

СИСТЕМА ТЕХНИЧЕСКОГО ЗРЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ НОМЕРОВ ЖЕЛЕЗНОДОРОЖНЫХ ЦИСТЕРН С ИСПОЛЬЗОВАНИЕМ МОДИФИЦИРОВАННОГО КОРРЕЛЯТОРА В МЕТРИКЕ ХАУСДОРФА

Волотовский С.Г., Казанский Н.Л., Попов С.Б., Хмелев Р.В.
Институт систем обработки изображений РАН
Самарский государственный аэрокосмический университет

Введение

В настоящее время на нефтеналивных терминалах актуальной становится задача автоматизации всего технологического процесса наполнения железнодорожных составов цистерн. Поступающие цистерны имеют уникальные восьмизначные номера, двух- или трехзначные коды типа цистерны (определяющие габариты и вместимость) и двузначные коды принадлежности стране (рис. 1). Эта информация для всех цистерн прибывающих составов в кратчайшие сроки должна быть введена в автоматизированную систему управления технологическим процессом.



Рис. 1. Информация о цистерне: восьмизначный номер, тип (62) и код принадлежности стране (20)

Ручной ввод информации (на основании рукописных заметок оператора или с использованием видеозаписи слежения) занимает достаточно много времени и служит источником ошибок. Именно поэтому желательно максимально автоматизировать ввод информации о составах, используя систему технического зрения. Ввиду сильного загрязнения значительного количества цистерн (рис. 2) задача стопроцентного автоматического распознавания не ставилась, требовалось лишь добиться как можно большего количества автоматически распознанных номеров цистерн и, в конечном счете, снизить нагрузку на оператора, предоставив ему удобную среду для визуального контроля и редактирования автоматически сформированного списка цистерн.

Для решения данной задачи была разработана система технического зрения (СТЗ) «Система регистрации железнодорожных составов цистерн» [1, 2, 3].

Нефтеналивной терминал состоит из наливной платформы, с обеих сторон которой имеются подъездные пути, что позволяет одновременно обслуживать два состава. Эти пути соединяются в один въездной путь, по нему осуществляется въезд и выезд всех поездов. В месте соединения путей расположена площадка видеонаблюдения, на которой с обеих сторон пути расположены две цифровые видеокамеры, направленные на фоновые щиты с уникальным контрастным рисунком, не встречающимся на вагонах (рис. 3). Использование двух видеокамер

повышает надежность распознавания номеров на цистернах, имеющих сильное загрязнение.

Система работает круглосуточно в условиях естественного освещения в дневное время и при искусственном – в ночное.



Рис. 2. Примеры номеров цистерн с разной степенью загрязнения: а) чистые, б) умеренно загрязненные, в) сильно загрязненные

СТЗ обеспечивает непрерывное наблюдение за въездными путями, автоматическое обнаружение и видеозапись проходящих составов, определение направления их движения, подсчет количества цистерн в составе, локализацию и распознавание номеров цистерн, формирование отчета о составе. Сформированные отчеты предоставляются оператору, который проверяет результаты машинного распознавания и имеет возможность добавить дополнительную информацию о цистернах. Дополнительная информация о типе и принадлежности цистерны либо выбирается из базы данных известных цистерн, либо вводится вручную (и при этом автоматически добавляется в базу и используется при повторном появлении данной цистерны). Проверенные и отредактированные данные передаются в автоматизированную систему верхнего уровня.

Основные задачи системы технического зрения

Рассмотрим основные задачи, которые были решены в процессе создания системы регистрации железнодорожных составов цистерн, и возникающие при этом проблемы.

Ввод изображений. СТЗ осуществляет непрерывный ввод данных с двух видеокамер для оперативного контроля при постоянной автоматической подстройке выдержки камер. Работа при естественном освещении сопровождается значительными изменениями условий освещенности в течение суток, что приводит к появлению на изображениях значительного контраста по яр-

кости и даже засветке отдельных участков при определенном положении солнца. На одной из камер для подавления возникающих на цистернах солнечных бликов был установлен поляризационный фильтр.

В таких условиях получение качественных бинарных изображений, формируемых в процессе работы СТЗ, становится нетривиальной задачей. Для ее решения была разработана итерационная процедура бинаризации с использованием локальных порогов.

Обнаружение состава и разбиение на вагоны

Система автоматически определяет момент появления поезда. Для решения данной задачи используются фоновые щиты, расположенные напротив каждой камеры (рис. 3).

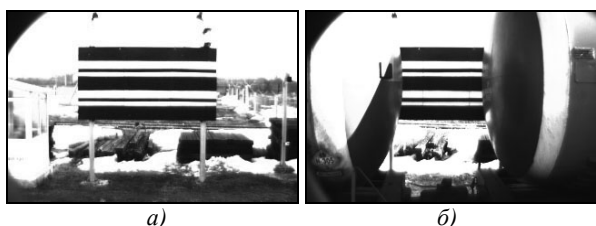


Рис. 3. а) Изображение полного контрастного щита, б) тот же щит, частично закрытый цистернами

Если состава нет, то фоновый щит виден целиком. Как только часть фонового щита перекрывается поездом, система начинает запись кадров проходящего поезда. При этом первые кадры обрабатываются особо. На этом этапе определяется наличие поезда и направление его движения, отсеиваются различные помехи, которые могут частично закрывать фоновый щит, например: проходящий рабочий с высокой лестницей, тень от рядом стоящего столба и т.п. При обработке первых кадров выявляются также случаи, когда поезд состоит из одного локомотива. Обработка такого состава не выполняется.

При анализе степени перекрытия щита он делится на 30 участков (вертикальных полос), высота которых равна высоте всего щита, а ширина – 1/30-ой всей ширины. Видимость каждой такой полосы определяется с помощью сравнения ее бинарного изображения с эталоном. При перекрытии крайних полос система регистрирует начало движения поезда, порядок последующего перекрытия полос определяет направление движения поезда.

Последовательный анализ видимости полос фонового щита (порядка, в котором полосы перекрываются и появляются) позволяет зафиксировать промежутки между вагонами, изменение направления движения состава, выявить то, что перекрытие щита вызвано не наличием на путях состава цистерн, а другими мешающими факторами.

Локализация и распознавание номеров цистерн

Объекты наблюдения – цистерны, имеют различную длину и высоту. Промежутки между цистернами не одинаковы.

Цистерны могут иметь, как светлую окраску с темными номерами, так и наоборот, темную со светлыми надписями (рис. 2а).

Цистерны имеют цилиндрическую форму, и положение номера относительно оси цистерны не фиксировано – номер может быть как посередине, так и существенно выше или ниже середины. В последних случаях цифры искажаются – сжимаются по вертикали, перекашиваются. Искажения вносит также широкоугольный объектив камеры (бочкообразная дисторсия) (рис. 4). Иногда номер в целом имеет небольшое отклонение от горизонтали (до 5°).

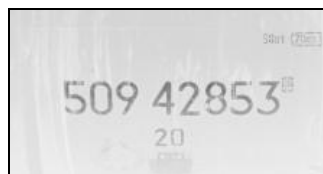


Рис. 4. Искаженный номер цистерны – цифры сжаты по вертикали, чем ближе к краю, тем больше перекося цифр

Шрифты цифр, используемые на вагонах, не являются стандартизированными, хотя в большинстве своем похожи и имеют близкие размеры символов.

Сильное загрязнение значительного числа цистерн делает задачу чтения номеров таких вагонов проблематичной даже для человека. Характер загрязнения цистерн – самый разнообразный: от небольшого естественного, до практически сплошного залива нефтяными потёками. Встречаются цистерны с мелкими контрастными пятнами различного размера.

В таких условиях в качестве основного метода распознавания был выбран метод сравнения объекта и эталона в среднеквадратической метрике Хаусдорфа [4, 5] с использованием вспомогательных алгоритмов преобразования расстояния [6, 7, 8]. В данной системе использовалась модификация алгоритма сравнения объекта и эталона в метрике Хаусдорфа, учитывающая направление обхода контуров [9]. Для преобразования расстояния использовался алгоритм, описанный в [8] как «возвратный». При переборе эталонов учитывались искажения шрифтов, связанные с кривизной цистерн, для чего использовались наборы предискаженных эталонов.

Такой алгоритм модифицированного коррелятора в метрике Хаусдорфа является вычислительно очень тяжелым, поэтому для сужения области интереса использовался вспомогательный алгоритм быстрой локализации области номера по максимуму контраста.

Данный метод распознавания работает с бинарными изображениями. Для их получения используется упомянутая процедура локальной пороговой обработки с итерационным определением порога по локальному контрасту. Из-за загрязнений область, содержащая номер, может иметь не два, а три и более различных цвета. С помощью дополнительной пороговой обработки на таких изображениях, как правило, удается выделить удовлетворительные бинарные образы цифр.

Более подробно основные методы, применяемые в СТЗ, будут рассмотрены ниже.

Структура системы и организация программного обеспечения

Аппаратная часть СТЗ состоит из двух цифровых видеокамер, компьютера на базе процессора *Intel Pentium 4* с двумя контроллерами ввода изображений на шине PCI. Компьютер подключен к локальной сети нефтеналивного терминала.

Программное обеспечение системы регистрации железнодорожных составов цистерн выполнено как многопоточное приложение и допускает запуск в нескольких режимах, в том числе для работы с одной и двумя камерами. В приложении одновременно работают от трех до пяти программных потоков команд. Главный управляющий программный поток при запуске осуществляет ввод необходимых параметров системы из конфигурационных файлов и инициализацию фабрик для рабочих потоков. Каждая камера обслуживается собственным рабочим потоком команд. После прохождения поезда программные потоки, работающие с камерами (рабочие потоки), параллельно запускают программные потоки, которые выполняют распознавание номеров цистерн (потоки-распознаватели). Главный управляющий программный поток осуществляет координацию работы всех остальных потоков, получает от них результаты работы и формирует результирующий отчет о прошедшем составе цистерн.

Алгоритм функционирования рабочих потоков

Рабочие программные потоки команд осуществляют управление видеокамерами и непрерывный ввод данных в память для оперативного контроля.

Рабочий поток может находиться в одном из трех режимов: в режиме ожидания поезда, в режиме обработки кадров проходящего поезда и в режиме ожидания улучшения освещения.

После запуска рабочий поток выполняет инициализацию видеокамеры и начинает непрерывный ввод кадров с заданной частотой с автоматической подстройкой выдержки. Если на введенном кадре фоновый щит представлен целиком, то поток переходит в режим ожидания поезда.

Корректная запись кадров поезда и разделение вагонов считается невозможным, если контрастность рисунка на фоновом щите падает ниже некоторого порога. Обычно это означает, что щит слабо освещен, или, наоборот, засвечен, или загрязнено защитное стекло камеры. В этом случае рабочий поток камеры переходит в режим ожидания улучшения освещения.

Как только часть фонового щита перекрывается поездом, рабочий поток переходит в режим обработки кадров проходящего поезда. Еще раз напомним, что первые кадры в этом режиме обрабатываются по отдельному алгоритму.

В режиме обработки кадров все вводимые видеоданные сохраняются в отдельном файле кадров поезда. Одновременно формируется промежуточный отчет о структуре поезда для данной камеры (число цистерн, индексы кадров каждой цистерны в файле кадров поезда, направление движения цистерны). Разделение состава на вагоны выполняется с помощью выделения

подпоследовательности кадров, которые содержат фоновый щит частично (рис. 3). По чередованию на них участков фона, не перекрытых поездом, определяется промежуток между цистернами и направление движения поезда в этот момент. При остановке состава запись в файл кадров поезда прекращается, она возобновляется вновь при начале движения.

Как только фоновый щит вновь появляется в кадре целиком, рабочий поток команд запускает поток-распознаватель, передавая ему сформированный промежуточный отчет и файл кадров поезда, а затем переходит в режим ожидания поезда.

Алгоритм работы распознавателя

Распознаватель на основе данных промежуточного отчета последовательно выделяет из файла кадров поезда подпоследовательности кадров, относящихся к одной цистерне. Для каждой цистерны кадры упорядочиваются в порядке убывания вероятности присутствия номера. Кадр обрабатывается в два этапа.

На первом этапе выполняется быстрая локализация области, в которой расположено номер цистерны. Локализация основана на сравнении локальных статистических параметров предполагаемой области номера и некоторого набора ее окрестностей.

Второй этап – собственно распознавание номера цистерны, то есть обнаружение в локализованной области восьми цифр, наиболее похожих на заданные эталоны.

Для этого выполняется сканирование выделенной области с целью обнаружения объектов с допустимым отклонением от заданных эталонов цифр.

Распознанным цифрам присваивается флаг достоверности – цифра считается распознанной достоверно, если взаимное отклонение объекта и эталона не превышает некоторого порога. Если отклонение выше порога достоверности, но ниже допустимого порога, то цифра считается распознанной недостоверно.

В конечном итоге достоверность распознавания целого восьмизначного номера определяется не только достоверностью каждой цифры, но и структурой номера (допустимые расстояния между цифрами, допустимые отклонения от горизонтали и пр.).

Поток-распознаватель дополняет промежуточный отчет, вписывая туда для каждой цистерны номер и признак его достоверности.

Скорректированный отчет передается главному управляющему программному потоку команд, который формирует объединенный отчет по результатам работы распознавателей на основании данных, полученных с каждой стороны. При этом если результаты распознавания для конкретной цистерны не совпадают, то выбирается лучший из двух по числу достоверных цифр и структуре номера.

Рассмотрим основные методы обработки подробнее.

Вычисление локального контраста и локального порога для бинаризации

На всех этапах обработки кадров используется итерационная процедура бинаризации с использова-

нием локальных порогов, вычисляемых по характеристикам скользящих окон различных размеров.

Рассмотрим, как происходит получение одного значения локального порога. Введем некоторые обозначения характеристик окна с центром в точке (x, y) :

- $S(x,y)$ – сумма пикселей в окне;
- $N(x,y)$ – количество пикселей в окне;
- $L(x,y)$ – уровень пороговой обработки;
- $N'(x,y)$ – число темных пикселей (не выше L);
- $N''(x,y)$ – число светлых пикселей (выше L);
- $S'(x,y)$ – сумма темных пикселей (не выше L);
- $S''(x,y)$ – сумма светлых пикселей (выше L);
- $A'(x,y)$ – средняя яркость темных пикселей;
- $A''(x,y)$ – средняя яркость светлых пикселей;
- $C(x,y)$ – контраст (разность между средними значениями яркости темных и светлых областей).

В самом благоприятном случае, когда изображения цифр являются контрастными двухцветными, для вычисления порога проще всего использовать среднее по окну

$$L(x,y) = \frac{S(x,y)}{N(x,y)}$$

Однако так как в большинстве случаев число пикселей фона существенно больше, чем число пикселей, относящихся к цифрам, то значение порога, вычисляемое как простое среднее, часто оказывается очень близко к значению средней яркости фона. Поэтому часть фона присоединяется к объекту, и цифры на бинарном изображении получаются утолщенными, с неровными краями (рис. 5а, б).

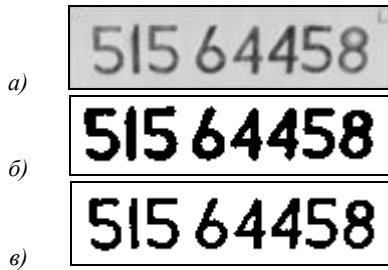


Рис. 5. а) Исходное полутоновое изображение; б) результат пороговой обработки по локальному среднему; в) результат четвертой итерации повторной бинаризации

Более приемлемый результат получается, если использовать полученное бинарное изображение как первичную маску областей "темный/светлый", затем посчитать локальные средние значения яркости темных и светлых пикселей, после чего снова выполнить бинаризацию изображения с локальными порогами, равными полусумме средних значений темных и светлых областей:

$$L_0(x,y) = \frac{S(x,y)}{N(x,y)}$$

$$A'_0(x,y) = \frac{S'_0(x,y)}{N'_0(x,y)}, \quad A''_0(x,y) = \frac{S''_0(x,y)}{N''_0(x,y)}$$

$$L_1(x,y) = \frac{A'_0(x,y) + A''_0(x,y)}{2}$$

$$L_{n+1}(x,y) = \frac{A'_n(x,y) + A''_n(x,y)}{2}$$

В этом случае относительное количество пикселей фона и цифр играет менее существенную роль при вычислении значений порога. Эту операцию можно выполнить несколько раз.

В системе при локализации области номера данная процедура выполнялась один раз (поскольку при локализации точность нахождения контуров не очень важна). Для устойчивой работы алгоритма модифицированного коррелятора в метрике Хаусдорфа бинарные изображения формировались в результате четырех итераций описанной процедуры (рис. 5в).

Локализация области номера

Для локализации расположения номера цистерны желательно использовать вычислительно простую и надежную процедуру. В СТЗ для этой цели использован поиск максимума разности контраста в области предполагаемого номера и некоторой прилегающей к ней окрестности. Упрощенно это представлено на рис. 6.

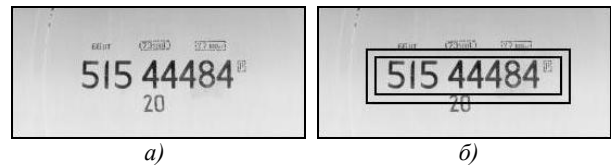


Рис. 6. а) Исходный вид номера цистерны; б) контраст в области номера (внутренняя рамка) выше, чем в ближайшей окрестности (внешняя рамка); локализация производится по максимальной разнице контрастов этих двух областей

Выше была описана процедура итерационной пороговой обработки с локальными порогами, вычисляемыми по характеристикам скользящего окна. При локализации формируется поле локальных контрастов со значениями

$$C_0(x,y) = A''_0(x,y) - A'_0(x,y)$$

Для успешной локализации большое значение имеет выбор размера окна. Желательно, чтобы в области цифр контраст был максимальным, для этого необходимо, чтобы при всех или почти всех положениях скользящего окна захватывались и пиксели фона, и пиксели цифр.

Наиболее естественным кажется выбор прямоугольного окна с размерами, примерно равными среднему размеру цифр, однако данное окно, как показала практика, является довольно неудачным, поскольку на цистернах выше и ниже номера часто есть дополнительные надписи, имеющие тот же размер и цвет, что и цифры (рис. 6). Справа и слева также могут присутствовать контрастные надписи, элементы конструкции цистерн. При этом локализация смещается в область этих дополнительных надписей или элементов конструкций.

Наилучшие результаты получены при использовании окна с высотой, несколько меньшей, чем расстояние от номера до дополнительных надписей (рис. 7а). Такой выбор окна приводит к тому, что скользящее окно внутри области номера почти везде захватывает и фоновые, и объектные пиксели, что дает в результате хороший контраст. В то же время, в небольшой внешней окрестности в скользящее окно почти всегда попадают только фоновые пиксели, что обеспечивает низкий контраст (рис. 7б).



Рис. 7. Вычисление локальных контрастов:

- а) показаны различные положения скользящего окна;
 б) поле локальных контрастов, получаемое с помощью такого локального окна;
 в) локализация выполняется по максимуму разности контраста внутри окна и контраста в четырех линиях

Для каждой точки поля локальных контрастов вычисляется разность между средним контрастом в окне с размерами, равными максимальному размеру номера цистерны за вычетом размеров маленького скользящего окна, и средним контрастом в прилегающей окрестности, обрамляющей границы окна номера на величину маленького скользящего окна (рис. 7в). Область локализации определяется как точка максимума разности контрастов.

Этот алгоритм обеспечивает правильную локализацию в большинстве случаев даже для грязных цистерн и при наличии солнечных бликов – кривизна цистерн такова, что вертикальный размер бликов существенно больше размера маленького окна, используемого при вычислении локального контраста, поэтому внутри бликов контраст низкий.

Распознавание номера цистерны

После того, как локализована область номера цистерны, внутри нее производится поиск объектов, наиболее похожих на заданные эталоны цифр. Предварительно выполняется бинаризация фрагмента с помощью описанной выше итерационной процедуры (4 итерации) с размером скользящего окна, равным среднему размеру цифры. После пороговой обработки бинарные изображения приводятся к виду «черные цифры на белом фоне», при этом используется априорная информация о том, что пикселей фона больше, чем пикселей цифр.

При поиске сканируется выделенная область с целью обнаружения объектов с допустимым отклонением от заданных эталонов цифр. Данная процедура выполняется с помощью модификации алгоритма сравнения объекта и эталона учетом направления обхода контуров. Детально данная модификация описана в [9], базовый алгоритм – в [5], возвратный алгоритм преобразования расстояния – в [8], поэтому здесь опишем лишь специфику применения в данной системе.

Специфика состоит в том, что отклонение эталона от объекта считается как отклонение контуров с учетом направления обхода, а вот отклонение объекта от эталона считается как отклонение бинарного образа объекта от бинарного образа эталона. Суммарное взаимное отклонение вычисляется по формуле:

$$d = \sqrt{\frac{\Sigma_{std} + \Sigma_{obj}}{N_{std} + N_{obj}}},$$

где Σ_{std} – сумма квадратов отклонений контуров эталона от контуров объекта с привязкой, учитывающей допустимость направления; N_{std} – число контурных точек эталона; Σ_{obj} – сумма квадратов отклонений бинарного образа объекта от бинарного образа эталона; N_{obj} – число точек бинарного образа объекта, попавшего в значимую область.

Такая схема обеспечивает наибольший процент правильных распознаваний. Использование контуров объекта для подсчета отклонения объекта от эталона ухудшает результаты, что объясняется тем, что на реальных изображениях помимо линий цифр часто присутствуют шумы небольшого размера и толщены. Эти шумы дают много контурных точек, у которых, кроме того, еще и хаотическая ориентация направления обхода, и в подсчетах отклонений объекта от эталона эти шумовые контуры играют большую роль, чем малочисленные точки контуров толстых линий цифр.

Для комбинированного взаимного отклонения существуют два порога – порог допустимости и порог достоверности (который меньше порога допустимости). Если отклонение не превышает порога достоверности, цифра считается распознанной с высокой степенью достоверности. Если отклонение больше порога достоверности, но не превышает порога допустимости, цифра считается распознанной недостоверно.

Существование двух порогов связано с тем, что шрифты цифр на вагонах не стандартизированы и многовариантны, с разной степенью отклонения от основного шрифта. Сравнительно высокий порог допустимости чаще всего позволяет правильно распознать эти нестандартные цифры, однако при этом нельзя говорить о достоверности.

Борьба с загрязнениями номеров.

Цистерны часто сильно загрязнены, и области номеров не являются двухцветными. После пороговой обработки может оказаться, что в качестве объ-

екта для сравнения с эталоном выделились грязные пятна вместо цифр или цифры, искаженные прилепленными грязными пятнами. К счастью, во многих случаях загрязнение отличается по цвету от цифр, и после дополнительной обработки от него можно избавиться или хотя бы значительно уменьшить его влияние.

Для этого после формирования бинарного изображения можно провести еще одну процедуру пороговой обработки внутри области, соответствующей цифрам, т.е. объектные пиксели еще раз разделяются на фон и объект. При этом если есть загрязнение, которое отличается по цвету от цифр, то таким образом часто удается его удалить (рис. 8).

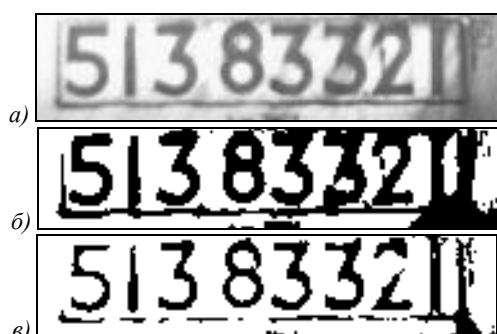


Рис. 8. Использование повторной пороговой обработки объектных пикселей для борьбы с загрязнениями:

- а) исходное полутоновое изображение;
 б) результат первой итерационной пороговой обработки;
 в) результат повторной пороговой обработки объектных пикселей

Использование повторной пороговой обработки объектных пикселей может привести как к улучшению, так и к ухудшению бинарного изображения (если загрязнения нет, то пропадают фрагменты цифр). Поэтому при распознавании цифр сканированию подвергаются оба фрагмента, и в каждую точку записывается лучший результат.

Использование предсказанных эталонов.

Как уже упоминалось, цифры часто имеют искажения в связи с цилиндрической формой цистерн и aberrациями объектива.

В СТЗ учитывался только один вид искажений – вертикальное сжатие цифр при существенном смещении номера вверх или вниз относительно середины цистерны. Для этого были сформированы наборы сжатых по вертикали эталонов в зависимости от расположения номера. После локализации области номера определялось ее вертикальное смещение, и выбирался соответствующий набор сжатых эталонов.

Применение данной техники позволило повысить число достоверных распознаваний среди расположенных таким образом на цистернах номеров.

Формирование результата распознавания.

При сканировании предполагаемой области номера цифры, как правило, обнаруживаются в нескольких соседних точках. В результате необходимо

оставить лишь одну в некоторой окрестности. У цифр существует небольшая допустимая область перекрытия, т.к. цифры могут быть расположены плотно и их области могут пересекаться краями. В случае если перекрытие существенно, менее достоверный результат удаляется.

Процедура удаления менее достоверных результатов осложняется тем, что существует особая цифра «1», которая на цистернах изображается в виде вертикального отрезка и может быть частью других цифр (часто обнаруживается на цифрах «0», «4», «7», иногда «8»). Основной принцип удаления состоит в том, что сначала удаляются менее достоверные перекрывающиеся широкие цифры, единицы при этом игнорируются. Только затем происходит финальная очистка от менее достоверных цифр, когда единицы перекрываются более достоверными единицами и оставшимися широкими цифрами.

Следует отметить, что из-за того, что единицы являются структурной частью других цифр, у них иначе, чем у прочих цифр определяется достоверность. А именно, единица может быть признана достоверной, только если она имеет отклонение в пределах достоверности и плотно зажата с обеих сторон двумя широкими цифрами. Таким образом, все распознанные номера с широкими расстояниями между единицами, с несколькими единицами, идущими друг за другом, или с крайними единицами (которые не могут быть окружены с обеих сторон) считаются недостоверными.

Проверка структуры номера цистерны.

После того, как найдены все цифры, из них формируется восьмизначный номер. Если найдено меньше восьми цифр, это означает неудачу распознавания. Если больше или ровно восемь, запускается процедура формирования из данных цифр номера с допустимой структурой.

Проверка на допустимость структуры включает в себя проверку расстояний между цифрами, проверку допустимого отклонения средних цифр от прямой, соединяющей крайние цифры, допустимость начальной цифры (все цистерны имеют номера, начинающиеся с «5» или «7»).

Оптимизация процесса распознавания

Организация рассматриваемой СТЗ не требует выполнения распознавания номеров в реальном времени. Тем не менее, ограничения по времени обработки были заданы. В связи с этим для ускорения процесса распознавания использованы следующие технологические решения.

Настройка на путь.

Перед камерами расположены два железнодорожных пути, которые проходят на различном расстоянии от камер, поэтому для каждого пути средние размеры цифр и размеры предполагаемой области номера различны, что учитывается при распознавании.

Для уменьшения времени распознавания поезда выполняется настройка распознавателя на ближний

или дальний путь. Для первых вагонов поезда производится распознавание номеров с объединенным набором эталонов, соответствующих как ближнему, так и дальнему пути, затем результаты сравниваются, и выбирается лучший по структуре и суммарной достоверности.

После того, как число достоверных результатов для одного пути существенно превысит число достоверных результатов для другого пути, один из путей более не рассматривается. "Существенно превысит" в данном случае означает, что число достоверных результатов для одного пути больше на несколько штук и в несколько раз, чем число достоверных результатов для другого:

$$r_1 > r_2 + a \text{ и } r_1 > b \cdot r_2,$$

где r_1 и r_2 – число достоверных результатов для каждого пути, a и b – некоторые константы.

Оптимизация сканирования.

Особенность используемого алгоритма модифицированного коррелятора в метрике Хаусдорфа такова, что если в текущей точке отклонение эталона от объекта слишком велико, то и в соседних оно также не войдет в допустимые пределы, поэтому ближайших соседей текущей точки можно не рассматривать. На этом основана следующая оптимизация – если в центре некоторой области размером 3×3 пиксела отклонение от эталона слишком велико, то в остальных точках области поиск не выполняется.

Кроме того, в данной задаче цифры допускают лишь небольшие вариации по размеру и углу поворота, и если отклонение одного варианта эталона со средним размером и средним углом поворота слишком велико, то никакие небольшие вариации размера и угла не позволят достичь допустимого отклонения. Поэтому поиск цифры в данной точке можно свести к проверке одного варианта.

При использовании этих оптимизаций применялись дополнительные пороговые значения отклонений, определяемые экспериментально таким образом, чтобы с одной стороны, уменьшалось время сканирования предполагаемой области номера, с другой – не ухудшались результаты распознавания тестовых выборок из-за слишком жесткого отсека. В конечном итоге с использованием этих оптимизаций удалось достичь выигрыша по времени в 3–4 раза.

Машинно-зависимая оптимизация. Все процедуры нижнего уровня системы являются машинно-оптимизированными. Следует отметить, что процедуры вычисления сумм квадратов отклонений хорошо оптимизируются под расширенные наборы команд Intel MMX и SSE2 (параллельная обработки целых чисел). При использовании расширенных наборов команд процессора удалось добиться примерно 40% ускорения, используя для хранения квадратов расстояний 16-битные целые числа вместо 32-битных. При типичных значениях расстояний

между пикселями в пределах от 10 до 30 это не вызвало потери точности.

Полная обработка одного изображения, включающая локализацию и распознавание, занимает 0.07-0.15 секунды на машине Athlon XP 2500+, чем грязнее изображение – тем дольше.

Результаты

При разработке алгоритмов системы, их доводке и оптимизации использовалась тестовая выборка из 820 цистерн, которая была отобрана таким образом, чтобы содержать большинство встречающихся особенностей написания номеров и характерных загрязнений.

Цистерны в составах имеют разную степень загрязнения, поэтому для улучшения интерпретируемости результатов распознавания вся исследуемая выборка была разделена экспертами на три группы. Первая группа содержит цистерны, номера которых, по мнению экспертов, подлежат безусловному распознаванию. Вторая группа содержит цистерны с номерами, распознавание которых желательно, но, на взгляд экспертов, очень проблематично из-за характера загрязнений или особенностей написания цифр номера. Третья группа состоит из цистерн, номера которых были прочитаны экспертами с существенными затруднениями.

Результаты исследования работы распознавателя с использованием модифицированного коррелятора в метрике Хаусдорфа представлены в таблице. Таким образом, у цистерн, относящихся к первой группе, правильно распознается порядка 96% номеров, ко второй – 68%, хотя часть из них помечается признаком недостоверности (таблица 1).

Таблица 1.

Результаты распознавания тестовой выборки

Группа	1	2	3
Всего цистерн	426	271	123
Распознано ВЕРНО , из них с признаком:	402	185	18
Достоверно	247	54	3
Недостоверно	155	131	15
Отказ	0	0	0
Распознано НЕВЕРНО , из них с признаком:	24	86	105
Достоверно	0	0	0
Недостоверно	11	27	17
Отказ	13	59	88

Алгоритм быстрой локализации для цистерн первой группы обрабатывает правильно почти во всех случаях (случаи неправильной локализации единичны), для второй группы – в 95% случаев.

С мая по сентябрь 2004 года система находилась в опытной эксплуатации, за это время накопилась некоторая статистика (таблица 2).

Таблица 2. Результаты опытной эксплуатации

Обработано 275 поездов, 8187 вагонов		
Результаты распознавания		
3557	43,4%	высокая достоверность
3804	46,5%	сомнительные
826	10,1%	отказ
Результаты проверки		
6884	84,1%	подтверждено
1303	15,9%	ошибочные
Результаты проверки по категориям		
Высокая достоверность		
3507	98,6%	подтверждено
50	1,4%	ошибочные
Сомнительные		
3377	88,8%	подтверждено
422	11,2%	ошибочные

Отдельного пояснения требует строка ошибочных распознаваний в категории номеров с заявленной высокой достоверностью (50 вагонов – 1,4%). На цистернах изредка встречаются шрифты, в которых цифры «6» и «9» очень похожи на цифры «8» (рис. 9). В тестовой выборке такого начертания шрифтов не было, что и явилось причиной ошибок «достоверного» результата распознавания. Одни и те же вагоны, появляясь на терминале много раз, сформировали данную статистику.



Рис. 9. Пример номера, в котором цифры «6» и «9» похожи на цифру «8»

По результатам анализа опытной эксплуатации набор базовых эталонов был доработан, что устранило возникшие ошибки.

Благодарности

Работа выполнена при поддержке российско-американской программы «Фундаментальные ис-

следования и высшее образование» (BRHE), грантов Президента России № НШ-1007.2003.01 и Российского фонда фундаментальных исследований № 04-07-90149 и № 04-07-96500.

Литература

1. Волоотовский С.Г., Казанский Н.Л., Попов С.Б. Система регистрации железнодорожных составов цистерн // Свидетельство об официальной регистрации программ для ЭВМ № 2004611969 по заявке № 2004611381 от 29 июня 2004 года. Зарегистрировано в Реестре программ для ЭВМ 26 августа 2004 года.
2. Волоотовский С.Г., Казанский Н.Л., Попов С.Б. Программное обеспечение распознавания номеров на основе анализа топологии контуров // Свидетельство об официальной регистрации программ для ЭВМ № 2004611970 по заявке № 2004611382 от 29 июня 2004 года. Зарегистрировано в Реестре программ для ЭВМ 26 августа 2004 года.
3. Хмелев Р.В., Казанский Н.Л., Попов С.Б. Программное обеспечение распознавания номеров на основе анализа взаимных отклонений геометрических форм объекта и эталона // Свидетельство об официальной регистрации программ для ЭВМ № 2004611971 по заявке № 2004611383 от 29 июня 2004 года. Зарегистрировано в Реестре программ для ЭВМ 26 августа 2004 года.
4. Huttenlocher D.P., Klauder G.A., Rucklidge W.J. Comparing images using the Hausdorff-distance // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993. Vol. 15. P.850-863.
5. Казанский Н.Л., Хмелев Р.В. Сравнение объекта и эталона по отклонению контуров // Компьютерная оптика, 2000. В. 20. С. 134-139.
6. Mauch S. A Fast Algorithm for Computing the Closest Point and Distance Transform // Technical Report at <http://www.acm.caltech.edu/~seanm/software/cpt/cpt.pdf>, 2000.
7. Felzenszwalb P.F., Huttenlocher D.P. Distance Transforms of Sampled Functions // Cornell Computing and Information Science Technical Report TR2004-1963, 2004.
8. Казанский Н.Л., Мясников В.В., Хмелев Р.В. Алгоритмы поиска расстояний до объектных пикселей на бинарных изображениях // Компьютерная оптика, 2000. В. 20. С. 128-133.
9. Хмелев Р.В. Совместное использование структурного анализа и метрики Хаусдорфа при сравнении объекта и эталона. // Компьютерная оптика, 2005. В. 27.

Vision system for the recognition of numbers of rail wagons using a modified correlation function in the Hausdorff metric

S.G. Volotovskiy^{1,2}, N.L. Kazanskiy^{1,2}, S.B. Popov^{1,2}, R.V. Khmelev²

¹Image Processing Systems Institute of RAS

²Samara State Aerospace University (SSAU)

Abstract:

The task of automating the entire technological process of filling the railway tank cars is currently taking priority at oil loading terminals. The incoming tank cars have unique eight-digit numbers, two- or three-digit codes of the tank type (they determine the dimensions and capacity) and two-digit codes of the country of origin. This information regarding all the tanks of the arriving trains should be entered into the automated process control system as soon as possible.

Keywords: Hausdorff metric, modified correlation, vision system, two-digit codes, automated process control system.

Acknowledgments: This work was supported by the Russian-American program Basic Research and Higher Education (BRHE), grants by the President of Russia No. HIII-1007.2003.01 and the Russian Foundation for Basic Research No. 04-07-90149 and No. 04-07-96500.

Citation: Volotovskiy SG, Kazanskiy NL, Popov SB, Khmelev RV. Vision system for the recognition of numbers of rail wagons using a modified correlation function in the Hausdorff metric. *Computer Optics* 2005; 27: 177-184.

References:

- [1] Volotovskiy SG, Kazanskiy NL, Popov SB. System for registration of oil tank wagons. Certificate of official software registration No 2004611969 based on the application No 2004611381 of June 29, 2004. Registered in the Software registry on August 26, 2004.
- [2] Volotovskiy SG, Kazanskiy NL, Popov SB. Software for the recognition of numbers based on the analysis of layout topology. Certificate of official software registration No 2004611970 based on the application No 2004611382 of June 29, 2004. Registered in the Software registry on August 26, 2004.
- [3] Khmelev RV, Kazanskiy NL, Popov SB. Software for the recognition of numbers based on the analysis of mutual deviations of geometric shapes of the object and sample. Certificate of official software registration No 2004611971 based on the application No 2004611383 of June 29, 2004. Registered in the Software registry on August 26, 2004.
- [4] Huttenlocher DP, Klauder GA, Rucklidge WJ. Comparing images using the Hausdorff distance. *IEEE Trans Pattern Anal Mach Intell* 1993; 15(9): 850-863. DOI: 10.1109/34.232073.
- [5] Kazanskiy NL, Khmelev RV. Comparison of an object and a sample for the deviation of the contours [In Russian]. *Computer Optics* 2000; 20: 134-139.
- [6] Mauch S. A fast algorithm for computing the closest point and distance transform. 2000. Source: <http://www.acm.caltech.edu/~seanm/software/cp`t/cpt.pdf>.
- [7] Felzenszwalb PF, Huttenlocher DP. Distance transforms of sampled functions. *cornell computing and information science Technical Report TR2004-1963* 2004.
- [8] Kazanskiy NL, Myasnikov VV, Khmelev RV. Algorithms for calculating distances to object pixels in binary images. *Computer Optics* 2000; 20: 128-133.
- [9] Khmelev RV. The combined use of structural analysis and Hausdorff metrics in comparison of an object to a sample [In Russian]. *Computer Optics*, 2005; 27: 174-176.