

На правах рукописи

Кривцов Олег Александрович

**МЕТОДЫ, АЛГОРИТМЫ И ПРОГРАММНАЯ СИСТЕМА
ТРЕКИНГА ГОЛОВЫ ЧЕЛОВЕКА НА ВИДЕОКАДРАХ,
ОСНОВАННЫЕ НА ГЕОМЕТРИЧЕСКИХ ТЕКСТУРНЫХ
МОДЕЛЯХ**

Специальность 05.13.11 – Математическое и программное
обеспечение вычислительных машин, комплексов и
компьютерных сетей

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Томск – 2010

Работа выполнена в ГОУ ВПО «Томский государственный университет систем управления и радиоэлектроники» и ГОУ ВПО «Национальный исследовательский Томский политехнический университет»

Научный руководитель: **Кориков Анатолий Михайлович,**
заслуженный деятель науки РФ,
доктор технических наук,
профессор

Официальные оппоненты: **Спицын Владимир Григорьевич,**
доктор технических наук,
профессор

Сущенко Сергей Петрович,
доктор технических наук,
профессор

Ведущая организация: **Новосибирский государственный
технический университет**

Защита диссертации состоится 1 декабря 2010 г. в 17 час. 00 мин. на заседании совета по защите докторских и кандидатских диссертаций **Д 212.269.06** при ГОУ ВПО «Национальный исследовательский Томский политехнический университет» по адресу: 634034, г. Томск, ул. Советская, 84/3, институт «Кибернетический центр» НИ ТПУ, ауд. 214.

С диссертацией можно ознакомиться в научной библиотеке ГОУ ВПО «Национальный исследовательский Томский политехнический университет» по адресу: 634034, г. Томск, ул. Белинского, 55.

Автореферат разослан 29 октября 2010 г.

Ученый секретарь совета
по защите докторских и
кандидатских диссертаций,
к.т.н, доцент

М.А. Сонькин

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Отслеживание головы человека (англ. head tracking, трекинг головы) – это решение задачи последовательного определения положения головы на видеокадрах, независимо от того, движется голова или камера. Трекинг головы на видеокадрах применяется или обладает потенциалом применения при создании систем трехмерной компьютерной анимации, разработке интерфейсов взаимодействия человек-компьютер, проведении видеоконференций в сети Интернет, стабилизации изображения лица с целью облегчения распознавания эмоций либо персональной идентификации и др.

Значительный вклад в развитие теории и практики трекинга головы внесли зарубежные ученые Брюс Д. Лукас (Bruce D. Lucas), Такео Канаде (Takeo Kanade), Цзин Сяо (Jing Xiao), Саймон Бэкер (Simon Baker), Иэн Мэтьюз (Iain Matthews), Тсуеши Морияма (Tsuyoshi Moriyama), Марко Ла Каскиа (Marco La Cascia), Стэн Скларофф (Stan Sclaroff), Пол Виола (Paul Viola), Майкл Дж. Джонс (Michael J. Jones), Ричард Зелиски (Richard Szeliski), Грегори Д. Хэгер (Gregory D. Hager), Петер Н. Белюме (Peter N. Belhumeur). Из работ отечественных ученых по этой теме отметим труды А.П. Вежневца и Р.С. Шоргина.

Несмотря на обилие научных трудов, опубликованных по теме трекинга головы за последние тридцать лет, в настоящее время проблему трекинга головы нельзя назвать решенной в объеме, позволяющем осуществлять выбор методов, алгоритмов и программных средств трекинга головы человека на видеокадрах в интересах разнообразных применений и, в частности, для человеко-машинных интерфейсов, расширяющих интеллектуальные возможности компьютеров и компьютерных систем. Это обусловлено многими факторами, усложняющими анализ изображения головы человека на видеокадрах, такими как неравномерность освещения сцены, непредсказуемость траектории и скорости движения головы, непостоянность (нестабильность) формы и текстуры головы за счет изменения мимики лица и т.д.

В наши дни большой популярностью пользуются недорогие веб-камеры, которые со времени их изобретения в начале 1990-х годов превратились во вполне сформировавшийся класс устройств, без которых невозможно представить себе общение в сети Интернет. Поэтому актуальным представляется исследование трекинга головы на видеоизображении с веб-камеры.

Среди существующих подходов к отслеживанию положения головы человека на видеокадрах, таких как методы, основанные на детектировании лица в разных ракурсах (англ. tracking-by-detection), мето-

ды, основанные на отслеживании «характерных точек» лица (англ. feature points) и методы, основанные на построении геометрической текстурной модели головы, автору представляется наиболее выгодным использование последних (модельных методов). Это обусловлено тем, что данный подход в отличие от других обеспечивает более высокую точность и имеет сравнительно малую вычислительную сложность. Однако данный подход обладает такими недостатками, как необходимость: выбора модели головы (у каждого человека своя (оригинальная) форма головы); ее первоначального позиционирования (на практике использование веб-камеры не обеспечивает возможность физического измерения расстояния до объекта и точного определения его первоначального положения); учета изменения освещенности и обновления текстуры модели (т. е. обновление шаблона, с которым сравнивается текущий видеокادر).

Изложенное выше доказывает актуальность и научную значимость решения проблемы инициализации модели головы при использовании веб-камеры, сравнительное исследование известных геометрических недеформируемых моделей головы и разработка на их основе методов, алгоритмов и программных средств трекинга головы человека на видеокдрах для расширения интеллектуальных возможностей вычислительных машин и систем.

Целью данной работы является исследование и разработка методов и алгоритмов отслеживания положения головы человека на поступающих с монокулярной веб-камеры видеокдрах, основанных на построении недеформируемых геометрических текстурных моделей головы, и разработка на их основе программных средств трекинга головы.

Для достижения поставленной цели необходимо решить следующие задачи:

- выполнить анализ состояния проблемы отслеживания положения головы на видеокдрах в целом и методов, основанных на построении геометрической текстурной модели головы в частности, определить возможные пути ее успешного решения;
- выполнить формализацию задачи инициализации модели головы и получить ее решение для первоначального позиционирования модели головы при использовании некалиброванной монокулярной веб-камеры;
- выделить основные сущности бизнес-логики, проведя декомпозицию задачи трекинга головы; реализовать алгоритм, основанный на построении модели головы, в составе программной системы (ПС);

- провести анализ производительности и точности разработанной системы; вывести критерии для сравнения моделей головы и провести сравнительный анализ различных моделей головы на основе ПС отслеживания головы;

Объектом исследования является проблема создания человеко-машинных интерфейсов и программных средств, расширяющих интеллектуальные возможности компьютерных систем на основе решения задач отслеживания положения головы человека по ее видеоизображению.

Предметом исследования являются методы отслеживания головы человека на поступающих с веб-камеры видеокадрах, основанные на применении геометрических текстурных моделей головы; методы построения жесткой модели головы и ее первоначального пространственного позиционирования; сравнительный анализ различных моделей головы.

Методы исследования, применяемые в данной работе, базируются на аппарате дифференциального исчисления, численных методах, методах оптимизации, компьютерной графике, математическом (имитационном) моделировании. Для сравнительного анализа моделей головы человека проводились модельные эксперименты с помощью ПС, созданной средствами языка программирования C++. В процессе программной реализации полученных результатов использовались методы теории алгоритмов, теории структур данных, объектно-ориентированного программирования. Для количественной обработки данных эксперимента применялись методы теории вероятностей и математической статистики.

Информационная база данной работы включает в себя доклады, представленные на конференциях по компьютерному зрению (главным образом это зарубежные источники), квалификационные работы, опубликованные по аналогичным темам, а также экспериментальные данные и системы, опубликованные в печати либо представленные в Интернет-ресурсах.

Новизна данной работы состоит в следующем:

1. Проведен анализ проблемы отслеживания головы человека на видеокадрах, определены факторы, усложняющие трекинг головы, классифицированы методы трекинга головы и разработаны рекомендации по выбору метода трекинга для реализации в ПС. Проведенный анализ отличается от известных аналогов указанными факторами, выделенными классами методов трекинга и рекомендациями по выбору методов для программной реализации.

2. Разработан оригинальный приближенный алгоритм автоматического пространственного позиционирования модели головы человека на первом видеокadre для некалиброванной видеокамеры, основанный на данных о приближенном положении области лица, получаемых от детектора лица, что отличает разработанный алгоритм от известных, предполагающих ручное задание положения головы.
3. Впервые применен инверсно-композиционный (ИК) алгоритм регистрации изображений для задачи трекинга головы как твердого тела, обладающего шестью степенями свободы движения.
4. Создана и зарегистрирована в Роспатенте программная система, реализующая модельный метод отслеживания головы человека и отличающаяся от известных аналогов реализованным в ней способом записи информации о положении головы человека на каждом видеокadre.
5. Предложены новые критерии для сравнительного анализа известных моделей головы (цилиндрической, эллипсоидальной, плоской, Candide-3), обеспечивающие выработку рекомендаций по областям применения сравниваемых моделей головы.

Положения, выносимые на защиту:

1. На основе анализа проблемы трекинга головы человека на видеокcadрах определено ее современное состояние и выполнена классификация методов трекинга головы.
2. Приближенный алгоритм позиционирования модели головы при использовании веб-камеры позволяет инициализировать процесс отслеживания головы в автоматическом режиме, т.е. без ручной разметки области лица на первом кадре, что является важным этапом создания полностью автоматических систем трекинга головы.
3. Выбранный ИК алгоритм регистрации изображений является более эффективным в вычислительном плане по сравнению с другими известными алгоритмами регистрации изображений.
4. Созданная программная система обеспечивает проведение экспериментальных исследований методов и алгоритмов отслеживания положения головы человека на видеокcadрах, необходимых для их тестирования и сравнения, в частности, для сопоставления по точности алгоритмов трекинга головы и обоснования рекомендаций по их реализации для расширения интеллектуальных возможностей компьютерных систем.
5. Экспериментальный сравнительный анализ четырех моделей головы позволяет сформулировать рекомендации по выбору мето-

дов трекинга головы человека для математического и программного обеспечения систем трекинга головы человека.

Достоверность результатов и выводов работы обеспечивается строгостью применения используемых математических методов, непротиворечивостью результатов и выводов автора с ранее полученными данными других исследований, совпадением результатов численных расчетов автора с расчетами и экспериментальными данными других авторов, а также успешным применением полученных результатов в объекте внедрения. Для всех разработанных программных модулей созданы автоматические тесты, обеспечивающие покрытие большей части программного кода.

Практическая ценность полученных результатов:

- Основные результаты диссертации реализованы (внедрены) в программном продукте ManuCam (<http://manucam.com>), разрабатываемым в компании ТомскСофт. Созданный программный модуль используется для отслеживания головы пользователя компьютера в трехмерном пространстве в режиме реального времени и наложения специальных трехмерных эффектов на видеоизображение лица пользователя.
- Разработана ПС «Head Tracking Demo», которая использовалась автором для оценки производительности и точности метода трекинга головы. Отметим возможность использования реализованной ПС с целью сравнения с аналогичными системами.

Внедрение.

- Полученные результаты реализованы (внедрены) в продукте ManuCam компании ТомскСофт, что подтверждается соответствующим актом о внедрении.
- Результаты диссертации использовались при выполнении госбюджетной НИР «Исследование методов и алгоритмов анализа изображений для задач идентификации личности» (рег. № 01200854086, инв. № 02201052300), выполнявшейся в Томском государственном университете систем управления и радиоэлектроники (ТУСУР) в 2005-2009 гг.
- Результаты исследований использованы в учебном процессе на кафедре АСУ ТУСУРа.

Апробация работы. Результаты работы докладывались и обсуждались на международных, всероссийских и региональных научно-практических конференциях, в том числе конференции «Современная техника и технологии», г. Томск, 2008 г. (доклад отмечен дипломом), 2009 г. (доклад отмечен дипломом I степени), 2010 г. (доклад участвовал в отборе на конкурс У.М.Н.И.К.), «Научная сессия ТУСУР»,

г. Томск, 2008 г., 2009 г. (доклад отмечен дипломом II степени), 2010 г. (доклад отмечен дипломом III степени), «Электронные средства и системы управления: итоги реализации программы развития электроники и ИТ-технологий в Томской области», г. Томск, 2008 г. (доклад участвовал в конкурсе У.М.Н.И.К), «Молодежь и современные информационные технологии», г. Томск, 2009 г., «Студент и научно-технический прогресс», г. Новосибирск, 2010 г, «СИБРЕСУРС-2010», г. Абакан, 2010 г.

Публикации. По теме диссертации было опубликовано 16 работ, в том числе пять статей в журналах, входящих в перечень ВАК («RSDN Magazine», «Информационные технологии» и «Доклады Томского государственного университета систем управления и радиоэлектроники») и 11 публикаций в сборниках трудов региональных, всероссийских и международных конференций. Публикации в электронном виде доступны по адресу <http://headtracking.googlepages.com>. Имеется свидетельство о регистрации разработанной ПС «Head Tracking Demo» в реестре федерального агентства по интеллектуальной собственности Роспатент.

Личный вклад. Автором самостоятельно получен алгоритм инициализации модели головы и создана реализация ПС «Head Tracking Demo». Совместно с научным руководителем разработаны критерии для проведения сравнительного анализа моделей головы, а также интерпретировались результаты этого анализа. Совместно с сотрудниками компании ТомскСофт проводилось внедрение полученных результатов в программный продукт компании ТомскСофт.

Структура и объем работы. Диссертационная работа состоит из введения, трех глав, заключения, списка литературы, и четырех приложений. Объем работы составляет 202 листа.

СОДЕРЖАНИЕ РАБОТЫ

В первой главе систематизируются результаты предыдущих исследований по решению проблемы отслеживания головы человека на видеокдрах: представлены основные области применения, изложена история развития методов трекинга головы, выделены основные классы методов трекинга (рис. 1), перечислены известные программы трекинга головы человека.

Во второй главе приведены теоретические основы используемого метода трекинга головы. *Дается обоснование выбора метода трекинга головы человека на видеокдрах с использованием геометрических текстурных моделей головы. Кратко рассмотрена используемая в работе терминология* компьютерной графики (хотя она и общеизвестна,

но требует разъяснения из-за наличия множества толкований одних и тех же терминов).

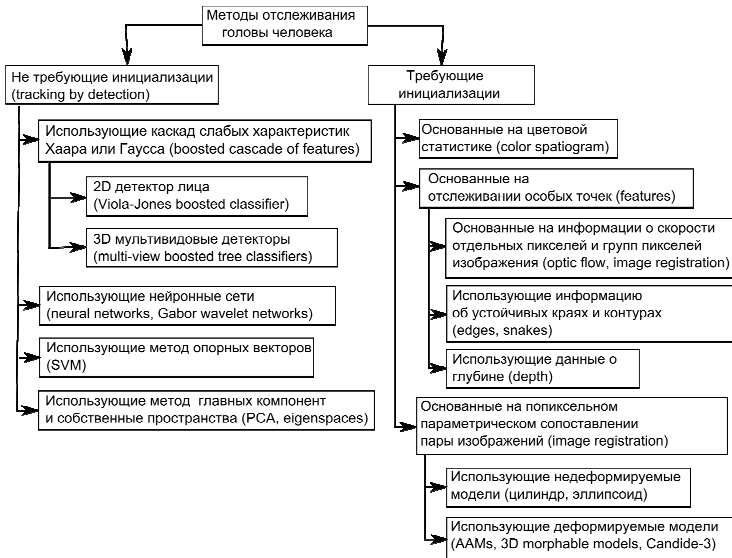


Рисунок 1 – Классификация методов трекинга головы

Голову человека как трехмерный объект в компьютерной 3D графике можно представить в виде модели, задающей геометрическую форму объекта и текстуру (рис. 2).

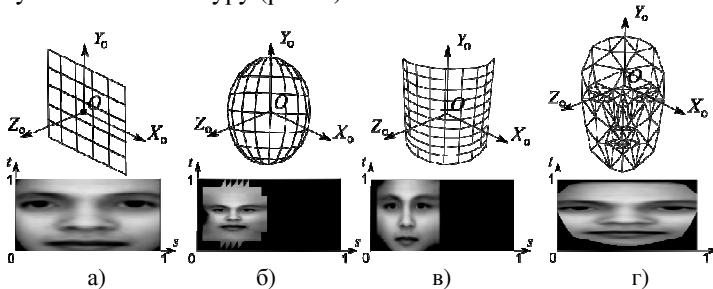


Рисунок 2 – Используемые модели головы и накладываемые на них текстуры: а) плоская модель; б) эллипсоид; в) цилиндр; г) модель Candide-3

Геометрическая поверхность модели головы аппроксимируется примитивными фигурами (треугольниками, полигонами). На рис. 3 показаны используемые системы координат.

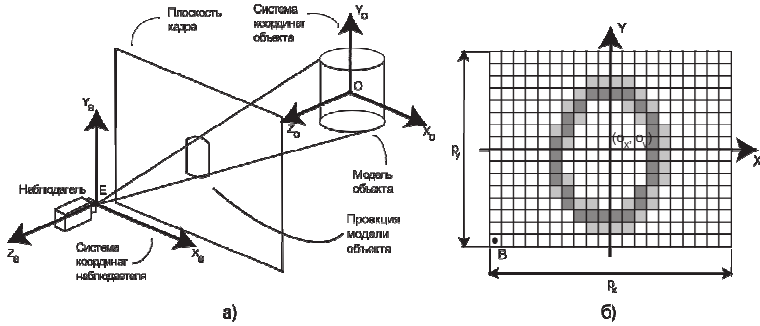


Рисунок 3 – Используемые системы координат: а) системы координат объекта (SKO) и наблюдателя (СКН); б) система координат окна

Рассмотрены операции и матрицы преобразования координат, которые позволяют моделировать движение твердого тела в трехмерном пространстве: *параллельный перенос* (задается матрицей $\mathbf{T}$), и *вращение* (задается матрицей $\mathbf{R}$). Комбинация данных преобразований называется матрицей модель-вид $\mathbf{M}$ (model-view).

Один из наиболее распространенных способов параметризации вращения – использование углов Эйлера, когда матрицу вращения $\mathbf{R}$ можно представить в виде произведения трех матриц $\mathbf{R}(\theta, \phi, \psi) = \mathbf{R}_y(\theta)\mathbf{R}_z(\phi)\mathbf{R}_x(\psi)$, где θ, ϕ, ψ – углы Эйлера, а $\mathbf{R}_x, \mathbf{R}_y$ и $\mathbf{R}_z$ – соответственно матрицы вращения вокруг осей X_e, Y_e и Z_e .

Вектор параметров, от которых зависит матрица модель-вид $\mathbf{M}$, будем называть вектором конфигурации $\mathbf{c} = (\theta, \phi, \psi, t_x, t_y, t_z)^T$, где компоненты t_x, t_y и t_z – параметры параллельного переноса. Имея вектор конфигурации, можем задать матрицу модель-вид $\mathbf{M}$ (и наоборот).

Известно, что малые перемещения и вращения удобнее задавать матрицей *twist-преобразования* $\hat{\xi}$ (англ. twist, винтовое движение):

$$\hat{\xi} = \begin{pmatrix} 1 & -\omega_z & \omega_y & t_x \\ \omega_z & 1 & -\omega_x & t_y \\ -\omega_y & \omega_x & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (1)$$

где компоненты $\omega_x, \omega_y, \omega_z$ задают параметры вращения, t_x, t_y и t_z – параметры параллельного переноса.

Точечная модель камеры определяется матрицей перспективного проецирования $\mathbf{P}$. В работе используется частный случай перспективной проекции – симметричная перспективная проекция:

$$\mathbf{P} = \begin{pmatrix} \frac{f_L}{\text{aspect}} & 0 & 0 & 0 \\ 0 & f_L & 0 & 0 \\ 0 & 0 & -\frac{f+n}{f-n} & -\frac{2fn}{f-n} \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad (2)$$

где $f_L = \text{ctg}(fovy/2)$ – фокальная длина объектива, $fovy$ – угол обзора по вертикали (англ. field of view), выраженный в градусах; $\text{aspect} = p_x / p_y$; p_x и p_y – соответственно ширина и высота кадра в пикселях, n и f – соответственно ближняя и дальняя плоскости отсечения.

Выделены и рассмотрены *четыре модели головы* (рис. 2): плоский прямоугольник, цилиндр, эллипсоид и модель Candide-3. При незначительных изменениях выражения лица можно упрощенно считать, что голова – *твердое недеформируемое тело*, и может иметь шесть степеней свободы движения в 3D пространстве (три степени свободы перемещения и три степени свободы вращения).

Рассмотрен метод отслеживания головы человека на видеокдрах. Пусть на каждом кадре видеопоследовательности отображена проекция головы пользователя, имеющей некую пространственную конфигурацию $\mathbf{c}_i$, где i – номер кадра. Необходимо: 1) инициализировать модель головы для первого кадра последовательности; 2) для каждого кадра, начиная со второго, определить параметры межкадрового движения модели головы и оценить ее новую конфигурацию.

Для решения первой из этих задач *предложен приближенный алгоритм инициализации модели головы*, основанный на детектировании лица и на том предположении, что при выбранных параметрах камеры, размере и положении модели головы, проекция модели головы должна совпадать с изображением головы на первом видеокadre. Кратко рассмотрим суть алгоритма.

Пусть пользователь расположил голову фронтально к камере (рис. 4, а). Получим с помощью детектора Виола-Джонса координаты прямоугольной области $F = \{x; y; w; h\}$, в которой предположительно находится лицо.

Выберем параметр $fovy$ (угол обзора по вертикали) для задания матрицы проецирования $\mathbf{P}$. Из характеристик анализируемых веб-камер (Trust WB-1400T, Trek 310 и Pico iMage Webcam) можно заключить, что этот параметр принимает значения в диапазоне от 27 до 40

градусов. Положим $\text{fovy} = 35^\circ$, так как это значение близко к среднему.

Координаты ближней и дальней плоскостей отсечения всегда выбирают произвольно, рекомендуется лишь, чтобы ближняя плоскость отсечения не была очень близка к нулю. Установим соответственно $n = 0,1$ м и $f = 10$ м.

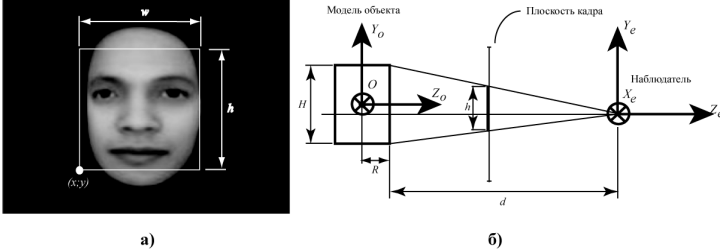


Рисунок 4 – Иллюстрация к способу инициализации модели головы

Выберем размер и положение модели головы таким образом, чтобы область проекции модели головы на видеокadre приблизительно соответствовала области F . Покажем это на примере цилиндрической модели головы. Пусть высота H цилиндра примерно соответствует вертикальному размеру области лица, заключенной между линией подбородка и линией волос. Считаем, что соотношение ширины и высоты области F примерно соответствует соотношению диаметра и высоты цилиндра: $H/2R = h/w$, откуда можно вычислить радиус цилиндра

$$R = Hw/2h \text{ (м)}.$$

Примем в качестве вертикального размера головы субъективную оценку $H = 0,15$ м.

Так как направления координатных осей Z_O и Z_e совпадают, то углы поворота цилиндра θ , φ и ψ вокруг осей координат СКН равны нулю. Определим координаты $(x_e^o, y_e^o \text{ и } z_e^o)$ центра O цилиндра (рис. 4, б) относительно СКН по следующим формулам:

$$x_e^o = -\frac{-z_e(x + w/2 - o_x) - R}{o_x p_{11}}; \quad y_e^o = -\frac{-z_e(y + h/2 - o_y) - H/2}{o_y p_{22}};$$

$$z_e^o = -\frac{Ho_y p_{22}}{h} - R,$$

где x, y, w, h – соответственно координаты нижнего левого угла, ширина и высота области F ; o_x, o_y – координаты центра окна, p_{11}, p_{22} – соответствующие элементы матрицы проецирования.

Инициализировав модель головы для первого кадра, будем искать параметры ее движения между первым и вторым кадрами, вторым и третьим кадрами, ..., i -м и $(i+1)$ -м кадрами. Эту задачу можно решить, используя инверсно-композиционный *алгоритм регистрации изображений*, предложенный С. Бэкером, Ф. Деллаертом и И. Мэтьюзом

Пусть имеем два grayscale-изображения I_1 и I_2 (рис. 5), первое из которых (либо его часть) может быть переведено во второе изображение (либо в часть второго изображения) посредством преобразования

$$\mathbf{x}' = W(\mathbf{x}, \Delta \mathbf{p}) = \begin{pmatrix} W_x(\mathbf{x}, \Delta \mathbf{p}) \\ W_y(\mathbf{x}, \Delta \mathbf{p}) \end{pmatrix}, \quad (3)$$

зависящего от набора параметров. Модель сопоставления пикселей $W(\mathbf{x}, \Delta \mathbf{p})$ – некая вектор-функция, ставящая в соответствие пикселю $\mathbf{x}=(x,y)^T$ на одном из изображений пиксель $\mathbf{x}'=(x',y')^T$ на другом изображении; $\Delta \mathbf{p} = (\Delta p_1, \Delta p_2, \dots, \Delta p_m)^T$ – m -мерный вектор приращений параметров.

Необходимо, задавшись начальным приближением, определить такие значения параметров преобразования (3), при которых устанавливается наилучшее соответствие между двумя изображениями.

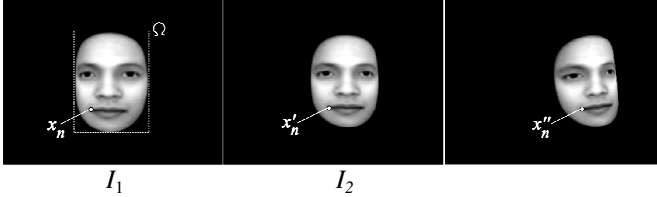


Рисунок 5 – Движение пикселя $\mathbf{x}_n$ проекции головы от видеокadra к видеокadру

Пусть известно, что некому конкретному пикселю $\mathbf{x}_n=(x_n, y_n)^T$ изображения I_1 семантически (по смыслу) соответствует пиксель $\mathbf{x}'_n$ изображения I_2 , причем интенсивности этих пикселей равны между собой:

$$I_1(W(\mathbf{x}_n, \Delta \mathbf{p})) = I_2(\mathbf{x}_n). \quad (4)$$

Линеаризуем уравнение (4), используя полный дифференциал для аппроксимации приращения функции:

$$I_1(\mathbf{x}_n) - I_2(\mathbf{x}_n) + \nabla I_1(\mathbf{x}_n) \mathbf{J}_w(\mathbf{x}_n, \Delta \mathbf{p}) = 0 \quad (5)$$

где $\nabla I_1(\mathbf{x}_n)$ – градиент изображения I_1 в точке $\mathbf{x}_n$; $\mathbf{J}_w(\mathbf{x}_n, \mathbf{0})$ – матрица Якоби функции $W(\mathbf{x}, \Delta \mathbf{p})$, вычисленная в точке $(\mathbf{x}_n, \mathbf{0})$.

В изображении I_1 можно выделить целевую область Ω , содержащую в себе N пикселей. Обычно это область, содержащая все пиксели визуального объекта, за исключением пикселей фона. Если предположить, что все N пикселей можно сопоставить со вторым изображением

ем, применяя функцию $W(\mathbf{x}, \Delta \mathbf{p})$, то можно записать систему из N линейных неоднородных уравнений:

$$I_1(\mathbf{x}_n) - I_2(\mathbf{x}_n) + \nabla I_1(\mathbf{x}_n) \mathbf{J}_w(\mathbf{x}_n, \Delta \mathbf{p}) = 0, \quad n = 1, 2, \dots, N. \quad (6)$$

Для поиска решения системы уравнений (6) по методу наименьших квадратов, составим функционал

$$\Phi(\Delta \mathbf{p}) = \sum_{n=1}^N w(\mathbf{x}_n) (I_1(\mathbf{x}_n) - I_2(\mathbf{x}_n) + \nabla I_1(\mathbf{x}_n) \mathbf{J}_w(\mathbf{x}_n, \Delta \mathbf{p}))^2 = 0, \quad (7)$$

где $w(\mathbf{x}_n)$ – скалярный вес n -го пикселя. Для минимизации функционала (7) применяется градиентный метод Ньютона.

Так как в формуле (5) использована лишь линейная часть приращения функции, то найденное на выходе цикла итераций решение $\Delta \mathbf{p}$ является приближенным. Чтобы улучшить точность поступают следующим образом. С помощью композиции моделей сопоставления пикселей $W^{(i)}(\mathbf{x}) = W^{-1}(\dots W^{-1}(W^{-1}(\mathbf{x}, \Delta \mathbf{p}^0), \Delta \mathbf{p}^1), \dots, \Delta \mathbf{p}^i)$ деформируют изображение I_2 , таким образом «сближая» его с первым изображением. Затем снова ищут вектор $\Delta \mathbf{p}$, подавая на вход цикла уже не исходное, а деформированное изображение.

Рассмотрена модель сопоставления пикселей для случая движения головы человека с шестью степенями свободы. Пусть точка поверхности модели головы с координатами $\mathbf{X}_e = (x_e, y_e, z_e, 1)^T$ проецируется в пиксель кадра с оконными координатами $\mathbf{X}_w = (x_w, y_w, z_w)^T$. Пусть за счет движения головы точка $\mathbf{X}_e$ перемещается в положение $\mathbf{X}_e' = \hat{\xi}(\Delta \mathbf{p}) \mathbf{X}_e$, где $\hat{\xi}(\Delta \mathbf{p})$ – матрица (1); $\Delta \mathbf{p} = (\omega_x, \omega_y, \omega_z, t_x, t_y, t_z)^T$ – приращение вектора параметров движения. Координаты точки окна, в которую проецируется точка модели $\mathbf{X}_e'$, обозначим через $\mathbf{X}_w'$. Используя формулы преобразований координат в процессе растеризации, можем установить соответствие между $\mathbf{X}_w$ и $\mathbf{X}_w'$:

$$\mathbf{X}_w' = WndTfm \left(PerspDiv \left(\mathbf{P} \hat{\xi}(\Delta \mathbf{p}) \mathbf{P}^{-1} \begin{pmatrix} \frac{2x_w}{p_x} - 1; \frac{2y_w}{p_y} - 1; 2z_w - 1; 1 \end{pmatrix}^T \right) \right), \quad (8)$$

где $WndTfm$ и $PerspDiv$ – операции оконного преобразования и перспективного деления соответственно. Матрица Якоби $\mathbf{J}_w(\mathbf{X}_w, \mathbf{0})$ для вектор-функции (3) примет вид

$$J_W(X_w, 0) = \begin{pmatrix} \tau_1 x_e y_e & \tau_1 (-x_e^2 - z_e^2) & \tau_1 y_e z_e & \tau_1 (-w_e z_e) & 0 & \tau_1 w_e x_e \\ \tau_2 (y_e^2 + z_e^2) & \tau_2 (-x_e y_e) & \tau_2 (-x_e z_e) & 0 & \tau_2 (-w_e z_e) & \tau_2 w_e y_e \end{pmatrix},$$

где $\tau_1 = \frac{p_x}{2} \frac{p_{11}}{z_e^2}$, $\tau_2 = \frac{p_y}{2} \frac{p_{22}}{z_e^2}$.

Для хранения в памяти композиции моделей (8) достаточно запомнить матрицу

$$W' = \xi^{-1}(\Delta p^1) \dots \xi^{-1}(\Delta p^1) \xi^{-1}(\Delta p^0). \quad (9)$$

Матрица (9) также является результатом работы алгоритма регистрации изображений.

Предложены две схемы трекинга головы человека: с динамически обновляемым шаблоном (на вход алгоритма регистрации подается предыдущий кадр) и со статическим шаблоном (подается проекция головы в том положении, в котором она была на предыдущем кадре). В работе используется вторая стратегия, так как она не страдает от накопления ошибки трекинга и «дрейфа» модели головы.

Вышеописанные методы и алгоритмы были реализованы в ПС для отслеживания головы человека на видеокдрах под названием «Head Tracking Demo». Приводится описание архитектуры ПС в соответствии со стандартом IEEE 1471. Рассмотрена миссия программной системы, заинтересованные лица и накладываемые их интересами ограничения и нефункциональные требования. Произведена идентификация точек зрения на архитектуру системы. Использован архитектурный фреймворк «4+1» Ф. Кратчена, включающий в себя пять видов: логический, процесса, разработки, физический и сценариев. Описаны основные сущности, выделенные при декомпозиции задачи трекинга головы, и получены составляющие основные блоки программной системы «Head Tracking Demo». Построены диаграмма вариантов использования, диаграммы классов и компонентов в нотации UML.

Предложен формат файла записи результатов трекинга головы человека на видеокдрах, позволяющий устранить недостатки известных форматов.

В третьей главе проведен модельный эксперимент на наборе из 45 реальных видеопоследовательностей, для которых известны эталонные оценки углов поворота головы человека (рис. 6).



Рисунок 6 – Кадры из видеоклипов, используемых в эксперименте

По результатам эксперимента *можно сделать выводы о производительности* ПС «Head Tracking Demo»: производительность системы составила примерно 20 кадров/с на компьютере с ЦП Intel CPU 1,6 GHz, 2 Гб памяти и видеокартой nVidia GeForce 9800 GTX+. Следовательно, ПС способна работать в режиме реального времени.

Предложены три количественных критерия оценки моделей головы: сумма квадратов разностей, сумма модулей разностей и коэффициент корреляции. На основе критериев проведен сравнительный анализ (таблица 1) четырех недеформируемых геометрических моделей головы: 3D цилиндра, 3D эллипсоида, плоской прямоугольной модели и модели Candide-3. Данный анализ отличается от аналогов используемыми критериями сравнения, а также набором сравниваемых моделей.

Таблица 1 – *Вычисленные значения критериев*

№ критерия	Candide-3	Цилиндр	Эллипсоид	Плоскость
1	91.72	106.48	130.52	495.73
2	10.98	11.35	11.46	22.44
3	0.67	0.64	0.58	0.50

Получены оценки ошибок углов поворота головы и построены гистограммы их распределения. Вычислены асимметрия и эксцесс; гипотеза о нормальности распределений не подтвердилась (по критерию Чебышева). Точечные оценки оказались смещенными, что можно объяснить систематической ошибкой, внесенной на этапе инициализации модели головы. О разбросе ошибок оценивания углов поворота ψ , θ , ϕ головы можно судить по «исправленным» среднеквадратичным отклонениям, которые, для модели Candide-3 составили соответственно 5, 7 и 3,5 градуса; для цилиндрической модели – 7, 6 и 4 градуса; для эллипсоидальной – 9, 5 и 4 градуса; для плоской – 13, 18 и 5 градусов.

Описано внедрение полученных результатов в программный продукт ManuCam, разрабатываемый в компании ТомскСофт. Рассмотрены перспективы использования разработанных алгоритмов и методов трекинга головы в компьютерных учебных программах.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

Основные результаты диссертационной работы состоят в следующем:

1. На основе анализа проблемы отслеживания головы и литературных источников сформулированы рекомендации по выбору метода трекинга головы для реализации в программной системе. Если требуется высокая точность определения положения головы, то предпочтение следует отдавать методам, основанным на построении геометри-

ческих моделей головы (отметим, что наихудшую точность имеют методы детектирования головы); однако если требуется автоматическая инициализация системы, то необходимо рассмотреть возможность использования детектирования для инициализации модели (например, мультимодового детектора), а из различных моделей головы предпочтение следует отдать более простым (цилиндр, эллипсоид), которые менее чувствительны к ошибкам инициализации модели. Для увеличения устойчивости и точности системы трекинга головы возможно комплексирование различных алгоритмов на основе методов фильтрации (например, на основе фильтра Калмана). Но этот путь может породить новые проблемы: высокую вычислительную сложность синтезированного алгоритма и т.п.

2. Разработан приближенный алгоритм инициализации модели головы. Данный способ основан на детектировании лица и на том предположении, что при выбранных параметрах камеры, размере и положении модели головы, проекция модели головы должна совпадать с изображением головы на первом видеокадре. Алгоритм пригоден для первоначального позиционирования модели головы без необходимости ручной разметки области лица при использовании некалиброванной монокулярной веб-камеры, что отличает его от методов, требующих ручное задание положения головы.
3. Впервые для задачи трекинга головы как твердого тела, обладающего шестью степенями свободы движения, применен ИК алгоритм регистрации изображений. ИК алгоритм является более эффективным в вычислительном плане, чем известные аналоги.
4. Разработана и зарегистрирована в Роспатенте ПС, реализующая модельный метод отслеживания головы человека и отличающаяся от известных автору аналогов реализованным в ней способом записи информации о положении головы человека на каждом видеокадре. Практическая ценность реализованной ПС состоит в возможности ее использования для получения экспериментальных данных о положении головы человека на видеокадрах, что делает возможным сравнение с данными, полученных другими исследователями, например при сравнении точности различных методов трекинга.
5. Предложены три количественных критерия оценки моделей: сумма квадратов разностей, сумма модулей разностей и коэффициент корреляции. На основе предложенных критериев проведен сравнительный анализ четырех недеформируемых геометрических моделей головы: 3D цилиндра, 3D эллипсоида, плоской прямоугольной модели и модели Candide-3. Данный анализ отличается от аналогов исполь-

зуемыми критериями сравнения, а также набором сравниваемых моделей.

6. Проведенный экспериментально сравнительный анализ четырех моделей головы позволил составить рекомендации по использованию данных моделей, а именно: наилучшую точность трекинга обеспечивают эллипсоидальная и цилиндрическая модели головы. Эти модели применимы при наличии значительных поворотов головы вне плоскости изображения. Модель Candide-3 также обеспечивает хорошую точность, но, так как она, по сути, моделирует лишь форму лица, а не всей головы, то при значительных поворотах головы использовать ее нецелесообразно. Плоская модель существенно уступает указанным трем моделям, особенно при наличии значительного поворота головы вне плоскости изображения.
7. Основные результаты диссертации внедрены в программный продукт ManyCam (<http://manycam.com>), разрабатываемый в компании ТомскСофт. Созданный программный модуль используется для отслеживания головы пользователя компьютера в трехмерном пространстве в режиме реального времени и наложения специальных трехмерных эффектов на видеоизображение лица пользователя, что подтверждается соответствующим актом о внедрении.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в журналах из перечня ВАК

1. Кривцов О.А. Инверсно-композиционный алгоритм регистрации изображений. Программная реализация алгоритма регистрации средствами языка C++ и библиотеки OpenCV / О.А. Кривцов, А.М. Кориков // RSDN Magazine – М.: К-Пресс. – 2010. – № 1 – С. 59–72.
2. Кривцов О.А. Отслеживание головы человека на видеокадрах: классификация и выбор / О.А. Кривцов, А.М. Кориков // Информационные технологии – М.: Новые технологии, 2010. – № 9 – С. 13–19.
3. Кривцов О.А. Отслеживание головы человека на видеокадрах: состояние проблемы / О.А. Кривцов, А.М. Кориков // Информационные технологии – М.: Новые технологии, 2010. – № 7 – С. 2–9.
4. Кривцов О.А. Моделирование головы человека для задачи отслеживания ее положения на видеокадрах // Доклады Томского государств. универ. систем управл. и радиоэлектр. – Томск: ТУСУР, 2009. – № 1(19), Ч. 1. – С. 109–115.
5. Кривцов О.А. Отслеживание положения головы человека в пространстве на основе анализа видеокадров / О.А. Кривцов, А.М. Кориков // Доклады Томского государств. универ. систем управл. и радиоэлектр. – Томск: ТУСУР, 2008. – № 2(18), часть 2. – С. 32–39.

Статьи в других изданиях

6. Кориков А.М. Развитие систем человек-компьютер на основе методов трекинга головы человека / А.М. Кориков, О.А. Кривцов // Материалы

- докладов 16-й Междунар. науч.-техн. конф-ии «Природные и интеллект. ресурсы Сибири» – Томск: САН ВШ; В-Спектр 2010. – С. 14–18.
7. Кривцов О.А. Исследование применимости метода Ньютона и метода Левенберга-Марквардта к задаче регистрации изображений // Материалы докладов Всероссийской науч.-техн. конф-ии студентов, аспиран. и молодых ученых «Научная сессия ТУСУР-2010» – Томск: В-Спектр, 2010 – Ч. 4 – С. 231–234.
 8. Кривцов О.А. Методы трекинга головы человека на видео // Материалы докладов Всероссийской науч.-техн. конф-ии студентов, аспиран. и молодых ученых «Научная сессия ТУСУР-2010» – Томск: В-Спектр, 2010 – Ч. 1 – С. 149–152.
 9. Кривцов О.А. Отслеживание жестких объектов на видеокадрах с помощью текстурных моделей // Материалы докладов VII Всероссийской науч.-практ. конф-ии студ., аспиран. и молодых ученых «Молодежь и современные информационные технологии» – Томск: СПб Графика, 2009 – Ч.1 – С. 149–150.
 10. Кривцов О.А. Программа отслеживания перемещения головы на видео // Труды XIV междунар. науч.-практ. конф-ии «Современная техника и технологии» – Томск: ТПУ, 2008 – Т. 2. – С. 320–321.
 11. Кривцов О.А. Система отслеживания перемещения головы на видео кадрах. Труды конф-ии «Научная сессия ТУСУР» – Томск: В-Спектр, 2008. – Т. 1. – С. 144.
 12. Кривцов О.А. Сравнительный анализ жестких моделей головы человека применительно к задаче отслеживания головы на видеокадрах / О.А. Кривцов, А.М. Корилов // Материалы докладов Всероссийской науч.-техн. конф-ии студ., аспиран. и молодых ученых «Научная сессия ТУСУР» – Томск: В-Спектр, 2009 – Ч. 4. – С. 218–221.
 13. Кривцов О.А. Тестовое видео для трекинга головы человека // Материалы XLVIII Международной научной студен. конф-ии «Студент и научно-технический прогресс» : Информационные технологии – Новосибирск, 2010. – С. 15.
 14. Кривцов О.А. Управление курсором мыши путем отслеживания наклона головы пользователя // Доклады науч.-практ. конф-ии «Электронные ср-ва и системы управл.: итоги реализ. прогр. развития электрон. и IT-технол. в Томской обл.» – Томск: В-Спектр, 2008. – С. 190–195.
 15. Krivtsov O. An Analysis of Performance and Accuracy of a Video Head Tracking System / O. Krivtsov, A. Korikov // Proc. 15th Intl. Conf. «Modern Technique and Technologies» – Томск: TPU Press, 2009. – P. 113–115.
 16. Krivtsov O. Applying Inverse-Compositional Image Registration Algorithm to Head Tracking with 6 DOF // Proc. 16th Intl. Conf. «Modern Technique and Technologies» – Томск: TPU Press, 2010. – P. 91–93.

Свидетельства об интеллектуальной собственности

17. Свидетельство о государственной регистрации программы для ЭВМ №2010610249 от 11 января 2010 г. «Head Tracking Demo» / Правообладатель и автор: Кривцов О.А.