

МЕТОДЫ УПРАВЛЕНИЯ ПРОЦЕССОМ РАСПОЗНАВАНИЯ ТЕКСТОВЫХ МЕТОК НА ИЗОБРАЖЕНИЯХ

Е.М. Воскресенский¹, В.А. Царев¹

¹Институт Менеджмента и Информационных Технологий (филиал)

Санкт-Петербургского Государственного Политехнического Университета, Санкт-Петербург, Россия

Аннотация

В статье предложен новый подход к проектированию структуры алгоритмов систем распознавания текстовых меток, основанный на идее автоматического управления параметрами алгоритмов в процессе функционирования системы. Выбор значений параметров осуществляется с использованием обратных связей от алгоритма распознавания символов. Представлены результаты экспериментов с алгоритмами реальной системы распознавания идентификационных номеров объектов подвижного состава железнодорожного транспорта, подтверждающие эффективность предложенного подхода.

Ключевые слова: распознавание символов, управление процессом распознавания.

Введение

Развитие вычислительной техники в последние годы обеспечило возможность получения, передачи и обработки больших объемов видеoinформации с использованием недорогого телекоммуникационного оборудования и обычных персональных компьютеров. В связи с этим значительно расширяется область применения систем технического зрения (СТЗ), в частности, систем распознавания текстовых меток (СРТМ). Наиболее важной и сложной задачей в процессе создания СРТМ является разработка методов и средств распознавания образов текстовых меток (ТМ) на изображениях. В общем случае текстовая метка – это строка букв и (или) цифр из некоторого априорно известного алфавита символов, нанесенная тем или иным образом на объект контроля. Примерами таких объектов могут служить наземные транспортные средства (рис.1), грузовые контейнеры, промышленные изделия, денежные купюры и др.



Рис. 1. Железнодорожная цистерна с идентификационным номером на борту

Принципам построения отдельных алгоритмов локализации, сегментации и распознавания текста посвящено множество работ, и уровень развития современных методов теории обработки и распознавания изображений задает некоторый предел эффективности отдельных алгоритмов. В то же время мало изученными остаются вопросы проектирования структуры алгоритмического обеспечения систем

расознавания. В основном в литературе описаны либо общие концептуальные схемы построения программного обеспечения СТЗ, либо частные реализации СРТМ, структура которых представлена разорванной (без обратных связей) последовательностью алгоритмов [1,2]. Такая структура интуитивно понятна и удобна в реализации, однако не позволяет в полной мере использовать потенциал имеющихся алгоритмов, в результате чего возникает потребность в создании более гибкой и эффективной структуры алгоритмов СРТМ.

В статье предложена новая концепция структуры алгоритмов СРТМ, основанная на идее автоматического управления поведением алгоритмов в процессе функционирования посредством изменения их управляемых параметров за счет использования обратных связей. Также на примере системы распознавания идентификационных номеров объектов подвижного состава железнодорожного транспорта описаны управляемые алгоритмы локализации и сегментации, обладающие более высокими показателями эффективности по сравнению с аналогичными реализациями алгоритмов без управления.

1. Алгоритмическое обеспечение систем распознавания текстовых меток

СРТМ в общем случае состоит из аппаратного и программного обеспечения. Как правило, структура ключевой составляющей СРТМ – модуля обработки и анализа изображений (МОАИ) – представляется последовательностью алгоритмов локализации образа текстовой метки (ОТМ), сегментации ОТМ, распознавания изображений отдельных символов и принятия решения.

Алгоритм локализации (АЛ) определяет границы образа текстовой метки на изображении. Входом АЛ является изображение $I_{m \times n}$, которое возможно содержит образ объекта контроля. Под изображением далее понимается прямоугольное растровое полутоновое ($I_{i,j} \in [0, 225]$) изображение, которое получено, как правило, посредством цифровой фотоили видеосъемки объекта. Выходом АЛ являются координаты прямоугольной зоны (x_1, y_1, x_2, y_2) на

исходном изображении, которое возможно содержит образ ТМ. Границы ОТМ на исходном изображении определяются геометрическими координатами образов символов текстовой метки. Количество символов обычно фиксировано и в большинстве случаев составляет от 2 до 10.

Алгоритм сегментации (АС) определяет границы каждого предполагаемого символа на локализованной зоне $zone_{h \times w}$, где $zone_{i,j} = I_{y1+i, x1+j}$, $i \in [y1, y1+h]$, $j \in [x1, x1+h]$, $h = y2 - y1 + 1$, $w = x2 - x1 + 1$. Входом АС является изображение и координаты зоны на нем. АС возвращает множество координат сегментов, каждый из которых возможно содержит изображение символа: $\{(x_{1,1}, y_{1,1}, x_{1,2}, y_{1,2}), \dots, (x_{n,1}, y_{n,1}, x_{n,2}, y_{n,2})\}$.

Алгоритм распознавания (АР) производит распознавание (классификацию) изображений отдельных символов. Выход АР – это код распознанного символа, то есть класс, к которому отнесено изображение. Помимо классов символов (цифр и букв) в МОАИ используется класс шумов, его составляют изображения, не содержащие корректно сегментированного символа.

Алгоритм принятия решений (АПР) формирует окончательный результат распознавания. На вход АПР поступают распознанные сегменты и их координаты. Выходом АПР является код ТМ, то есть строка символов, соответствующая ТМ, или решение об отсутствии метки на изображении.

Следует отметить, что на вход МОАИ могут поступать изображения двух видов: информативные и неинформативные. Информативные изображения содержат образ текстовой метки, в отличие от неинформативных. Будем называть информативными данными результаты корректной локализации ОТМ, а также корректной сегментации и распознавания. Как правило, алгоритмы МОАИ обрабатывают данные обоих видов единообразно и только АР «отличает» их. Выходами АС и АР являются списки объектов. Выход является полным, если найденных объектов достаточно для последующего принятия решения о наличии ОТМ на зоне (не обязательно верного), и неполным в противоположном случае.

2. Концепция управляемых алгоритмов модуля обработки и анализа изображений

Одна из ключевых проблем разработки алгоритмического обеспечения систем распознавания заключается в противоречивом характере основных показателей эффективности таких систем. К указанным показателям относятся: вероятность правильного распознавания образа текстовой метки P_{right} , вероятность «пропуска метки» P_{lost} (ОТМ не найден или неверно распознан), вероятность P_{err} распознавания «ложных» ТМ на неинформативных изображениях и время анализа отдельного изображения T . Оценка значений P_{right} и P_{lost} рассчитывается по изображениям, содержащим ОТМ (информативным

изображениям), причем $P_{lost} = 1 - P_{right}$, а значение P_{err} – по изображениям, не содержащим ОТМ (неинформативным).

Обычно в процессе разработки и настройки СРТМ задаются такие параметры алгоритмов, которые позволяют получить определенный компромисс показателей эффективности. Недостаток такого подхода заключается в том, что ко всем входным данным применяются одни и те же алгоритмы обработки с постоянными значениями параметров. Улучшить показатели эффективности возможно в том случае, если целенаправленно изменять параметры алгоритмов в зависимости от того, информативные данные поступили на вход алгоритма или неинформативные. Чтобы получить такую возможность, предложено осуществлять управление алгоритмами МОАИ, используя обратную связь от АР, способного определять, информативные данные обрабатываются или нет.

Управляемые параметры алгоритмов могут иметь значения, минимизирующие показатели T и P_{err} («жесткие» параметры) либо максимизирующие P_{right} («мягкие»). «Мягкие» параметры, как правило, задают более подробную процедуру анализа данных, которая требует больших вычислительных затрат. Например, при анализе изображения символа желательно предложить как можно больше вариантов сегментации, чтобы повысить вероятность правильного распознавания. «Жесткие» параметры, напротив, минимизируют затраты времени на анализ изображений, но при этом снижают вероятность P_{right} . Использование обратной связи от алгоритма распознавания позволяет принять решение о том, следует ли установить «мягкие» параметры алгоритма, чтобы максимально эффективно произвести дополнительный анализ данных.

Рассмотрим применение обратной связи на примере алгоритма сегментации (рис. 2).

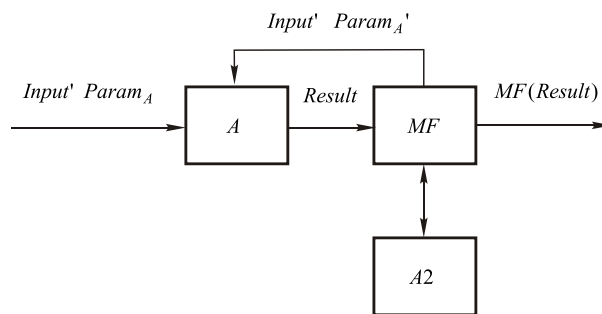


Рис. 2. Схема управления алгоритмом МОАИ

Обозначим алгоритм сегментации A , $Input = \{I_1, I_2, \dots, I_N\}$ – его входные данные, $Result = \{R_1, R_2, \dots, R_N\}$ – выходные, $Param_A$ – его параметры, тогда $A(Input, Param_A) = Result$. Выходные данные анализируются управляющей функцией MF с применением обратной связи от АР, обозначенного $A2$.

Задачами управляющей функции (УФ) MF являются:

1. Определение потребности вторичного вызова алгоритма A .
2. В случае необходимости выработка управляющих воздействий на алгоритм A , которые представляют собой изменения управляемых параметров алгоритма A .
3. В случае необходимости повторный вызов алгоритма A .

Обратная связь служит для определения информативности и полноты $Result$. При определенных условиях, например, выход $Result$ является неполным и $Result \neq \emptyset$, УФ повторно вызывает алгоритм A . Его входные данные при этом сокращаются, из списка $Input$ исключаются информативные (согласно выходу A_2) объекты $Input' \subset Input$. При этом УФ устанавливает для алгоритма A новые параметры $Param'_A$, которые максимизируют вероятность обнаружения информативных объектов. УФ осуществляет один или больше повторных вызовов алгоритма A , после чего передает управление следующему алгоритму МОАИ:

$MF(Result) =$

$$\left\{ \begin{array}{l} Result, \text{ если выход полный или вызов последний} \\ \{Result, MF(A(Input, Param'_A))\} \text{ иначе} \end{array} \right\}$$

Представленный вариант УФ является общим, на практике УФ может быть изменена. Во-первых, входные данные могут анализироваться в полном объеме с использованием «мягких» параметров для повышения надежности распознавания символов: $Input' = Input$. Во-вторых, УФ может выполнять роль фильтра выходных данных алгоритма, такой вариант описан ниже на примере управляемого АЛ. В-третьих, алгоритм, вызываемый УФ для дополнительной обработки данных, может быть иным, отличным от алгоритма A .

3. Управляемый алгоритм локализации

В соответствии с предложенной концепцией разработана обобщенная структура управляемого алгоритма локализации (УАЛ) (рис. 3). Задачей УФ АЛ является определение информативности найденных фрагментов изображения путем использования обратной связи от АС и АР. Фрагменты имеют ширину, достаточную для размещения образов 2-3 символов и высоту, примерно равную высоте символов.

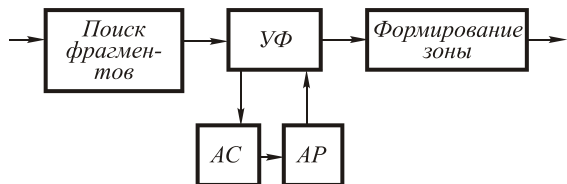


Рис.3. Структура управляемого АЛ

На первом шаге управляемой локализации осуществляется поиск фрагментов изображения, потенциально содержащих символы:

$$Frgm = \{frgm_1, frgm_2, \dots, frgm_k\},$$

где $frgm_i = \{x_1, y_1, x_2, y_2\}$ – координаты фрагмента, k – количество анализируемых фрагментов.

Управляющая функция ранжирует список фрагментов согласно некоторой оценке правдоподобия (ОП). Далее в порядке убывания ОП осуществляется сегментация фрагментов и распознавание сегментов, пока не будет найден информативный фрагмент. Правильно выбранный подход к вычислению ОП позволяет обнаружить информативный фрагмент достаточно быстро, то есть не анализировать все k фрагментов. После обнаружения информативного фрагмента АЛ вычисляет непосредственно координаты зоны.

Особенностью УФ АЛ является то, что она управляет процессом формирования зон, используя неявную обратную связь. То есть в момент образования управляющего воздействия на функцию формирования зоны УФ фактически уже имеет некоторую обратную связь от алгоритмов сегментации и распознавания.

4. Управляемый алгоритм сегментации

Одной из ключевых функций МОАИ является сегментация, поскольку для эффективного распознавания образов символов необходимо их как можно более точно сегментировать, что, как правило, достигается за счет анализа нескольких вариантов сегментации каждого символа. Обобщенная структура управляемого АС представлена на рис. 4. На первом этапе АС осуществляет поиск вертикальных разделителей – границ между образами символов. На втором этапе осуществляется уточнение вертикальных и горизонтальных границ каждого участка между разделителями. Вариантов границ может быть несколько, поэтому удобным средством повышения эффективности АС является управление функцией уточняющей сегментации.

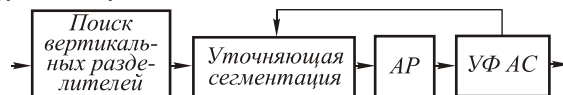


Рис.4. Структура управляемого АС

Задачами УФ АС являются:

1. Определение потребности вторичного вызова уточняющей сегментации и распознавания.
2. В случае необходимости реализация управляющих воздействий на АС, которые представляют собой изменения управляемого параметра АС, с целью увеличить количество вариантов уточняющей сегментации.
3. В случае необходимости повторный вызов алгоритмов АС или АР.

Управляемый АС в комплексе с АР анализируют информативное изображение несколько дольше, чем те же самые алгоритмы без УФ, однако позволяют достичь существенного выигрыша времени при обработке неинформативных изображений, что показывают приведенные ниже результаты экспериментов.

5. Результаты экспериментов

Приведенные эксперименты проводились в рамках решения задачи по созданию алгоритмического обеспечения системы идентификации объектов подвижного состава железнодорожного транспорта ARSCIS (www.artvision.mallenom.ru/). Исследуемые алгоритмы реализованы в математической среде MATLAB. Тесты производились на компьютере со следующей конфигурацией: частота двухъядерного микропроцессора 1,83 ГГц, объем оперативной памяти 1 Гб. Тестовая выборка состояла из 550 изображений (кадров) железнодорожных цистерн и вагонов с видимым идентификационным номером (текстовой меткой) на борту, а также 688 изображений цистерн и вагонов без номера в кадре.

Эксперименты показали, что эффективность АЛ за счет использования УФ практически не изменилась: без УФ $\tilde{P}(L) = 0,97475$, с УФ $\tilde{P}(L) = 0,97273$. В то же время, как показывает таблица 1, УАЛ позволил сократить среднее время \tilde{T}_{Empty} обработки неинформативного изображения в 2,8157 раз при незначительном увеличении времени \tilde{T}_{TM} анализа информативных изображений. Для эксперимента использовался АС без УФ с «мягкими» параметрами.

Таблица 1. Эффективность АЛ

Показатели эффективности	АЛ	Управляемый АЛ
$\tilde{P}(L)$	0,97	0,97
\tilde{T}_{TM} , сек	0,60	0,63
\tilde{T}_{Empty} , сек	0,71	0,25
\tilde{P}_{right}	0,68	0,68

При исследовании эффективности АС с УФ и без нее в составе МОАИ использовался АЛ без УФ. Исследования подтвердили эффективность предлагаемого метода – применение УФ в АС несколько увеличивает время анализа информативных изображений, однако при обработке неинформативных изображений выигрыш времени оказывается существенным (табл.2).

Таблица 2. Эффективность АС

Показатели эффективности	АС	Управляемый АС
\tilde{T}_{TM} , сек	0,60	0,87
\tilde{T}_{Empty} , сек	0,71	0,40
\tilde{P}_{right}	0,68	0,68
\tilde{P}_{err}	0,0014	0,0014
\tilde{C}	3,87	2,60

Также использование управляемого АС позволило снизить вероятность возникновения ложных рас-

познаваний. На относительно небольшой выборке оценить вероятность P_{err} сложно, поэтому в таблице также приведен показатель, во многом определяющий P_{err} – среднее количество распознанных на неинформативной зоне символов \tilde{C} . Применение УФ позволяет снизить этот показатель.

Показатели эффективности МОАИ, полученные с использованием указанной тестовой выборки изображений, являются точечными. Чтобы получить представление о точности, например, $\tilde{P}_{right} = 0,68$, построим доверительный интервал этой оценки. При надежности оценки в 0,95 доверительный интервал равен (0,65, 0,72). В этих границах с вероятностью 0,95 заключена доля изображений, которые будут распознаны, если подать их на вход МОАИ, который обеспечивает правильное распознавание ОТМ на 378 изображениях из 550.

Заключение

Таким образом, задача проектирования, реализации и настройки алгоритмического обеспечения СРТМ определяется сложным процессом поиска компромиссов между показателями эффективности СРТМ. Для повышения эффективности СРТМ предложена новая концепция структуры алгоритмов МОАИ, основанная на идее управления поведением алгоритмов в процессе функционирования посредством изменения их управляемых параметров за счет использования обратных связей.

В рамках решения задачи по разработке алгоритмического обеспечения системы идентификации объектов подвижного состава железнодорожного транспорта были разработаны методы управления алгоритмами локализации и сегментации ОТМ, использующие полученную посредством обратной связи информацию о результатах распознавания анализируемых объектов.

Результаты экспериментов показали, что при заданных вероятностных показателях МОАИ использование УФ позволяет значительно снизить время анализа изображений, давая возможность обрабатывать за единицу времени большее количество кадров, что, в конечном итоге, повышает эффективность СРТМ. В рассмотренном примере наиболее эффективной оказалась конфигурация с управляемым алгоритмом локализации и алгоритмом сегментации без управления. При прочих равных показателях такая конфигурация позволяет обрабатывать информативные и неинформативные изображения за меньшее время.

Литература

1. **Балыков, Е.А.** Разработка и обеспечение качества программных средств обработки и анализа изображений / Е.А. Балыков, В.А. Царев – М.: Компания Спутник+, 2006.- 178 с.
2. 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.

METHODS OF TEXT LABEL RECOGNITION PROCESS CONTROL

E.M. Voskresensky¹, V.A. Tsarev¹

*¹Institute of Management and Information Technologies (branch) of the St.-Petersburg State Polytechnical University,
St.-Petersburg, Russia*

Abstract

In article a new approach to text labels recognition systems algorithms designing is offered. Method bases on idea of automatic control of algorithms parameters. The choice of parameters values is carried out with use of feedback from symbols recognition algorithm. Results of experiments with algorithms of real system of railway cars identification numbers recognition are submitted.

Key words: recognition of symbols, recognition process control.

Citation: Voskresensky EM, Tsarev VA. Methods of text label recognition process control. Computer Optics 2008; 32(4): 413-6.

References

- [1] Balykov EA, Tsarev VA. Development and quality control of image processing and analysis software [In Russian]. Moscow: Sputnik+ Company 2006; 178 p.
- [2] Tsarev VA. Basic principles of vehicles optoelectronic identification system development. Pattern Recognition and Image Analysis 2005; 15(2): 454-457.