

ПАРАМЕТРИЧЕСКАЯ И СТРУКТУРНАЯ АДАПТАЦИЯ МНОГОУРОВНЕВЫХ СИСТЕМ РАСПОЗНАВАНИЯ ТЕКСТОВЫХ МЕТОК НА ВИДЕОИЗОБРАЖЕНИЯХ

Воскресенский Е.М., Царёв В.А.

*Институт менеджмента и информационных технологий (филиал)
Санкт-Петербургского государственного политехнического университета*

Аннотация

В статье предложены новые методы параметрической и структурной адаптации многоуровневых систем распознавания текстовых меток на видеоизображениях, структура которых представляется в виде вертикальной иерархической композиции алгоритмов. Методы позволяют настраивать параметры алгоритмов анализа видеоизображений, а также параметры подсистемы принятия решений на промежуточных этапах анализа видеоизображений. Предложена методика повышения эффективности многоуровневых систем распознавания текстовых меток за счёт их структурной адаптации без сопутствующей модификации алгоритмов.

Ключевые слова: распознавание символов, адаптация систем распознавания.

Введение

В настоящее время в связи с ростом производительности вычислительной техники значительно расширяется область применения систем технического зрения, в частности, систем распознавания текстовых меток (СРТМ) на видеоизображениях. Обычно СРТМ функционируют в составе оптоэлектронных систем контроля и используются для идентификации объектов, имеющих регистрационные надписи, например, промышленных изделий, упакованных продуктов, наземных транспортных средств, грузовых контейнеров, денежных купюр. 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). Объединяет системы распознавания текста и СРТМ одинаковый объект анализа – графическое изображение, похожие объекты распознавания – текст и текстовые метки, и некоторые общие подзадачи анализа изображений. Вместе с тем СРТМ имеют ряд отличительных особенностей. Во-первых, системы распознавания текста обрабатывают изображения с относительно высоким разрешением, полученные путём сканирования бумажных носителей. СРТМ же, как правило, функционируют в составе оптоэлектронных систем контроля и анализируют последовательность видеокадров, обладающих относительно низким разрешением. Во-вторых, системы распознавания текста малоэффективны при распознавании текста на сложном, неоднородном фоне, присущем входным изображениям СРТМ. В-третьих, условия функционирования СРТМ сопряжены с множеством факторов, затрудняющих визуальное восприятие текстовых меток, например, солнечные блики, тени, дождь, снег, грязь и т.п. Системы распознавания текста, в свою очередь, успешно обрабатывают в

основном лишь качественные изображения, поэтому часто внесение шума на изображение используется как средство защиты текста от автоматического распознавания (но не от чтения человеком). В-четвёртых, работая в составе оптоэлектронных систем контроля, СРТМ должны обеспечивать заданный режим быстрого действия. К системам распознавания текста жёсткие требования такого рода, как правило, не предъявляются.

Достаточно большое количество публикаций посвящено описанию различных прикладных СРТМ, показывающих, что в настоящее время сложился общепринятый подход к представлению алгоритмической части СРТМ в виде многоуровневой системы алгоритмов.

Обычно алгоритмы используемых в промышленности и на транспорте СРТМ обладают множеством параметров, настройка которых позволяет адаптировать систему к заданным условиям эксплуатации при первичном внедрении СРТМ, при существенных изменениях условий эксплуатации (оптической схемы, освещения, фона и пр.) и при изменениях требований пользователя к показателям эффективности системы. Эффективность СРТМ принято характеризовать, с одной стороны, качеством распознавания, с другой – длительностью рабочего цикла, которая часто ограничена требованием распознавания в реальном времени. На современном этапе эффективность большинства эксплуатируемых систем далека от идеальной, что обусловлено не только несовершенством используемых алгоритмов, но и нередко неудачно настроенными параметрами.

Количество параметров СРТМ, как правило, довольно велико, например, распространённые российские системы распознавания автомобильных номеров обладают более чем тремя десятками параметров алгоритмов каждая. 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 значе-

ний 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 исполнителя высокой квалификации, сравнимой с квалификацией разработчика, и не гарантирует нахождение оптимальных параметров.

В статье рассматриваются многоуровневые системы распознавания текстовых меток, которые возможно представить в виде вертикальной иерархической системы алгоритмов, решающих отдельные подзадачи анализа видеоизображений. Выделен класс таких многоуровневых алгоритмов, предложена их математическая модель, на основе которой разработаны методы поиска оптимальных (в рамках заданной экзаменационной последовательности видеоизображений и дискретизации значений параметров) значений параметров алгоритмов таких систем. В качестве экспериментальной базы выбрана система распознавания идентификационных номеров объектов подвижного состава железнодорожного транспорта, описанная в [1].

1. Задача повышения эффективности систем распознавания текстовых меток на видеоизображениях

Главной составной частью СРТМ является её алгоритмическая часть – программный модуль распознавания. Обычно модуль анализирует видеоизображения как содержащие образ текстовой метки (информативные), так и не содержащие (неинформативные). 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 учётом 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).

Алгоритмы локализации, сегментации и распознавания, как правило, эвристические и не всегда дают правильное решение, в связи с чем эффективность анализа информативных изображений характеризуется вероятностью правильного распознавания текстовой метки P_{right} и вероятностью ошибки

P_{err} , а неинформативных – вероятностью ложного решения P_{false} . Четвёртым 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 функционирования алгоритма A 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 критериев \hat{Q} обычно связано с существенными вычислительными затратами. Всё это приводит к тому, что на современном этапе настройка параметров модуля распознавания, как правило, выполняется вручную самими разработчиками, исходя из знания алгоритмов, опыта и интуиции. Это не гарантирует нахождения оптимальной конфигурации и требует существенного объёма трудозатрат при каждом внедрении СРТМ, поэтому актуальной задачей является создание методов, позволяющих рассчитывать оптимальные в заданных условиях эксплуатации значения параметров СРТМ без привлечения разработчиков.

Кроме того, эффективность современных СРТМ далека от идеальной по причине несовершенства современных методов распознавания образов. Например, для современных российских систем распознавания государственных регистрационных знаков автомобилей вероятность правильного распознавания в благоприятных условиях, как правило, составляет 0,8-0,9. Данная проблема ставит задачу совершенствования алгоритмического обеспечения СРТМ, в особенности подсистемы принятия решений СРТМ.

2. Модель многоуровневой системы распознавания текстовых меток на видеоизображениях

Современные методы оптического распознавания текстовой информации достаточно подробно описаны в литературе, однако малоисследованной остаётся задача синтеза модуля распознавания с заданными характеристиками из готовых алгоритмов с учётом условий эксплуатации. Большинство существующих моделей систем распознавания являются

моделями алгоритмов классификации и не подходят для описания многоуровневых алгоритмов распознавания текстовых меток. В связи с этим выделен класс A^C многоуровневых эвристических алгоритмов распознавания текстовых меток, в который попадает большинство описанных в литературе СРТМ:

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

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.

Для СРТМ, описываемых классом A^C , разработана математическая модель, которая задаёт способы декомпозиции СРТМ и устанавливает функциональную зависимость критериев эффективности Q от критериев эффективности алгоритмов-компонентов. Два базовых (неделимых) алгоритма являются смежными, если выход одного из них (вышележащего) является входом другого (нижележащего). Композиция смежных базовых алгоритмов образует алгоритм-компонент $A_i = A_k^b \circ A_{k+1}^b \circ \dots \circ A_l^b$, алгоритмы-компоненты (далее алгоритмы), в свою очередь, образуют $A = A_1 \circ A_2 \circ \dots \circ A_n$, 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 способы декомпозиции 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, 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 :

$$\begin{aligned} P_{right} &= P(C_1 | I)P(C_2 | IC_1) \times \dots \\ &\times P(C_n | IC_1 C_2 \dots C_{n-1}), \end{aligned}$$

$$\begin{aligned} P_{err} &= 1 - P_{right} - P_{empty}, \\ P_{empty} &= P(\bar{F}_1 | I) + P(\bar{F}_2 | IF_1) + \dots \\ &+ P(\bar{F}_n | IF_1 F_2 \dots F_{n-1}), \\ P_{false} &= P(F_1 | N)P(F_2 | NF_1) \times \dots \\ &\times P(F_n | NF_1 F_2 \dots F_{n-1}). \end{aligned} \quad (3)$$

Здесь события I и N указывают на различные (информативные и неинформативные) входы, их вероятности равны 1. Среднее время анализа входных данных T рассчитывается с учётом вероятности пустого выхода каждого алгоритма-компонента:

$$\begin{aligned} T &= T_1 + T_2 P(F_1) + T_3 P(F_2 | F_1) + \dots \\ &+ T_n P(F_n | F_1 F_2 \dots F_{n-1}). \end{aligned} \quad (4)$$

Особенности представления процесса распознавания текстовой метки в виде последовательности событий описаны в статье [2].

3. Метод параметрической адаптации многоуровневых систем распознавания текстовых меток на видеоизображениях

Обычно задаётся ограниченное множество допустимых значений каждого параметра $u_j \in \{v_{j,1}, \dots, v_{j,m}\}$, 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 учётом ограничений на значения остальных критериев:

$$u^* = \arg \max_{u \in Conf} (\hat{P}_{right}(u)), \hat{P}_{right}(u^*) \leq P_{right}^{tr}, \quad (5)$$

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

Декомпозиция позволяет производить поэтапную настройку алгоритма A . Способ декомпозиции предлагается выбирать с учётом 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 настройки алгоритма A_i можно свести к формированию подмножества $Conf^{i,filter}$:

$$Conf^{i,filter} = \{u^i | P(C_i | u^i) \geq C_i^{tr}, T_i(u_i) \leq T_i^{tr}\}, \quad (6)$$

где пороги C_i^{tr}, T_i^{tr} задаются с учётом (5). Множество $Conf^{i,filter}$ может быть довольно большим. Его сокращение до множества оптимальных по Парето конфигураций сопряжено с возможностью потерять оптимальную конфигурацию u^* . Это обусловлено тем, что выходы A_i образуют экзаменационную по-

следовательность 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.

Прецедентный критерий – это вектор $V = (V_1, V_2, \dots, V_{sz})$, $V_j \in \{0, 1\}$, $V_j = 1 \Leftrightarrow A_i(Inv_j) \sim E_j$, иначе $V_j = 0$. Здесь E_j – эталонный результат анализа экзаменационного входа Inv_j , « \sim » – знак эквивалентности, sz – количество экзаменационных примеров.

Необходимое условие глобальной оптимальности конфигурации u^i , которой соответствует прецедентный критерий V^i : 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 необходимому условию конфигураций $Conf^{i, opt}$. Если $Conf^{i, filter}$ не пусто, то формируется множество оптимальных по Парето конфигураций:

$$Conf^{i, opt} = \{u^i \mid u^i \in Conf^{i, filter}, Par(q^i(u^i)) = true\}, \quad (7)$$

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

Метод вычисления u^* заключается в последовательном – от вышележащих алгоритмов к нижележащим – вычислении множеств $Conf^{i, opt}$, образующих в результате множество потенциально оптимальных конфигураций $Conf^{opt}$, если $\forall i \in 1..n: Conf^{i, opt} \neq \emptyset$. Далее вычисляются критерии эффективности конфигураций $Conf^{opt}$ и выбирается u^* .

Настройка параметров описанного в [1] модуля распознавания идентификационных номеров объектов подвижного состава железнодорожного транспорта предложенным методом заняла менее 16 часов. При тех же условиях настройка методом перебора потребовала бы более 6 лет.

Возможна ситуация, при которой экзаменационная последовательность B содержит подмножества примеров, для эффективного анализа которых следует применять различные конфигурации. Формировать такие конфигурации предлагается с использованием предложенного выше

метода, исключив при этом ограничения $\hat{P}(C_i) \geq C_i^{tr}$, $\hat{T}_i \leq T_i^{tr}$.

4. Параметрическая адаптация подсистемы принятия решений многоуровневых систем распознавания текстовых меток на видеоизображениях

Вместе с тем, настройка параметров алгоритмов не всегда позволяет решить задачу (5). В таком случае возможно применение известного в области распознавания текста подхода, который заключается в формировании алгоритмами списков вариантов решений (СВР) и позволяет снизить вероятность потери истинного решения на промежуточных этапах анализа.

Каждому решению сопоставляется некоторая численная оценка его правдоподобия, в соответствии с которой упорядочен СВР. Предлагается алгоритм усеечения списков (АУС), принимающий решения об отклонении. Параметрами АУС являются пороги двух видов: K и E . Порог K_i ограничивает длину СВР алгоритма A_i . Решение о принятии или отклонении СВР выполняется по результатам сравнения осреднённой оценки правдоподобия элементов усечённого списка с порогом E_i .

Задачей настройки системы принятия решений модуля распознавания является поиск оптимальных значений порогов $u^{AVC} = \{K, E\}$. Обычно множество значений параметров АУС $Conf^{AVC}$ конечно, что позволяет настраивать их совместно с параметрами алгоритмов-компонентов. Однако часто требуется произвести настройку АУС при фиксированной конфигурации u , для решения этой задачи разработан метод расчёта оптимальных (при заданных A , B , h , $Conf^{AVC}$) значений параметров u^{AVC} . Пусть $A^{AVC} \in A^C$, при этом $A^{AVC} = \{A_1, A_2, \dots, A_n\}$, где каждый A_i возвращает СВР.

Для расчёта 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 определённую позицию $p(d_j^i)$ в соответствующем СВР, характеризуемым оценкой правдоподобия СВР $e(d_j^i)$,

а его корректность устанавливается с использованием массива прецедентов. Характеристики решений d_j :

$$Pos(d_j, u^{AVC}) = \{p(d_j^1), p(d_j^2), \dots, p(d_j^n)\}, \quad (8)$$

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

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

$$Conf^{AVC, filter} = \left\{ u^{AVC} \mid \begin{array}{l} u^{AVC} \in Conf^{AVC}, \\ \hat{P}_{right}(u^{AVC}) \geq P_{right}^{tr}, \\ \hat{P}_{err}(u^{AVC}) \geq P_{err}^{tr}, \\ \hat{P}_{false}(u^{AVC}) \geq P_{false}^{tr} \end{array} \right\}. \quad (9)$$

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

$$Conf^{AVC, opt} = \{u^{AVC} \mid u^{AVC} \in Conf^{AVC, filter}, T(u^{AVC}) \leq T^{tr}\}, \quad (10)$$

где $T(u^{AVC})$ либо измеряется экспериментально, либо оценивается предложенным методом, который основан на статистическом расчёте среднего количества вызовов каждого алгоритма-компонента на одно видеоизображение при заданном u^{AVC} .

Экспериментальной базой для исследования метода послужил разработанный модуль распознавания идентификационных номеров объектов подвижного состава железнодорожного транспорта, описанный в [1]. В ходе эксперимента настройка предложенным методом параметров АУС указанной системы распознавания при количестве конфигураций её параметров $|Conf^{AVC}| = 40000$ заняла около 20 минут (тесты производились в среде MATLAB на компьютере со следующей конфигурацией: частота двухъядерного микропроцессора 1,83 ГГц, объём оперативной памяти 1 Гб). При аналогичных условиях непосредственное вычисление значений критериев эффективности каждой конфигурации (перебор) займёт 31 сутки.

5. Структурная адаптация многоуровневых систем распознавания текстовых меток на видеоизображениях

Помимо настройки параметров повышать эффективность многоуровневых СРТМ, функционирующих в режиме реального времени, предлагается за счёт их структурной адаптации.

Среднее время анализа видеоизображения складывается из времени анализа информативных и неинформативных изображений:

$$T = T_I \cdot p_I + T_N \cdot p_N, \quad (11)$$

где $p_I + p_N = 1$, T_I – время анализа информативного изображения, T_N – неинформативного, p_I – частота

появления информативных изображений, p_N – неинформативных.

Следовательно, увеличить затраты времени T_I возможно за счёт уменьшения затрат времени T_N . Для этого в состав системы включается алгоритм, синтезированный из АС и АР, который производит классификацию локализованных зон на содержащие и не содержащие текстовую метку посредством анализа их фрагментов. Отсев неинформативных зон позволяет сократить средние затраты времени на анализ видеоизображений, существенная доля которых не содержит текстовых меток. В общем случае такая возможность существует, если:

1. Выходом некоторого алгоритма A_i является решение a , образованное множеством элементов $A_i(b) = a = \{a_1, a_2, \dots, a_k\}$, классифицированных как информативные.
2. Вероятность ошибки классификации достаточно низкая.
3. Существует возможность получить частичное решение $a' \subset a$, произведя частичный анализ b .

Примером такого A_i служит алгоритм распознавания сегментов, а также композиция алгоритмов сегментации и распознавания. В первом случае данные b представляют собой список сегментов, во втором – локализованную зону. В обоих случаях возможен частичный анализ b : подмножества сегментов и фрагмента зоны (полученного обрезкой левого и правого краев зоны), соответственно.

Если результаты частичного анализа b позволяют принимать решение о том, что данные b неинформативные, то, очевидно, следует использовать возможность «досрочного» завершения анализа b . Чтобы выяснить, является ли вход b информативным, достаточно получить подмножество решений $a' \subset \{a_1, a_2, \dots, a_k\}$. Алгоритм F выполняет данные действия и, исходя из полученных результатов (найденное необходимое количество информативных решений), принимает решение о классификации b .

При этом возникает задача выбора порога Tr на количество элементов в a' , например, на количество распознанных сегментов. Для решения данной задачи предлагается рассчитать (при заданном способе формирования a'):

1. Распределение вероятности количества элементов в a' , рассчитанное по экзаменационному множеству информативных входов B^I .
2. Аналогичное распределение, рассчитанное по множеству неинформативных входов B^N .

Зная данные распределения, можно оценить эффективность алгоритма F при любом заданном пороге Tr и выбрать подходящий порог.

Для проведения экспериментов по структурной адаптации использовалась система распознавания идентификационных номеров объектов подвижного

состава железнодорожного транспорта. В состав разработанного модуля распознавания был включён 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 её версии с алгоритмом классификации зон

Критерии эффективности СРТМ	Без АКЗ	АКЗ
Время анализа инф. кадра, с	0,60	0,63
Время анализа неинф. кадра, с	0,71	0,25
Среднее время анализа кадра, с	0,68	0,34

Эксперименты показали, что вероятностные показатели эффективности СРТМ с АКЗ идентичны показателям типовой СРТМ. В то же время использование АКЗ позволило сократить среднее время анализа видеоизображения на 50%, с учётом того, что в заданных условиях около 23% кадров содержат образ номера. Доверительный интервал для $\hat{P}_{right} = 0,65$ при надёжности оценки $\gamma = 0,95$ равен $(0,62, 0,68)$, оценка $\hat{t} = 0,6$ характеризуется $P(|t - \hat{t}| \leq 0,02) = 0,923$.

Заключение

Таким образом 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 предложенной в статье моделью СРТМ, и к ним применимы описанные методы параметрической и структурной

адаптации СРТМ к заданным условиям эксплуатации. Для большинства современных СРТМ данные методы позволяют, во-первых, вычислять за приемлемое время оптимальный (в рамках заданной экзаменационной последовательности видеоизображений и дискретизации значений параметров) набор значений параметров системы. Во-вторых, они позволяют в значительной мере автоматизировать процесс параметрической адаптации СРТМ.

На практике это даёт 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 трудоёмкость процесса внедрения и его длительность.

Предложенная в статье методика повышения эффективности функционирования СРТМ за счёт их структурной адаптации позволяет существенно снизить среднее время анализа видеоизображений, большая часть которых не содержит образ текстовой метки. При этом не требуется сопутствующая модификация составляющих СРТМ алгоритмов.

Литература

1. **Воскресенский, Е.М.** Метод параметрической оптимизации процесса принятия решений в системах распознавания текстовых меток на видеоизображениях / Е.М. Воскресенский, В.А. Царёв // Компьютерная оптика. – 2009. – Т. 33, № 2. – С. 202-209.
2. **Воскресенский, Е.М.** Метод оценки эффективности систем распознавания текстовых меток на сложном фоне с использованием дерева вероятностных характеристик / Е.М. Воскресенский, В.А. Царёв // Компьютерная оптика. – 2008. – Т. 32, № 3. – С. 283-290.

References

1. **Voskresensky, E.M.** Method of solution lists cut algorithms in text label recognition systems parametric optimization / E.M. Voskresensky, V.A. Tsarev // Computer Optics. – 2008. – V. 33, N 2. – P. 202-209. – (in Russian).
2. **Voskresensky, E.M.** Method of estimation of text labels recognition systems efficiency parameters with use of likelihood characteristics tree / E.M. Voskresensky, V.A. Tsarev // Computer Optics. – 2008. – V. 32, N 3. – P. 283-290. – (in Russian).

PARAMETRIC AND STRUCTURAL ADAPTATION OF TEXT LABELS RECOGNITION SYSTEMS

*E.M. Voskresensky, V.A. Tsarev
Institute of Management and Information Technologies (branch)
of the St.-Petersburg State Polytechnical University*

Abstract

In article a new methods of multilevel text label recognition systems parametric and structural adaptation is offered. Such systems is described as hierarchical composition of algorithms. Methods allow adapting parameters of algorithms and parameters of decision-making subsystem of multilevel text label recognition systems. Method of multilevel text label recognition systems efficiency increase with their structural adaptation without necessity of algorithms modification is offered.

Key words: recognition of symbols, recognition systems adaptation.

Сведения об авторах



Воскресенский Евгений Михайлович, 1983 года рождения. В 2005 году окончил СПбГПУ по специальности «Программное обеспечение вычислительной техники и автоматизированных систем». Аспирант кафедры информационных и управляющих систем факультета технической кибернетики СПбГПУ, научный сотрудник лаборатории систем технического зрения и экспертных систем Института менеджмента и информационных технологий (филиала) Санкт-Петербургского государственного политехнического университета в г. Череповце (ИМИТ СПбГПУ).

Область научных интересов: обработка графических изображений, программирование, проектирование программного обеспечения.

E-mail: v.e.m@mail.ru .

Evgeny Mikhailovich Voskresensky (b. 1983) graduated (2005) from the Saint-Petersburg State Polytechnical University (SPbSPU). Post-graduate student of information and control systems chair of technical cybernetics faculty of SPbSPU, research assistant of vision systems and expert systems laboratory of Institute Of Management And Information Technologies (branch) of the SPbSPU in Cherepovets (IMIT SPbSPU).

Research interests are computer graphics processing, programming, and software designing.

E-mail: v.e.m@mail.ru



Царёв Владимир Александрович, 1972 года рождения, в 1993 г. с отличием окончил МГУ им. М.В. Ломоносова по кафедре «Математической теории интеллектуальных систем» механико-математического факультета. Кандидат технических наук (1998), заведующий кафедрой программного обеспечения вычислительной техники и автоматизированных систем ИМИТ СПбГПУ.

Область научных интересов: обработка изображений и распознавание образов, математическое моделирование, системный анализ сложных систем обработки информации.

E-mail: vats @imit.ru .

Vladimir Aleksandrovich Tsarev (b. 1973) graduated with honours from Moscow State University in 1993. Doctor in Technical (1998) degrees, head of chair of computer and automated system software of IMIT SPbSPU.

Research interests are image processing and pattern recognition, mathematical modeling, systems analysis of complicated information processing systems.

E-mail: vats @imit.ru.

Поступила в редакцию 28 мая 2010 г.