

На правах рукописи

Михайлов Иван Александрович

**РАЗРАБОТКА И ИССЛЕДОВАНИЕ МЕТОДОВ
ОБРАБОТКИ И РАСПОЗНАВАНИЯ ОБЪЕКТОВ
НА ПОСЛЕДОВАТЕЛЬНОСТИ ИЗОБРАЖЕНИЙ**

Специальность 05.13.18

**Математическое моделирование, численные методы и
комплексы программ**

Автореферат

на соискание учёной степени

кандидата физико-математических наук

Ярославль — 2011

Работа выполнена на кафедре теоретической информатики Ярославского государственного университета им. П.Г. Демидова.

Научный руководитель — доктор физико-математических наук,
профессор
Тимофеев Евгений Александрович

Официальные оппоненты: доктор физико-математических наук,
профессор
Дольников Владимир Леонидович

доктор физико-математических наук,
профессор
Дмитриев Михаил Геннадиевич

Ведущая организация — Ярославский государственный техни-
ческий университет

Защита состоится «25» ноября 2011 г., в __ часов на заседании диссер-
тационного совета Д 212.002.05 при Ярославском государственном уни-
верситете им. П.Г. Демидова по адресу: 150000, Ярославль, ул. Советская,
14.

С диссертацией можно ознакомиться в библиотеке Ярославского госу-
дарственного университета им. П.Г. Демидова.

Автореферат разослан «__» октября 2011 г.

Учёный секретарь
диссертационного совета



С.Д. Глызин

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

К настоящему моменту область обработки изображений уже прошла длинный путь развития. Важным шагом на этом пути явилось изобретение фотографии, которая позволила наглядно документировать результаты научных исследований. Следующим шагом послужило появление вычислительных машин, достаточно мощных, чтобы работать с изображениями — это позволило автоматизировать и ускорить их обработку. Сегодня обработка изображений представляет собой активно развивающуюся область знания, методы которой используются во многих естественных науках и технических дисциплинах.

Действительно, изображение является объектом исследования или его результатом в таких областях человеческой деятельности, как космонавтика, астрономия, биология, медицина, физика, геология, криминалистика, дефектоскопия и других. Изображения получают не только в диапазоне частот электромагнитного излучения, отвечающего видимому свету, но и в диапазоне частот акустическом, инфракрасном, ультразвуковом, ультрафиолетовом, рентгеновских и гамма-лучей. Большой объём видеоинформации ведёт к необходимости автоматизации процессов её обработки. Среди направлений исследований в данной области, приведённых в работе [1, стр. 5–6], выделим следующие два.

- Восстановление искажённых изображений, а также улучшение их визуального качества заключается в его преобразовании, обратном преобразованию, вызвавшему искажение. Соответствующие алгоритмы восстановления позволяют скомпенсировать искажения, обусловленные рефракцией или турбулентностью атмосферы, недостатками оптических приборов, взаимным расположением устройства съёмки и объектов наблюдения, деформацией фотоматериала, помехами и т. д. Улучшение визуального качества изображений облегчает их восприятие и анализ человеком.
- Автоматическое чтение текста, дешифрование аэрокосмических снимков, диагностика заболеваний — примеры задач, относящихся к проблеме распознавания образов. Интерес, проявляемый к её разработке в настоящее время, связан с открывающимися возможностями лучшего понимания процессов обработки информации в живых системах, создания более эффективных систем «человек-машина», а также автоматических систем, позволяющих быстрее, точнее, надёж-

нее или экономичнее человека решать различные задачи обработки визуальной информации.

Выделим отдельно задачу обнаружения на видеопоследовательности движущихся объектов и отслеживание их траектории. Для решения данной задачи (точнее, подзадачи обнаружения) оказываются полезными методы распознавания образов. Алгоритмы, разрабатываемые в процессе решения задачи обнаружения и отслеживания, предназначены для использования во всевозможных системах наблюдения и контроля.

Цели исследования

Целями диссертационной работы являются:

- 1) разработка методов распознавания чёрно-белых растровых изображений алфавитно-цифровых символов малого размера;
- 2) разработка алгоритма восстановления изображений, искажённых перспективным преобразованием;
- 3) разработка алгоритма разбиения железнодорожного состава на отдельные вагоны.

Результаты, выносимые на защиту

1. Разработаны три метода распознавания чёрно-белых растровых изображений алфавитно-цифровых символов малого размера; произведена оценка их трудоёмкости; изучены некоторые свойства расстояний между изображениями, лежащих в основе предложенных методов; предложено четыре модели искажения чёрно-белых растровых алфавитно-цифровых символов; разработана программная реализация методов; проведён ряд экспериментов для проверки эффективности предложенных методов с использованием искусственно построенных и реальных тестовых изображений, в ходе которых выполнено сравнение данных методов с аналогичным.
2. Разработан алгоритм восстановления изображений, искажённых перспективным преобразованием; приведено теоретическое обоснование данного алгоритма; разработана программная реализация алгоритма; выполнено тестирование алгоритма с использованием набора изображений.
3. Разработан алгоритм разбиения железнодорожного состава на отдельные вагоны; разработана программная реализация данного алгоритма; проведено тестирование алгоритма с использованием набора видеопоследовательностей.

Методы исследования

В диссертационной работе использованы методы обработки изображений, распознавания образов (в том числе обнаружения объектов), проективной геометрии, математического анализа.

Научная новизна работы

1. В проведённых экспериментах с использованием реальных зашумлённых цифровых символов, а также цифровых символов с искажениями каркаса результаты предложенных методов распознавания оказались близкими к результатам корреляционного алгоритма или лучшими по сравнению с ними.
2. Преимущество алгоритма восстановления изображений, искажённых перспективным преобразованием, над аналогичными состоит в том, что он не использует каких-либо параметров оптической системы, с помощью которой было получено изображение. В некоторых случаях восстановление изображений может быть полностью автоматизировано.
3. Преимущества алгоритма разбиения железнодорожного состава на отдельные вагоны по сравнению с аналогичными состоят в том, что он в одном случае является более автоматизированным, а в другом — не требует наличия вспомогательных технических средств.

Практическая значимость работы

Методы и алгоритмы, разработанные в ходе работы над диссертацией, могут быть использованы для создания систем наблюдения за движением автомобильного или железнодорожного транспорта (в том числе для решения задачи идентификации); предложенные методы распознавания могут найти применение также и в системах автоматического чтения текста.

Апробация работы

Основные положения работы докладывались и обсуждались на перечисленных ниже научных конференциях и семинарах.

1. Шестидесят вторая региональная научно-техническая конференция студентов, магистрантов и аспирантов высших учебных заведений с международным участием «Молодёжь. Наука. Инновации — 2009», 2009, г. Ярославль.

2. IV Международная научно-практическая конференция «Современные информационные технологии и ИТ-образование», 2009, г. Москва.

3. XLVIII Международная научная студенческая конференция «Студент и научно-технический прогресс»: Информационные технологии, 2010, г. Новосибирск.

4. XV Всероссийская научно-техническая конференция студентов, молодых ученых и специалистов, 2010, г. Рязань.

5. II Международная молодёжная научно-практическая конференция, 2010, г. Елабуга.

5. Шестьдесят четвёртая региональная научно-техническая конференция студентов, магистрантов и аспирантов с международным участием, 2011, г. Ярославль.

6. Семинар института программных систем им. А.К. Айламазяна РАН, 2011, г. Переславль-Залесский.

7. Семинар «Моделирование и анализ информационных систем», 2011, г. Ярославль.

8. Семинар «Нелинейная динамика и синергетика», 2011, г. Ярославль.

Публикации

По теме диссертационной работы опубликовано 15 печатных работ (в том числе 5 публикаций в ведущих рецензируемых научных изданиях, рекомендованных Высшей аттестационной комиссией), а также получено свидетельство о государственной регистрации программы для ЭВМ. Из пяти статей три написаны без соавторов, две — тремя авторами (Малков А.Н., Штерн Г.П., Михайлов И.А.). Все совместные работы написаны в нераздельном соавторстве. Список работ приведён в конце автореферата.

Структура и объём работы

Диссертация состоит из введения, трёх разделов, заключения, списка литературы и приложения. Работа изложена на 112 страницах, содержит 2 таблицы, 6 рисунков, библиография включает 60 наименований.

СОДЕРЖАНИЕ РАБОТЫ

Введение

Во введении обосновывается актуальность тем диссертационной работы и даются постановки решаемых задач.

Первый раздел

В первом разделе приводится краткий обзор методов распознавания символьных изображений, даётся постановка задачи, описывается принятый в работе общий подход к её решению. Далее последовательно определяются три предложенных метода распознавания. Эксперименты, проведённые для оценки эффективности методов, описываются в конце раз-

дела.

Постановка задачи и общий подход к её решению

Рассматривается задача распознавания чёрно-белых растровых изображений цифровых символов. Пусть имеется набор классов, для каждого из которых существует, по крайней мере, одно изображение-эталон. Пусть дано также тестовое изображение (тест). Размеры изображений: как эталонов, так и теста, — могут быть различны. Задача состоит в том, чтобы определить, к какому классу относится тестовое изображение.

Общий подход к решению задачи распознавания состоит в том, чтобы для классификации тестового изображения использовать правило ближайшего соседа: вычислим расстояние между тестом и каждым эталоном, сопоставим тесту класс того эталона, на котором реализуется минимум.

Таким образом, поставленная задача распознавания сводится к выбору меры близости на множестве символьных изображений. Поэтому описание каждого метода будет ограничено определением соответствующего расстояния.

Метод, основанный на модификации метрики Хаусдорфа

Метрика Хаусдорфа и её модификации находят своё применение в задаче распознавания изображений в качестве меры близости множеств [3, 4]. При использовании данной метрики изображения рассматриваются как конечные множества точек, причём точки одного изображения привязываются к точкам другого.

1. *Расстоянием* от множества из одной точки $\{z'\}$ до конечного множества G будем называть

$$\sigma(z', G) = \min_{z \in G} \|z' - z\|,$$

где $\|z' - z\|$ — расстояние между точками z' и z .

2. *Отклонением* множества G_1 от G_2 будем называть

$$\tilde{\sigma}(G_1, G_2) = \frac{1}{|G_1|} \sum_{z' \in G_1} \sigma(z', G_2).$$

3. *ht-расстоянием* будем называть

$$\tilde{\rho}(G_1, G_2) = \tilde{\sigma}(G_1, G_2) + \tilde{\sigma}(G_2, G_1).$$

Теорема 2.3.1 *Предложенное ht-расстояние обладает свойствами неотрицательности, эквивалентности и симметричности.*

Теорема 2.3.2 *ht-расстояние не является метрикой, так как для него не выполняется свойство $\tilde{\rho}(G_1, G_3) \leq \tilde{\rho}(G_1, G_2) + \tilde{\rho}(G_2, G_3)$.*

Для того чтобы применить описанный подход к чёрно-белым изображениям, каждому чёрному (принадлежащему символу) пикселю с координатами (x, y) сопоставим одну точку с теми же координатами. Алгоритм сравнения изображений с помощью hm -расстояния, используемый нами в экспериментах, устроен следующим образом.

1. Привести тест к размеру эталона, используя алгоритм масштабирования.
2. Вычислить hm -расстояние между эталоном и тестом для двух взаимных расположений: в первом случае совмещаются края изображений, во втором — их центры тяжести.
3. Минимум из двух вычисленных расстояний возратить в качестве результата.

Теорема 2.3.3 *Сравнение изображений с использованием hm -расстояния может быть выполнено за время $O(mn)$ (m, n — ширина и высота изображений соответственно), если в качестве расстояния между точками на плоскости применяется величина $a_x|z'_x - z''_x| + a_y|z'_y - z''_y|$, где a_x, a_y — константы.*

Метод срезов

Под *бесконечным вектором* V будем понимать вектор бесконечной длины, обладающий свойством: $\exists n : (\forall i \leq n v_i \geq 0) \wedge (\forall i > n v_i = 0)$.

Пусть V', V'' — бесконечные векторы. *Расстоянием* на множестве бесконечных векторов будем называть $\rho(V', V'') = \sum_{p=1}^{\infty} |v'_p - v''_p|$.

Теорема 2.4.1 *Предложенное расстояние на множестве бесконечных векторов $\rho(V', V'')$ обладает свойствами метрики.*

Пусть имеется матрица изображения $A = (a_{ij})$ размера $n \times m$ (n строк, m столбцов).

Для каждого $i : 1 \leq i \leq n$ определим множество чисел $S(A, i)$:

$$S(A, i) = \{s_1, s_2, \dots, s_k\} = \{0\} \cup \{j : a_{i,j} = 0 \wedge a_{i,j+1} = 1, 1 \leq j \leq m-1\} \cup \{m\}.$$

Будем считать, что $s_1 < s_2 < \dots < s_k$.

Вектором разности для тройки (A, i, N) , где $1 \leq i \leq n$, N — натуральное, будем называть бесконечный вектор действительных чисел, построенный следующим образом:

$$V_p(A, i, N) = \begin{cases} (s_{p+1}(A, i) - s_p(A, i)) \cdot N/m, & 1 \leq p \leq |S(A, i)| - 1; \\ 0, & |S(A, i)| \leq p < \infty. \end{cases}$$

Таким образом, вектор разности $V_p(A, i, N)$ содержит в себе расстояния между соседними переходами с белого на чёрный на высоте i , а

также расстояния от крайних переходов до границ изображения (первая и последняя компоненты). Параметр N играет роль коэффициента масштабирования.

Линейным образом матрицы изображения $A = (a_{ij})$ размера $n \times m$ с частотой сечения N , где N — натуральное, будем называть упорядоченное мультимножество векторов разности для этой матрицы

$$L(A, N) = \{l_p : l_p = V(A, \lceil n \cdot p / N \rceil, N), 1 \leq p \leq N\}.$$

Таким образом, линейный образ содержит векторы разности (*срезы*) для всех n строк матрицы изображения (a_{ij}) . При этом для масштабирования каждой строке сопоставляется несколько одинаковых срезов в зависимости от параметра N .

Пусть $L(A_1, N), L(A_2, N)$ — линейные образы матриц изображений A_1 и A_2 с частотой сечения, равной N .

Расстоянием на множестве линейных образов матриц изображений с частотой сечения N будем называть

$$\rho(L(A_1, N), L(A_2, N)) = \sum_{1 \leq i \leq N} \rho(l_i(A_1, N), l_i(A_2, N)).$$

Теорема 2.4.2 *Расстояние на множестве линейных образов $\rho(L(A', N), L(A'', N))$ обладает свойствами метрики.*

Алгоритм сравнения изображений с применением предложенного расстояния на множестве линейных образов, используемый нами в экспериментах, устроен следующим образом: на первом этапе происходит построение линейных образов теста и эталона, на втором вычисление значения расстояния между ними; полученное значение возвращается в качестве результата.

Теорема 2.4.3 *Сравнение изображений с использованием предложенного расстояния на множестве линейных образов может быть выполнено за время $O(mN)$, где m — ширина изображений эталона и теста (в пикселях), N — частота сечения.*

Метод радиальных окрестностей

Пусть имеется изображение A , заданное матрицей (a_{ij}) размера $n \times m$. Под *радиальной окрестностью* точки с координатами (i, j) и коэффициентом нормирования s будем понимать упорядоченный набор из четырех бинарных векторов (v^l, v^r, v^t, v^b) , построенных следующим образом:

$$\bullet v_k^l = a_{i, j'}, j' = \left\lceil m \cdot \frac{k-1}{s} \right\rceil, 2 \leq k \leq \left\lfloor \frac{s \cdot (j-0,5)}{m} + 1 \right\rfloor;$$

- $v_k^r = a_{i',j'}, j' = \left\lceil m \cdot \frac{s-k+1}{s} \right\rceil, 2 \leq k \leq \left\lfloor \frac{s \cdot (m-j+0,5)}{m} + 1 \right\rfloor$;
- $v_k^t = a_{i',j}, i' = \left\lceil n \cdot \frac{k-1}{s} \right\rceil, 2 \leq k \leq \left\lfloor \frac{s \cdot (i-0,5)}{n} + 1 \right\rfloor$;
- $v_k^b = a_{i',j}, i' = \left\lceil n \cdot \frac{(s-k+1)}{s} \right\rceil, 2 \leq k \leq \left\lfloor \frac{s \cdot (n-i+0,5)}{n} + 1 \right\rfloor$.

При этом $v_1^l = v_1^r = v_1^t = v_1^b = 1$. Каждый из этих векторов соответствует одному из четырёх основных направлений *от границы окаймляющего прямоугольника к центру окрестности* (элементу с координатами (i, j)): v^l — слева направо, v^r — справа налево, v^t — сверху вниз, v^b — снизу вверх. Каждый вектор включает в себя масштабированную (с помощью параметра s) часть строки или столбца вдоль соответствующего направления от границы изображения до центра окрестности. Исключением являются первые компоненты векторов — они всегда равны 1 независимо от цвета пикселей.

Под *радиальным образом* изображения A будем понимать множество радиальных окрестностей, построенных для каждого элемента соответствующей матрицы (a_{ij}) .

bt-расстоянием на множестве бинарных векторов будем называть

$$\tilde{\rho}_{bt}(a, b) = \sigma(a, b) + \sigma(b, a), \text{ где } \sigma(a, b) = \sum_{a_i=1} \min_{b_j=1} |i - j|.$$

Введём расстояние на множестве радиальных окрестностей:

$$\tilde{\rho}(e', e'') = \sum_{i=1}^4 \tilde{\rho}_{bt}(e'_i, e''_i),$$

где e'_i, e''_i — i -е компоненты окрестностей e', e'' соответственно.

Определим *pp-расстояние* на множестве радиальных образов:

$$\tilde{\rho}_{pp}(A, B) = \sigma(A, B) + \sigma(B, A), \text{ где } \sigma(A, B) = \frac{1}{|A|} \cdot \sum_{i=1}^{|A|} \min_j \tilde{\rho}(A_i, B_j).$$

Здесь A_i — i -я окрестность образа A , B_j — j -я окрестность образа B .

Теорема 2.5.1 *bt-расстояние на множестве бинарных векторов обладает свойствами неотрицательности и симметричности.*

Теорема 2.5.2 *bt -расстояние не является метрикой, так как для него не выполняется свойство неравенства треугольника.*

Теорема 2.5.3 *Расстояние на множестве радиальных окрестностей $\tilde{\rho}(e', e'')$ обладает свойствами неотрицательности и симметричности; свойство неравенства треугольника не выполняется для указанного расстояния.*

Теорема 2.5.4 *pp -расстояние на множестве радиальных образов изображений обладает свойствами неотрицательности и симметричности.*

Алгоритм сравнения изображений с помощью pp -расстояния, используемый нами в экспериментах, состоит из двух этапов: на первом из них для эталона и теста строятся радиальные образы, на втором — вычисляется расстояние между ними, которое и возвращается в качестве результата.

Теорема 2.5.6 *Сравнение изображений с использованием pp -расстояния на множестве радиальных образов может быть выполнено за время $O(m^2n^2s)$, где m, n — ширина и высота изображений соответственно, s — коэффициент нормирования.*

Описание экспериментов

Изображения-эталонны

В ходе проведённых экспериментов в качестве эталонов использовались два набора изображений. Первый из них включает в себя 10 изображений символов '0' – '9' высотой 14 пикселей, полученных с помощью шрифта Times New Roman. Данный набор будем обозначать CS.

Второй набор (обозначим его ES) содержит изображения символов, вырезанные из бинаризованных фотографий цистерн. Этот набор использовался в работе [5] для тестирования. Средний размер эталонов — 8×12 . Набор включает в себя 55 изображений всех цифр за исключением единицы. На рис. 1 показаны примеры изображений-эталонов, входящих в его состав.

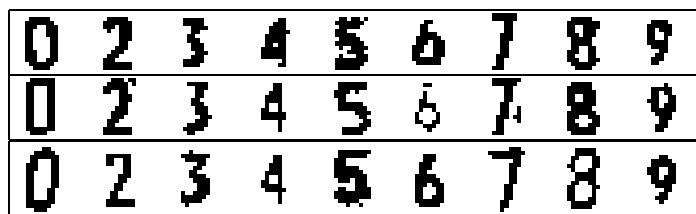


Рис. 1. Примеры изображений-эталонов

Изображения-тесты

В качестве тестов применялись пять наборов изображений. Для построения первых четырёх использовались 4 модели искажения. Первые две модели предназначены для шумовых искажений изображения, вторые две — для искажения формы символа. Таким образом, тестовые изображения были получены путём применения указанных моделей искажения к изображениям из набора CS. В результате было построено 4 набора тестов, каждый из которых содержит 5000 изображений (500 для каждого из 10 эталонов). Первые два набора с шумовыми искажениями будем обозначать NP и NL соответственно, вторые два с искажениями формы — SS и ST.

Наконец, пятый тестовый набор включает в себя изображения цифр, вырезанных с фотографий номеров цистерн после их бинаризации. Он содержит изображения всех цифр. Общее количество изображений в наборе — 1614. Средний размер тестов — 8×15 пикселей. Данный набор, некоторые изображения из которого показаны на рис. 2, будем обозначать EXP.

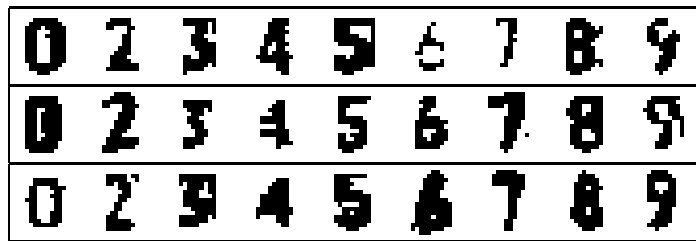


Рис. 2. Примеры тестовых изображений

Результаты экспериментов

Для сравнения представим также и результаты работы корреляционного алгоритма, описанного в работе [5]. При описании результатов будем использовать следующие обозначения:

- 1) *Corr* — корреляционный алгоритм;
- 2) *Slice* — метод срезов;
- 3) *MHaus* — метод, основанный на модификации метрики Хаусдорфа;
- 4) *Radial* — метод радиальных окрестностей.

Результаты экспериментов приведём в табличной форме: первый столбец содержит обозначения используемых наборов эталонов и тестов соответственно, разделённые косой чертой; в каждой строке приведено соответствующее число верно распознанных тестов для четырёх алгоритмов (в процентах).

Таблица 1. Сводные результаты алгоритмов распознавания

Наборы	Corr	Slice	MHaus	Radial
CS / NP	99,68	85,42	96,52	91,60
CS / NL	96,66	89,54	89,88	88,14
CS / SS	60,40	61,82	80,56	84,96
CS / ST	66,72	68,84	72,42	60,94
CS / EXP	54,34	49,81	65,86	78,44
ES / EXP	89,27	90,11	91,58	94,25

Второй раздел

Второй раздел начинается кратким обзором методов восстановления изображений, искажённых перспективным преобразованием, затем приводятся постановка задачи восстановления и дополнительные ограничения для её условия и решения. Далее описывается решение поставленной задачи. В конце раздела приведён алгоритм восстановления растровых изображений в оттенках серого, а также ряд изображений, на котором он был протестирован.

Обозначения и определения

1. Пусть в пространстве имеются две различные вертикальные плоскости p и p' , имеющие общую прямую.

2. Пусть I есть некоторое изображение в плоскости p и мы проектируем его из точки C_f , лежащей вне p и p' . В сечении проектирующих прямых плоскостью p' , мы получим новое изображение I' . Установленное таким образом соответствие между изображениями I и I' называется *перспективным преобразованием* [2].

3. Указанное в предыдущем пункте преобразование будем обозначать T_f , а точку C_f называть *фокусом перспективного преобразования* T_f .

4. Пусть имеется вертикальная плоскость p'' , также пересекающая p' .

5. Пусть преобразование T_b с фокусом в точке C_b сопоставляет изображению I' в плоскости p' изображение I'' в плоскости p'' .

6. Обозначим $l_f = |L_f C_f|$, $f_f = |F_f C_f|$, $l_b = |L_b C_b|$, $f_b = |F_b C_b|$, α_f — угол между p' и p , α_b — угол между p' и p'' (см. рис. 3).

На плоскостях p , p' , p'' введём систему координат. Для каждой из них ось Y будет направлена вверх, ось X — вправо. В плоскости p началом координат будет точка L_f , в плоскости p'' — точка L_b . В плоскости p' — точка F_f , когда будем говорить о прямом преобразовании, и F_b — когда будем говорить об обратном.

На изображениях I , I' , I'' также введём систему координат. Началом

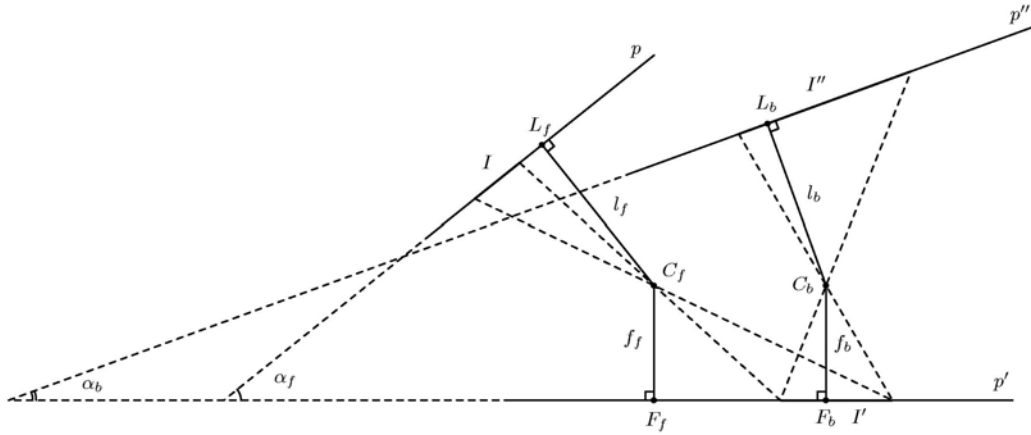


Рис. 3. Прямое и обратное преобразования. Вид сверху

координат для каждого изображения будет левый нижний угол ограничивающего прямоугольника, оси X , Y сонаправлены с осями соответствующих плоскостей.

Таким образом, вычисление преобразования T_b состоит в определении величин α_b , f_b , l_b , $\overrightarrow{(x'_s, y'_s)}$, где $\overrightarrow{(x'_s, y'_s)}$ — смещение левого нижнего угла изображения I' относительно начала координат p' . Все эти величины в совокупности будем называть также *параметрами преобразования T_b* .

Формальная постановка задачи

Пусть нам известны

1) преобразованное изображение I' , лежащее в плоскости p' , его ширина w' и высота h' ;

2) точки A'_1, A'_2, B'_1, B'_2 (их координаты относительно левого нижнего угла изображения I') такие, что если A_1, A_2, B_1, B_2 — их соответствующие прообразы, лежащие в I , то прямые A_1A_2 и B_1B_2 параллельны оси X (горизонтальны);

3) заданный размер восстановленного изображения I'' : ширина w'' и высота h'' .

Задача состоит в вычислении такого преобразования T_b (в определении его параметров), чтобы выполнялось следующее условие. Если

$$I'' = \{ (x'', y'')^T \mid (x, y)^T \in I, (T_b \circ T_f): (x, y)^T \rightarrow (x'', y'')^T \},$$

то

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} k_x & 0 \\ 0 & k_y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} s_x \\ s_y \end{pmatrix} \quad \forall \begin{pmatrix} x \\ y \end{pmatrix} \in I, \quad (1)$$

$$\max |x_1'' - x_2''| = w'' \quad \forall (x_1'', y_1'')^T, (x_2'', y_2'')^T \in I'', \quad (2)$$

$$\max |y_1'' - y_2''| = h'' \quad \forall (x_1'', y_1'')^T, (x_2'', y_2'')^T \in I''. \quad (3)$$

Величины $k_x > 0$, $k_y > 0$, s_x , s_y в условии (1) — независимые от x , y константы, координаты даны относительно точек L_f и L_b .

Теорема 3.3.1 *Для выполнения условия (1) необходимо и достаточно, чтобы выполнялись следующие соотношения*

$$f_f \operatorname{ctg} \alpha_f = f_b \operatorname{ctg} \alpha_b - x_d, \quad (4)$$

$$y_d = 0, \quad (5)$$

$$\sin \alpha_f \cdot \sin \alpha_b > 0, \quad (6)$$

где $\overrightarrow{(x_d, y_d)} = \overrightarrow{F_f F_b}$. Первые два соотношения гарантируют сохранение горизонтальности прямых, последнее — отсутствие вертикальной инверсии.

Теорема 3.3.2 *Знак α_f противоположен знаку абсциссы точки пересечения прямых, прообразы которых были горизонтальны, в координатах изображения I' .*

Теорема 3.3.3 *Следующие соотношения являются необходимыми и достаточными для сохранения горизонтальности прямых:*

$$x'_s + f_b \operatorname{ctg} \alpha_b = \frac{c_2 - c_1}{k_2 - k_1}, \quad (7)$$

$$y'_s = \frac{c_2 k_1 - c_1 k_2}{k_2 - k_1}, \quad (8)$$

где

$$k_1 = \frac{y_{12}^i - y_{11}^i}{x_{12}^i - x_{11}^i}, \quad c_1 = \frac{x_{12}^i y_{11}^i - x_{11}^i y_{12}^i}{x_{12}^i - x_{11}^i}, \quad (9)$$

$$k_2 = \frac{y_{22}^i - y_{21}^i}{x_{22}^i - x_{21}^i}, \quad c_2 = \frac{x_{22}^i y_{21}^i - x_{21}^i y_{22}^i}{x_{22}^i - x_{21}^i}, \quad (10)$$

$\overrightarrow{(x'_s, y'_s)}$ — смещение левого нижнего угла изображения I' относительно начала координат p' , а точки $A'_1(x_{11}^i, y_{11}^i)$, $A'_2(x_{12}^i, y_{12}^i)$, $B'_1(x_{21}^i, y_{21}^i)$,

$B'_2(x_{22}^i, y_{22}^i)$ лежат на прямых в плоскости p' , прообразы которых горизонтальны (координаты точек указаны в системе координат изображения I').

Теорема 3.3.4 Пусть изображение I' имеет размеры (w', h') , а изображение $I'' = (w'', h'')$. Тогда всякое преобразование T_b такое, что угол α_b удовлетворяет соотношению

$$\begin{cases} \alpha_b \in (0, \pi/2), & \text{если } m > 0; \\ \alpha_b \in (-\pi/2, 0), & \text{если } m < -w'; \end{cases}$$

где

$$m = \frac{c_2 - c_1}{k_2 - k_1} \quad (\text{в обозначениях (9) - (10)}),$$

параметр y'_s находится из соотношения (8), параметры x'_s, f_b, l_b определяются следующими соотношениями:

$$x'_s = m - \frac{\lambda w'' \cos \alpha_b}{w' h''}, \quad f_b = (m - x'_s) \operatorname{tg} \alpha_b, \quad l_b = \frac{w'' m (w' + m) \sin^2 \alpha_b}{f_b w'}$$

где

$$\lambda = \begin{cases} mh' - y'_s w', & \text{если } (m > 0 \wedge y'_s \in (-\infty, -h')) \vee (m < -w' \wedge y'_s \in [0, +\infty)); \\ h'(m + w'), & \text{если } m > 0 \wedge y'_s \in [-h', 0); \\ h'm, & \text{если } m < -w' \wedge y'_s \in [-h', 0); \\ y'_s w' + h'(w' + m), & \text{если } (m > 0 \wedge y'_s \in [0, +\infty)) \vee (m < -w' \wedge y'_s \in (-\infty, -h')); \end{cases}$$

— удовлетворяет условиям (1), (2) и (3).

Третий раздел

В начале третьего раздела приведён краткий обзор методов решения задачи обнаружения объектов на видеопоследовательности и отслеживания их траекторий, а также двух методов решения задачи разбиения железнодорожного состава на отдельные вагоны, которую можно рассматривать как особый случай задачи обнаружения и отслеживания. Затем даётся постановка задачи и описывается специфика исходных данных. Основную часть раздела занимает подробное изложение предложенного алгоритма решения задачи разбиения состава. Раздел заканчивается описанием экспериментов, проведённых для тестирования алгоритма.

Входные данные алгоритма

- Фильм, представляющий собой последовательность полутоновых кадров. Каждая цистерна представлена на нескольких кадрах (примерно от 10 до 100). Кадр имеет размер 288×352 пикселя.
- Четыре точки на кадре, задающие горизонтальные прямые в его (кадра) *нормальном прообразе*, то есть изображении, которое было бы получено, если бы горизонтальный поворот видеокамеры относительно состава был равен нулю.

Выходные данные: для тех кадров, на которых запечатлён край некоторого вагона, — абсцисса вертикальной линии, отделяющей данный вагон от соседнего вагона (вертикальная линия должна попадать в просвет, если он виден на кадре) или от фона, если вагон является первым или последним; для иных кадров выходные данные отсутствуют.

Специфика задачи

1. Интенсивности объектов (т.е. цистерн и локомотива) и фона считаются неизвестными; фон может быть неоднородным. Поверхность цистерн может быть значительно загрязнена (например, нефтяными потёками).

2. Скорость движения состава в пределах одного фильма может изменяться; длина и высота цистерн, запечатлённых в одном и том же фильме, не одинакова.

3. Расстояние между видеокамерой и составом, горизонтальный поворот между ними, а также высота видеокамеры относительно состава могут быть различны для разных фильмов.

Краткое описание алгоритма

Задача разбиения состава решается путём обнаружения просветов между вагонами. Можно выделить следующие основные этапы алгоритма.

1. Выделение пометок на кадре. Пометки — это прямоугольные области, предположительно соответствующие просветам. Для выделения пометок используется следующее эвристическое правило: средняя яркость просвета несколько больше яркости областей на кадре, лежащих слева и справа от данного просвета.
2. Определение истинных пометок. Не все пометки, выделенные на первом этапе, соответствуют просветам (т.е. не лежат в области просвета и не располагаются в непосредственной близости от неё). Для того чтобы отсеять такие пометки, выполняются следующие шаги.

- (a) Обнаруженные пометки группируются в серии таким образом, чтобы все те пометки на разных кадрах, которые соответствуют одной и той же области на поверхности вагона (или просвету), оказались в одной серии.
- (b) Серии объединяются в последовательности таким образом, чтобы любые две соседние в некоторой последовательности серии могли ограничивать собой некоторый вагон (для этого расстояние между областями на поверхности вагонов, которым соответствуют эти серии, должно лежать в определённых пределах). Серия может являться элементом одной или нескольких различных последовательностей.
- (c) Каждой серии сопоставляется вес тем больший, чем:
 - 1) больше области пометок данной серии сходны с фоном;
 - 2) больше пометок она содержит;
 - 3) в большей степени изменяется ширина пометок, составляющих эту серию.
- (d) Среди различных последовательностей выбирается та, которая предположительно в большей степени, чем другие удовлетворяет следующим критериям.
 - 1. Серии, составляющие данную последовательность, обладают большим весом.
 - 2. Расстояния между соседними сериями в этой последовательности примерно одинаковы.

После того, как выбрана искомая последовательность, считается, что составляющие её серии (и только они) соответствуют просветам (первая и последняя соответствуют началу и концу всего ж/д состава), а промежутки между ними — вагонам.

- 3. Вычисление позиции разделяющей вертикальной линии для всех кадров, на которых виден некоторый просвет или край всего состава. Для расчёта координаты вертикали используются координаты пометок той серии, которая соответствует текущему просвету, а также приближительная величина скорости состава в данный период времени.

Результаты экспериментов

Предложенный алгоритм был протестирован с использованием 16 фильмов, в каждом из которых изображено движение состава, в среднем состоящего из 26,5 вагонов. Среднее количество кадров в фильмах —

1960. Во время обработки данных фильмов алгоритм не допустил ошибку типа пропуска и ложного обнаружения края, однако на части кадров позиция отделяющей вертикальной линии была определена неточно.

Заключение

В заключении перечислены разработанные методы и алгоритмы, отмечены некоторые их достоинства и недостатки и указаны возможные области применения.

Список цитируемой литературы

[1] Анисимов, Б. В. Распознавание и цифровая обработка изображений: Учеб. пособие для студентов вузов / Б. В. Анисимов, В. Д. Курганов, В. К. Злобин. – М.: Высш. шк., 1983. – 295 с.

[2] Юнг, Дж. В. Проективная геометрия / Дж. В. Юнг. Перевод с английского под редакцией проф. В. Ф. Кагана. – М.: ИЛ, 1949. – 184 с.

[3] Kara, L. V. Sketch Understanding for Engineering Software / L. V. Kara // Ph.D. Thesis Proposal. – 2003. – 68 p.

[4] Хмелев, Р. В. Совместное использование структурного анализа и метрики Хаусдорфа при сравнении объекта и эталона / Р. В. Хмелев // Компьютерная оптика. – 2005. – В. 27. – С. 174–176.

[5] Распознавание номеров железнодорожных цистерн с использованием корреляционного алгоритма / А. К. Карлин, А. Н. Малков, Е. А. Тимофеев, Г. П. Штерн // Математика, кибернетика, информатика. Труды международной научной конференции, посвящённой памяти профессора А. Ю. Левина (Ярославль, 25 – 26 июня, 2008) / Яросл. гос. ун-т. – Ярославль: ЯрГУ, 2008. – С. 103–110.

Работы автора по теме диссертации

Публикации в изданиях, рекомендованных ВАК РФ

1. Михайлов, И. А. Об одном методе распознавания изображений / И. А. Михайлов // Моделирование и анализ информационных систем. – 2007. – Т. 14, № 4. – С. 7–12.

2. Михайлов, И. А. Некоторые методы распознавания изображений / И. А. Михайлов // Моделирование и анализ информационных систем. – 2008. – Т. 15, № 4. – С. 56–64.

3. Малков, А. Н. Восстановление изображений, искажённых перспективным преобразованием / А. Н. Малков, Г. П. Штерн, И. А. Михайлов // Моделирование и анализ информационных систем. – 2009. – Т. 16, № 2. – С. 88–102.

4. Михайлов, И. А. Распознавание изображений с помощью метода радиальных окрестностей / И. А. Михайлов // Компьютерная оптика. –

2010. – Т. 34, № 3. – С. 399–407.

5. Малков, А. Н. Разбиение железнодорожного состава на отдельные вагоны / А. Н. Малков, Г. П. Штерн, И. А. Михайлов // Моделирование и анализ информационных систем. – 2011. – Т. 18, № 2. – С. 77–112.

Другие публикации

6. Михайлов, И. А. Train Decomposition — научно-исследовательская программа для разбиения железнодорожного состава на отдельные вагоны // Свидетельство о государственной регистрации программы для ЭВМ. М.: РОСПАТЕНТ. – 2011. № 2011616542.

7. Михайлов, И. А. Об одном подходе к распознаванию изображений / И. А. Михайлов // Студенческие заметки по информатике и математике. Вып. 1. – Ярославль: Изд-во Яросл. гос. ун-та им. П. Г. Демидова, 2007. – С. 42–43.

8. Михайлов, И. А. Восстановление изображений, искажённых перспективным преобразованием / И. А. Михайлов // Шестьдесят вторая региональная научно-техническая конференция студентов, магистрантов и аспирантов высших учебных заведений с международным участием «Молодёжь. Наука. Инновации — 2009». 15 апреля 2009 г., Ярославль: Тезисы докладов. – Ярославль: Изд-во ЯГТУ, 2009. – С. 179.

9. Михайлов, И. А. Некоторые методы распознавания изображений [Электронный ресурс] // Материалы IV Международной научно-практической конференции «Современные информационные технологии и ИТ-образование»: [тез. докл.] / Моск. гос. ун-т им. М. В. Ломоносова, фак-т ВМиК. – Режим доступа: http://2009.it-edu.ru/docs/Sekzii_4-5/4_Mihaylov_ImgRec2.doc.

10. Михайлов, И. А. Некоторые методы распознавания изображений / И. А. Михайлов // Материалы XLVIII Международной научной студенческой конференции «Студент и научно-технический прогресс»: Информационные технологии / Новосиб. гос. ун-т. Новосибирск, 2010. С. 235.

11. Михайлов, И. А. Восстановление изображений, искажённых перспективным преобразованием / И. А. Михайлов // Новые информационные технологии в научных исследованиях и образовании: материалы XV Всероссийской научно-технической конференции студентов, молодых ученых и специалистов. Рязанский государственный радиотехнический университет, 2010. С. 347–348.

12. Михайлов, И. А. Распознавание изображений с помощью метода радиальных окрестностей / И. А. Михайлов // Новые информационные технологии в научных исследованиях и образовании: материалы XV Всероссийской научно-технической конференции студентов, молодых ученых

и специалистов. Рязанский государственный радиотехнический университет, 2010. С. 305–306.

13. Михайлов, И. А. Распознавание изображений с помощью метода, основанного на модификации метрики Хаусдорфа / И. А. Михайлов // Научно-практические исследования и проблемы современной молодежи: Труды II Международной молодёжной научно-практической конференции. Т. 1. Елабуга, 23 – 24 декабря 2010 г. Казань: Изд-во Казан. гос. техн. ун-та, 2010. С. 131–135.

14. Михайлов, И. А. Распознавание изображений с помощью метода радиальных окрестностей / И. А. Михайлов // Научно-практические исследования и проблемы современной молодежи: Труды II Международной молодёжной научно-практической конференции. Т. 1. Елабуга, 23 – 24 декабря 2010 г. Казань: Изд-во Казан. гос. техн. ун-та, 2010. С. 136–140.

15. Михайлов, И. А. Восстановление изображений, искажённых перспективным преобразованием / И. А. Михайлов // Научно-практические исследования и проблемы современной молодежи: Труды II Международной молодёжной научно-практической конференции. Т. 1. Елабуга, 23 – 24 декабря 2010 г. Казань: Изд-во Казан. гос. техн. ун-та, 2010. С. 141–143.

16. Михайлов, И. А. Разбиение железнодорожного состава на отдельные вагоны (цистерны) / И. А. Михайлов // Шестьдесят четвёртая региональная научно-техническая конференция студентов, магистрантов и аспирантов с международным участием. 20 апреля 2011 г., Ярославль. Ч. 2: Тез. докл. – Ярославль: Изд-во ЯГТУ, 2011. – С. 309.