

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

Федеральное государственное бюджетное учреждение науки
Институт проблем управления им. В. А. Трапезникова
Российской академии наук



На правах рукописи

Домулян Андрей Александрович

**Аппаратные и программные средства для анализа жестов
руки, наблюдаемых с помощью видеокамеры**

Автореферат диссертации на соискание учёной степени кандидата
технических наук по специальностям:

05.13.11 - Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

05.13.05 – Элементы и устройства вычислительной техники и
систем управления

Москва – 2012

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте проблем управления им. В. А. Трапезникова Российской академии наук.

Научный руководитель:

к.т.н., с.н.с. лаб. №27 ИПУ РАН

А.М.Михайлов

Официальные оппоненты:

д.т.н., в.н.с. лаб. №56 ИПУ РАН

В.П. Морозов

д.т.н., учёный секретарь ОАО НИИ "Субмикрон"

А.В. Лобанов

Ведущая организация: Федеральное государственное бюджетное учреждение науки Вычислительный центр им. А. А. Дородницына Российской академии наук.

Защита состоится 17 мая 2012 г. в 11 ч. в большом конференц-зале на заседании диссертационного совета Д.002.226.03 Федерального государственного бюджетного учреждения науки Института проблем управления им. В. А. Трапезникова Российской академии наук по адресу: 117997, Москва, ул. Профсоюзная, 65.

С диссертацией можно ознакомиться в библиотеке Федерального государственного бюджетного учреждения науки Института проблем управления им. В. А. Трапезникова Российской академии наук.

Автореферат разослан 4 апреля 2012 г.

Ученый секретарь диссертационного совета Д.002.226.03

кандидат технических наук



А.А. Кулинич

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность диссертационной работы определяется мировой тенденцией развития вычислительных средств в направлении, связанном с их миниатюризацией и с разработкой естественных человеко-машинных интерфейсов.

Существует широкий круг технических и бытовых приложений, автоматизация которых сдерживается отсутствием удобной и дешевой вычислительной платформы. Этот круг приложений может быть охарактеризован условным термином не-трогай-меня (НТМ). Так, например, замена кнопочных выключателей комнатного света на умные выключатели, управляемые с помощью жестов, требует использования такой, относительно дорогой, вычислительной платформы как персональный компьютер. Очевидно, что подобный подход не может быть использован для разработки коммерчески приемлемых умных выключателей минимальной стоимости. Другим примером могут служить бесконтактные способы управления самыми различными устройствами – от аудиосистем автомобилей до детских игрушек. В этих случаях умные управляющие устройства должны понимать наборы самых разнообразных команд.

Создание вычислительной платформы, которая на порядок дешевле, чем персональный компьютер и мобильный телефон открывает дорогу в мир простых, удобных и умных устройств, которые обязательно найдут широкое применение в повседневной жизни.

Целью диссертационной работы является исследование математических и алгоритмических аспектов задач распознавания жестов, разработка математического и алгоритмического обеспечения, формулировка требований к аппаратным характеристикам НТМ-платформы, зависящим от вычислительной сложности предлагаемых алгоритмов анализа видеоинформации и от существующей элементной базы. Заключительная часть работы посвящена натурному тестированию разработанных алгоритмов в реальном масштабе времени.

Для достижения поставленных целей решаются следующие задачи:

1) Исследование существующих методов распознавания жестов для выявления их достоинств и недостатков, а также определение вычислительной сложности таких методов.

2) Исследование и разработка алгоритмов почти минимальной вычислительной сложности (МВС-алгоритмов), достаточной для распознавания произвольных жестов, представленных последовательностями изображений, поступающими в реальном масштабе времени (видеоклипами).

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

3) Разработка требований к вычислительной платформе минимальной сложности, достаточной для реализации разработанных МВС-алгоритмов.

4) Апробация практической реализуемости разработанной спецификации вычислительной платформы путем передачи ее в компанию, специализирующуюся на разработке вычислительных устройств.

5) Апробация качества разработанных МВС-алгоритмов путем натурального моделирования их работы при распознавании произвольных жестов с использованием персонального компьютера.

Методы исследования

Для решения поставленных задач использовались: теория множеств, теория распознавания образов, методы компьютерного зрения, язык программирования Visual Basic 6, язык программирования JAVA и его видео библиотека (Java Media Framework API) для работы с устройствами ввода изображениями в реальном масштабе времени, видеочамера Logitech Webcam Pro 9000.

Научная новизна

1) Исследованы особенности задачи распознавания жестов. На основе проведенных исследований показана целесообразность использования метода разностных изображений для выделения признаков объектов при обработке потоков видеокладов, позволяющего существенно снизить общий объем вычислений.

2) Разработаны алгоритмы почти минимальной вычислительной сложности для распознавания произвольных жестов в реальном масштабе времени.

Практическая значимость работы

1) Программно реализованы алгоритмы почти минимальной сложности для распознавания произвольных жестов в реальном масштабе времени с помощью персональных компьютеров.

2) Разработаны требования к автономной вычислительной платформе, а также спецификация такой платформы, ориентированной на распознавания произвольных жестов, и характеризуемой минимальной сложностью и, соответственной, минимальной стоимостью, абсолютная величина которой зависит от существующей элементной базы.

Реализация результатов работы

Результаты теоретических и экспериментальных исследований, выполненных в диссертационной работе, были использованы компанией FirstCortex, Inc, Maryland, при создании системы MicroVision. Имеется акт о внедрении.

Апробация работы

Положения диссертационной работы докладывались и обсуждались на 53^{ой} и 54^{ой} конференциях МФТИ, научно-технической конференции “Техническое зрение в системах управления мобильными объектами-2010”, международной конференции MEDIAS 2010, международной конференции ANNIE 2007, на международном конгрессе WASET 2009: World Congress on Science, Engineering and Technology, Singapore 2009, на международной конференции ICMMM 2011, Zhengzhou, China.

Разработанные требования к автономной вычислительной платформе, ориентированной на распознавания произвольных жестов, были положены в основу технического задания, переданного компании FirstCortex, которая провела анализ практической реализуемости проекта. Полученное заключение прилагается.

Публикации. Результаты проведенных автором научных исследований опубликованы в 6 работах, среди которых 1 статья в ведущих рецензируемых журналах из списка ВАК России.

Структура и объем работы. Диссертационная работа состоит из введения, 6 глав, заключения и приложения и содержит 124 страницы машинописного текста, включая 28 рисунков, а также список литературы (90 наименований).

На защиту выносятся следующие основные результаты и положения:

1) Метод разностных изображений для выделения признаков объектов при обработке потоков видеок кадров, позволяющий существенно снизить общий объем вычислений.

2) Алгоритмы почти минимальной вычислительной сложности для распознавания произвольных жестов руки в реальном масштабе времени и их программная реализация.

3) Результаты экспериментального моделирования разработанных алгоритмов.

4) Требования к автономной вычислительной платформе, спроектированной по результатам моделирования, а также спецификация такой платформы, ориентированной на распознавание произвольных жестов.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность работы, сформулирована основная цель, научная новизна и практическая значимость результатов, приведено краткое описание структуры диссертации.

В первой главе дается обзор существующих методов распознавания жестов. Проводится сравнительный анализ различных подходов, их

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

- работает с внешним источником питания в 3.6 вольт.
- интерфейс для подключения цифровой камеры (8^{ми} и 14^{ти} разрядный параллельный интерфейс, поддерживает передачу до 27 Мбайт в сек при 27 МГц или 48 Мбайт в сек. при 48 МГц, поддерживает 8-ми разрядный монохромный видеосигнал или сигнал в формате Байера).

- цифро-аналоговые и аналого-цифровые преобразователи.

Ориентировочная стоимость – \$5 при размере партии более 100000 штук.

Микропроцессор STM32F207 по своим характеристикам отвечает требованиям, сформулированным в разделе «Требования к вычислительной платформе».

Микрокамера (камера на одной микросхеме) vs6724 имеет следующие характеристики:

- размер изображения 1600 x 1200 пикселей.
- 30 кадров в сек.
- размер устройства 3 x 3 мм с линзой.
- 8-ми разрядный параллельный видеоинтерфейс .
- диапазон напряжения питания 3-5 вольт.
- стоимость \$2 при заказе больших партий.

Микрокамера vs6724 по своим характеристикам отвечает требованиям, сформулированным в разделе «Требования к вычислительной платформе». Кроме того, суммарная стоимость выбранных компонентов равна \$7, что меньше проектной стоимости всего устройства, равной \$10.

Реализация проекта МикроЗрение

Проект МикроЗрение является результатом сотрудничества с компанией

FirstCortex, Inc
5817 Edson Lane
Rockville, MD 20852
Tel: (301) 792-1354
e-mail: vfe@mailworks.org

Разработанные технические требования и алгоритмы были переданы в компанию FirstCortex с целью оценки технической реализуемости проектируемого устройства, которое будет служить вычислительной платформой для реализации разработанных программных средств распознавания жестов.

Анализ, проведенной компанией FirstCortex подтвердил техническую реализуемость проекта и возможность обеспечения заявленной стоимости устройства в массовом производстве (\$10).

Был составлен список работ, проведение которых необходимо для

(ОЗНАКОМИТЕЛЬНЫЙ ФРАГМЕНТ)

реализации проекта, то есть для разработки полной технической документации и 10 образцов устройства. Этот список включает:

Задачи

Разработка детальной технической спецификации устройства
Разработка технической документации
Сборка прототипа устройства
Тестирование и отладка
Контроль за ходом работ и консультации с клиентом

Стоимость

| | |
|---|----------|
| Инженерно-управленческие работы | \$142000 |
| Компоненты, программные средства, производство и сборка | \$8410 |
| Общая стоимость | \$150410 |

Акт об использовании результатов компанией FirstCortex, Inc вынесен в приложение к диссертации.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ РАБОТЫ

1) Исследованы особенности задачи распознавания жестов. На основе проведенных исследований показана целесообразность использования метода разностных изображений для выделения признаков объектов при обработке потоков видеок кадров, позволяющего существенно снизить общий объем вычислений.

2) Разработаны алгоритмы почти минимальной вычислительной сложности для распознавания произвольных жестов в реальном масштабе времени.

3) Программно реализованы алгоритмы почти минимальной сложности для распознавания произвольных жестов в реальном масштабе времени с помощью персональных компьютеров.

4) Разработаны требования к автономной вычислительной платформе, а также спецификация такой платформы, ориентированной на распознавания произвольных жестов, и характеризуемой минимальной сложностью и, соответственной, минимальной стоимостью, абсолютная величина которой зависит от существующей элементной базы.

5) Разработанные требования к автономной вычислительной платформе, ориентированной на распознавания произвольных жестов, переданы компании FirstCortex. Анализ требований, выполненный компанией, подтвердил практическую реализуемость проекта. Стоимость реализации оценена компанией в размере \$150000 тысяч долларов США. Полученное заключение прилагается.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. *Домунян А.А.* Кортексный сканер // Труды МФТИ том 3, № 2 (10) (2011). ISSN 2072-6759. – С. 80 – 83.
2. *Домунян А. А.* Программно-аппаратная реализация кортексного сканера. Техническое зрение в системах управления мобильными объектами-2010 // Труды научно-технической конференции-семинара Выпуск 4. – С.310 – 316.
3. *Домунян А. А.* Программно-аппаратная реализация кортексного сканера // MEDIAS2010 Труды Международной научной конференции. ISBN 978-5-88835-030-0. – С.28 – 31.
4. *М.Н. Сергейчук, А.А. Домунян.* Решение задачи эталонного поиска методом нейронного кортекса // Труды 53-й научной конференции МФТИ «Современные проблемы фундаментальных и прикладных наук». Часть IX. Инновации и высокие технологии.— М.: МФТИ, 2010. — 132 с. ISBN 978-5-7417-0391-5. – С. 103 – 104.
5. *А.А. Домунян, М.Н. Сергейчук.* Кортексный сканер // Труды 53-й научной конференции МФТИ «Современные проблемы фундаментальных и прикладных наук». Часть I. Радиотехника и кибернетика. Том 1.— М.: МФТИ, 2010. — 178 с. ISBN 978-5-7417-0322-9. – С. 45 – 49.
6. *А.А. Домунян, А. М. Михайлов, М.Н. Сергейчук.* Поисковая система Кортекс // Труды 54-й научной конференции МФТИ «Проблемы фундаментальных и прикладных естественных наук в современном информационном обществе». Управление и прикладная математика. Том 2. — М.: МФТИ, 2011. — 148 с. ISBN 978-5-7417-0404-2. – С. 121 – 122.

Личный вклад диссертанта в публикациях, выполненных в соавторстве: в [4] обосновал метод нейронного кортекса и привёл пример его применения, в [5] автор научно обосновывает применение используемого алгоритма, в [6] автор описывает поисковую систему и описывает метод определения расстояний.