

На правах рукописи

УДК 519.68; 681.513.7;
612.8.001.57; 007.51/.52

ЛОБИВ
Игорь Васильевич

**ПРОГРАММНЫЕ СИСТЕМЫ
ДЛЯ
ИДЕНТИФИКАЦИИ И ЛОКАЛИЗАЦИИ ОБЪЕКТОВ
В ИЗОБРАЖЕНИЯХ**

05.13.11 – математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико – математических наук

Красноярск 2004

Работа выполнена в Институте систем информатики СО РАН

Научный руководитель: Мурзин Федор Александрович,
кандидат физико – математических
наук

Официальные оппоненты: Кашкин Валентин Борисович
доктор технических наук, профессор

Пестунов Игорь Алексеевич
кандидат физико – математических
наук

Ведущая организация: Институт вычислительной
математики и математической
геофизики СО РАН

Защита состоится _____ в _____ ч. _____ мин. на заседании
диссертационного совета Д 212.098.03 в Красноярском
государственном техническом университете по адресу:

666074, г.Красноярск, ул. Киренского, 26, ауд 4-17,
тел. (8-3912) – 49-73-81, факс (8-3912) – 49-79-90

С диссертацией можно ознакомиться в научной библиотеке
Красноярского государственного технического университета

Автореферат разослан _____ 2004 г.

Ученый секретарь
Диссертационного совета,
кандидат технических наук, профессор

Вейсов Е.А.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы

Важнейшей задачей в ряде областей – промышленности, научных исследованиях, медицине, системах безопасности – является анализ изображений.

Решение сложных и разнообразных проблем, стоящих перед исследователями, зачастую требует достоверных опытных данных, для получения которых необходимы быстро работающие программно-аппаратные системы для автоматической локализации и идентификации объектов в изображениях.

Одна из важных задач – это быстрый поиск фрагментов в изображениях. Яркими примерами являются задачи подсчета количества бактерий на фотоснимке, поиска заданного участка местности на аэрофотоснимке, сравнение выпускаемой продукции с эталоном и т.д.

На данный момент в мире существует всего несколько программных продуктов, осуществляющих быстрый поиск образца на изображении для случая, когда образец может быть повернут внутри данного изображения. Еще меньше программ, позволяющих выполнить поиск образца, искаженного аффинным преобразованием.

Один из лидеров по времени поиска среди таких систем является продукт американской фирмы Datacube – MaxVision Toolkit [<http://www.datacube.com>], который обладает рядом неоспоримых достоинств: возможность поиска образца, искаженного аффинным преобразованием, нелинейная коррекция освещенности. Но основным недостатком метода нормализованной корреляции оттенков серого, на котором основан продукт, является большая вычислительная сложность. Это стимулирует поиск новых идей и алгоритмов для решения данной задачи.

Нельзя не отметить вклад фирмы ABBY Software House в создание систем автоматического распознавания текстов. Продукт FineReader является одним из лучших в мире для этой задачи. Одним из недостатком данного продукта является чувствительность к большим (больше 15 градусов) углам поворота сканируемого изображения. Таким же недостатком обладают различные системы локализации и распознавания

этикеток и штрих кодов. На сегодняшний день требуются быстро работающие алгоритмы для решения данных задачи.

Цель работы

Цель работы – создание алгоритмов и программных средств, расширяющих интеллектуальные возможности компьютеров и компьютерных систем машинного зрения, повышающих эффективность процессов обработки данных в вычислительных машинах и комплексах. Под данными здесь понимается фотографические изображения и видеоинформация. Проведение комплексных теоретических исследований по созданию новых алгоритмов в данной области.

Основные задачи состоят в создании алгоритмов и программных систем, позволяющих осуществлять следующие функции:

- 1) быстрый поиск (локализация) фрагмента фотографического изображения в другом изображении;
- 2) распознавание и локализация движущегося объекта в видеопотоке в режиме реального времени;
- 3) локализация на фотографическом изображении шаблона и распознавание содержащегося в нём текста;
- 4) восстановление положения движущегося тела по информации, получаемой с множества одномерных видеокамер.

Научная новизна

В результате проведенных исследований реализован ряд конкурентно-способных программных систем для локализации и идентификации объектов в изображениях.

Предложен высокоэффективный алгоритм для решения задачи о локализации фрагмента в изображении, базирующийся на использовании множества опорных точек.

Программная система созданная на основе этого метода успешно локализует фрагменты, искаженные преобразованием масштаба и поворота. По времени локализации она не уступает лидерам в данной области, а в некоторых случаях и превосходит их.

В задаче о локализации объектов в видео потоке применен метод «сверхадаптивного фильтра» и метод фильтрации одномерных шумов, с помощью которых удалось построить систему устойчивую к небольшим резким изменениям угла зрения видеокамеры. Созданная система работает в режиме реального времени и по времени локализации не уступает своим зарубежным аналогам.

В рамках решения задачи о локализации маркера и распознавания содержащегося в нем текста, реализована система, которая успешно локализует маркеры специального вида, повернутые на произвольный угол, и имеющие произвольный размер. По времени локализации маркеров, искаженных преобразованием поворота, превосходит широко известную систему распознавания текста FineReader.

В задаче о восстановлении положения тела по информации, получаемой с множества одномерных видеокамер, предложен новый подход, основанный на использовании многомерного метода Ньютона.

Практическая ценность

В результате проведенных исследований создан ряд алгоритмов и программных систем, которые переданы зарубежным заказчикам и применяются в области машинного зрения.

Апробация работы

Результаты работы докладывались на международной конференции «Современные проблемы прикладной математики и механики: теория, эксперимент и практика», посвященной 80-летию академика Н.Н. Яненко в 2001 г., на Пятой международной конференции памяти академика А.П. Ершова «Перспективы систем информатики» в 2003 г., на Международной конференции «Вычислительные технологии – 98», а также в Институте систем информатики СО РАН на семинаре «Конструирование и оптимизация программ» (председатель семинара д.ф.-м.н. Касьянов В.Н.) и на семинаре в Институте вычислительной математики и математической геофизики СО РАН (председатель семинара д.т.н. Пяткин В.П.).

Публикации. Автором опубликовано 20 печатных работ, из них по теме диссертации – 14 работ.

Структура и объем работы

Диссертационная работа состоит из введения, четырех глав и списка литературы. Объем диссертации – 116 стр. Список литературы содержит 60 наименований. Работа включает 45 рисунков и графиков, полученных в результате расчетов на ЭВМ.

СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обосновывается актуальность темы исследований и приводится краткое содержание работы.

В первой главе представлены алгоритмы и дано краткое описание программной системы для быстрого поиска фрагмента изображения в другом изображении. При этом фрагмент может быть повернут и/или изменен его масштаб.

Речь идет об обработке цветных изображений. Более того, в предложенном алгоритме для цветного изображения поиск происходит быстрее. Это является отличием от типичных алгоритмов, описанных в литературе, которые работают с цветными изображениями так же медленно, как и с черно-белыми. Особенностью данного алгоритма является также и то, что яркость и контрастность всего изображения и фрагмента могут быть различными.

На данный момент в мире существует всего несколько программных продуктов, осуществляющих быстрый поиск образца на изображении для случая, когда образец внутри данного изображения может быть повернут. Еще меньше программ, позволяющих выполнить поиск образца, искаженного аффинным преобразованием.

Один из таких продуктов — MaxVision Toolkit фирмы Datacube [<http://www.datacube.com>].

Существуют и другие программные и программно-аппаратные комплексы, осуществляющие поиск объектов теми или иными методами [<http://www.optimas.com>, <http://www.way2c.com>, <http://www.impuls-imaging.com>, <http://www.binary-technologies.de>].

В реализованной программной системе искомый фрагмент может быть повернут, растянут или иметь яркость, контраст и насыщенность цвета, отличные от аналогичных параметров образца.

Система состоит из двух частей — вычислительного ядра и управляющей оболочки. Вычислительное ядро можно с минимальными изменениями, связанными с особенностями конкретной аппаратной платформы, использовать при разработке других программ. Программный продукт разработан для операционной системы Microsoft Windows.

Пользовательский интерфейс обеспечивает удобное управление функциями системы и вывод результатов в понятной форме. При обработке изображений в диссертации используется система цветовых координат RGB.

Цветное изображение размером $n \times m$ задается тремя матрицами $S_R = S_R(i, j)$, $S_G = S_G(i, j)$ и $S_B = S_B(i, j)$, где $0 \leq i \leq n - 1$, $0 \leq j \leq m - 1$.

Значения элементов матриц $S_R(i, j)$, $S_G(i, j)$ и $S_B(i, j)$ изменяются в пределах от 0 до 255.

Пусть точки $p = (i, j)$ и $p' = (i', j')$ имеют цвета (r, g, b) и (r', g', b') соответственно, $S_R(i, j) = r$, $S_G(i, j) = g$, $S_B(i, j) = b$, $S_R(i', j') = r'$, $S_G(i', j') = g'$, $S_B(i', j') = b'$.

Тогда цветовое расстояние определяется, как

$$cd(p, p') = \max\{|S_R(p) - S_R(p')|, |S_G(p) - S_G(p')|, |S_B(p) - S_B(p')|\}.$$

Вводится такая константа C_V , что если выполнено условие $cd(p, p') \leq C_V$, то считается, что p, p' имеют одинаковый цвет.

Евклидово расстояние определяется, как $\rho(p, p') = \sqrt{(i - i')^2 + (j - j')^2}$.

Пусть $S = (S_R, S_G, S_B)$ — некоторое фотографическое изображение, $S^c = (S_R^c, S_G^c, S_B^c)$ — образец для локализации.

Задача состоит в том, чтобы найти такой локализованный фрагмент

$$S' = (S'_R, S'_G, S'_B) \subset S, \quad \text{что} \quad Eq = |S' - F(S^c)| \leq Th(S^c), \quad \text{где}$$

$F = I \circ M_2 \circ M_1 \circ R$ — композиция функций:

$R: S^c \rightarrow S_1^c$ — поворот на некоторый угол относительно центра S^c ,

$M_1: S_1^c \rightarrow S_2^c$ — изменения масштаба вдоль направления \perp ширине S_1^c ,

$M_2: S_2^c \rightarrow S_3^c$ — изменения масштаба вдоль направления \perp высоте S_2^c ,

$I: S_3^c \rightarrow S_4^c$ — коррекция интенсивности (яркости и/или контрастности),

$Th(S^c)$ — адаптивный порог для S^c ,

$$Eq = |S - S'| = \frac{\sum_{p \in S, p' \in S', \rho(p, p')=0} cd(p, p')}{\Omega}, \text{ где } \Omega - \text{ площадь } S \cap S'.$$

Для решения данной задачи в диссертации предложен метод, основанный на использовании опорных точек.

Он состоит из двух этапов – предварительной обработки и собственно поиска.

Предварительная обработка включает в себя несколько подэтапов: низкочастотную фильтрацию, медианную фильтрацию.

Далее следует подэтап выделения контуров.

Пусть p', p'' – две точки исследуемого изображения:

$$S(p') = (r', g', b'), S(p'') = (r'', g'', b'').$$

Используется обозначение $B_n(p) = B_n(i, j)$ – квадрат размером $n \times n$ с центром в точке $p = (i, j)$, где n – нечетное.

Пусть $cd[B_m(p)] = \max\{cd(p', p'') : p', p'' \in B_m(p)\}$ – максимальное цветовое расстояние в окрестности $B_m(p)$ исследуемой точки.

Заключение о том, проходит контур через точку или нет, фиксируется соответствующими значениями функции f (0 или 1 соответственно).

$$f(p) = \begin{cases} 0, & \text{if } cd[B_m(p)] \geq C_V, \\ 1, & \text{if } cd[B_m(p)] < C_V. \end{cases}$$

Следующий этап – выбор опорных точек, производится с помощью вспомогательной функции

$$g(p, p') = \alpha \cdot \rho(p, p') + (1 - \alpha) \cdot cd(p, p'), \quad 0 \leq \alpha \leq 1.$$

Наиболее сложным этапом алгоритма является построение древообразной поисковой структуры, ассоциированной с данным изображением.

На построение данной структуры затрачивается некоторое время, но после ее создания появляется возможность находить опорные точки за логарифмическое время, т.е. за $C \cdot \log(n \times m)$ шагов, где C – константа, характеризующая время выполнения одного шага поиска и не зависящая от размеров исходного изображения.

Деревообразная структура порождается в процессе разбиения данного изображения на прямоугольники.

Для нумерации прямоугольников используются пары (i, j) , кодирующие уровень разбиения, и порядковый номер прямоугольника.

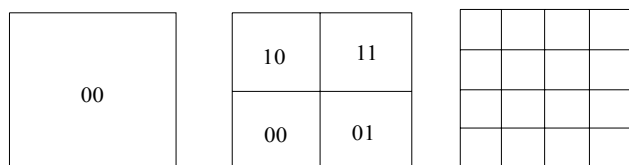


Рисунок 1 – Рекурсивное разбиение изображения.

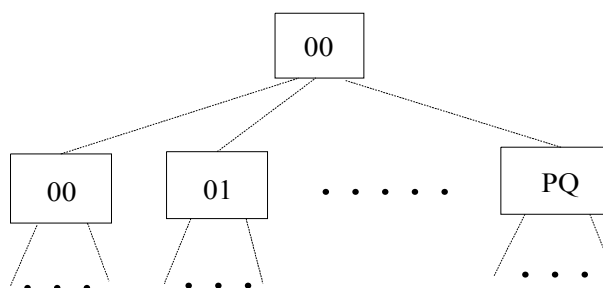


Рисунок 2 – Деревообразная структура.

Пусть $R_{r,s}(i, j)$ – соответствующий прямоугольник размера $r \times s$. Далее с каждым $R_{r,s}(i, j)$ ассоциируется определенная на рассматриваемой палитре характеристическая функция $\chi_{k,i,j}$. Функция $\chi_{k,i,j}$ показывает, существует ли точка, имеющая цвет (r, g, b) в грубой палитре, в прямоугольнике $R_{r,s}(i, j)$ на уровне k .

Метод построения грубых палитр состоит в том, что последние три бита значений r, g, b для любой точки полагаются равными нулю.

В итоге, вместо 16М цветов получается $2^{15} = 32K$ цветов.

Множество связанных с каждым узлом дерева характеристических функций $\chi = \{\chi_{kij} : 0 \leq i \leq \dots, 0 \leq j \leq \dots\}$ является той самой специальной структурой данных, упоминавшейся выше.

Поиск опорных точек в исходном изображении осуществляется с помощью построенной деревообразной структуры рекурсивно в лексикографическом порядке с возвратами на верхние уровни разбиения.

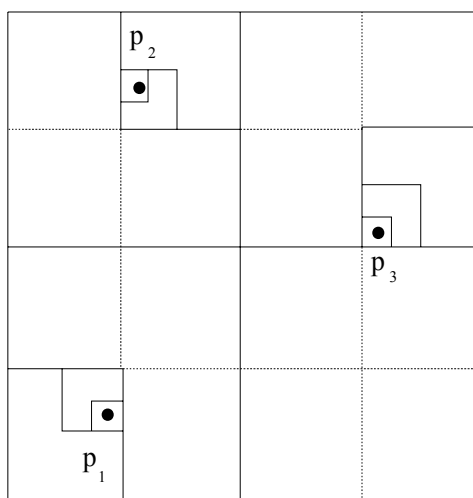


Рисунок 3 – Рекурсивный поиск опорных точек.

Дополнительное использование метрических характеристик, а именно, расстояний между опорными точками на образцах позволяет пропускать большое число узлов при обходе дерева.

После локализации опорных точек можно однозначно восстановить фрагмент, соответствующий данному положению опорных точек. Далее проводится контрольное сравнение локализованного фрагмента с образцом, а также коррекция яркости и контрастности.

Если при сравнении в указанной мере близости образца и локализованного фрагмента с помощью адаптивного порога сделан вывод, что, несмотря на близость опорных точек, фрагменты различны – продолжается поиск следующих положений опорных точек, удовлетворяющих цветовым и метрическим соотношениям. Иначе, локализация образца успешно завершена.

Разработанный алгоритм допускает различные модификации, в том числе распространение на аффинный случай. Аффинный случай не реализовывался ввиду отсутствия заказчика на подобный программный продукт. Реализация же на инициативной основе затруднена ввиду высокой трудоемкости проекта.

Во второй главе описан программный комплекс для локализации и идентификации движущихся объектов в видеопотоке в рамках задачи контроля дорожного трафика.

В последнее время в связи с ростом мощности компьютеров появилась возможность создания относительно недорогих программно-аппаратных систем, реализующих обработку большого количества информации, поступающей в реальном времени. Примером может служить система отслеживания движения объектов, на вход которой поступают изображения, полученные с видеокамеры, а на выходе имеются данные об объектах, попавших в поле зрения, а также об их динамических характеристиках, например скорости и ускорения.

Такая система успешно реализована, в ней в качестве аппаратной части выступает обычный компьютер Pentium-3, а в качестве программной части используется логический модуль собственной разработки, устройству которого и посвящена данная глава.

К особенностям системы можно отнести следующие моменты:

- 1) камера должна обзирать от кадра к кадру одну и ту же область пространства;
- 2) допускаются цветовые шумы и небольшое дрожание видеокамеры;
- 3) отслеживаемые объекты могут плавно менять форму и размер;
- 4) не допускается пересечение движущихся объектов одного цвета на продолжительное время.

Метод можно условно разбить на три части:

- 1) выделение адаптивного фильтра;
- 2) локализация объектов;
- 3) идентификация локализованных объектов (идентификация одного и того же объекта в различных кадрах);

Выделение адаптивного фильтра – фильтр R моделируется как медленно изменяемая последовательность изображений, которая позволяет адаптироваться к изменениям освещения и погодным условиям.

Локализация объектов.

Пусть F_i – текущий фрейм (кадр), а R – фильтр. Вводится цвета точки p с координатами (x, y) на этих фреймах, т.е. c_i, c_R соответственно.

Пусть Ω_i – объединение всех областей на фрейме F_i , в которых могут находиться объекты.

Вводится $\Omega'_{i-1} \subseteq \Omega_{i-1}$ – объединение всех областей, в которых находятся идентифицированные объекты на предыдущем фрейме F_{i-1} .

Если $\Omega'_{i-1} = 0$, то производится сравнение только с фильтром R , то есть $p \in \Omega_i, \Leftrightarrow c_i \neq c_R$

В противном случае, решение о том, что точка p принадлежит Ω_i принимается следующим образом:

если $p \in \Omega'_{i-1}$, то $p \in \Omega_i, \Leftrightarrow c_i \neq c_R$,

если $p \notin \Omega'_{i-1}$, то $p \in \Omega_i, \Leftrightarrow c_i \neq c_{i-1}$.

Такой комбинированный метод применения адаптивного фильтра отличается от встречающихся в литературе подходов, он позволяет учитывать изменения погодных условий уже при обработке следующего фрейма.

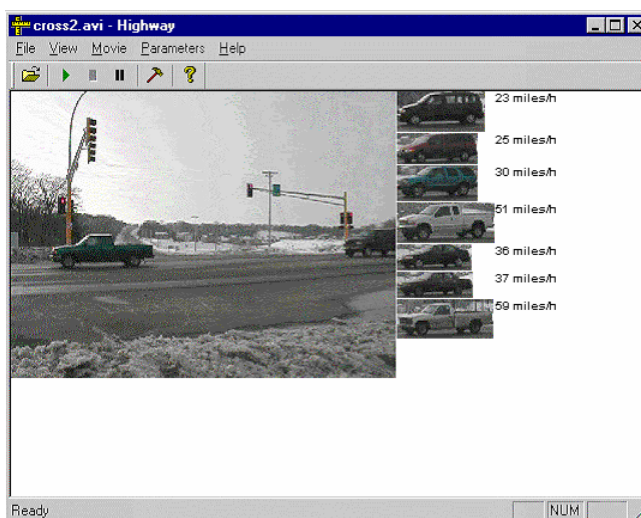


Рисунок 4 – Внешний вид программной системы

В третьей главе описана программная система быстрого поиска маркеров и распознавания текста, содержащегося в них. Такие программные продукты применяются, например, в обслуживании обычной почты. Трудность задачи в том, что допускается поворот маркера на любой произвольный угол. При использовании знаменитых систем (например, ABBY FineReader) значительно замедляются локализация и

распознавание символов при наличии поворота фрагмента. А также возрастает ошибка распознавания.

Содержательно постановка задачи заключается в следующем. На изображении имеется прямоугольная рамка (маркер), разделенная на верхнюю и нижнюю половину. В каждой половине есть набор цифр. Нужно быстро найти маркер на изображении и распознать цифры в нем. Рамка может быть повернута относительно изображения на произвольный угол и может иметь произвольный размер. В главе описан эффективный алгоритм для решения этой задачи.

Дополнительные требования:

- 1) фон должен быть примерно одного цвета;
- 2) прямоугольник и текст также должны быть одного цвета, отличающегося от фонового.

Геометрия маркера представлена как набор связанных векторов. Следовательно, геометрия может быть легко изменена. Алгоритмы универсальны, т.е. работают с любыми символами.

Обработка изображения делится на две стадии: МРМ (Модуль Распознавания Маркера) и (МИМ) Модуль Интерпретации Маркера.

В четвертой главе представлено программное средство быстрого нахождения точного положения тела по данным, поступающим с одномерных камер для 3d оптической системы анализа движения объектов. Система работает в реальном времени.

Содержательно постановка задачи заключается в следующем. В пространстве имеется твердое тело с известной геометрией, на котором в определенных местах закреплены светоизлучающие диоды, в дальнейшем называемые маркерами, которым присвоены уникальные номера. Также имеется набор одномерных детекторов, расположенных вокруг твердого тела, которые достаточно часто (200 – 300 раз в секунду) фиксируют положение маркеров. В силу одномерности детектора, о каждом маркере, находящимся в его поле видимости, можно получить лишь уравнение плоскости, в которой этот маркер содержится.

Задача состоит в том, чтобы по имеющимся данным в реальном времени восстанавливать положение твердого тела с максимальной точностью. Осложняющие факторы:

- 1) в нескольких последовательных фреймах может быть недостаточно информации для однозначного восстановления положения тела (недоопределенность);
- 2) в силу несовершенства детектора, возможна ситуация, когда он выдает ошибочный номер маркера.

Математическая постановка.

Положение тела в пространстве можно представить в виде $Ax + b$, где $b = (u, v, w)$ – вектор сдвига, A – ортогональная матрица поворота тела относительно начального положения.

Задача будет решена, если будут найдены величины A, b , при которых отклонение положения маркеров от соответствующих плоскостей будет минимально в некоторой норме.

Для эффективного решения этого класса задач необходимо выявить проблемы, которые возникают у исследователей в данной области (Vicon, Acty, Advanced Mechanical Technology, DKH, Gsport, Motion Analysis, Peak Performance, PhaseSpace, Phoenix Technologies, Polhemus и Qualisys). Для этого дается оценка объема информации, который должен обрабатываться в реальном времени для обычных видеокамер и для одномерных.

Рассматривается профессиональная 2d камера среднего класса, с характеристиками: 1024x1024 True Color размер изображения, 300 fps - рабочая частота. (Так как частота промышленных мониторов и телеприемников 24-30 fps, соответственно, для того, чтобы изображение на экране было непрерывным, нужна частота в 10 раз большая, т.е. 300 fps.)

Получается поток 900Mb в секунду лишь с одной видеокамеры. Но одной камеры недостаточно для того, чтобы восстановить положение тела в пространстве. В решениях фирмы «Vicon» применяются наборы из 4, 6, 12, 24 видеокамер. Если взять самую простую конфигурацию из 4х камер, получается информационный поток объемом более 3GB в секунду.

Эта цифра огромна даже для сегодняшнего уровня вычислительной техники как для передачи такого объема информации в реальном времени по компьютерным сетям, так и для обработки. Поэтому исследователи стремятся уменьшить этот показатель, например, применение методов сжатия, увеличение количества вычислительных узлов и их мощности, привлечение суперкомпьютеров.

Предлагается новое решение этого вопроса – использование одномерных видеокамер. Далее оценивается объем информационного потока для одномерных камер. 1024x1 – размер изображения, 300 fps – условие на частоту то же самое. С одной камеры 900Кб в секунду. Так как камер теперь нужно больше, то пусть будет 40 камер, итого около 40Мб в секунду, эта цифра вполне реальна как для обработки, так и для передачи. Более того, за счет применения одномерных камер можно увеличить скорость обработки и точность.

Матрица A в явном виде представляется, как:

$$A = A(p, q, r) = \begin{pmatrix} c & -r & q \\ r & c & -p \\ -q & p & c \end{pmatrix} + \frac{1}{1+c} \begin{pmatrix} p^2 & pq & pr \\ qp & q^2 & qr \\ rp & rq & r^2 \end{pmatrix},$$

$$\text{где } c^2 = 1 - (p^2 + q^2 + r^2) = \cos^2 \phi.$$

Пусть на данном временном шаге известны N плоскостей, заданных в виде $(n_i, X) = d_i$, $i = 1 \dots N$, где $n_i = (a_i, b_i, c_i)$ – вектор нормали к плоскости, (\circ, \circ) – скалярное произведение, а также N точек $X_i = (x_i, y_i, z_i)$, характеризующих координаты маркера в системе координат твердого тела (предполагается, что эта информация считывается из файла с описанием твердого тела).

Метод наименьших квадратов приводит к следующей задаче:

$$\sum_{i=1}^N [(n_i, AX_i + b) - d_i]^2 \rightarrow \min,$$

В итоге остается система трех нелинейных уравнений с тремя неизвестными p, q, r :

$$\begin{cases} f_1(p, q, r) = 0 \\ f_2(p, q, r) = 0, \text{ или в векторном виде} \\ f_3(p, q, r) = 0 \end{cases}$$

$$F(P) = 0, \text{ где } F = (f_1, f_2, f_3), P = (p, q, r).$$

Для ее решения воспользуемся методом последовательных приближений Ньютона. Пусть на n -ом шаге известно приближение $P_n = (p_n, q_n, r_n)$, тогда следующее приближение P_{n+1} находится по формуле: $P_{n+1} = P_n - J^{-1}F(P_n)$, где J – матрица Якоби преобразования F .

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial p} & \frac{\partial f_1}{\partial q} & \frac{\partial f_1}{\partial r} \\ \frac{\partial f_2}{\partial p} & \frac{\partial f_2}{\partial q} & \frac{\partial f_2}{\partial r} \\ \frac{\partial f_3}{\partial p} & \frac{\partial f_3}{\partial q} & \frac{\partial f_3}{\partial r} \end{pmatrix}.$$

Как известно, метод Ньютона очень быстро сходится в случае хорошего начального приближения.

В алгоритме для начального приближения используются данные, полученные в результате экстраполяции по предыдущим временным шагам (с предыдущих фреймов), – это обеспечивает хорошее начальное приближение и позволяет обойтись в методе Ньютона двумя-тремя итерациями для достижения нужной нам точности.

О ЛИЧНОМ ВКЛАДЕ АВТОРА

Реализация описанных выше алгоритмов является достаточно трудоемкой задачей, и выполнялась коллективно. Наибольший вклад в первой задаче автор диссертации внес в разработку и реализацию следующих алгоритмов: быстрый поворот фрагмента изображения, метод адаптивных порогов, оптимизация древовидной поисковой структуры, контрольное сравнение, метод обратной связи между алгоритмом поиска и контрольного сравнения. Во второй задаче диссертанту полностью принадлежит реализация всех методов. Совместно с Мехонцевым Д.Ю. осуществлялась отработка алгоритмических деталей. При разработке третьей задачи диссертант был ответственным за реализацию алгоритма локализации маркера, программу распознавания текста реализовывали другие соавторы. В четвертой задаче автор диссертации участвовал в

разработке алгоритмов, проводил проектирование и вывод наиболее сложных формул в системе символьных преобразований Maple vR4 и частично реализовывал программу на C++ (фильтрация и интерполяция данных).

ОСНОВНЫЕ РЕЗУЛЬТАТЫ

Проведены комплексные теоретические исследования и созданы алгоритмы и программные средства для локализации объектов в системах распознавания изображений для следующих задач.

1. Быстрый поиск (локализация) фрагмента фотографического изображения в другом изображении. Программное средство для локализации объектов по своим характеристикам не уступает лидеру в данной области, продукту MaxVision Toolkit фирмы Datacube и в некоторых случаях и превосходят его.
2. Распознавание и локализация движущегося объекта в видеопотоке. Основным результатом является метод «сверх» адаптивного фильтра и метод фильтрации одномерного шума, которые позволили решить две важные задачи: небольшие быстрые изменения угла зрения («дрожание») видеокамеры и устойчивость всей системы при внесении различных помех, например, MPEG-сжатие видеопотока с потерей качества.
3. Локализация на фотографическом изображении шаблона и распознавание содержащегося в нём текста. Время локализации маркера быстрее, чем у продукта лидера в области распознавания текста Fine Reader ABBY Software House. Речь идет о задаче, осложненной тем, что допускается поворот на произвольный угол и изменение масштаба маркера.
4. Восстановление положения движущегося тела по информации, получаемой с множества одномерных видеокамер. Предложена и обоснована новая идея использования одномерных камер для решения этой задачи. В этой постановке предложен и реализован эффективный алгоритм, основанный на многомерном методе Ньютона.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Лобив, И. В. Алгоритмы быстрого поиска фрагментов фотографических изображений / А. А. Дунаев, И. В. Лобив, Д. Ю. Мехонцев, Ф. А. Мурзин, О. Н. Половинко, Д. Ф. Семич, А. В. Чепель, К. А. Ярков // В кн.: Современные проблемы конструирования программ. – Новосибирск, 2002. – С. 88–109.
2. Лобив, И. В. Слежение и определение скорости движущихся на плоскости объектов в реальном времени / Д. Ю. Мехонцев, И. В. Лобив, К. С. Селезнев // В кн.: Современные проблемы конструирования программ. – Новосибирск, 2002. – С. 243–246.
3. Лобив, И. В. Решение задачи нахождения оптимального положения тела в пространстве по данным, поступающим с одномерных камер для трехмерной оптической системы анализа движения объектов / Д. Ю. Мехонцев, И. В. Лобив // Материалы междунар. конф. молодых ученых по математическому моделированию и информатике. – Новосибирск. – 2002. – С. 19–20.
4. Лобив, И. В., Слежение и определение скорости движущихся на плоскости объектов в реальном времени / И. В. Лобив, Д. Ю. Мехонцев, К. С. Селезнев // Материалы междунар. конф. Современные проблемы прикладной математики и механики: теория, эксперимент и практика. – Новосибирск. – 2001. – С. 112–114.
5. Черемушкин, Е. С. Алгоритмы визуализации генетической информации / А. А. Дунаев, А. Э. Кель, И. В. Лобив, Ф. А. Мурзин, О. Н. Половинко, Е. С. Черемушкин // Материалы междунар. конф. памяти академика А.П. Ершова «Перспективы систем информатики». – Новосибирск, 2003. – С. 43–47.
6. Семич, Д. Ф. Программное обеспечение для поддержки процесса ядерного каротажа нефтяных скважин / А. А. Винокуров, И. В. Ильин, И. В. Лобив, Ф. А. Мурзин, О. Н. Половинко, Д. Ф. Семич. // Материалы междунар. конф. памяти академика А. П. Ершова, «Перспективы систем информатики». – Новосибирск, 2003. – С. 40–43.
7. Мурзин, Ф. А. О распараллеливании РС - метода / И. В. Лобив, Ф. А. Мурзин // В кн.: Проблемы систем информатики и программирования. – Новосибирск, 1999. – С. 146–155.
8. Вшивков, В. А. Параллельный алгоритм решения задачи о взаимодействии потоков разряженной плазмы / В. А. Вшивков, И. В. Лобив, Ф. А. Мурзин // В кн.: Проблемы систем информатики. – Новосибирск, 2001. – С. 76–82.

9. Лобив, И.В. О распараллеливании метода частиц в ячейках и метода дискретных вихрей / И. В. Лобив, Ф. А. Мурзин // Материалы междунар. конф. Distributed Data Processing. – Новосибирск, 1998, – С. 43–44.
10. Мурзин, Ф. А. О распараллеливании метода дискретных вихрей / Ф. А. Мурзин, И. В. Лобив // Тр. XVI Междунар. школы - семинара по численным методам механики вязкой жидкости. – Новосибирск, 1998. – С. 37–38.
11. Мурзин, Ф. А. Параллельный алгоритм на основе PIC-метода для расчета задач о взаимодействии потоков плазмы / В. А. Вшивков, И. В. Лобив, Ф. А. Мурзин // Тр. XVI Междунар. школы - семинара по численным методам механики вязкой жидкости. – Новосибирск, 1998. – С. 38–39.
12. Лобив, И. В. Исследование вопросов, связанных с распараллеливанием PIC-метода / Ф. А. Мурзин, И. В. Лобив // Тр. XVI Междунар. школы - семинара по численным методам механики вязкой жидкости. – Новосибирск, 1998. – С. 44–46.
13. Cheremushkin, E. S. Genetic code and Visualization // Bioinformatics of Genom Regulation and Structure / E. S. Cheremushkin, A. E. Kel, I. V. Lobiv, F. A. Murzin, O. N. Polovinko – Novosibirsk, – 2002, – Vol. 1, – P. 94–97.
14. Черемушкин, Е. С. Визуализация генетической информации / А. А. Дунаев, А. Э. Кель, И. В. Лобив, Ф. А. Мурзин, О. Н. Половинко, Е. С. Черемушкин // Материалы междунар. конф. Новые информационные технологии в науке и образовании. – Новосибирск, 2003. – С. 147–156.
15. Мурзин, Ф. А. О некоторых задачах, связанных с автоматизацией процесса ядерного каротажа нефтяных скважин / А. А. Винокуров, И. В. Ильин, И. В. Лобив, Ф. А. Мурзин, О. Н. Половинко, Д. Ф. Семич // Материалы междунар. конф. Новые информационные технологии в науке и образовании. – Новосибирск, 2003. – С. 112–123.
16. Лобив, И.В. Алгоритмы быстрого поиска повернутых и масштабированных образов внутри данного изображения / А. А. Дунаев, И. В. Лобив, Д. Ю. Мехонцев, Ф. А. Мурзин, О. Н. Половинко, Д. Ф. Семич, К. А. Ярков // Материалы междунар. конф. «Перспективы систем информатики». – Новосибирск, 2003. – С. 50–53.
17. Лобив И.В. Слежение и определение скорости движущихся на плоскости объектов в реальном времени / И. В. Лобив,

Д. Ю. Мехонцев, К. С. Селезнев // Доклады междунар. конф. молодых ученых по математическому моделированию и информатике. – Новосибирск 2001. – С. 12–13.

18. Лобив, И. В. Восстановление положения тела в пространстве для системы реального времени анализа движения объектов / И. В. Лобив, Д. Ю. Мехонцев, Ф. А. Мурзин // Тр. школы – конкурса молодых ученых ИСИ СО РАН «Теоретические и прикладные задачи информатики: новые подходы и решения». – Новосибирск, 2003. – С. 54–62.
19. Семич, Д. Ф. Программный комплекс, предназначенный для обработки результатов, полученных методом ядерного каротажа нефтяных скважин / А. А. Винокуров, И. В. Ильин, И. В. Лобив, Ф. А. Мурзин, О. Н. Половинко, Д. Ф. Семич // Тр. школы – конкурса молодых ученых ИСИ СО РАН «Теоретические и прикладные задачи информатики: новые подходы и решения». – Новосибирск, 2003. – С. 23–31.
20. Половинко, О.Н. Распознавание текстур по пространственным закономерностям / Ф. А. Мурзин, О. Н. Половинко, И. В. Лобив // Материалы междунар. конф. Новые информационные технологии в науке и образовании. – Новосибирск, 2003. – С. 256–268.

Лобив И.В.

ПРОГРАММНЫЕ СИСТЕМЫ
ДЛЯ
ИДЕНТИФИКАЦИИ И ЛОКАЛИЗАЦИИ ОБЪЕКТОВ
В ИЗОБРАЖЕНИЯХ

Автореферат

Подписано в печать

Объем 1,1 уч.-изд. л.

Формат бумаги 60 × 90 1/16

Тираж 100 экз.

Отпечатано в ЗАО РИЦ «Прайс-курьер»

630090, г. Новосибирск, пр. Ак. Лаврентьева, 6, тел. 34-22-02

Заказ №135