заключении** подводятся итоги работы, формулируются основные выводы.

### ОБЩИЕ ВЫВОДЫ ПО РАБОТЕ

Установлено, что в настоящее время большинство описанных в литературе систем распознавания текстовых меток на видеоизображениях являются многопараметрическими системами и имеют однотипную структуру, представленную композицией алгоритмов локализации, сегментации, распознавания и принятия решений.

Такие системы могут быть описаны представленной в работе моделью СРТМ, и к ним применимы предложенные автором методы параметрической и структурной адаптации СРТМ с учетом заданных условий эксплуатации. Для большинства современных СРТМ данные методы позволяют:

- Находить за приемлемое время оптимальный (в рамках заданной экзаменационной последовательности видеоизображений и дискретизации значений параметров) набор значений параметров системы.
- Снизить трудозатраты на параметрическую адаптацию системы.

На практике это дает возможность повышать эффективность не только самих СРТМ, но и процессов их разработки и внедрения, поскольку предложенные методы позволяют производить внедрение и вторичную настройку СРТМ без привлечения алгоритмистов-разработчиков, снижая трудоемкость процесса внедрения и его длительность.

Предложенный в работе способ повышения эффективности функционирования СРТМ за счет их структурной адаптации позволяет существенно снизить среднее время анализа видеоизображений в тех условиях, когда значительная часть поступающих на вход СРТМ видеоизображений не содержит образ текстовой метки. При этом не требуется сопутствующая модификация базовых алгоритмов СРТМ.

### СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

В изданиях, рекомендованных ВАК:

1. **Воскресенский, Е. М.** Метод декомпозиции и вычисления показателей эффективности систем распознавания текстовых меток на видеоизображениях [Текст] / Е. М. Воскресенский, В. А. Царев // Научно-технические ведомости СПбГПУ. – 2008. – №3. – С. 140–145.
2. **Воскресенский, Е. М.** Метод оценки эффективности систем распознавания текстовых меток на сложном фоне с использованием дерева вероятностных характеристик [Текст] / Е. М. Воскресенский, В. А. Царев // Компьютерная Оптика. – 2008. – Т.32. – №3. – С. 283–290.

3. **Воскресенский, Е. М.** Методы управления процессом распознавания текстовых меток на изображениях [Текст] / Е. М. Воскресенский, В. А. Царев // Компьютерная Оптика. – 2008. – Т.32. – №4. – С. 413–416.
4. **Воскресенский, Е. М.** Метод параметрической оптимизации процесса принятия решений в системах распознавания текстовых меток на видеоизображениях [Текст] / Е. М. Воскресенский, В. А. Царев // Компьютерная Оптика. – 2009. – Т. 33. – №2. – С. 202–209.

В иных изданиях:

5. **Воскресенский, Е. М.** Моделирование и адаптация систем распознавания текстовых меток на видеоизображениях [Текст] : монография / Е.М. Воскресенский, В.А. Царев. – Череповец: ИНЖЭКОН-Череповец, 2009. – 154 с. – ISBN 978-5-902459-06-4.
6. **Воскресенский, Е. М.** Классификация фрагментов изображений в задаче поиска и локализации графических образов текстовых меток на сложном фоне [Текст] / Е. М. Воскресенский, В. А. Царев // Образование, наука, бизнес: особенности регионального развития и интеграции: Сборник трудов Всероссийской научно-методической конференции. – Череповец: ИМИТ СПбГПУ, 2004. – С. 262–270.
7. **Воскресенский, Е. М.** Программная реализация нейронной сети прямого распространения на языке С# и ее применение в задачах классификации изображений [Текст] / Е. М. Воскресенский, В. А. Царев // Технологии Microsoft в теории и практике программирования: Материалы межвузовского конкурса-конференции студентов, аспирантов и молодых ученых Северо-Запада. – СПб.: Изд-во Политехн. ун-та, 2005. – С.19–20.
8. **Воскресенский, Е. М.** Распознавание буквенно-цифровых символов с использованием нейронных сетей на примере изображений цифр регистрационных номеров транспортных средств [Текст] / Е. М. Воскресенский, Е. Н. Веснин // Образование, наука, бизнес: особенности регионального развития и интеграции: Сборник трудов Всероссийской научно-практической конференции. – Череповец: ИМИТ СПбГПУ, 2005. – С. 246–250.
9. **Воскресенский, Е. М.** Исследование процедур сегментации цифровых изображений идентификационных номеров железнодорожных вагонов [Текст] / Е. М. Воскресенский // Образование, наука, бизнес: особенности регионального развития и интеграции: Сборник трудов Всероссийской научно-практической конференции. – Череповец: ИМИТ СПбГПУ, 2006. – С. 283–286.
10. **Воскресенский, Е. М.** Методы и средства автоматизации разработки систем распознавания [Текст]/ Е. М. Воскресенский, В. А. Царев // Технологии Microsoft в теории и практике программирования: Материалы межвузовского конкурса-конференции студентов, аспирантов и молодых ученых Северо-Запада. – СПб.: Изд-во Политехн. ун-та, 2007. – С. 156–157.
11. **Voskresenskiy, E. M.** Design automation of text labels recognition systems [Text] / E. M. Voskresenskiy // Pattern Recognition and Information Processing (PRIP-2007): Proceedings of the Ninth International Conference. – Minsk: United

Institute of Informatics Problems of National Academy of Science of Belarus, 2007. – Vol. II. – P. 236–239.

12. **Voskresenskiy, E. M.** Text label recognition systems efficiency control using localization and segmentation algorithms output likelihood [Text] / E. M. Voskresenskiy, V. A. Tsarev // International Conference «Pattern Recognition and Image Analysis: New Information Technologies» (PRIA-8-2007): Conference Proceedings. – Yoshkar-Ola: Mari State Technical University, 2007. – Vol. 3. – P. 57–61.
13. **Воскресенский, Е. М.** Метод оценки эффективности систем распознавания текстовых меток на сложном фоне с использованием дерева вероятностных характеристик [Текст] / Е. М. Воскресенский, В. А. Царев // Технологии Microsoft в теории и практике программирования: Материалы межвузовского конкурса-конференции студентов, аспирантов и молодых ученых Северо-Запада. – СПб.: Изд-во Политехн. ун-та, 2008. – С. 193–194.
14. **Воскресенский, Е. М.** Метод расчета эффективности систем распознавания текстовых меток на видеоизображениях [Текст] / Е. М. Воскресенский // Сборник трудов всероссийской научной конференции студентов и аспирантов «Молодые исследователи – регионам». – Вологда: ВоГТУ, 2008. – С. 47–48.
15. **Voskresenskiy, E. M.** Methods of searching for quasioptimal configurations of text label recognition systems [Text] / E. M. Voskresenskiy // International Conference «Pattern Recognition and Image Analysis: New Information Technologies» (PRIA-9-2008): Conference Proceedings. – Nizhni Novgorod: N. I. Lobachevski State University of Nizhni Novgorod, 2008. – Vol. 2. – P. 307–311.
16. **Воскресенский, Е. М.** Модель системы распознавания текстовых меток на видеоизображениях с использованием событийного представления процесса анализа видеокадров [Текст] / Е. М. Воскресенский // Материалы II (вторых) ежегодных смотров-сессий аспирантов и молодых ученых Вологодской области по отраслям наук. – Вологда: ВоГТУ, 2008. – Т. 1. – С. 13–20.