

## МЕТОД ПАРАМЕТРИЧЕСКОЙ ОПТИМИЗАЦИИ ПРОЦЕССА ПРИНЯТИЯ РЕШЕНИЙ В СИСТЕМАХ РАСПОЗНАВАНИЯ ТЕКСТОВЫХ МЕТОК НА ВИДЕОИЗОБРАЖЕНИЯХ

*Евгений Михайлович Воскресенский (научный сотрудник, e-mail: [v.e.m@mail.ru](mailto:v.e.m@mail.ru)),*

*Владимир Александрович Царев (зав. каф. ПО ВТ и АС, e-mail: [vats@imit.ru](mailto:vats@imit.ru))*

*Институт менеджмента и информационных технологий*

*(филиал Санкт-Петербургского государственного политехнического университета в г. Череповце)*

### **Аннотация**

В статье предложен новый метод параметрической оптимизации систем распознавания текстовых меток на видеоизображениях. Оптимизируются параметры, управляющие размерами списков решений, передаваемых между алгоритмами системы распознавания. Метод требует невысоких вычислительных затрат по сравнению с непосредственным расчетом критериев оптимальности каждой комбинации параметров. Приведены результаты исследований по оптимизации системы распознавания идентификационных номеров железнодорожного транспорта, подтвердившие целесообразность применения метода на практике.

**Ключевые слова:** распознавание символов, оптимизация систем распознавания.

### **Введение**

В настоящее время распознавание текстовых меток на видеоизображениях является одной из важных прикладных задач компьютерного зрения. Обычно оно применяется для идентификации движущихся объектов контроля, имеющих регистрационные надписи. Примерами таких объектов могут служить наземные транспортные средства, грузовые контейнеры, промышленные изделия и др.

Распространение систем распознавания текстовых меток (СРТМ) выявило проблему, связанную со сложностью настройки параметров алгоритмов СРТМ для наиболее эффективного функционирования в заданных условиях эксплуатации. Сложность заключается в том, что алгоритмы СРТМ, как правило, весьма нетривиальны и обладают большим количеством настраиваемых параметров, каждый из которых принимает значения из определенного диапазона. Конкретный набор значений параметров алгоритма будем называть конфигурацией. Количество конфигураций часто чрезвычайно велико, что существенно затрудняет поиск такой конфигурации, которая бы обеспечивала требуемые значения критериев эффективности СРТМ. На современном этапе эта трудоемкая работа, как правило, производится разработчиками вручную, исходя из опыта, интуиции и знания алгоритмов. Это усложняет разработку и удорожает такие системы, ограничивая их применение. Поэтому важной задачей является создание новых методов и средств параметрической оптимизации алгоритмов СРТМ с учетом заданных условий эксплуатации.

Известен подход, при котором с целью повышения эффективности распознавания текстовой информации между алгоритмами передаются списки возможных решений, что позволяет снизить вероятность потери истинного решения на промежуточных этапах анализа изображения [3]. Каждому сопоставляется некоторая численная оценка его правдоподобия, список решений упорядочен по убыванию оценок правдоподобия.

Чем больше элементов допускается в списке, тем выше вероятность того, что он содержит верное решение, но при этом увеличивается время анализа и возрастает вероятность ошибки второго рода. Отсюда возникает задача усечения списка либо его отклонения согласно выбранным порогам на длину списка и оценку его правдоподобия. В статье описывается метод поиска оптимальных значений порогов, а в качестве экспериментальной базы выбрана система распознавания идентификационных номеров железнодорожных вагонов.

### **1. Алгоритмы системы распознавания текстовых меток на видеоизображениях на примере идентификационных номеров железнодорожных вагонов**

Основная часть алгоритмического обеспечения систем распознавания текстовых меток реализуется в виде отдельного программного модуля – модуля распознавания. Он представляет собой последовательность взаимосвязанных алгоритмов. В общем случае модуль распознавания состоит из алгоритмов локализации текстовой метки (далее под текстовой меткой понимаем ее графический образ), сегментации локализованных зон изображения, распознавания сегментов и принятия решений. Ввиду неопределенности условий формирования входных изображений, создание алгоритмов, гарантирующих правильное решение, затруднено. Поэтому на практике применяются эвристические алгоритмы, ответы которых являются правильными с некоторой вероятностью.

Алгоритм локализации (АЛ) производит поиск зон – прямоугольных фрагментов изображения, предположительно содержащих образ текстовой метки. На вход модуля могут поступать изображения  $I_{m \times n}$  двух видов: информативные и неинформативные. Информативные изображения, в отличие от неинформативных, содержат образ текстовой метки. На рис. 1 представлены примеры входных видеоизображений (кадров с разрешением 288×384).



Рис. 1. Примеры входных видеоизображений (а – неинформативное; б – при дневном освещении; в – при ночном освещении)

На первом этапе локализации рассчитывается матрица оценок правдоподобия принадлежности каждой точки образу текстовой метки:  $E_{m \times n}$ , где  $E_{i,j}$  – оценка правдоподобия принадлежности пикселя с координатами  $i, j$  образу метки. Матрица  $E_{m \times n}$  является исходными данными для поиска зоны, содержащей текстовую метку. Рассмотрим метод ее построения.

В теории обработки изображений одним из основных методов преобразования изображений к виду, удобному для анализа, является применение фильтров. Реализуем вычисление матрицы  $E_{m \times n}$  как наложение фильтра  $H_{D \times D}$ :

$$E_{i,j} = \left| \sum_{k1=i-d}^{i+d} \sum_{k2=j-d}^{j+d} I_{k1,k2} \cdot H_{k1-i+1,k2-j+1} \right|. \quad (1)$$

Здесь  $d = \frac{D-1}{2}$  определяет размер фильтра, который представляется в виде квадратной матрицы с нечетным количеством элементов в строке (столбце).

Как правило, особенным свойством участка изображения, содержащего текстовую метку, по сравнению с другими участками, является высокий уровень градиента (перепадов яркости). Поэтому в роли фильтра можно взять, к примеру, матрицу, применение которой позволяет вычислить величину перепада яркости в заданной окрестности и направлении, например:

$$H = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}. \quad (2)$$

На рис. 2 представлен результат наложения фильтра (2) на изображение железнодорожной цистерны с идентификационным номером на борту.



Рис. 2. Результат наложения фильтра

Представим матрицу  $E_{m \times n}$  в виде последовательности столбцов шириной  $s = \left\lfloor \frac{n}{parts} \right\rfloor$  элементов, где  $parts$  – параметр АЛ. Для каждого столбца строится функция среднего значения оценки правдоподобия в строке:

$$A_{i,k} = \frac{1}{s} \sum_{j=s*k}^{s*(k+1)} E_{i,j}, \quad (3)$$

где  $k \in 0..parts-1$  – номер столбца. После этого выбирается координата  $i$ , максимизирующая отношение:

$$EA_{i,k} = \frac{\frac{1}{h} \sum_{j=i+1}^{i+h-1} A_{j,k}}{0.5(A_{i,k} + A_{i+h,k})}. \quad (4)$$

По результатам анализа каждого столбца формируется зона, попадающая в список решений АЛ. Из (4) определяются координаты  $k$ -ой зоны  $SZ^k$ :

$$y_1^{(k)} = e = \arg \max_{i=1..m} EA_{i,k}, \quad y_2^{(k)} = y_1 + h,$$

$$x_1^{(k)} = \max\{ks - (w - s), 0\},$$

$x_2^{(k)} = \min\{(k+1)s + (w-s), n-1\}$ , где  $h$  – высота текстовой метки,  $w$  – ширина. Координаты  $x_1, x_2$  рассчитываются с некоторым «запасом»  $x_2 - x_1 > w$ , а также с учетом возможного выхода за пределы границ изображения  $(0, n-1)$ . Оценкой правдоподобия зоны является  $EA_{e,k}$ .

Алгоритм сегментации (АС) производит поиск сегментов – координат предполагаемых символов на локализованной зоне. На первом этапе сегментации вычисляются оценки правдоподобия принадлежности пикселей линиям символов, в простейшем случае используются готовые результаты наложения фильтра  $H$  либо непосредственно яркость пикселей. Результатом первого этапа является массив оценок правдоподобия  $EZ_{h \times w}$ .

Второй этап – это вычисление вектора средней оценки правдоподобия по столбцам:

$$A_j = \frac{1}{h} \sum_{i=1}^h EZ_{i,j}. \quad (5)$$

Вектор  $A$  позволяет обнаружить промежутки фона между символами, которые проявляются в виде экстремумов (рис.3).

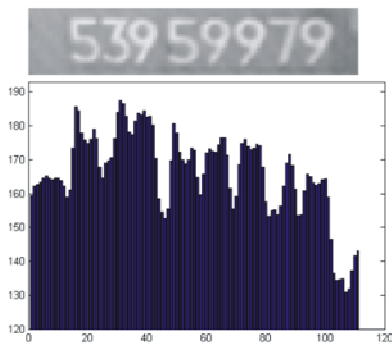


Рис. 3. Вектор средней оценки правдоподобия по столбцам локализованной зоны

На третьем этапе производится обнаружение вертикальных разделителей – границ между символами  $Dev$ . Для этого каждому элементу  $A_i$  сопоставляется оценка правдоподобия  $H_i$ . Чем выше эта оценка, тем выше вероятность, что в данном месте находится промежуток фона между символами. Индексы  $i$  выбранных  $H_i$  – это координаты вертикальных разделителей (границы) между символами:  $Dev = \{dev_1, dev_2, \dots, dev_k\}$ , где  $k$  – количество найденных разделителей.

На четвертом этапе осуществляется уточнение вертикальных и горизонтальных границ символов. Для этого методом Отсу вычисляется пороговое значение яркости для фрагмента изображения между двумя разделителями. Порог позволяет бинаризовать фрагмент и в результате найти координаты выделенного на фрагменте объекта. Оценкой правдоподобия сегмента является отношение его самого яркого пикселя к самому темному. Сегменты, упо-

рядоченные по убыванию оценки правдоподобия, образуют список решений АС.

Алгоритм распознавания (АР) производит распознавание (классификацию) изображений отдельных сегментов. Сегмент, отнесенный к одному из классов символов, называем распознанным, а отнесенный к классу шумов – нераспознанным. АР реализован в виде многослойного перцептрона, обученного классифицировать изображения символов. Выход нейронной сети есть оценка правдоподобия принадлежности изображения конкретному классу символов. На данном этапе не производится формирование списка решений, поскольку это приводит к комбинаторному взрыву количества решений, формируемых по результатам распознавания сегментов зоны, и при этом, возможно, имеет смысл лишь при наличии априорной информации о кодах текстовых меток.

Множество распознанных сегментов анализируется алгоритмом формирования решений (АФР), который производит поиск всех возможных последовательностей распознанных сегментов, потенциально являющихся образами символов метки (решениями). Такие последовательности образуют список решений, упорядоченный по убыванию оценки правдоподобия, которая рассчитывается как сумма оценок правдоподобия соответствующих распознанных сегментов.

АФР реализован на основе алгоритма поиска в глубину. Исходными данными для такого поиска является матрица смежности сегментов, которая содержит информацию о парах сегментов, потенциально являющихся изображениями соседних символов метки. При отсутствии какой-либо априорной информации о текстовых метках, например, списка текстовых меток, в качестве итогового выбирается решение, обладающее максимальной оценкой правдоподобия.

## 2. Задача параметрической оптимизации алгоритмов управления списками решений

В работах [1-3] с целью повышения эффективности распознавания текстовой информации предложено передавать между алгоритмами списки решений, что позволяет снизить вероятность потери истинного решения на промежуточных этапах анализа изображения. Каждому решению сопоставляется некоторая численная оценка его правдоподобия, которая используется для упорядочивания списка и последующего усечения списка либо его отклонения как вероятно неинформативного [2]. Список решений некоторого алгоритма модуля распознавания анализируется алгоритмом усечения списков (АУС). Для этого используются два порога:  $K$  и  $E$ . Решение о принятии или отклонении списка выполняется согласно результату сравнения оценки правдоподобия списка с порогом  $E$ . Далее предполагаем, что параметры  $E$  целочисленные и принимают значения из некоторого ограниченного диапазона. Порог  $K$  ограничивает длину принятого списка.

Обозначим параметры АУС алгоритмов локализации, сегментации, распознавания и формирования решений, соответственно,  $\{K_L, E_L\}$ ,  $\{K_S, E_S\}$ ,  $\{K_K, E_K\}$ ,  $\{K_R, E_R\}$ . На рис. 4 показан пример – структурная схема модуля распознавания с алгоритмами усечения списков.

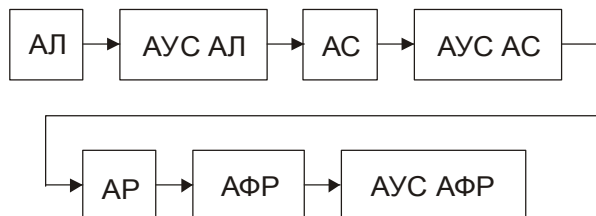


Рис. 4. Структурная схема модуля распознавания с алгоритмами усечения списков

Чем больше элементов допускается в списке решений, тем выше вероятность того, что он содержит верное решение, но при этом увеличивается время анализа и возрастает вероятность ошибки второго рода. Данные противоречия приводят к необходимости выбора оптимального набора значений порогов  $K$  и  $E$  для каждого из алгоритмов модуля распознавания.

Критериями эффективности модуля распознавания являются вероятности возможных исходов анализа изображения, а также время его анализа:

- $P_{right}$  – вероятность правильного распознавания текстовой метки на изображении;
- $P_{err}$  – вероятность неправильного распознавания текстовой метки на изображении;
- $P_{false}$  – вероятность распознавания ложной метки на изображении, не содержащем ее образ;
- $T$  – среднее время анализа изображения.

Таким образом, задача настройки порогов АУС формулируется как задача параметрической оптимизации, где множество параметров состоит из  $param = \{K_L, E_L, K_S, E_S, K_K, E_K, K_R, E_R\}$  либо из некоторого подмножества  $param$ , а критериями эффективности анализа информативного изображения являются:  $P_{right}$ ,  $P_{err}$ ,  $P_{false}$ ,  $T$ . Некоторый конкретный набор значений параметров АУС называем конфигурацией АУС.

В работах [1,2] предложено производить спецификацию алгоритмов модуля распознавания с использованием массивов прецедентов – структурированной априорной информации о тестовой выборке изображений. Массив прецедентов включает информацию о координатах текстовой метки и отдельных ее символов, а также о классах символов по каждому тестовому изображению. Помимо возможности спецификации, он позволяет вычислять критерии эффективности модуля и его отдельных алгоритмов, что необходимо при проведении оптимизации модуля распознавания.

Если количество конфигураций невелико, то выбор оптимальной конфигурации сводится к вычис-

лению критериев оптимальности каждой из них путем анализа выхода модуля по некоторой экзаменационной последовательности изображений. Однако часто на практике вычисление критериев оптимальности всех конфигураций неприемлемо из-за высоких вычислительных затрат ввиду большого количества конфигураций и (или) существенных затрат времени на вычисление критериев отдельной конфигурации.

### 3. Метод оптимизации процесса управления списками решений

Для сокращения вычислительных затрат на проведение параметрической оптимизации АУС предлагается следующий метод. Он использует то, что параметры АУС являются порогами, определяющими количество элементов в списке решений – 0 или  $K$ . Это отличает параметры АУС от параметров прочих алгоритмов модуля. Особенность таких параметров-порогов заключается в том, что при задании им различных значений выход алгоритма-компонента изменяется не качественно (становится иным), а количественно – принимается некоторая часть списка решений согласно порогу, сам же список не изменяется. Это позволяет, зная полный список решений алгоритма, определить результат при любых параметрах АУС, лишь единожды вычислив выход алгоритма по заданному изображению, что и позволяет значительно сократить вычислительные затраты.

Метод оптимизации АУС производит анализ изображений экзаменационной последовательности, статистически рассчитывая вероятности исходов анализа изображений соответственно каждой конфигурации АУС. Опишем основные шаги метода:

#### 1. «Отключение» АУС.

Предварительно алгоритмы усечения списков решений «отключаются», т.е. их пороги устанавливаются в такие значения, при которых все списки решений пропускаются и не производится их усечение (либо производится до некоторого максимально допустимого размера). Поскольку АУС «отключены», то выходом модуля распознавания является полный список решений, сформированных по результатам анализа изображения.

2. Цикл по изображениям экзаменационной последовательности и определение исхода анализа изображения при каждой конфигурации АУС.

2.1. Анализ экзаменационного изображения и получение полного списка решений модуля распознавания.

Каждый код текстовой метки  $kod_i$  из списка решений АФР  $\{kod_1, kod_2, \dots, kod_n\}$  характеризуется набором признаков, которые определяют, в каких конфигурациях АУС данный вариант является решением.

Код  $kod_i$  сформирован из совокупности элементов, каждый из которых характеризуется некоторой

оценкой правдоподобия и местом в списке решений. Эти данные определяют значения порогов АУС, при которых решение может быть сформировано. Например, последовательность сегментов, составляющих решение, распознана на зоне, которая находится на третьем месте списка решений АЛ. Это значит, что при прочих равных условиях данная последовательность распознается при пороге на размер списка решений АЛ, большем либо равном трем:  $K_L \geq 3$ .

Помимо этого для каждой последовательности определяется, правильно ли сегменты распознаны и являются ли они образами символов, для чего используется массив прецедентов.

## 2.2. Анализ списка решений модуля распознавания.

Полный список решений по каждому экзаменационному изображению анализируется следующим образом: для каждой конфигурации АУС устанавливается исход обработки изображения и инкрементируется переменная, соответствующая исходу.

Пусть списки решений возвращаются алгоритмами локализации, сегментации и формирования решений, а в качестве итогового решения принимается первый элемент списка АФР. Тогда каждый код текстовой метки обладает следующими характеристиками:

- решение получено посредством анализа зоны, занимающей позицию  $k_L$  списка решений АЛ;
- список решений АЛ имеет оценку правдоподобия, равную  $e_L$  (оценка первого элемента списка);
- сегменты, из которых сформирован ответ, располагаются в пределах  $k_S$  позиций списка решений АС;
- список решений АС имеет оценку правдоподобия, равную  $e_S$  (средняя оценка первых  $n$  элементов списка).

При правильном распознавании текстовой метки в выходном массиве *Right* инкрементируются значения ячеек, соответствующих конфигурациям, при которых данное решение занимает первую позицию в списке:  $Right_{i,j,k,m} = Right_{i,j,k,m} + 1$ , где  $i \in k_L..m_L$ ,  $j \in 1..e_L$ ,  $k \in k_S..m_S$ ,  $m \in 1..e_S$ ,  $m_L$  – максимально допустимое количество элементов в списке решений АЛ,  $m_S$  – максимально допустимое количество элементов в списке решений АС. Здесь предполагается, что оценки правдоподобия нормированы таким образом, что принимают значения из некоторого диапазона целых чисел от 1 до  $n_L$  и  $n_S$ , соответственно, для АЛ и АС. Размеры массива *Right* равны количеству соответствующих порогов –  $m_L \times n_L \times m_S \times n_S$ .

Для вычисления вероятностей ошибки распознавания и ложного распознавания аналогичным образом используются соответствующие массивы.

## 3. Вычисление значений критериев оптимальности модуля распознавания.

По результатам анализа экзаменационной последовательности изображений по каждой конфигура-

ции АУС вычисляются значения  $P_{right}$ ,  $P_{err}$ ,  $P_{false}$  как доля соответствующих исходов от общего количества информативных и неинформативных экзаменационных изображений.

Объем памяти, необходимый для использования предложенного метода, зависит от количества конфигураций АУС. Исходя из этого, предъявляются соответствующие требования к аппаратному обеспечению компьютера, на котором проводятся вычисления по оптимизации АУС.

Приведем в качестве примера описанные выше алгоритмы системы распознавания идентификационных номеров железнодорожного транспорта. В процессе разработки алгоритмов были определены возможные значения порогов  $K_L$ ,  $E_L$ ,  $K_S$ ,  $E_S$ , количество которых равно 5, 10, 80, 10 соответственно. При размерах массива  $5 \times 10 \times 80 \times 10$  количество ячеек в таком массиве равно 40000, в памяти компьютера такой массив занимает менее 160Кб (при использовании типа *integer*), что не является ограничением для современных компьютеров. Тем не менее, перед использованием предложенного метода в прикладных задачах, которые могут обладать другими особенностями, следует заранее учитывать объем памяти, необходимый для хранения соответствующей информации.

Таким образом, предложенный метод позволяет вычислить вероятностные критерии оптимальности каждой конфигурации АУС. Однако не менее значимым критерием является среднее время анализа изображения.

## 4. Модель зависимости времени анализа изображения от параметров алгоритмов усечения списков

Среднее время  $T$  анализа изображения, поступающего на вход модуля распознавания, является значимым критерием оптимальности модуля. Оно определяет количество изображений, которое система распознавания «успевает» анализировать за единицу времени.

В отличие от вероятностных критериев, время анализа изображения аппаратно зависимо. Достаточно сложно измерить его с высокой точностью. Как правило, системы распознавания создаются на платформе распространенных операционных систем Windows, что позволяет упростить разработку и снизить ее стоимость. Однако такие операционные системы не являются системами реального времени, что позволяет лишь приблизительно оценивать время выполнения функций тех или иных алгоритмов модуля.

Однако в большинстве случаев высокая точность не требуется и достаточно лишь приблизительной оценки, поскольку время анализа изображения влияет на эффективность системы достаточно опосредованно и незначительные изменения времени мало сказываются на эффективности системы. Вследствие этого предлагается использовать сле-



дующую модель, которая позволяет вычислять оценку  $T$ , учитывая лишь основные факторы, влияющие на время анализа.

Обозначим  $T_Z^I$  – оценку среднего времени анализа информативной зоны и  $T_Z^N$  – неинформативной. Сначала вычислим  $T_Z^I$ . Время анализа зоны определяется временем сегментации, распознавания сегментов и формирования форм. Среднее время сегментации зоны  $T^S$  и время распознавания отдельного сегмента  $T^K$  рассчитываются статистически.

Время распознавания всех сегментов зоны зависит главным образом от количества сегментов. Оно оценивается с учетом порога  $K_S$  и распределения вероятности количества сегментов ( $P(C^I)$  для информативных зон и  $P(C^N)$  для неинформативных), поскольку определенная доля зон содержит количество сегментов меньше  $K_S$ , это следует учитывать при оценке среднего времени распознавания сегментов зоны. Оценка времени распознавания сегментов зоны равна:

$$Trz(K_S, Y) = \sum_{i=1}^{K_S-1} i \cdot T^K \cdot P(C_i^Y) + \sum_{i=K_S}^{m_S} K_S \cdot T^K \cdot P(C_i^Y),$$

где  $Y \in \{I, N\}$  обозначает информативная или неинформативная зона.

Среднее время работы АФР пропорционально среднему количеству решений, которое алгоритм формирует по зоне, в сумме с некоторым константным временем, которое затрачивает алгоритм, даже если ни одного решения не сформировано. Данная зависимость может быть установлена средствами вычислительной математики. В представленном далее примере она моделируется линейной функцией:  $Tf(K_S, Y) = c_1 \cdot FCnt(K_S, Y) + c_2$ . Функция  $FCnt(K_S, Y)$  рассчитывается статистически. Зависимость количества решений от порога на число сегментов может быть вычислена для каждой комбинации  $K_L$  и  $E_L$ , однако данные параметры оказывают, как правило, весьма косвенное влияние на функцию  $FCnt(K_S, Y)$ .

Найдем среднее время анализа  $i$ -ой зоны списка решений АЛ. Пусть среднее время сегментации равно  $T_S$  рассчитывается статистически. Среднее время распознавания сегментов зависит от вероятностей событий «зона информативна (неинформативна)» и «список решений АС принят». Первая вероятность определяется позицией  $i$ , она равна  $P(L_i)$ . Вероятность события «список решений АС принят» определяется порогом  $E_S$  и распределением вероятности оценки правдоподобия списка решений АС

$P(L^Y)$ , она равна  $\sum_{j=E_S}^{m_S} P(E_j^{S,Y})$ . Отсюда среднее

время распознавания сегментов  $i$ -ой зоны вычисляется следующим образом:

$$Trs(i, K_S, E_S) = P(L_i) \cdot \left( \sum_{j=E_S}^{m_S} P(E_j^{S,I}) \right) \cdot Trz(K_S, I) + (1 - P(L_i)) \cdot \left( \sum_{j=E_S}^{m_S} P(E_j^{S,N}) \right) \cdot Trz(K_S, N). \quad (6)$$

Среднее время формирования решений по  $i$ -ой зоне с учетом вероятностей указанных событий равно:

$$Tf(i, K_S) = P(L_i) \cdot \left( \sum_{j=E_S}^{n_S} P(E_j^{S,I}) \right) \cdot Tf(K_S, I) + (1 - P(L_i)) \cdot \left( \sum_{j=E_S}^{n_S} P(E_j^{S,N}) \right) \cdot Tf(K_S, N). \quad (7)$$

Таким образом, среднее время анализа  $i$ -ой зоны равно  $Tz(i, K_S, E_S) = T_S + Trs(i, K_S, E_S) + Tf(i, K_S)$ , а среднее время анализа всего информативного изображения равно:

$$Tii(K_L, E_L, K_S, E_S) = T_L + \sum_{j=E_L}^{n_L} P(E_j^{L,I}) \cdot \sum_{i=1}^{K_L} Tz(i, K_S, E_S).$$

Аналогичным образом рассчитывается среднее время анализа неинформативного изображения:

$$Tin(K_L, E_L, K_S, E_S) = T_L + \sum_{j=E_L}^{n_L} P(E_j^{L,N}) \cdot \sum_{i=1}^{K_L} Tz(i, K_S, E_S).$$

Зная частоты появления информативного и неинформативного кадров  $p^I$  и  $p^N$ , найдем оценку среднего времени анализа произвольного кадра:

$$T = p^I \cdot Tii(K_L, E_L, K_S, E_S) + p^N \cdot Tin(K_L, E_L, K_S, E_S).$$

Предложенная модель учитывает лишь основные факторы, влияющие на среднее время анализа изображения, однако, как показывают дальнейшие эксперименты, позволяет достаточно точно оценить  $T$ .

## 5. Практические эксперименты

Продемонстрируем применение предложенного метода оптимизации АУС на практике. Для проведения исследований были использованы алгоритмы оптоэлектронной системы идентификации объектов железнодорожного транспорта, описанные в п.1. Тестовая выборка состояла из 550 информативных изображений (видеокадров), содержащих образ железнодорожных вагонов с видимой текстовой меткой на борту. Для выборки предварительно был сформирован массив прецедентов. Тесты производились на компьютере со следующей конфигурацией: частота двухъядерного микропроцессора 1,83 ГГц, объем оперативной памяти 1 Гб.

В качестве примера зависимости эффективности модуля от параметров АУС на рис. 5 представлены

графики зависимостей вероятностных критериев оптимальности модуля распознавания  $P_{right}$ ,  $P_{err}$  от порога  $K_S = 1..50$  при  $K_L = 1$ ,  $E_S = 1$ ,  $E_L = 1$ ,  $K_R = 1$ ,  $E_R = 1$ .

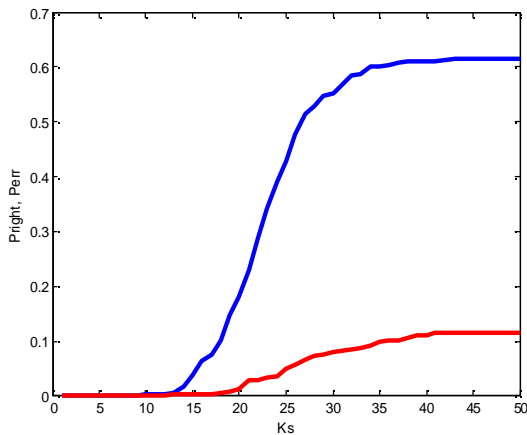


Рис. 5. Зависимость вероятностных критериев оптимальности модуля распознавания от порога на количество сегментов

Следует заметить, что вероятность распознавания идентификационного номера железнодорожного вагона на отдельно взятом кадре превышающая 0,6 является достаточно высоким показателем. При практическом использовании модуля распознавания решение об идентификации объекта принимается, как правило, по результатам анализа последовательности видеок кадров, полученных от нескольких видеоисточников (объект железнодорожного транспорта имеет четыре номера, нанесенные в разных местах). Поэтому вероятность идентификации объекта распознавания значительно выше.

На рис. 6 представлены графики зависимости среднего времени анализа отдельного информативного изображения (в среде MATLAB) от порога  $K_S = 20..45$  при тех же параметрах.

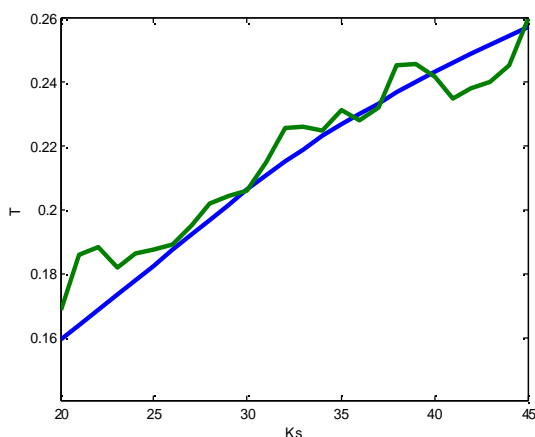


Рис. 6. Зависимость среднего времени анализа изображения от порога на количество сегментов

Гладкий график характеризует оценку  $T$ , вычисленную с использованием предложенной выше модели, второй график отображает время  $T$ , непо-

средственно измеренное в процессе анализа модулем изображений тестовой выборки.

Замеры времени производились в среде MATLAB, которая является интерпретатором, и в операционной среде Windows XP, которая не является операционной системой реального времени. В силу этих факторов результаты измерений  $T$  также являются лишь приблизительной оценкой. Тем не менее, как показывает рисунок 4, предложенный метод позволяет получить оценку  $T$ , достаточно близкую к результатам непосредственных измерений, что говорит о возможности использования предложенного метода оценки  $T$  в процессе параметрической оптимизации АУС.

Оптимизируем параметры:  $K_L$ ,  $E_L$ ,  $K_S$ ,  $E_S$  алгоритмов, описанных в п.1. Применение предложенного метода дало следующие результаты. На рис. 7 представлено множество точек, соответствующих конфигурациям АУС. Координатами точек являются значения критериев оптимальности конфигураций:  $P_{right}$  и оценка времени  $\hat{T}$ .

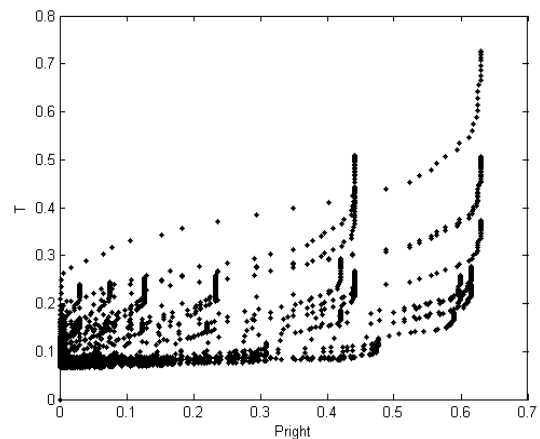


Рис. 7. Критерии оптимальности  $P_{right}$ ,  $P_{err}$  и  $\hat{T}$  множества конфигураций АУС

Из рисунка видно, что точки образуют характерные фигуры, которые отражают рост вычислительных затрат по мере увеличения длины списка решений АС и прекращение по достижению некоторого значения  $K_S$  роста вероятности правильного распознавания образа текстовой метки.

Руководствуясь целью максимизации  $P_{right}$ , примем в качестве итоговой следующую конфигурацию АУС:  $K_L = 2$ ,  $E_L = 1$ ,  $K_S = 43$ ,  $E_S = 2$ . Эта конфигурация обеспечивает показатели эффективности:  $P_{right} = 0,53$ ,  $P_{err} = 0,07$ ,  $P_{false} = 0$ ,  $\hat{T} = 0,33$ . Затраты времени на вычисление критериев оптимальности всех конфигураций АУС, составления множества оптимальных по Парето конфигураций и выбора конечной конфигурации составили 20 минут 40 секунд.

Оценим время, затрачиваемое на получение того же самого множества критериев оптимальности каж-

дой конфигурации, если непосредственно вычислять эффективность каждой конфигурации. Количество конфигураций равно  $3 \times 10 \times 80 \times 10 = 24000$ . Примем время вычисления эффективности отдельной конфигурации равным времени вычисления эффективности конфигурации  $\{K_L = 1, E_L = 1, K_S = 20, E_S = 1\}$ , которой соответствует небольшой объем вычислений по сравнению с большинством прочих конфигураций (значительно ниже среднего). Это время равно 67 секундам, отсюда следует, что для непосредственного вычисления критериев оптимальности всех конфигураций потребуется более чем  $67 \times 24000 = 1608000$  секунд (18 суток). Данный пример наглядно показывает, что применение предлагаемого метода позволяет производить параметрическую оптимизацию АУС модуля распознавания в тех задачах, где применение непосредственного вычисления критериев оптимальности модуля сопряжено с неприемлемыми затратами времени.

### Заключение

Системы распознавания текстовых меток на видеоизображениях находят все более широкое практическое применение. Их основным компонентом является программное обеспечение, реализующее алгоритмы локализации, сегментации, распознавания и принятия решений. С целью повышения эффективности алгоритмического обеспечения таких систем в работах [1,2] предложено между промежуточными этапами передавать списки возможных вариантов решений, что дает возможность не потерять на промежуточных этапах истинное решение. Чем больше список, тем выше вероятность того, что он содержит верное решение, но при этом растет время анализа и увеличивается вероятность ошибки. Данные противоречия приводят к необходимости выбора оптимального набора параметров (конфигурации) алгоритмов усечения списков, т.е. ставят задачу параметрической оптимизации этих алгоритмов.

Обычно на практике вычисление критериев оптимальности всех конфигураций неприемлемо из-за высоких вычислительных затрат.

Для решения этой проблемы в статье предложен метод оптимизации параметров алгоритмов усечения списков, требующий невысоких вычислительных затрат.

Метод использует тот факт, что параметры АУС являются порогами, определяющими количество элементов в списке решений. Особенность таких параметров-порогов заключается в том, что полный список решений алгоритма позволяет определить результат при любых параметрах АУС, запустив алгоритм единожды, чтобы рассчитать этот список. Это свойство отличает параметры АУС от обычных параметров алгоритмов.

Экспериментальные исследования по оптимизации системы распознавания идентификационных номеров железнодорожного транспорта подтвердили эффективность предлагаемого метода.

### Литература

1. **Tsarev, V. A.** Basic principles of vehicles optoelectronic identification system development / V. A. Tsarev // Pattern Recognition and Image Analysis – 2005. - Vol. 15, N 2. – P. 454-457. – ISSN 1054-6618.
2. **Царев, В.А.** Проектирование и оптимизация оптоэлектронных приборов контроля передвижения наземных транспортных средств. / В.А. Царев // Приборы и системы. Управление, контроль, диагностика, 2008.- №7.- С. 16-23.
3. **Горский, Н.Д., Анисимов, В.А., Горская, Л.М.** Распознавание рукописного текста: от теории к практике / Горский Н.Д., Анисимов В.А., Горская Л.М.// – СПб.: Политехника, 1997. - 126 с.

### References

1. Basic principles of vehicles optoelectronic identification system development / Vladimir A. Tsarev // Pattern Recognition and Image Analysis – 2005. – Vol. 15, N 2. – P. 454-457. – ISSN 1054-6618.
2. **Tsarev V.A.** Design and optimization of land-based vehicle relocation optoelectronic inspection tool. // Devices and systems. Direction, control, diagnostics, 2008. – №7. – С. 16-23.
3. Handwritten recognition: from theory to practice / Gorsky N.D., Anisimov V.A., Gorskaya L.M. – Saint-Petersburg: Polytechnic, 1997. - 126 с.

## METHOD OF SOLUTION LISTS CUT ALGORITHMS IN TEXT LABEL RECOGNITION SYSTEMS PARAMETRIC OPTIMIZATION

*E.M.Voskresensky (research assistant, e-mail: v.e.m@mail.ru),  
V.A.Tsarev (deputy director on research, e-mail: vats@imit.ru)  
Institute of Management and Information Technologies (branch)  
of the St.-Petersburg State Polytechnical University in Cherepovets*

### Abstract

In article a new method of text label recognition systems parametric optimization is offered. Algorithms lists of solutions control parameters are optimized. The method requires low computational capability in comparison with direct calculation of optimality criteria of each parameters combination. Experiment results of working recognition system optimization are described.

**Key words:** recognition of symbols, recognition systems optimization.

*В редакцию поступила 27.03.2009 г*