

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»

А. А. ЗАХАРОВ А. Е. БАРИНОВ
А. Л. ЖИЗНЯКОВ

МОДЕЛИ И АЛГОРИТМЫ
ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ
И ОРИЕНТАЦИИ ГОЛОВЫ ЧЕЛОВЕКА
ПО ВИДЕОИЗОБРАЖЕНИЯМ
С ИСПОЛЬЗОВАНИЕМ ГРАФОВ

Монография



Владимир 2019

УДК 004.932
ББК 32.972.13
3-38

Рецензенты:

Доктор технических наук, профессор
профессор кафедры авиационных радиоэлектронных комплексов
Военно-воздушной академии имени профессора
Н. Е. Жуковского и Ю. А. Гагарина
А. В. Леньшин

Доктор технических наук, профессор
профессор кафедры информационных систем и технологий
Юго-Западного государственного университета
С. В. Дегтярев

Захаров, А. А. Модели и алгоритмы определения положения и ориентации головы человека по видеоизображениям с использованием графов : монография / А. А. Захаров, А. Е. Баринов, А. Л. Жизняков ; Владим. гос. ун-т им. А. Г. и Н. Г. Столетовых. – Владимир : Изд-во ВлГУ, 2019. – 124 с.
3-38 ISBN 978-5-9984-0985-1

Рассмотрены модели и алгоритмы определения положения и ориентации головы человека по видеоизображениям с использованием графов, охарактеризованы методы и системы определения положения и ориентации объектов в пространстве. Показаны преимущества использования видеоизображений для поставленной задачи. Предложена система алгоритмов определения положения и ориентации головы человека на основе изображений, рассмотрено применение разработанных алгоритмов. Приведены результаты исследований.

Предназначена для специалистов в области обработки и распознавания изображений.

Табл. 4. Ил. 70. Библиогр.: 165 назв.

УДК 004.932
ББК 32.972.13

ISBN 978-5-9984-0985-1

© Захаров А. А., Баринов А. Е.,
Жизняков А. Л., 2019

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
Глава 1. АНАЛИЗ ЗАДАЧИ, МЕТОДОВ И СИСТЕМ ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ И ОРИЕНТАЦИИ ГОЛОВЫ ЧЕЛОВЕКА ДЛЯ ТРАНСПОРТНЫХ ТРЕНАЖЕРОВ.....	7
1.1. Актуальность задачи определения положения и ориентации головы человека для транспортных тренажеров	7
1.2. Анализ систем и технологий определения положения и ориентации головы человека	12
1.3. Анализ методов определения положения и ориентации головы человека на основе изображений	21
Глава 2. РАЗРАБОТКА МОДЕЛИ И АЛГОРИТМОВ ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ И ОРИЕНТАЦИИ ГОЛОВЫ ЧЕЛОВЕКА ПО СТЕРЕОИЗОБРАЖЕНИЯМ С ИСПОЛЬЗОВАНИЕМ ГРАФОВ.....	32
2.1. Модель стереоустановки.....	32
2.2. Разработка модели определения положения и ориентации головы человека с использованием стереоизображений	36
2.3. Разработка алгоритма кластеризации особенностей на основе спектральной теории графов и ограничений для выделения лица человека на изображениях	47
2.4. Разработка алгоритма нахождения соответствий на изображениях с использованием тепловых ядер на графах	57
2.5. Разработка алгоритма обнаружения головы человека на изображениях на основе вложения графа особенностей в векторное пространство.....	65

Глава 3. ИССЛЕДОВАНИЕ РАЗРАБОТАННЫХ МОДЕЛИ И АЛГОРИТМОВ ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ И ОРИЕНТАЦИИ ГОЛОВЫ ЧЕЛОВЕКА ПО СТЕРЕОИЗОБРАЖЕНИЯМ С ИСПОЛЬЗОВАНИЕМ ГРАФОВ.....	73
3.1. Исследование алгоритма кластеризации особенностей на основе спектральной теории графов и ограничений для выделения лица человека на изображениях	73
3.2. Исследование алгоритма нахождения соответствий на изображениях с использованием тепловых ядер на графах	81
3.3. Исследование алгоритма обнаружения головы человека на изображениях на основе вложения графа особенностей в векторное пространство.....	85
3.4. Исследование модели определения положения и ориентации головы человека с использованием стереоизображений	88
 Глава 4. ПРИМЕНЕНИЕ РАЗРАБОТАННЫХ МОДЕЛИ И АЛГОРИТМОВ В ТРЕНАЖЕРАХ ТРАНСПОРТНЫХ СРЕДСТВ	92
4.1. Структура системы определения положения и ориентации головы человека по стереоизображениям	92
4.2. Методика настройки разработанной системы определения положения и ориентации головы человека по стереоизображениям.....	94
4.3. Внедрение системы определения положения и ориентации головы человека по стереоизображениям.....	102
 ЗАКЛЮЧЕНИЕ.....	105
 БИБЛИОГРАФИЧЕСКИЙ СПИСОК.....	107

ВВЕДЕНИЕ

На сегодняшний день неотъемлемым элементом качественной подготовки водителей стало использование автомобильных тренажеров. Применение тренажеров позволяет сократить сроки подготовки водителей, снизить расход моторесурса техники, уменьшить затраты на горючее и обслуживание транспорта. Тренажеры позволяют осуществлять индивидуальный подход к обучению. С помощью тренажерных технологий легче прививать необходимые водительские навыки путем многократного повторения заданных действий для управления машиной. При использовании тренажеров процесс подготовки водителей представляется как набор определенных упражнений. Кроме того, можно воспроизвести любые аварийные ситуации, отработка которых в реальном транспортном средстве сопряжена с серьезной опасностью.

Важной частью современных тренажеров является система визуализации окружающего пространства. На протяжении всего периода развития тренажеров разработке и совершенствованию системы визуализации уделялось большое внимание. Но даже в настоящее время создаваемый эффект присутствия в тренажере сильно отличается от ощущений водителя в реальном транспортном средстве. Одной из причин этого является то, что при синтезе окружающей обстановки в транспортном тренажере система визуализации не изменяет отображаемую картину мира в зависимости от положения и ориентации головы человека в кабине. В свою очередь, это приводит к невозможности реалистичного имитирования зеркал заднего вида и всей области вокруг транспортного средства. Это также отрицательно влияет на подготовку водителей транспортных средств. Для решения этой проблемы необходимо при моделировании виртуального пространства учитывать положение и ориентацию головы водителя.

В настоящее время для определения положения и ориентации объектов в пространстве используются маркерные и безмаркерные подходы. Маркерные методы основаны на использовании различного рода датчиков, которые крепятся к голове человека. Основным недо-

статком маркерных методов является невозможность применения датчиков при постоянном наблюдении множества сцен, что ограничивает их использование. Также подобные технологии характеризуются высокой стоимостью. Поэтому важна разработка методов и программных средств определения положения и ориентации головы человека на основе безмаркерных технологий.

Безмаркерные технологии подразумевают использование методов на основе компьютерного зрения. Однако вопросы определения положения и ориентации объектов на основе пассивных оптических методов в настоящее время изучены недостаточно. Использование компьютерного зрения ограничено следующими факторами: малый размер сцен, наличие шумовых и оптических эффектов, изменяющийся текстурированный фон, эффекты резкой смены освещения, взаимное перекрытие объектов, низкая скорость обработки данных [2, 15, 36, 37, 41]. Однако эти методы являются наиболее перспективными с точки зрения простоты использования, универсальности и стоимости аппаратного обеспечения. В последнее время было разработано большое количество методов и систем определения положения и ориентации головы на основе изображений. Однако использование подобных систем ограничено конкретными приложениями, поэтому их затруднительно применять в распространенных тренажерах транспортных средств.

Создание системы определения положения и ориентации головы человека на основе стереоизображений позволит увеличить эффект присутствия обучаемого в синтезируемой среде за счет изменения области видимости.

Глава 1. АНАЛИЗ ЗАДАЧИ, МЕТОДОВ И СИСТЕМ ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ И ОРИЕНТАЦИИ ГОЛОВЫ ЧЕЛОВЕКА ДЛЯ ТРАНСПОРТНЫХ ТРЕНАЖЕРОВ

1.1. Актуальность задачи определения положения и ориентации головы человека для транспортных тренажеров

На сегодняшний день неотъемлемым элементом качественной подготовки водителей транспортных средств является использование тренажеров, с помощью которых можно получить навыки вождения на основе комплекса отдельных упражнений. Основная задача обучения на тренажере – выработка у учащегося начальных зрительно-двигательных навыков управления транспортным средством и восприятия дорожной обстановки в безопасных для жизни условиях [30, 38].

Современные тренажеры являются единой системой, состоящей из множества устройств, каждое из которых имеет строго определенное назначение. В зависимости от степени сложности тренажер может имитировать либо определенные функции транспортного средства, либо полностью весь процесс управления. Тренажеры транспортных средств позволяют проводить качественную подготовку, воспроизводя сложные ситуации, которые отрабатывать в действительности опасно.

По назначению современные тренажеры транспортных средств обычно делятся на функциональные и комплексные [30, 38]. Функциональные тренажеры вырабатывают отдельные навыки. С помощью комплексных тренажеров развивается целый набор водительских навыков одновременно.

Тренажеры могут быть статическими и динамическими. Платформа в статических тренажерах неподвижна. Динамические тренажеры включают в себя подвижную платформу, которая имитирует процесс перемещения транспортного средства: разгон и торможение, повороты, неровности дорожного покрытия и вибрации двигателя. В подобных тренажерах реалистично передается движение по пересеченной

местности, наезды на препятствия, повороты и т.д. В кабине тренажера присутствуют многие элементы управления транспортного средства: педали, рулевая колонка, коробка переключения передач и т.д.

Одной из основных частей тренажера является система визуализации. На протяжении всей истории развития тренажеров вопросам создания систем визуализации уделялось большое внимание. В первых образцах тренажеров моделирование окружения выполнялось в виде стендов.

С развитием вычислительной техники для имитации визуальной обстановки используются компьютерные технологии [1]. Инструктор может самостоятельно выбирать маршрут в рамках программы обучения.

На сегодняшний день тренажеры с помощью программного обеспечения могут достаточно точно и реалистично синтезировать визуальную обстановку. Однако в настоящее время системы визуализации местности в транспортных тренажерах все еще имеют некоторые недостатки.

Одним из недостатков многих тренажеров является то, что при изменении положения и ориентации головы человека на экранах не изменяется область видимости [1]. Таким образом, снижается эффект присутствия обучаемого в синтезируемой среде, что не лучшим образом сказывается на процессе подготовки.

В реальной кабине транспортного средства происходит изменение области видимости при перемещении головы в пространстве. В этом случае область видимости в пространстве можно представить усеченной пирамидой. При отображении сцены сверху область видимости упрощенно можно описать равнобедренной трапецией (рис. 1.1).

В то же время, например в автомобильном тренажере, существует проблема имитации зеркал заднего вида. В автомобиле использование зеркал заднего вида является необходимым условием безопасного вождения.

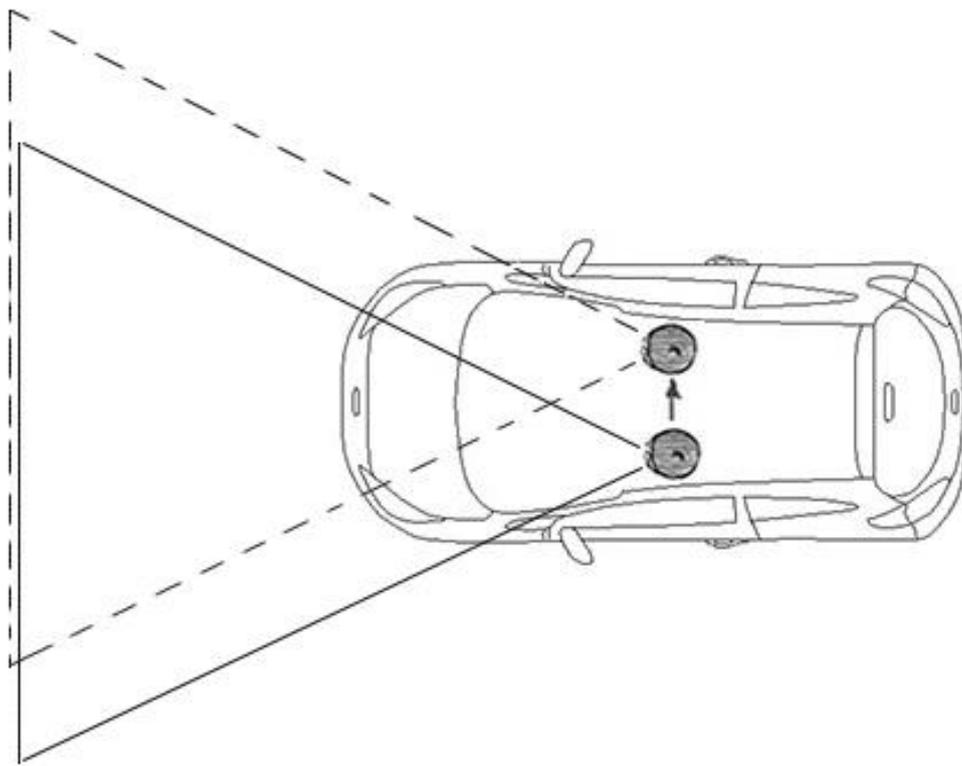


Рис. 1.1. Изменение области видимости при изменении положения и ориентации головы человека

При обучении на автотренажере без системы определения положения и ориентации головы человека привычка контролировать обстановку вокруг транспортного средства с использованием зеркал заднего вида не вырабатывается [13, 20, 23].

При любой настройке зеркал невозможно избежать появления слепых зон (рис. 1.2). Слепая зона автомобиля – пространство вокруг машины, которое водитель не может увидеть, располагаясь на своем месте. Такая область в зависимости от конструкции автомобиля может присутствовать в разных частях сцены. Часто происходят ситуации, когда проезжающий мимо автомобиль не видно ни в одном из зеркал заднего вида, так как он находится непосредственно в слепой зоне. Наличие слепых зон является одной из причин, по которой происходят аварии на дорогах.

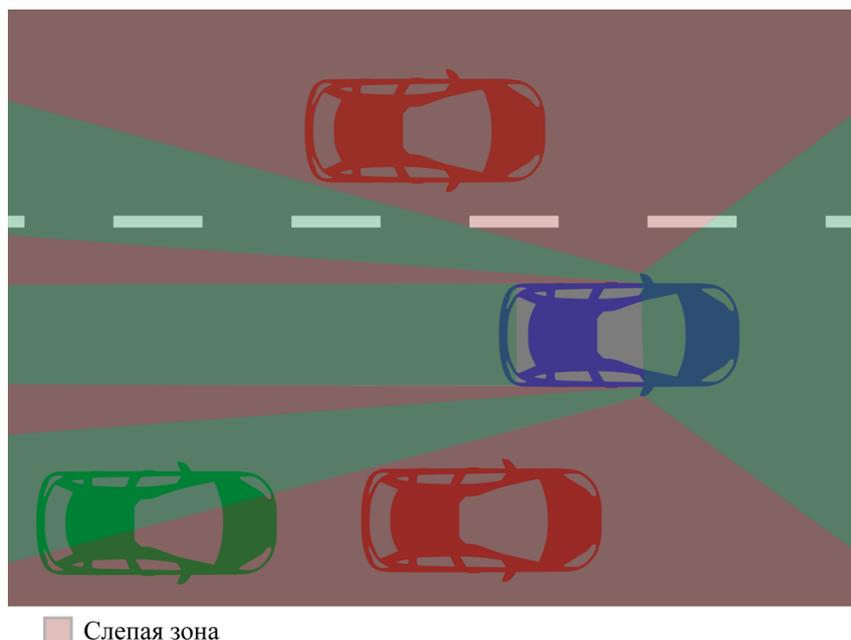


Рис. 1.2. Слепые зоны автомобиля

Чтобы избавиться от эффекта слепых зон, водителю необходимо крутить головой, изменять положение тела для нахождения более выгодного ракурса (рис. 1.3). Таким образом, регулярный обзор зеркал заднего вида – важный навык в управлении автомобилем, который в большинстве современных тренажеров не развивается.

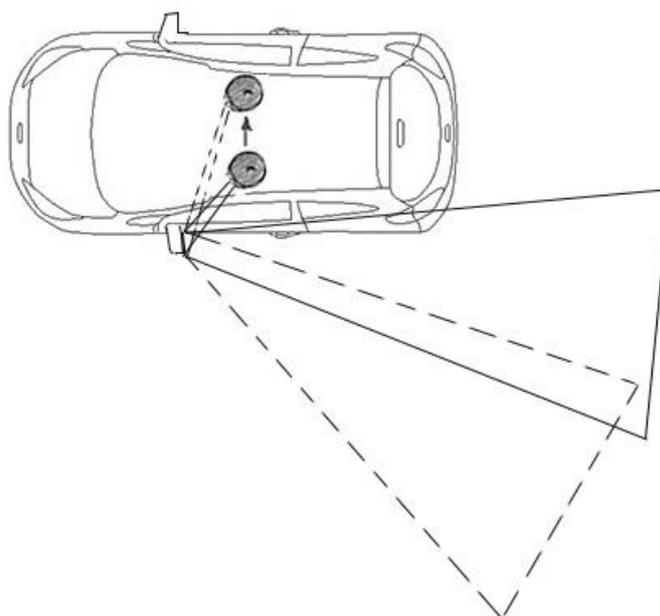


Рис. 1.3. Изменение области видимости в зеркале заднего вида в зависимости от текущего положения головы водителя

В некоторых современных моделях тренажеров на месте зеркал заднего вида установлены специальные дисплеи, которые показывают ситуацию на дороге. Однако они не изменяют изображение в зависимости от положения и ориентации головы водителя.

Для решения описанной проблемы в существующие тренажеры необходимо интегрировать систему, которая будет определять положение и ориентацию головы водителя. В соответствии с текущим положением головы системе моделирования визуального пространства тренажера будут передаваться управляющие команды. Для этого предлагается использовать методы компьютерного зрения. Слежение за головой человека в пространстве будет осуществляться с помощью оптических камер, объединенных в стереоустановку. Это позволит определять трехмерные координаты головы в пространстве. Разработка новых алгоритмов компьютерного зрения позволит решить поставленную задачу в реальном масштабе времени с высокой точностью.

С внедрением данной системы изменится отображение внешнего пространства транспортного средства. Система даст возможность контролировать обстановку и лучше чувствовать габариты транспортного средства, так как в зависимости от положения и ориентации головы будет меняться отображение всей сцены. Таким образом, тренажеры, использующие данную систему, смогут повысить эффект присутствия обучаемого в синтезированной среде транспортного средства. Кроме повышения реалистичности отображения пространства, у обучающегося появится возможность освоения такого важного навыка, как контроль дорожной обстановки посредством обзора зеркал заднего вида. Это повысит качество подготовки водителей и в дальнейшей перспективе сократит количество дорожно-транспортных происшествий [31].

Таким образом, определение положения и ориентации головы человека в пространстве является актуальной научно-технической задачей при создании тренажеров транспортных средств.

1.2. Анализ систем и технологий определения положения и ориентации головы человека

На сегодняшний день задача определения положения и ориентации объектов решается с помощью различных технологий (рис. 1.4).

Технологии захвата движения обозначаются термином MoCap (Motion Capture). По принципу работы с отслеживаемым объектом можно выделить две большие группы: маркерные и безмаркерные технологии [111, 148].



Рис. 1.4. Технологии определения положения и ориентации объектов

В маркерных системах принцип отслеживания состоит в использовании специальных датчиков, которые размещаются на объекте (табл. 1.1). По характеру физического взаимодействия маркерные методы можно разделить на следующие группы: оптические, магнитные, инерциальные, звуковые, механические [111, 148].

Таблица 1.1. *Маркерные системы определения положения и ориентации трехмерных объектов в пространстве*

Название / Производитель	Технология	Количество измерений в секунду
Smart-DX (BTS Bioengineering, США)	Оптические пассивные	120
Raptor-12HS Digital RealTime System (Motion Analysis, США)	Оптические пассивные	300
Bonita (Vicon-Peak, Великобритания)	Оптические пассивные	250
Certus HD (Northern Digital, Канада)	Оптические активные	90
Visualeyez system (PTI Phoenix Technologies Inc., Канада)	Оптические активные	10000
trakSTAR (Ascension technology, США)	Магнитные	80
Polhemus system (Polhemus, США)	Магнитные	100
IGS-Cobra (Synertial, Великобритания)	Инерциальные	180
MVN Awinda (Xsens, Нидерланды)	Инерциальные	60/240
Oculus VR (Oculus VR, LLC, США)	ШВР	60
PlayStation VR (Sony, Япония)	ШВР	60
TrackIR (Natural Point Inc., США)	Оптически пассивные	30
RUCAP UM-5 (РУКЭП, Россия)	Звуковые	160
HeadJoy (Gadget63, Россия)	Оптические активные	125
A.R.T. (Advanced Realtime Tracking GmbH, Германия)	Оптические активные	120
Gypsy 7 (MetaMotion, США)	Механические	30/60/120

В зависимости от принципа работы со светом оптические подходы могут быть пассивными и активными. В оптических пассивных методах используются датчики, которые отражают падающий на них свет. В качестве источника света используются высокочастотные стробоскопы. Отражения маркеров фиксируются специальными камерами. Такие методы имеют следующие недостатки: необходимость оснащения объекта маркерами, уязвимость к взаимным перекрытиям объектов, высокая стоимость [111, 148].

В оптических активных системах свет испускается непосредственно маркерами, которые срабатывают в определенный момент времени. Используются синхронизируемые светодиоды. Каждому светодиоду назначается идентификатор. Это дает возможность опознавать маркеры при взаимном перекрытии. Недостатки подобных систем: необходимость размещения маркеров, относительно высокая стоимость технологии [111, 148].

При использовании магнитных технологий маркерами являются магниты, а вместо камер применяются сенсоры измерения магнитного поля. Магниты крепятся на объекте, движение которого отслеживается. Система определяет положение объекта по значению искажения магнитного поля. Недостатки подобных систем: зависимость от магнитных и электрических помех, малое рабочее пространство, необходимость размещения маркеров, высокая стоимость [111, 148]. Также электромагнитный сенсор, устанавливаемый на объект, не может быть беспроводным из-за большого количества потребляемой энергии.

Инерциальные системы для сбора информации о положении и ориентации объекта используют гироскопы и инертные сенсоры. Подобные устройства оценивают изменение углов ориентации объекта относительно инерциальной системы отсчета. Недостатки инерциальных систем: необходимость размещения маркеров, высокая стоимость [111, 148]. Шлемы виртуальной реальности (ШВР) также относятся к инерциальным системам [3]. Но из-за высокой стоимости, закрытости архитектуры, а также наличия индивидуальной непереносимости у некоторых людей применение ШВР в тренажерах часто является затруднительным.

Звуковые системы подобны оптическим системам с активными маркерами. Датчики формируют звук, вместо камер используются специальные микрофоны. Методы, основанные на анализе звуковых волн, можно разделить на две группы:

- подходы, основанные на времени прохождения сигнала;
- подходы, учитывающие разность фаз сигнала.

В системах, оценивающих время прохождения сигнала, датчики расположены на отслеживаемом объекте. Большое количество сенсоров, определяющих время прохождения сигнала от передатчика, позволяет получить целостную трехмерную картину перемещения объекта в пространстве. Недостатком подобных систем является невысокая скорость звуковой волны, а также уменьшение точности из-за таких факторов, как температура и влажность [111, 148].

В системах, основанных на получении разности фаз сигнала, используются два датчика. Первый датчик размещается на движущемся объекте, второй – в контрольной точке. Основным недостатком данных систем состоит в том, что известно не положение отслеживаемого объекта, а изменения его координат в отдельные моменты времени.

Механические системы представляют собой каркас, который крепится к объекту. Движение составных частей объекта приводит в движение элементы каркаса. В этом случае анализируются углы между скелетными элементами каркаса. Однако перемещение всего объекта в пространстве с помощью этой системы отслеживать невозможно [111, 148].

Компактная система Smart-DX от компании BTS Bioengineering для захвата движения объекта использует 4 инфракрасные камеры. Система имеет возможности автоматического выделения маркеров и быстрой калибровки. При отслеживании используются оптические пассивные маркеры. Система сохраняет свою функциональность при наличии слабого и интенсивного освещения. Захват движения осуществляется с частотой 120 кадров в секунду при разрешении 4 мегапикселя [117].

В системе Raptor-12HS используются одновременно двенадцать камер. Камеры имеют встроенный процессор, предназначенный для захвата движения. Система может работать в условиях плохой освещенности. Используются оптические пассивные маркеры. Частота при захвате движения достигает 300 кадров в секунду при разрешении 12 мегапикселей [117].

Система Vonita имеет оптическую камеру, которая позволяет осуществлять захват движения. При разрешении 1 мегапиксель частота захвата движения достигает 250 кадров в секунду. Калибровка камеры осуществляется в течение двух минут. Отслеживаемые маркеры являются оптическими пассивными [116].

Система определения положения и ориентации объектов Certus HD основана на использовании оптических активных маркеров. Позволяет отслеживать одновременно до 170 объектов по 6 степеням свободы. Камера работает со специальными маркерами Optotral Smart Markers. Маркеры в количестве до 50 штук размещаются на объекте и взаимодействуют с 8 стробоскопами. Частота захвата движения достигает 90 кадров в секунду [117].

Система Visualeyzez компании Phoenix Technologies обладает самой большой рабочей зоной среди аналогов, равной 190 м³. Калибровка проводится полностью в автоматическом режиме. Для отслеживания используются оптические активные светодиодные маркеры. Система позволяет отслеживать одновременно до 512 маркеров. Захват движения осуществляется с частотой 10000 кадров в секунду [117].

Система trakSTAR используется в медицине для обучения персонала обращению с инструментами. На объект прикрепляются 16 электромагнитных сенсоров различного размера, которые могут отслеживаться одновременно. Используется проводное подключение. Обработка данных о положении и ориентации объекта осуществляется со скоростью 80 измерений в секунду. Движения объекта анализируются по 6 степеням свободы. Одновременно могут отслеживаться до 4 объектов [125].

Система Polhemus используется для обучения хирургов, врачей функциональной диагностики, рентгенологов и т.д. В основе работы лежит использование электромагнитных датчиков. Определение положения и ориентации объекта выполняется по 6 степеням свободы. Движение объекта фиксируется со скоростью до 100 измерений в секунду [140].

Система IGS-Cobra основана на использовании инерциальных маркеров. Система обладает высокой точностью и скоростью обработки. Передача информации осуществляется беспроводным способом. Отслеживаемый объект оснащается акселерометрами в количестве до 47 штук. Калибровка осуществляется в автоматическом режиме. Система работает со скоростью 180 измерений в секунду [98].

Компания Xsens поставляет систему MVN Awinda, в основе которой лежит использование инерциальных датчиков. Есть два типа подключения: беспроводное и проводное. В первом случае скорость обработки составляет 60 измерений в секунду, во втором – 240. Система использует 17 инерциальных датчиков и позволяет отслеживать движения объекта по 6 степеням свободы [100].

Система Oculus Rift представляет собой очки виртуальной реальности, где определение положения и ориентации головы осуществляется по 3 степеням свободы с помощью гироскопов и акселерометров. Используется проводной тип подключения. В зависимости от персонального компьютера, к которому подключены очки, производительность системы сильно изменяется и может достигать 60 кадров в секунду [143].

Очки виртуальной реальности PlayStation VR от компании Sony предназначены для проводного подключения к игровой консоли PlayStation 4. Характеристики этой системы аналогичны параметрам системы Oculus Rift [143].

В системе TrackIR применяется инфракрасная камера с диодной подсветкой, оценивающая положение объекта на основе оптических пассивных маркеров. Слежение происходит по 6 степеням свободы со скоростью обработки 30 кадров в секунду. Система используется при создании авиасимуляторов [39].

В системе RUCAP UM-5 отслеживание положения и ориентации головы человека осуществляется за счет использования ультразвукового маркера. Система позволяет выполнять до 160 измерений в секунду. Слежение происходит по 5 степеням свободы [39].

Система HeadJoy имеет оптический активный маркер с четырьмя светодиодами, закрытыми специальной шторкой. Приемник срав-

нивает интенсивность света от каждого из них и вычисляет поворот головы пользователя. Скорость обработки видеоданных составляет 125 кадров в секунду [64].

Система A.R.T. применяется в научных исследованиях, а также в системах виртуальной реальности. Система основана на использовании оптических активных инфракрасных датчиков. Система A.R.T. позволяет определять положение и ориентацию объекта с частотой 120 кадров в секунду [126].

Механическая система отслеживания положения и ориентации тела человека Gypsy 7 от компании MetaMotion представляет собой каркас, на котором расположено 14 датчиков. В зависимости от предоставляемого программного обеспечения система обеспечивает от 30 до 120 измерений в секунду. Вес скелета составляет около 5 кг [99].

Общим недостатком описанных выше систем является необходимость оснащения отслеживаемого объекта специальными датчиками, что во многих практических приложениях может быть затруднительно. Также подобные системы ограничивают свободу действий пользователя, сковывая движения и вызывая дискомфорт. Кроме того, маркерные системы имеют высокую стоимость, поэтому их использование ведет к удорожанию разрабатываемых тренажеров.

Безмаркерные системы основаны на использовании методов компьютерного зрения и распознавания образов. Безмаркерные технологии определения положения и ориентации не ограничивают движения объекта, давая свободу для совершения различных действий. В то же время аппаратная часть таких систем имеет более низкую стоимость по сравнению с маркерными подходами.

Безмаркерные системы используют RGB-D и цифровые оптические камеры. RGB-D камеры являются устройствами, в состав которых входят цифровые камеры и сенсоры глубины. Датчик глубины состоит из инфракрасного проектора и светочувствительной матрицы. Полученные значения глубины позволяют формировать трёхмерное изображение сцены при любом естественном освещении. Самыми популярными моделями таких камер являются Microsoft Kinect и Asus Xtion. RGB-D камеры подходят для решения задачи определения по-

ложения и ориентации объекта, так как имеют возможности для получения трехмерных координат. В то же время они стоят дороже своих оптических аналогов [143].

Использование цифровых камер предполагает применение методов компьютерного зрения и распознавания образов. В этом случае для получения трехмерных координат отслеживаемого объекта может использоваться стереосистема. Цифровые камеры широко распространены и обладают относительно низкой стоимостью.

На сегодняшний день на рынке присутствует целый ряд систем, отслеживающих положение и ориентацию головы человека на основе безмаркерного подхода (табл. 1.2). Часто подобные системы применяются для контроля состояния водителя (КСВ). Главным недостатком таких продуктов является возможность работать только в автомобиле фирмы-производителя. Кроме того, точность определения положения и ориентации головы в таких системах ограничена, так как принцип их действия часто основан на слежении за глазами. Также подобные системы используются для управления компьютером при помощи движений головы (УКГ). Однако они не отличаются высокой точностью и не могут в полной мере заменить компьютерную мышь и клавиатуру [14].

Система мониторинга ComSense фирмы Saab отслеживает сонливость и потерю концентрации водителя. Она состоит из двух инфракрасных камер. Используется интеллектуальная система SmartEye для анализа движения век, направления взгляда и ориентации головы. Точность и скорость распознавания в данной системе заметно падает в ночное время суток [75].

Система контроля усталости водителя Toyota Driver Monitoring System использует одну инфракрасную камеру, с помощью которой отслеживается положение и ориентация головы человека. Если голова водителя долгое время не направлена в сторону движения автомобиля, то производится звуковое оповещение. Если голова водителя не возвращается в исходное положение, система может активировать режим торможения. Также система может определить состояние сонливости водителя на основе анализа движения век [75].

Система мониторинга усталости водителя фирмы Delphi использует одну камеру и два источника инфракрасного света. Система осуществляет слежение за особыми точками головы на изображении и анализирует частоту моргания. Система определяет состояние водителя во время управления автомобилем [77].

Таблица 1.2. *Безмаркерные системы определения положения и ориентации головы человека*

Название / Производитель	Тип	Особенности	Недостатки
ComSense (Saab Automobile AB, Швеция)	КСВ	Контроль направления взгляда инфракрасными камерами	Неустойчивость к перепадам освещения
Toyota Driver Monitoring System (Toyota, Япония)	КСВ	Определение положения головы по инфракрасным изображениям	Неустойчивость к перепадам освещения
Delphi's Driver State Sensor (Delphi, США)	КСВ	Анализ частоты моргания и положения головы	Чувствительность к быстрому движению головы
faceLAB (Seeing Machines, США)	КСВ	Слежение за положением головы и направлением взгляда	Используется только в научных целях. Существуют ограничения для практического применения
Smart Eye (Smart Eye AB, Швеция)	КСВ	Использование двух независимых камер	Чувствительность к быстрому движению головы. Необходимость инициализации и калибровки
InSight (SensoMotoric Instruments, Германия)	КСВ	Слежение за положением головы и направлением взгляда	Необходимость инициализации и калибровки
Enable Viacam (Crea Software Systems, Испания)	УКГ	Эмуляция нажатия клавиш с помощью взгляда	Низкая точность
FaceTrackNoIR (Open-source)	УКГ	Управление обзором в компьютерных играх	Низкая скорость и точность
Cam2Pan (Open-source)	УКГ	Эмуляция нажатия клавиш с помощью взгляда	Низкая точность

Система контроля состояния водителя faceLAB от компании Seeing Machines отслеживает положение головы, век, направление взгляда. Применяются две камеры, образующие стереосистему. Система используется как инструмент для проведения различных исследований, но еще не готова для практического применения в реальных приложениях [101].

Система компании Smart Eye AB отслеживает движение головы и направление взгляда водителя. Система имеет функции обнаружения лица, определения направления взгляда, анализа положения век. Одним из преимуществ системы является то, что используются две независимые камеры. Тем самым устраняется проблема, связанная с возможными взаимными перекрытиями лица другими объектами [78].

Система InSight немецкой компании SensoMotoric Instruments может определять положение и ориентацию головы, направление взгляда, частоту моргания, продолжительность закрытия век, положение и диаметр зрачков. Недостатком системы является необходимость долгого процесса начальной и повторной калибровки для каждого водителя [145].

Отдельно стоит выделить программы, которые позволяют с помощью обычной веб-камеры отслеживать положение и ориентацию головы человека: Enable Viacam, FaceTrackNoIR, Cam2Pan [14].

Enable Viacam является программой, которая управляет компьютером при помощи движений головы пользователя [14]. Реализация алгоритмов компьютерного зрения осуществляется с помощью библиотеки OpenCV [63].

FaceTrackNoIR используется для управления в компьютерных играх. Система Cam2Pan обладает аналогичными с Enable Viacam функциями, но для большей точности разработчики рекомендуют использовать специальный светоотражающий маркер [14].

1.3. Анализ методов определения положения и ориентации головы человека на основе изображений

На сегодняшний день разработан ряд методов для определения положения и ориентации головы человека на основе компьютерного зрения. В таблице 1.3 приведена сравнительная характеристика существующих подходов.

Методы шаблонов. Определение положения и ориентации головы с помощью методов шаблонов основано на сравнении текущего изображения головы с заранее подготовленными эталонами (рис. 1.5). На этапе обучения для каждого характерного положения головы создается свой собственный шаблон. В некоторых реализациях для сопоставления с шаблонами используется нормализованная корреляция или среднеквадратичное отклонение по скользящему окну [121]. Для сравнения шаблонов могут использоваться локальные бинарные детекторы [113]. Для повышения точности алгоритма могут применяться различные техники уточнения лицевых характеристик: фильтры Гаусса и Лапласа [81], вейвлеты Габора [141]. Применение рассматриваемого подхода требует предварительного обучения.

Таблица 1.3. Методы определения положения и ориентации головы человека, основанные на использовании компьютерного зрения

Метод	Диапазон углов (поворот, кивок, наклон)	Достоинства	Недостатки	Скорость (кадров/сек)
Методы шаблонов (Moore, 2011)	$\pm 80^\circ$; $\pm 25^\circ$; $\pm 10^\circ$.	Простота реализации; возможность расширения базы эталонов	Низкая точность; медленная обработка	15
Методы массивов детекторов (Zhang, 2007)	$\pm 90^\circ$; $\pm 30^\circ$; $\pm 30^\circ$.	Не требуется отдельный этап локализации головы; работа с изображениями низкого разрешения	Сложность обучения; медленная обработка	15
Методы нелинейной регрессии (Drouard, 2015)	$\pm 90^\circ$; $\pm 90^\circ$; 0.	Высокое быстродействие; высокая точность	Сложность формирования базы эталонов; производительность сильно зависит от размерности входных данных	60

Окончание табл. 1.3

Метод	Диапазон углов (поворот, кивок, наклон)	Достоинства	Недостатки	Скорость (кадров/сек)
Методы вложенных многообразий (Sundararajan, 2015)	$\pm 50^\circ$; $\pm 45^\circ$; $\pm 20^\circ$.	Обработка взаимных перекрытий в сцене	Чувствительность к оптическим искажениям	30
Гибкие модели (Chen, 2014)	$\pm 55^\circ$; $\pm 45^\circ$; $\pm 45^\circ$.	Высокое быстродействие; высокая точность	Чувствительность к взаимным перекрытиям и быстрому вращению головы	60
Геометрические методы (Hatem, 2015)	$\pm 45^\circ$; $\pm 45^\circ$; $\pm 45^\circ$.	Высокое быстродействие; высокая точность при малых углах поворота головы	Чувствительность к взаимным перекрытиям и шумам. Возможны ошибки при локализации особых точек лица	45
Гибридные методы: статические шаблоны и отслеживание (Murphy-Chutorian, 2008)	$\pm 40^\circ$; $\pm 40^\circ$; $\pm 20^\circ$.	Высокое быстродействие; автоматическая инициализация	Чувствительность к взаимным перекрытиям	30
Гибридные методы: динамические шаблоны и отслеживание (Baltrusaitis, 2011)	$\pm 55^\circ$; $\pm 40^\circ$; $\pm 25^\circ$.	Высокое быстродействие; автоматическая инициализация	Чувствительность к взаимным перекрытиям	60
Гибридные методы: вероятностная модель и отслеживание (Cabrerera, 2015)	$\pm 40^\circ$; $\pm 30^\circ$; $\pm 20^\circ$.	Автоматическая инициализация; высокая точность	Малый диапазон углов; чувствительность к взаимным перекрытиям	30

Методы шаблонов обладают простотой реализации и возможностью обработки изображений низкого качества. Коллекция шаблонов может быть расширена. Недостаток подхода состоит в возможности распознавания только того положения головы, которое описано в базе эталонов. Если изображение текущего положения объекта имеет сходство сразу с несколькими шаблонами, система не сможет найти правильное решение. При экспериментах с большой базой эталонов описанный подход имеет высокую долю ошибок [84].

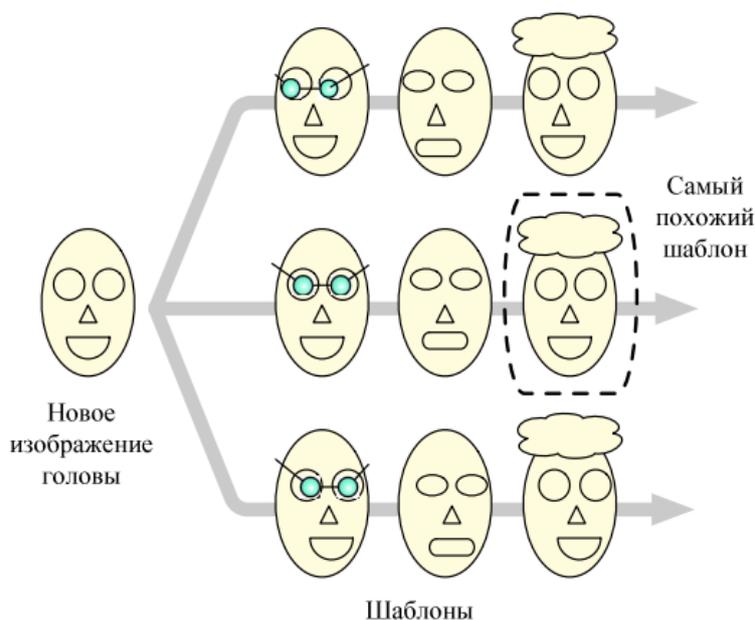


Рис. 1.5. Методы шаблонов

Методы массивов детекторов. В методах массивов детекторов используются наборы бинарных классификаторов [35, 61, 124, 150, 151] (рис. 1.6). Каждый классификатор настроен на распознавание одного положения головы. В процессе обучения для каждого положения головы используется множество изображений из базы данных [97]. Методы массивов детекторов могут быть реализованы с помощью нейронных сетей [130], сетей Байеса [52], скрытых цепей Маркова [162], опорных векторов [105, 137].

Методам на основе массивов детекторов не требуется отдельный этап локализации головы на изображении. Одновременное обнаружение и определение положения головы может быть выполнено для многих изображений. Данные методы могут работать с изображениями низкого и высокого разрешений. Недостатком методов является

ся необходимость обучения для каждого конкретного положения, что является сложным и долгим процессом. При этом вычислительные требования значительно увеличиваются с ростом числа детекторов, что делает данный подход неприменимым в системах реального времени.

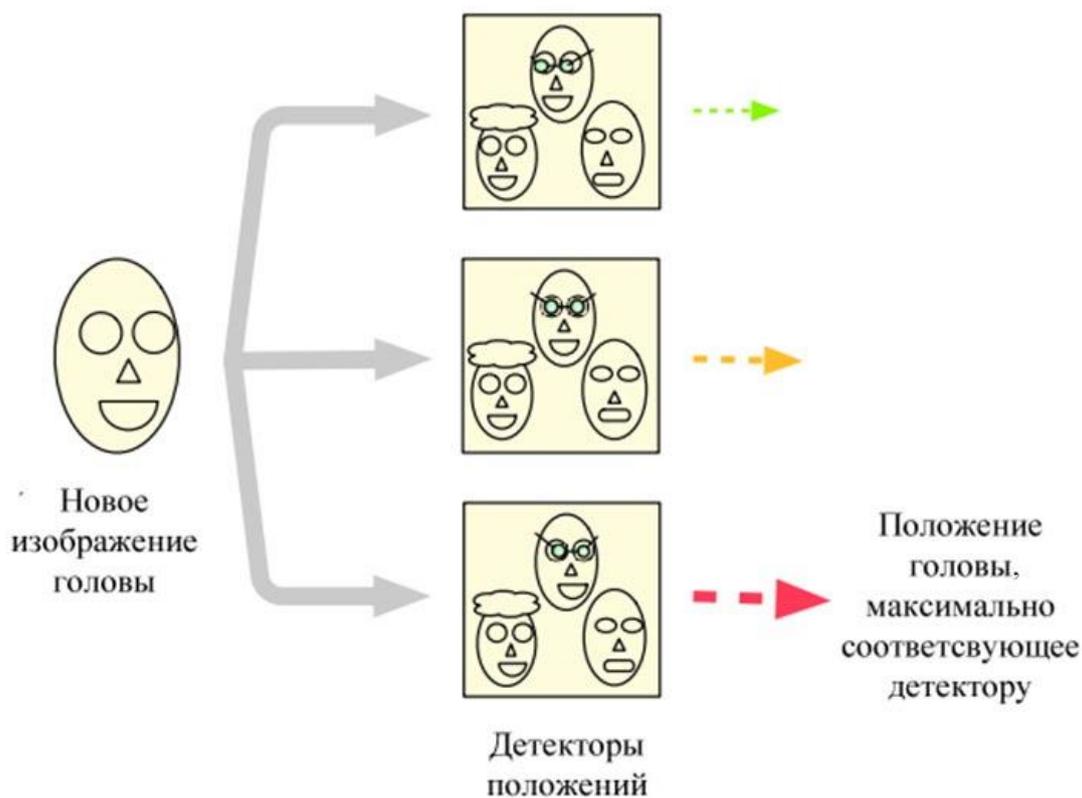


Рис. 1.6. Методы детекторов

Методы нелинейной регрессии. Идея, лежащая в основе методов нелинейной регрессии, состоит в том, чтобы связать вариации положения головы с векторным пространством, которое формируется по набору входных изображений (рис. 1.7). Полученная связь служит для обработки новых изображений, по которым можно определить положение и ориентацию головы [108, 112]. В подходах на основе нелинейной регрессии могут использоваться метод опорных векторов [105, 112] или нейронные сети [136]. При использовании нейронных сетей часто применяется многослойный персептрон [65, 146, 164].

Нейронные сети обладают высоким быстродействием. Для их обучения требуются маркированные изображения лиц. Нейронные сети дают самую точную оценку положения и ориентации, которая сильно зависит от этапа локализации головы на изображении. Так как метод основан на использовании конечного набора эталонов, то правильное вычисление возможно только для определенных положений головы. Производительность методов сильно зависит от размерности входных данных. Добавление нового эталонного положения головы требует полного переобучения всей сети.

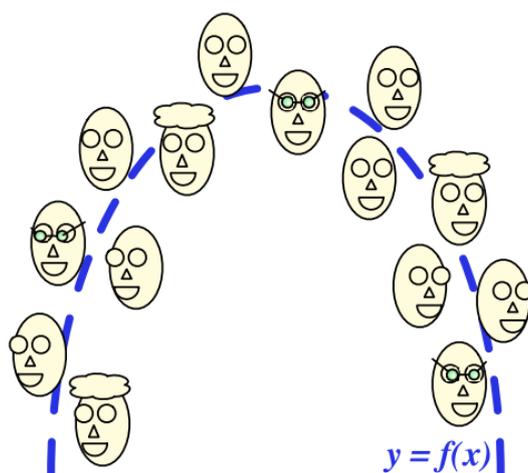


Рис. 1.7. Методы нелинейной регрессии

Методы вложенных многообразий. Методы вложенных многообразий рассматривают изображение головы как экземпляр данных в многомерном пространстве, ограниченном допустимыми вариациями положения [51, 144] (рис. 1.8). Для определения положения и ориентации головы на основе некоторых особенностей изображений формируется экземпляр данных, который проецируется в векторное пространство с использованием методов понижения размерности [52]. В качестве таких методов используются метод главных компонент (PCA – principal component analysis) [47], ядерный метод главных компонент (KPCA – kernel PCA) [109], локально-линейные вложения (LLE – Locally Linear Embedding) [59, 79, 83], анализ локальных вложений (LEA – Locally Embedded Analysis) [79], изокарты (Isomap) [92, 128].

При работе с изображениями головы методы вложенных многообразий показывают плохие результаты, так как они сильно зависят от геометрической структуры лица конкретного человека, а также от искажений при изменении положения головы.

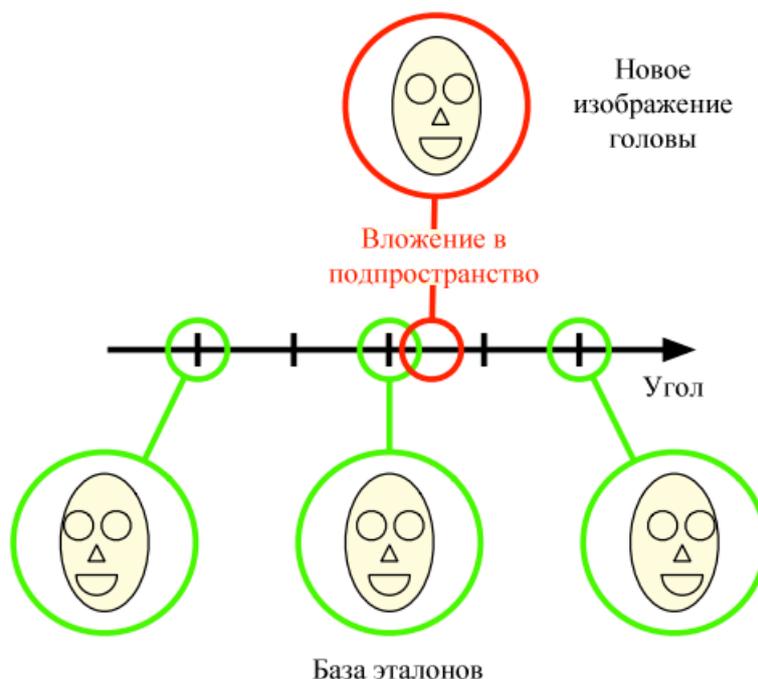


Рис. 1.8. Методы вложенных многообразий

Методы гибких моделей. Методы на основе гибких моделей описывают структуру лица с помощью графов (рис. 1.9). Для определения положения и ориентации головы анализируемое изображение накладывается на каждый шаблон. Используется пиксельная метрика для сопоставления изображений. Для обучения такой системы интерактивно указываются особые точки и вычисляются локальные дескрипторы [27, 32, 40]. Далее осуществляется сопоставление графа шаблона с графом, построенным по новому изображению лица. Находится минимальное расстояние между особенностями в каждом узле графа. Для определения положения головы создаются графы для каждого отдельного изображения. Граф с максимальным сходством свидетельствует о конкретной ориентации головы [103, 104]. Недостатком этого метода является то, что при определении положения и ориентации головы требуется сравнение множества графов. Кроме того, точное сравнение множества графов имеет высокую вычислительную сложность.

Часто используются активные модели формы (ASM – Active Shape Model) [33, 65, 71, 72, 158]. Если на изображении известно расположение особых точек, то задача может быть решена посредством проецирования координат в пространство меньшей размерности для

вычисления характеристик, соответствующих определенному состоянию. Известна также активная модель внешнего вида (ААМ – Active Appearance Model), которая формируется на основе базовых состояний. На основе детекторов изображения строится вектор особенностей. Если векторы вычислить для множества положений, то они могут использоваться для решения задачи [70, 165]. Модели ASM и ААМ показывают плохие результаты при возникновении взаимных перекрытий или быстрого вращения головы. Для повышения точности могут использоваться трехмерные активные модели внешнего вида [82]. Данный подход также чувствителен к взаимным перекрытиям. Этап обучения при создании ASM и ААМ является достаточно трудоемким и требует ручной разметки сотен и тысяч изображений. Главное ограничение гибких моделей состоит в том, что все лицевые точки должны присутствовать в каждом кадре видеопоследовательности. Данное условие делает подход слабо применимым на практике [155].

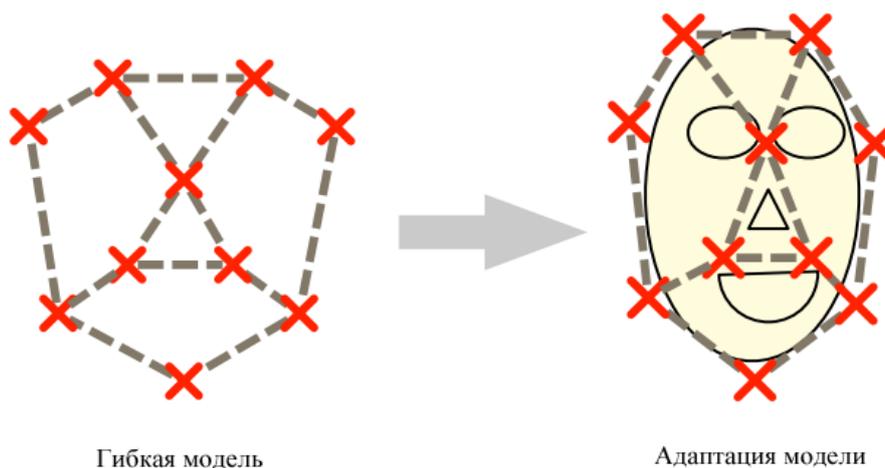


Рис. 1.9. Методы гибких моделей

Геометрические методы. В геометрических методах используется форма головы и данные точного расположения отдельных деталей: уголки глаз и рта, кончик носа (рис. 1.10). Используется ось лицевой симметрии. Направление лица может быть определено с помощью перспективной геометрии [24, 73, 85, 132]. Одним из стандартных подходов является использование шаблона для отслеживания основных частей лица. Определение положения головы на основе геометрических методов возможно с использованием проекции изобра-

жения головы человека [90, 153]. Положение головы можно определить в том случае, если заранее известно расстояние между глазами и уголками рта.

При сравнении с другими подходами геометрические методы обладают более высокой скоростью, так как требуется найти только несколько особенностей лица и провести ряд геометрических вычислений. Данные методы имеют высокую точность при небольших углах поворота головы. Однако вычисления основаны на малом количестве характеристик лица. Геометрические методы показывают плохие результаты в ситуациях, когда присутствуют взаимное перекрытие объектов и шумы. Ограничением геометрических подходов является то, что все отслеживаемые характеристики должны быть видны на всех кадрах видеопоследовательности, что ограничивает вычисления. Любая ошибка локализации особенностей приводит к ошибке определения положения и ориентации головы.

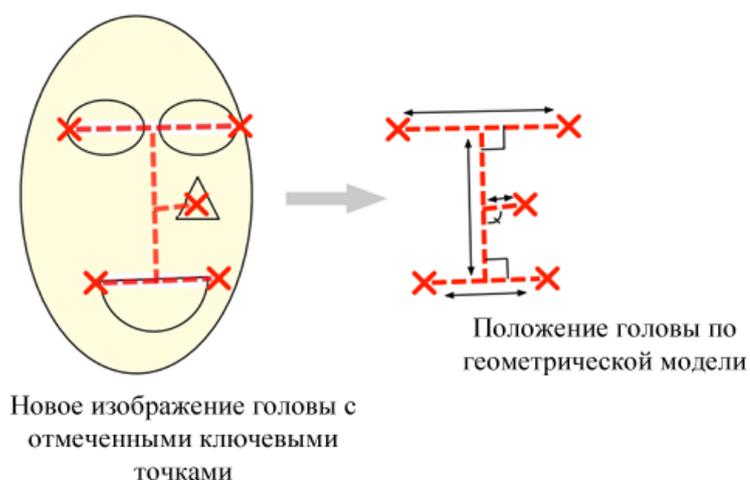


Рис. 1.10. Геометрические методы

Методы слежения. Методы слежения основаны на оценке относительного перемещения головы между последовательными кадрами видеоряда и установления соответствий между ними (рис. 1.11). Некоторые из этих методов обнаруживают особенности на каждом кадре и по ним вычисляют поворот и смещение головы [16, 17, 96, 123]. Методы слежения находят относительное смещение текущего положения головы от начального положения путем анализа изменений между кадрами видеопоследовательности [139, 152, 159, 163]. Ча-

сто используются также особенности, которые присутствуют на всех кадрах видеопоследовательности [2, 46, 134, 156]. Также могут применяться трехмерные модели, реконструируемые на основе стереоизображений [95, 154].

Многие методы слежения рассматривают голову человека как недеформируемый объект, что в действительности не совсем верно. Деформации, вызываемые эмоциями, жестами, различными движениями, могут приводить к ошибкам определения положения и ориентации. Методы слежения уязвимы к ошибке локализации начального положения головы. Чаще всего в методах слежения начальное положение определяется пользователем. Если положение головы на первом кадре определено неверно, то ошибки будут возникать на всех следующих кадрах. Также при использовании методов слежения может накапливаться ошибка из-за неправильной локализации особенностей [122]. На сегодняшний день методы слежения используются в основном только в составе гибридных методов.

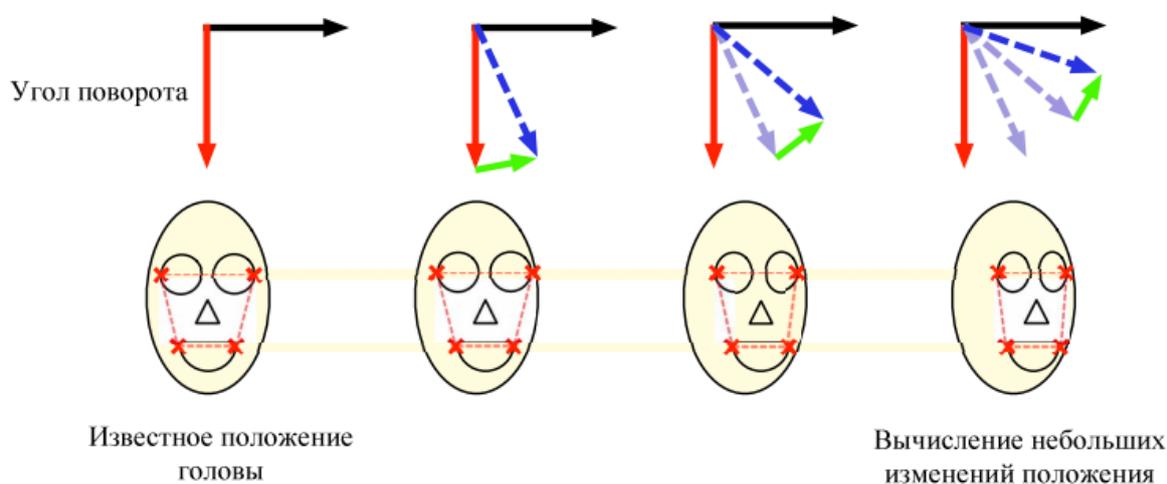


Рис. 1.11. Методы слежения

Гибридные методы. Гибридные методы комбинируют вышеперечисленные подходы, способствуя преодолению ограничений, возникающих при их использовании (рис. 1.12).

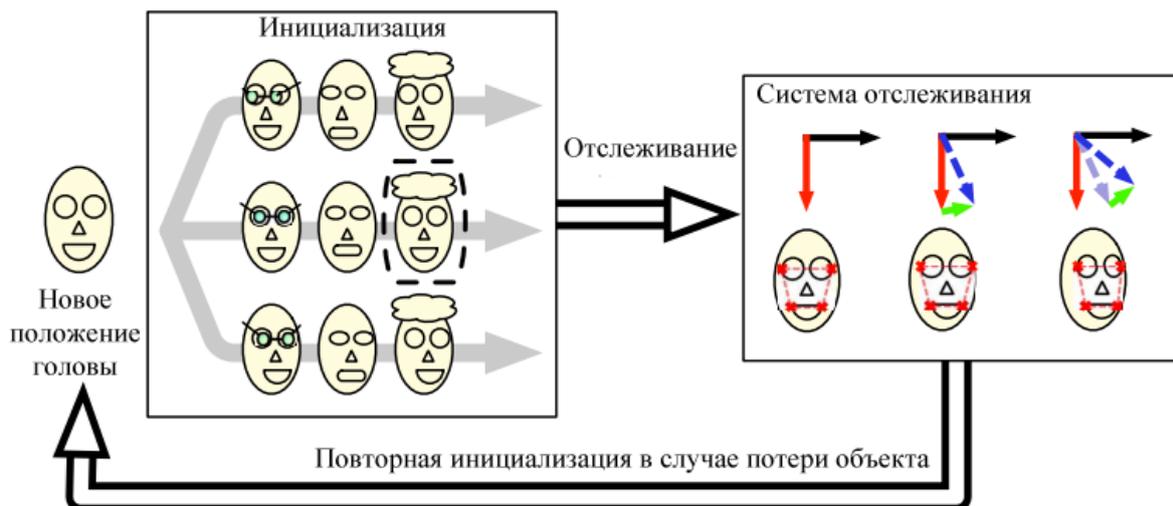


Рис. 1.12. Гибридные методы

В самой распространенной реализации статичная оценка положения головы дополняется методами слежения. Для начального определения положения могут использоваться геометрические методы или методы с использованием шаблонов. Этот этап необходим для начальной инициализации, а также для уточнения значений на протяжении работы всего алгоритма [26, 48, 49, 93, 114, 115, 119].

В гибридных подходах часто совместно используются геометрические методы и методы слежения [21, 22, 25, 89, 94]. Геометрические методы также могут комбинироваться с методами нелинейной регрессии [131], вычислением трехмерной деформируемой модели [66] или с нейронными сетями [120]. Нейронные сети могут применяться совместно с базой шаблонов для различных положений головы [138]. Также могут использоваться классические методы слежения совместно с подходами на основе нелинейной регрессии для уточнения модели [106, 118]. Существуют более сложные комбинации методов, включающие в себя фильтры Габора, скрытые цепи Маркова и сети Байеса [53]. Для отслеживания лица на протяжении всей видеопоследовательности может использоваться деформируемая модель вместе с нейронной сетью [102]. Этот метод показывает хорошие результаты для фронтальных изображений лица. Однако подход на основе деформируемой модели не может работать в ситуациях, при которых возникает взаимное перекрытие объектов сцены или осуществляется быстрое движение головы. Также использование нейронной сети делает алгоритм очень медленным.

Глава 2. РАЗРАБОТКА МОДЕЛИ И АЛГОРИТМОВ ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ И ОРИЕНТАЦИИ ГОЛОВЫ ЧЕЛОВЕКА ПО СТЕРЕОИЗОБРАЖЕНИЯМ С ИСПОЛЬЗОВАНИЕМ ГРАФОВ

2.1. Модель стереоустановки

В процессе определения положения и ориентации головы человека необходимо постоянно анализировать динамическую сцену. Для оценки положения и ориентации головы в пространстве нужно знать ее трехмерные координаты. Определение трехмерных координат наблюдаемых точек головы возможно при использовании двух камер. Для этого предлагается использовать стереоустановку [43]. Она состоит из двух камер, расположенных на одной горизонтальной оси.

Пусть имеются две камеры, заданные своими матрицами параметров P_l и P_r в некоторой системе координат. Для получения матриц необходимо произвести процесс калибровки [41, 45]. Система координат выбирается таким образом, чтобы начало координат совпадало с центром первой камеры, а ось Z была направлена вдоль ее оптической оси. В таком случае матрицы параметров будут иметь вид [86]:

$$\begin{aligned} P_l &= K_l [I/0]; \\ P_r &= K_r [R/tr], \end{aligned} \quad (2.1)$$

где K_l, K_r – матрицы внутренней калибровки соответствующих камер;

I – единичная матрица размером 3×3 ;

R – матрица поворота;

tr – вектор перемещения.

В ходе калибровки будут получены внутренние и внешние параметры камеры. Калибровка камер обычно выполняется за счет многократной съемки некоторого шаблона, на изображении которого можно легко выделить особые точки. Шаблон представляет собой набор чередующихся черных и белых ячеек. Далее составляются и решаются системы уравнений, связывающие координаты проекций, матрицы камер и положения точек шаблона в пространстве [41, 45].

На сегодняшний день существуют общедоступные реализации алгоритмов калибровки. При разработке алгоритмов определения по-

ложения и ориентации использовалась библиотека Matlab Calibration Toolbox.

Взаимосвязь между координатной системой изображения $O - xy$ и координатной системой камер $O - X_c Y_c Z_c$ описывается следующим образом:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z_c} K \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}, \quad (2.2)$$

где K – матрица внутренней калибровки.

Матрица внутренней калибровки будет иметь следующий вид:

$$K = \begin{bmatrix} f \cdot m_x & 0 & x_0 & 0 \\ 0 & f \cdot m_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (2.3)$$

где f – фокусное расстояние камеры;

m_x, m_y – ширина и высота пикселя камеры;

x_0, y_0 – координаты центра камеры.

В некоторых случаях искажения, вносимые линзами, могут сильно влиять на результат. Поэтому нелинейные параметры внутренней калибровки, такие как коэффициенты дисторсии, также имеют значение [43]. Они учитываются в функции применения дисторсии и представляются в виде вектора:

$$D = [k_1 \quad k_2 \quad p_1 \quad p_2 \quad k_3]^T, \quad (2.4)$$

где k_1, k_2, k_3 – коэффициенты радиальной дисторсии;

p_1, p_2 – коэффициенты тангенциальной дисторсии.

Большинство современных алгоритмов калибровки камер определяют коэффициенты вместе с матрицами внешней и внутренней калибровки [43, 45, 86].

Взаимосвязь между координатной системой камеры $O - X_c Y_c Z_c$ и глобальной системой координат $O - X_g Y_g Z_g$ описывается следующим выражением:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = K \times Dist \left(\begin{bmatrix} R & tr \\ 0_3^T & 1 \end{bmatrix} \times \begin{bmatrix} X_g \\ Y_g \\ Z_g \\ 1 \end{bmatrix} \right), \quad (2.5)$$

где K – матрица внутренней калибровки;

$Dist$ – функция вычисления дисторсии с учетом коэффициентов;

R – матрица поворота;

$\begin{bmatrix} R & tr \\ 0_3^T & 1 \end{bmatrix}$ – матрица внешней калибровки;

tr – вектор перемещения.

Таким образом, процесс калибровки стереоустановки состоит из следующих этапов:

Шаг 1. Обе камеры, входящие в стереоустановку, размещаются на одной горизонтальной оси.

Шаг 2. Калибровочный шаблон размещается в области обзора двух камер.

Шаг 3. Каждой камерой формируется последовательность снимков калибровочного шаблона, на которых выделяются ключевые точки.

Шаг 4. Вычисляются внутренние и внешние параметры камер, а также коэффициенты дисторсии.

Вычисление трехмерных координат особых точек головы основано на эпиполярной геометрии. На рисунке 2.1 изображена координатная система используемой стереоустановки. На изображении показаны центр левой камеры C_l и центр правой камеры C_r . Точка пространства X проецируется на плоскость изображения левой камеры в точку x_l и на плоскость изображения правой камеры в точку x_r . Луч $x_l X$ проецируется на плоскость изображения правой камеры в прямую, которая параллельна горизонтальной оси стереоустановки. Она называется эпиполярной линией. Каждой точке x_l на изображении левой камеры соответствует эпиполярная линия на изображении правой камеры. Аналогично каждой точке x_r на правом изображении соответствует эпиполярная линия на левом изображении [43, 45, 86].

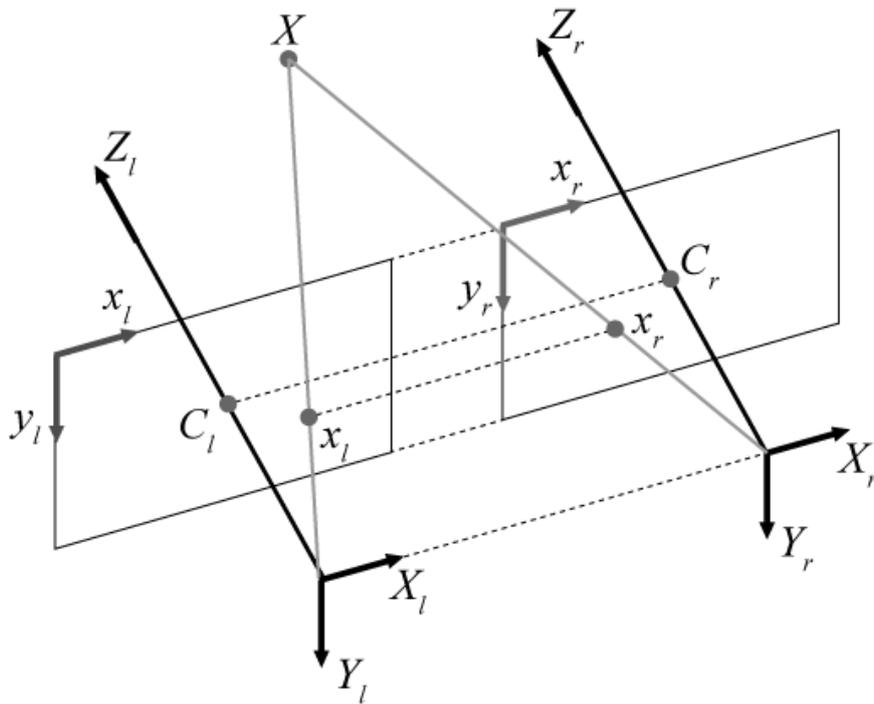


Рис. 2.1. Координатная система стереоустановки для определения положения и ориентации головы человека

С помощью эпилюлярной геометрии рассчитываются координаты точек в пространстве. Две точки, присутствующие на изображениях разных камер, будут являться стереопарой, когда верно следующее выражение:

$$x_l^T F x_r = 0, \quad (2.6)$$

где F – фундаментальная матрица;

x_l, x_r – однородные координаты точек на изображениях левой и правой камер соответственно.

Фундаментальная матрица зависит только от исходных матриц камер, находящихся в стереоустановке. На основе фундаментальной матрицы вычисляются уравнения эпилюлярных линий. Для точек x_l и x_r векторы, задающие эпилюлярную линию, будут иметь следующий вид:

$$\begin{cases} l_r = F x_l; \\ l_l = F^T x_r, \end{cases} \quad (2.7)$$

где l_l, l_r – векторы, эпилюлярных линии для левой и правой камер соответственно.

Для непосредственного определения трехмерных координат точек в стереоустановке проводится процесс триангуляции. Имея матрицы параметров откалиброванных камер P_l и P_r , а также однородные координаты проекций некоторой точки пространства X на плоскости левого и правого изображений стереоустановки, можно составить следующую систему уравнений:

$$\begin{cases} x_l = P_l X; \\ x_r = P_r X. \end{cases} \quad (2.8)$$

Результатом решения данной системы уравнений являются трехмерные координаты некоторой точки X в пространстве. В рассматриваемом случае, когда оптические оси камер являются параллельными, нахождение координат выглядит следующим образом:

$$\begin{cases} z = \frac{(bf)}{(x_l - x_r)}; \\ x = \frac{x_l z}{f}; \\ y = \frac{y_l z}{f}, \end{cases} \quad (2.9)$$

где z, x, y – трехмерные координаты точки;

b – расстояние между центрами камер;

f – фокусное расстояние камеры.

Для осуществления калибровки и триангуляции используются алгоритмы, описанные в [41, 45, 86].

2.2. Разработка модели определения положения и ориентации головы человека с использованием стереоизображений

Для определения положения и ориентации головы человека необходимо разработать математическую модель, описывающую вычисления в рамках предлагаемого подхода [10, 19].

Изменения положения и ориентации головы человека могут быть описаны с использованием шести степеней свободы (рис. 2.2). Вращение головы относительно осей X, Y, Z обозначаются терминами кивок, поворот и наклон. Возможны также перемещения головы относительно осей X, Y, Z .

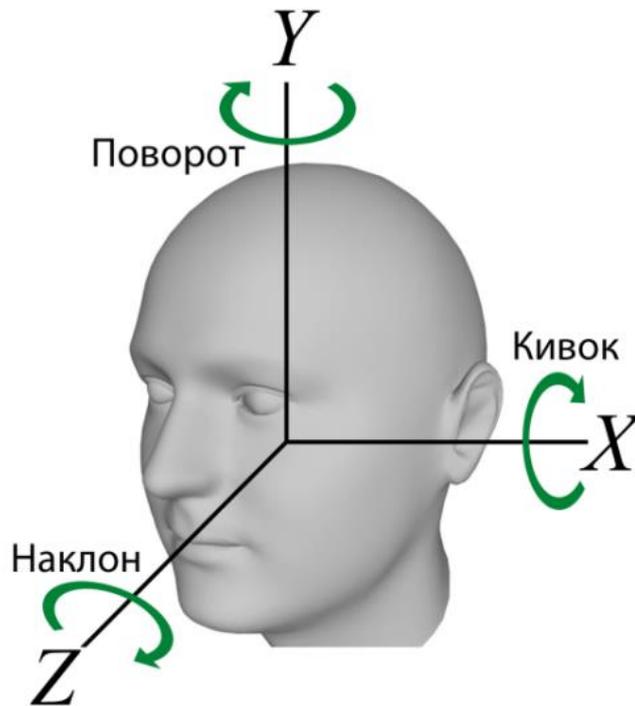


Рис. 2.2. Вращение головы относительно осей X , Y , Z :
кивок, поворот, наклон

Для отслеживания головы необходимо работать с особыми точками, выделенными на изображении. В модели подразумевается, что голова человека является эллиптическим недеформируемым твердым телом, а особые точки, которые отмечены на ней, остаются на своих местах в процессе движения. В таком случае углы кивка $\alpha_{\text{кивок}}$, поворота $\alpha_{\text{поворот}}$ и наклона $\alpha_{\text{наклон}}$ рассчитываются относительно осей глобальных координат X , Y , Z соответственно.

В модели принимается, что начальное положение головы соответствует анфасному. Также предполагается, что в начальном положении направление взгляда перпендикулярно плоскости, образованной осями X и Y .

Для того чтобы определить искомые углы, необходимо выделить ключевые точки на изображении головы человека. На данный момент выделять ключевые точки наиболее целесообразно с помощью таких методов, как SIFT, SURF [29, 56, 144]. Для определения положения головы используется следующий набор ключевых точек (рис. 2.3): крайние боковые точки (A, C); крайние верхняя и нижняя точки (B, D); центральная точка (E). В предлагаемой модели ориентация головы в пространстве определяется на основе этих точек.

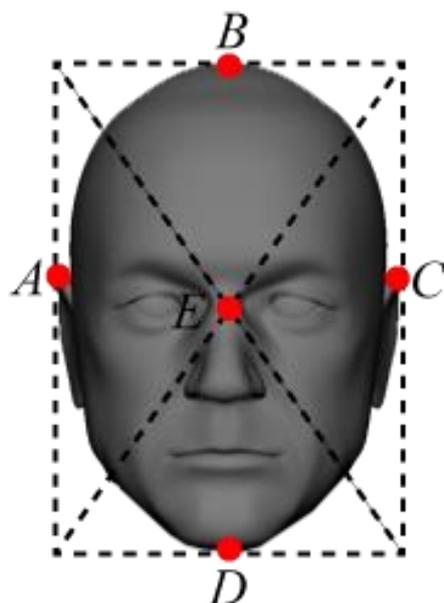


Рис. 2.3. Ключевые точки для определения положения и ориентации головы

Получение трехмерных координат ключевых точек возможно после проведения процесса калибровки, в ходе которого формируются матрицы параметров для камер в стереоустановке. Далее выполняется трехмерная реконструкция модели головы, осуществляемая при помощи алгоритма вычисления суммы абсолютных разностей (Sum of Absolute Differences, SAD) [68, 136]. Согласно результатам исследований точность данного алгоритма при трехмерной реконструкции головы достигает 86%. Также высокая скорость построения трехмерной модели позволяет использовать алгоритм SAD на практике [136]. Примеры проведенной реконструкции представлены на (рис. 2.4).



Рис. 2.4. Трехмерная реконструированная модель головы на основе данных стереоустановки

Определение углов кивка, поворота и наклона сводится к вычислению угла между векторами. Данные векторы образованы соответствующими ключевыми точками в начальном и последующих положениях.

Кивок является вращением головы вокруг оси X в плоскости, образованной осями Y и Z . В начальном положении имеется два вектора \overline{BE} и \overline{ED} . Кивок может осуществляться как вниз, так и вверх. В этом случае вычисляются векторы, характеризующие текущее положение головы во время кивка $\overline{B_c E_c}$ и $\overline{E_c D_c}$ (рис. 2.5).

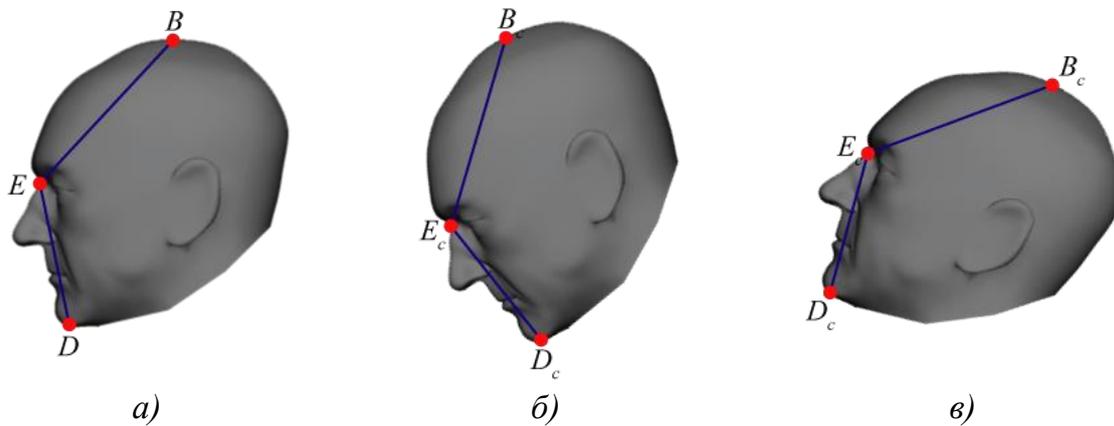


Рис. 2.5. Векторы, используемые при вычислении кивка: а – начальное положение; б – кивок вниз; в – кивок вверх

При кивке вниз необходимо найти угол между векторами \overline{BE} и $\overline{B_c E_c}$:

$$\cos \alpha = \frac{\overline{BE} \cdot \overline{B_c E_c}}{|\overline{BE}| |\overline{B_c E_c}|}, \quad (2.10)$$

где $\overline{BE} \cdot \overline{B_c E_c}$ – скалярное произведение векторов;

$|\overline{BE}|$, $|\overline{B_c E_c}|$ – длины векторов.

Таким образом, угол кивка вниз относительно начального положения вычисляется следующим образом (рис. 2.6, а):

$$\alpha_{\text{кивок}} = \arccos \frac{\overline{BE} \cdot \overline{B_c E_c}}{|\overline{BE}| |\overline{B_c E_c}|}. \quad (2.11)$$

Угол кивка вверх относительно начального положения (рис. 2.6, б):

$$\alpha_{\text{кивок}} = \arccos \frac{\overline{ED} \cdot \overline{E_c D_c}}{|\overline{ED}| |\overline{E_c D_c}|}. \quad (2.12)$$

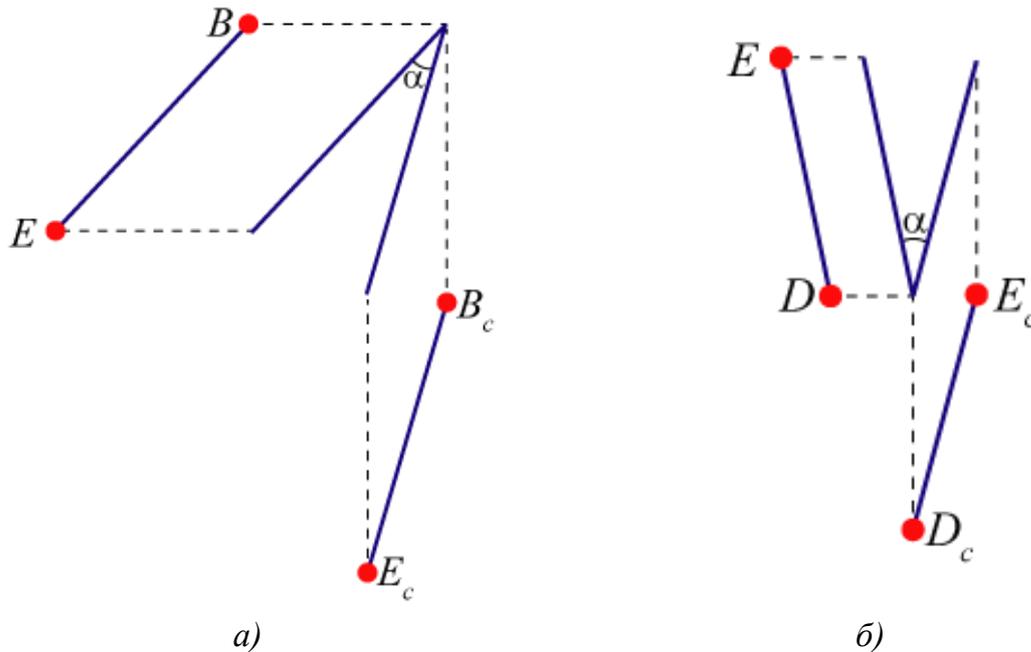


Рис. 2.6. Вычисление угла между векторами при кивке головы:
а – кивок вниз; б – кивок вверх

Углы при кивке вниз и кивке вверх могут быть найдены с помощью одинакового набора векторов. Таким образом, для любой точки, участвующей в рассмотрении, можно получить координаты в любой момент времени. Однако использование в расчетах двух формул позволит выполнить дополнительную проверку точности вычислений. Если разность между значениями углов, рассчитанных по формулам (2.11) и (2.12), превышает допустимую величину, то в вычислениях имеется ошибка. Поэтому текущие значения углов не используются в системе на данном этапе.

Поворот осуществляется при вращении головы относительно оси Y в плоскости, образованной осями X и Z . В начальном положении имеется два вектора \overline{AE} и \overline{EC} . Поворот может осуществляться в двух направлениях: вправо и влево. Отсюда получают векторы,

описывающее текущее положение головы во время поворота $\overline{A_c E_c}$ и $\overline{E_c C_c}$ (рис. 2.7).

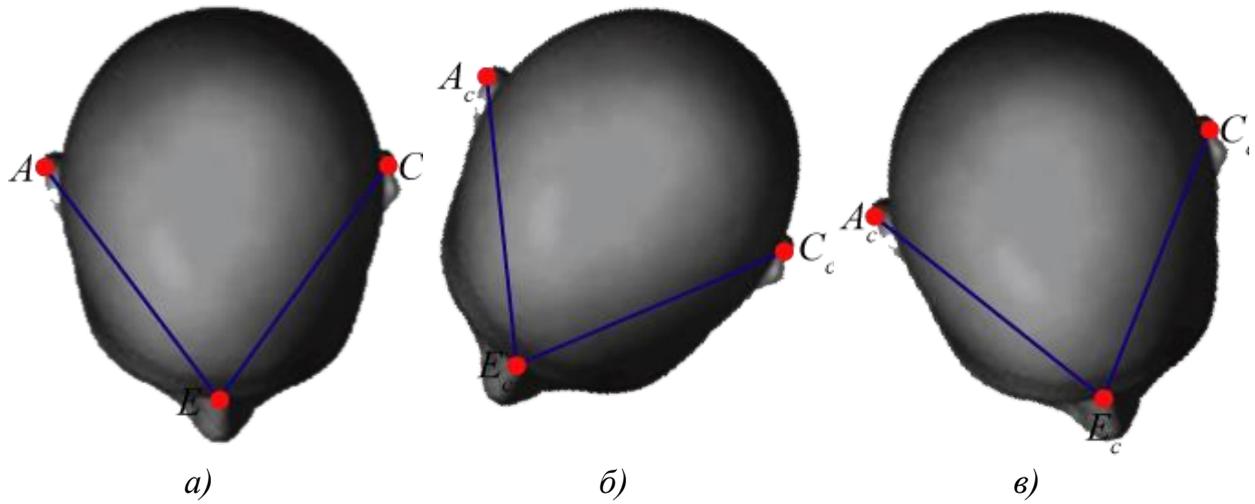


Рис. 2.7. Векторы, используемые при вычислении угла поворота: а – начальное положение; б – поворот вправо; в – поворот влево

Для определения угла поворота необходимо найти углы между рассматриваемыми векторами (рис. 2.8).

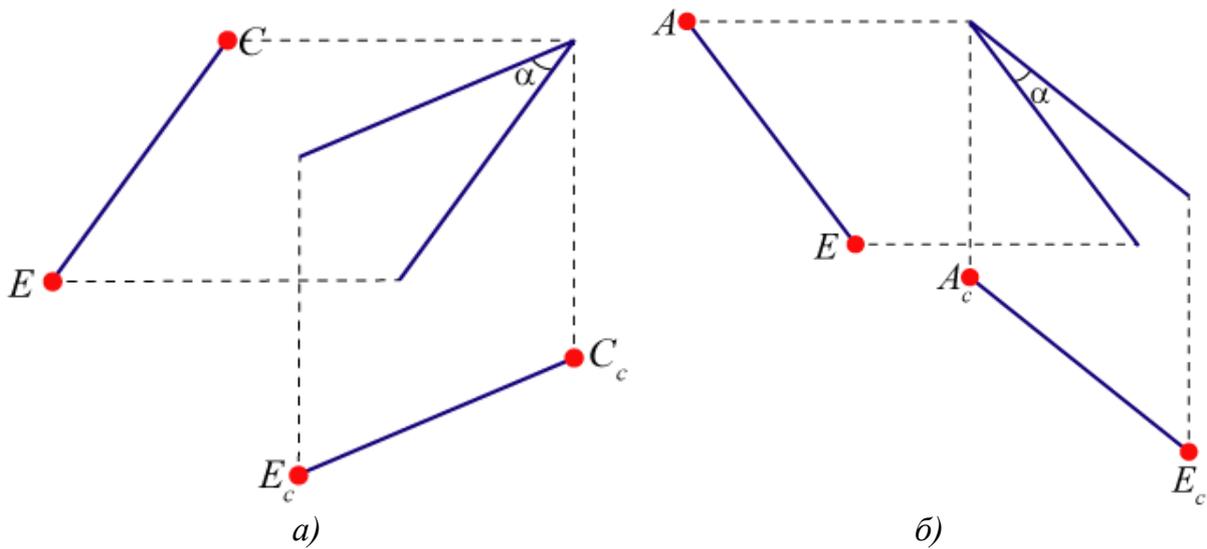


Рис. 2.8. Вычисление угла между векторами при повороте головы: а – поворот вправо; б – поворот влево

Угол поворота вправо относительно начального положения:

$$\alpha_{\text{поворот}} = \arccos \frac{\overline{CE} \cdot \overline{C_c E_c}}{|\overline{CE}| |\overline{C_c E_c}|}. \quad (2.13)$$

Угол поворота влево относительно начального положения:

$$\alpha_{\text{поворот}} = \arccos \frac{\overline{AE} \cdot \overline{A_c E_c}}{|\overline{AE}| |\overline{A_c E_c}|}. \quad (2.14)$$

Совместное использование формул (2.13) и (2.14), как и при нахождении угла кивка, позволяет выявить возможные ошибки.

Наклон осуществляется при вращении головы вокруг оси Z в плоскости, образованной осями X и Y .

Для вычислений используется вектор \overline{AC} , характеризующий текущее положение головы $\overline{A_c C_c}$ (рис. 2.9).

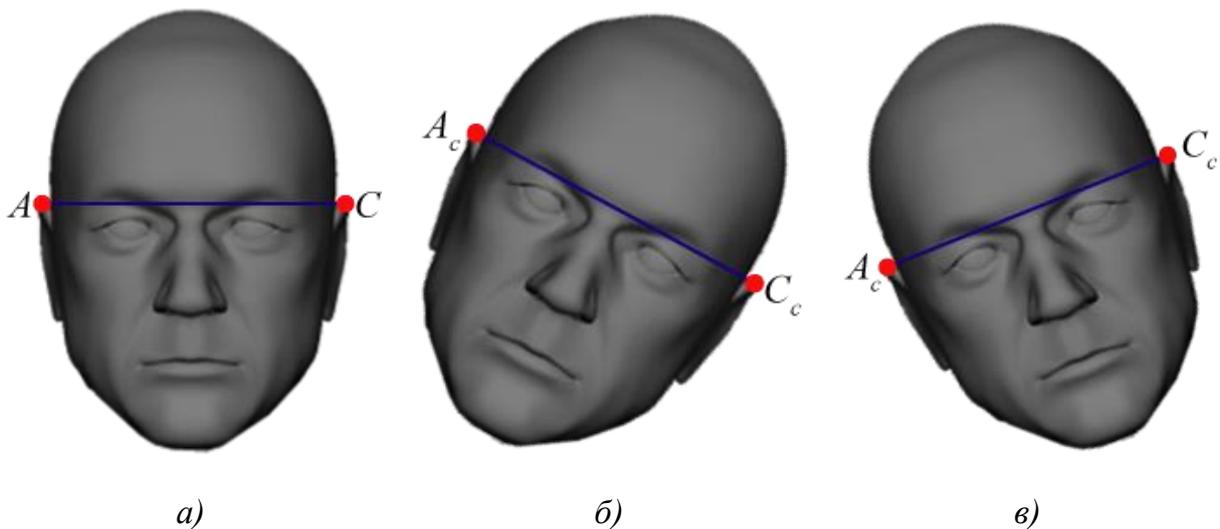


Рис. 2.9. Вычисление угла наклона:
а – начальное положение; *б* – наклон влево; *в* – наклон вправо

Для определения угла наклона необходимо найти углы между рассматриваемыми векторами (рис. 2.10).

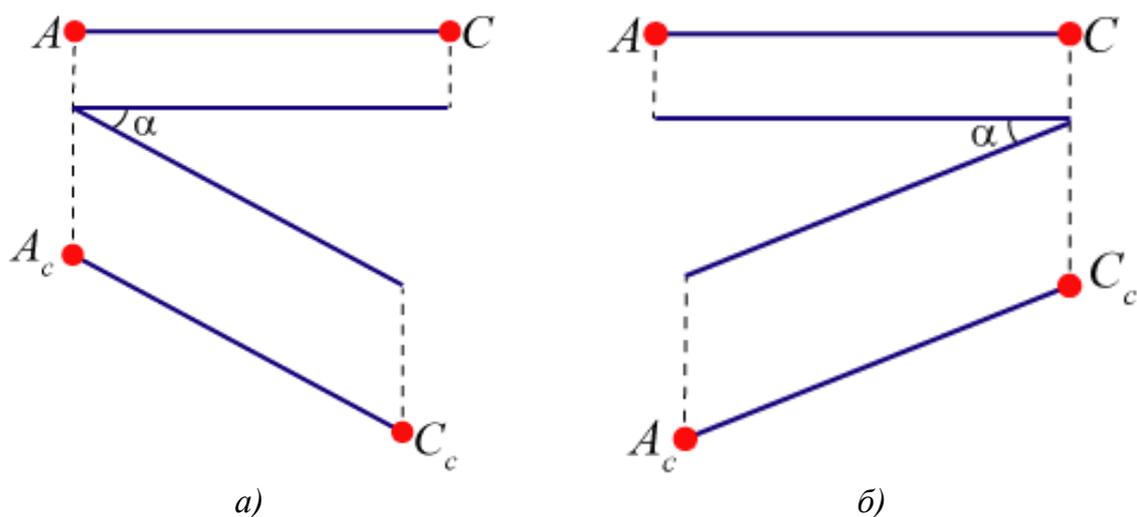


Рис. 2.10. Вычисление угла между векторами при наклоне головы:
а – наклон влево; *б* – наклон вправо

Таким образом, угол наклона влево и вправо относительно начального положения вычисляется следующим образом:

$$\alpha_{\text{наклон}} = \arccos \frac{\overline{AC} \cdot \overline{A_c C_c}}{|\overline{AC}| |\overline{A_c C_c}|}. \quad (2.15)$$

Для определения того, в какую сторону происходит наклон головы, сравнивается координата по оси Y для начальной и конечной точек использующихся векторов.

Модель определения положения и ориентации головы человека является моделью действий и состоит из следующих шагов: инициализация изображения головы путем кластеризации особенностей на основе спектральной теории графов и ограничений; отслеживание головы на изображениях путем нахождения соответствий с использованием тепловых ядер на графах; обнаружение головы на изображении на основе вложения графа особенностей в векторное пространство; вычисление углов кивка, поворота и наклона [5, 7, 9, 54-57, 160, 161].

В основе модели определения положения и ориентации головы человека лежит идея отслеживания особых точек на всей видеопоследовательности. Особые точки выделяются на всем изображении головы. Однако на кадрах видеопоследовательности предлагается сопоставлять наиболее стабильные особенности. Это позволит сократить количество ложных соответствий и повысит точность определения положения и ориентации головы человека [5, 7, 9, 54-57, 160, 161].

Предлагается использовать структурные методы компьютерного зрения с использованием спектральной теории графов [44]. Преимуществом структурных подходов является то, что они позволяют анализировать большое множество элементов на основе малого количества простых составляющих и правил формирования графической модели. Также структурные методы позволяют описать те характеристики объекта, которые исключают его отнесение к другому классу, что повышает надежность распознавания [34, 42].

Наблюдение за динамической сценой будет осуществляться с помощью стереоустановки, включающей две оптические камеры. Для реализации алгоритмов определения положения и ориентации головы человека предлагается использовать графы, в которых вершины представлены особыми точками.

Для определения положения и ориентации головы, используются предложенные выше выражения. Трехмерные координаты ключевых точек рассчитываются на основе внешних и внутренних параметров заранее откалиброванной стереоустановки.

Для реализации общего алгоритма определения положения и ориентации головы необходимо выполнить следующие действия:

- инициализация изображения головы. Вычисление структурного дескриптора для последующего обнаружения головы на изображении;
- отслеживание головы по видеопоследовательности. Расчет углов кивка, поворота и наклона головы;
- обнаружение головы на изображении с использованием дескриптора.

Блок-схема общего алгоритма определения положения и ориентации головы человека в пространстве представлена рисунке 2.11.

Инициализация изображения головы. Дескриптор представляет собой метод, который идентифицирует некоторую область изображения на основе набора признаков. Наиболее известными дескрипторами изображений являются SIFT, SURF, BRIEF, BRISK и др. Недостатком подобных методов является то, что они позволяют идентифицировать только отдельную область, определяемую несколькими пикселями. Предлагается разработать структурный дескриптор, который идентифицирует область, состоящую из десятков и сотен пикселей. Такая область сопоставима с изображением головы.

Для выделения головы на изображении предлагается оценивать спектральные характеристики графа, узлами которого являются особые точки. Для вычисления дескриптора необходимо выделить особые точки на изображении в процессе инициализации.



Рис. 2.11. Блок-схема общего алгоритма определения положения и ориентации головы человека в пространстве

Чтобы выделить область изображения, в которой содержится голова человека, необходимо произвести кластеризацию особенностей. На данном этапе изображение сцены представляется в виде взвешенного графа, где в качестве вершин выступают особые точки. Вершины графа соединены ребрами, веса которых рассчитываются на основе гауссовой функции.

Так как осуществляется поиск головы человека, то можно использовать априорную информацию для задания ограничений при кластеризации особенностей. Таким образом, выделение области головы на изображении осуществляется путем кластеризации особенностей и наложения условий, основанных на априорной информации.

Отслеживание головы по видеопоследовательности. Отслеживание головы осуществляется по видеопоследовательности. Этот процесс необходим, чтобы получить положение и ориентацию головы человека в любой момент времени. Так как получение видеопоследовательности осуществляется с частотой 25 – 60 кадров/сек, то перемещение головы между кадрами будет небольшим.

Таким образом, при отслеживании происходит сопоставление особенностей области головы, полученных на этапе инициализации, и особенностей, выделенных в текущем кадре видеопоследовательности.

При сопоставлении особенностей могут быть найдены ложные соответствия. Для уменьшения количества ложных соответствий предлагается сопоставлять только самые стабильные области изображения. Предполагается, что такие особенности будут оставаться при значительном изменении ракурса головы. Для их выделения предлагается использовать тепловые ядра на графах.

Обнаружение головы на изображении с использованием дескриптора. На протяжении всей работы системы определения положения и ориентации головы человека возможны ситуации, когда объект пропадет из поля зрения. Причиной этого могут быть следующие события: объект отслеживания покинул область видимости, возникли взаимные перекрытия объектов в сцене. В этом случае необходимо найти область головы на изображении на основе дескриптора, вычисленного при инициализации.

Для сравнения изображения-эталона с текущим изображением предлагается оценивать спектральные характеристики графа. При этом необходимо осуществить проецирование характеристик графа в пространство меньшей размерности. В результате этого можно получить векторы признаков, которые можно сравнивать. Таким образом, если в процессе отслеживания голова пропадает в поле зрения, то выполняется процесс поиска. Поиск будет вестись до тех пор, пока текущий дескриптор головы не будет найден на изображениях. Если область головы найдена, то возобновляется отслеживание движения головы.

2.3. Разработка алгоритма кластеризации особенностей на основе спектральной теории графов и ограничений для выделения лица человека на изображениях

Выделение лица человека на видеопоследовательности является важным этапом при решении задачи определения положения и ориентации головы в пространстве. В предлагаемом подходе на этапе инициализации выделяется область лица человека.

Несмотря на наличие множества исследований в области выделения лиц на изображениях, до сих пор остаются нерешенными проблемы, связанные с низкой скоростью обработки, с недостаточной точностью, слабой устойчивостью к шумовым и оптическим эффектам. Для преодоления проблем предлагается использовать структурный подход, в основе которого лежит применение спектральной теории графов и ограничений при кластеризации.

На вход алгоритма поступает изображение сцены, на котором выделяются особенности. Объединение особенностей в граф, позволит проанализировать структуру сцены и точнее обнаружить искомый объект. Предлагается использовать кластеризацию на основе спектральной теории графов. Таким образом, будут проанализированы не отдельные пиксели изображения, а структура особенностей, представленная в виде графа. Так как сцена характеризуется тем, что может содержать большое количество разнообразных объектов, появляющихся или исчезающих на протяжении всей видеопоследовательности, то не всегда результаты кластеризации будут корректными. Для повышения точности кластеризации особенностей лица человека на

изображениях предлагается использовать ограничивающие условия. В алгоритме применяется априорная информация о форме, пропорциях и взаимном расположении частей головы человека [6, 55].

Используется метод на основе случайного обхода графа. Случайный обход графа основан на теории марковских цепей, где переход из одного состояния в другое обусловлен вероятностью. Случайный обход – процесс случайных изменений, происходящих в определенные моменты времени на графе. Пусть имеется ненаправленный граф, который не содержит циклов. Даны вершина, из которой начинается путь (начало пути), и вершина, в которой путь должен быть завершен (конец пути). Путь от начала до конца будет проходить через выбранные случайным образом смежные вершины. Общее количество проделанных таким образом шагов будет случайным обходом графа. Вероятность перехода из одной вершины в другую зависит от весов ребер и описывается матрицей состояний [107].

Случайный обход графа обладает такими важными свойствами, как время прохождения случайного пути и время обхода [107, 110]. Время прохождения случайного пути $H(A,B)$ между вершинами A и B определяется количеством вершин, которые будут случайно пройдены, если путь начнется из вершины A и закончится в вершине B . В процессе обхода уже посещенные вершины могут быть пройдены вновь. Это зависит от вероятности перехода.

Время обхода графа $CT(A,B)$ (Commutе Time) между вершинами A и B равно количеству вершин, которые случайно будут пройдены, если путь начнется из вершины A , пройдет через вершину B и снова закончится в вершине A .

Информация обо всех случайных путях, которые могут быть в графе, содержится в матрице времени обхода. Важным моментом при анализе данной матрицы является тот факт, что у вершин, которые расположены рядом, время обхода гораздо меньше, чем у вершин, расположенных на значительном удалении друг от друга [107]. Таким образом, матрица времени обхода содержит значения времени обхода между любой парой вершин графа и представляет информацию о кластерной структуре графа. Эта матрица обладает следующими важными свойствами, повышающими надежность кластеризации: устойчивость к перестановке, удалению и добавлению вершин графа [107].

Матрица времени обхода вычисляется по спектру нормализованной матрицы Лапласа. Матрица Лапласа является одним из представлений графа и может использоваться для нахождения многих его важных свойств.

Структура графа может быть охарактеризована с помощью собственных векторов и собственных значений матрицы. Матрица Лапласа рассчитывается путем вычитания матрицы смежности из матрицы степеней графа. В случае, когда нужно охарактеризовать случайный обход по графу, используется нормализованная версия матрицы Лапласа [127].

Совокупность собственных значений матрицы, отсортированных в порядке убывания или возрастания, является спектром. Спектр матрицы играет важную роль в теории графов, так как позволяет вычислить индивидуальные характеристики графа [44, 133, 135]. При расчете матрицы времени обхода также применяется дискретная функция Грина. Она содержит информацию о среднем количестве посещений какой-либо вершины во время случайного обхода [69].

Для получения более точных результатов при кластеризации предлагается использовать ограничивающие условия. В этом случае пропорции головы используются в качестве априорной информации, которая накладывает ограничения на результаты кластеризации. Так как известна информация о форме головы и взаимном расположении частей лица, то сделано предположение, что особые точки на голове будут присутствовать в следующих местах: по контуру, в области глаз, рта, носа, ушей. Также в качестве ограничивающих условий должны быть учтены пропорции головы. Таким образом, для ограничения результатов кластеризации предлагается использовать прямоугольную сетку, где для каждой ячейки будет задано условие возможного наличия или отсутствия особых точек. Сетка накладывается на изображение кластера-кандидата, чтобы принять окончательное решение о его принадлежности к области головы на изображении.

Так как разрабатываемый алгоритм решает проблему выделения области головы на этапе инициализации, то на вход поступает изображение сцены, полученное с одной из камер стереосистемы. С помощью метода SURF (Speeded Up Robust Features) [58] выделяются особые точки изображения. Предлагается на основе выделенных особенностей построить взвешенный граф. В этом графе каждая вершина

соответствует одной из полученных особых точек, а каждая пара вершин соединена ребром, вес которого равен евклидовому расстоянию между этими точками. Матрица смежности рассчитывается с помощью метрики Гаусса и имеет следующий вид:

$$W_{ij} = e^{-\frac{\|x_i - x_j\|^2}{2\sigma}}, \quad (2.16)$$

где x_i, x_j – особые точки сцены, между которыми рассчитывается вес ребра;

σ – параметр, который регулирует чувствительность метрики Гаусса.

Параметр σ позволяет описать степень близости между точками. Параметр σ должен выбираться в соответствии с максимальным размером изображения, на котором выделяются особые точки.

На основе взвешенной матрицы смежности вычисляется матрица степеней графа D . Она является диагональной матрицей, в которой по главной диагонали расположены степени вершин графа, а остальные элементы заполняются нулями. В соответствии с матрицей степеней графа вычисляется сумма степеней вершин:

$$vol = \sum_{v \in V} d_v, \quad (2.17)$$

где d_v – степень соответствующей вершины графа.

Для расчета случайного обхода графа и получения матрицы времени обхода используется нормализованная матрица Лапласа. Так как применяется взвешенный граф, то нормализованная матрица Лапласа имеет следующий вид:

$$L_{nw} = \begin{cases} 1, & \text{если } u = v; \\ -\frac{w(u,v)}{\sqrt{d_u d_v}}, & \text{если } u \neq v; \\ 0, & \text{в другом случае,} \end{cases} \quad (2.18)$$

где $w(u,v)$ – вес ребра, которое соединяет вершины u и v ;

d_u, d_v – степени вершин u и v соответственно.

Для вычисления собственных значений и собственных векторов нормализованной матрицы Лапласа необходимо провести сингулярное разложение матрицы:

$$L_{nw} = \Phi \Lambda \Phi^T, \quad (2.19)$$

где L – матрица собственных значений $\lambda_1, \lambda_2, \dots, \lambda_{|V|}$;

Φ – матрица собственных векторов $\phi_1, \phi_2, \dots, \phi_{|V|}$.

Для представления матрицы времени обхода используется функция Грина. Она связана с матрицей Лапласа и характеризует связи между вершинами графа.

На основе нормализованной матрицы Лапласа и ее спектра с помощью функции Грина строится матрица:

$$G_n(u, v) = \sum_{i=2}^{|V|} \frac{1}{\lambda_i} \phi_i(u) \phi_i(v), \quad (2.20)$$

где λ_i – собственные значения нормализованной матрицы Лапласа L_n ;

ϕ_i – собственные векторы нормализованной матрицы Лапласа L_n .

Пусть $H(u, v)$ – время прохождения случайного пути между вершинами u и v . Тогда время обхода будет иметь вид $CT(u, v) = H(u, v) + H(v, u)$. Матрица времени обхода может быть рассчитана с использованием матрицы Грина (2.20) и примет следующий вид:

$$CT(u, v) = vol \sum_{i=2}^{|V|} \frac{1}{\lambda_i} \left(\frac{\phi_i(u)}{\sqrt{d_u}} - \frac{\phi_i(v)}{\sqrt{d_v}} \right)^2. \quad (2.21)$$

Элементами этой матрицы являются значения времени обхода между каждой парой вершин. Далее осуществляется спектральная декомпозиция матрицы времени обхода. Элементы собственного вектора $CT\phi_0$, соответствующего минимальному собственному значению этой матрицы, содержат информацию о кластерной структуре графа. Каждое значение этого вектора характеризует соответствующую точку исходного пространства данных. Отнесение точек к различным кластерам происходит согласно следующему условию [107]:

$$Point\ s(v) \in \begin{cases} Cluster\ 1, & CT\phi_0(v) \geq 0; \\ Cluster\ 2, & \text{в другом случае.} \end{cases} \quad (2.22)$$

Для управления процессом кластеризации вводятся ограничивающие условия, основанные на априорной информации о пропорциях

голова человека. Предлагается использовать сетку пропорций головы. Данное понятие заимствовано из искусства для изображения головы человека. В ходе исследований в данной области, было установлено, что независимо от пола и расы пропорции лиц различных людей одинаковы [28, 73]. В этом случае рассматривается взаимное расположение различных частей головы человека: носа, рта, глаз, ушей.

Для использования сетки пропорций при кластеризации и последующего выделения головы человека необходимо рассчитать усредненный шаблон. Сама сетка состоит из 7 строк и 5 столбцов. Это обусловлено выбором графических методов, которые применяются в изобразительном искусстве при изображении головы человека [73]. Было отобрано 500 изображений лиц различных людей. На этих изображениях были выделены особые точки. Для каждого лица на группу особых точек накладывалась сетка пропорций. Затем вычислялось количество особых точек, находящихся в каждой ячейке. Далее эти 500 измерений были объединены, чтобы получить среднее количество точек по всем ячейкам сетки пропорций. Пример выделенных особенностей и наложенных на них сеток пропорций представлен на рис. 2.12, а. Результирующий усредненный шаблон приведен на рис. 2.12, б.

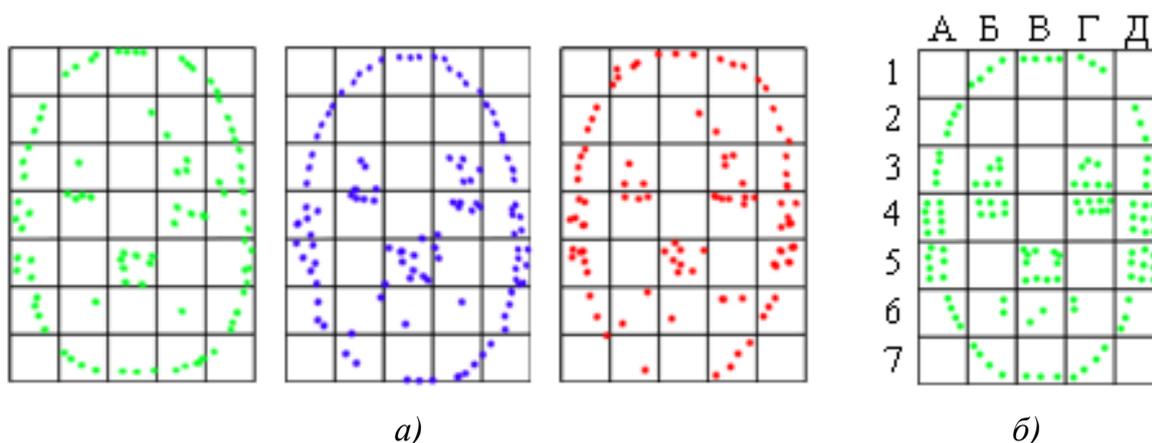


Рис. 2.12. Сетка пропорций для выделения особенностей, принадлежащих изображению головы человека: а – примеры выделения особенностей на различных лицах; б – усредненный шаблон сетки пропорций

Чтобы принять решение о том, какой кластер является областью изображения головы, необходимо воспользоваться априорной информацией. В разрабатываемом подходе должны выполняться следующие условия.

Изображение лица может быть ограничено прямоугольником. Отношение ширины к высоте в сетке пропорций равно $\frac{5}{7}$. Тогда условие, ограничивающее размер кластера на основе пропорций лица имеет вид:

$$Cluster = \begin{cases} true, & \text{если } \frac{5}{7} - \delta \leq \frac{W_{cluster}}{L_{cluster}} \leq \frac{5}{7} + \delta; \\ false, & \text{в другом случае,} \end{cases} \quad (2.23)$$

где $W_{cluster}$, $L_{cluster}$ – ширина и длина кластера;

δ – порог отклонения, равный 0,1.

Опираясь на априорную информацию о возможной форме головы, можно сказать, что точки кластера не могут принадлежать к угловым ячейкам сетки пропорций. Если при наложении сетки пропорций в кластере все-таки будут присутствовать особые точки в соответствующих угловых ячейках, значит это точки заднего плана, которые должны быть исключены из рассмотрения.

Условие вхождения особенностей $ClusterPoint s(n)$ в угловые ячейки сетки пропорций для задания кластера $Cluster$ имеет вид:

$$Cluster = \begin{cases} true, & \text{если } ClusterPoint s(n) \notin \{Grid_{A,1}, Grid_{D,1}, Grid_{A,7}, Grid_{D,7}\}; \\ false, & \text{в другом случае.} \end{cases} \quad (2.24)$$

Так как речь идет об изображении головы человека, то в ячейках, соответствующих областям глаз, рта, ушей и носа, особые точки должны обязательно присутствовать. Пусть множество ячеек сетки пропорций, в которых присутствуют особенности глаз, рта, ушей и носа имеет вид:

$$FF = \left\{ \begin{array}{l} Grid_{A,4}, Grid_{A,5}, Grid_{B,3}, Grid_{B,4}, Grid_{B,6}, Grid_{B,5}, \\ Grid_{B,6}, Grid_{\Gamma,3}, Grid_{\Gamma,4}, Grid_{\Gamma,6}, Grid_{D,4}, Grid_{D,5} \end{array} \right\}. \quad (2.25)$$

Условие вхождения особенностей $ClusterPoint s(n)$ в ячейки, относящиеся к областям глаз, рта, ушей, носа:

$$Cluster = \begin{cases} true, & \text{если } ClusterPoint s(n) \in FF; \\ false, & \text{в другом случае.} \end{cases} \quad (2.26)$$

Также предполагается, что в большинстве случаев особые точки будут выделены по всему контуру головы на изображении. Таким образом, будет выделена форма головы. Контур лица на изображении описывается множеством ячеек:

$$CF = \left\{ \begin{array}{l} Grid_{A,2}, Grid_{A,3}, Grid_{A,4}, Grid_{A,5}, \\ Grid_{A,6}, Grid_{B,1}, Grid_{B,7}, Grid_{B,1}, \\ Grid_{B,7}, Grid_{\Gamma,1}, Grid_{\Gamma,7}, Grid_{D,2}, \\ Grid_{D,3}, Grid_{D,4}, Grid_{D,5}, Grid_{D,6} \end{array} \right\}. \quad (2.27)$$

Следовательно, условие вхождения особенностей контура лица в соответствующие ячейки имеет вид:

$$Cluster = \begin{cases} true, & \text{если } ClusterPoint s(n) \in CF; \\ false, & \text{в другом случае.} \end{cases} \quad (2.28)$$

Описанные выше условия позволят более точно отнести группу точек, находящихся в кластере-кандидате, к голове пользователя.

Таким образом, учитывая описанные выше выражения для нахождения матрицы времени обхода, выделения кластерной структуры и наложения ограничений, алгоритм кластеризации особенностей на основе спектральной теории графов и ограничений для выделения лица человека на изображениях выглядит следующим образом.

Шаг 1. На изображении сцены выделяются особые точки с помощью метода SURF. Имеющийся набор особенностей описывается с помощью взвешенного графа $G=(V,E)$, где V – множество вершин, E – множество ребер. В этом графе все вершины, соответствующие особым точкам, соединены ребрами, вес которых определяется евклидовым расстоянием. Рассчитывается матрица смежности W (2.16).

Шаг 2. Вычисляется матрица степеней D и сумма степеней графа (2.17).

Шаг 3. Строится нормализованная матрица Лапласа (2.18). Вычисляется ее спектр с помощью сингулярного разложения (2.19).

Шаг 4. Рассчитывается матрица Грина (2.20). Строится матрица времени обхода (2.21).

Шаг 5. Находится минимальное собственное значение матрицы времени обхода и соответствующий ему собственный вектор. Происходит разбиение исходного набора точек на кластеры (2.22).

Шаг 6. Происходит наложение сетки пропорций. Проверяются условия, описываемые выражениями (2.23 – 2.28). Если ни один кластер не удовлетворяет условиям, происходит рекурсивное деление кластеров до достижения минимально допустимого размера.

Пример работы алгоритма изображен на рис. 2.13.

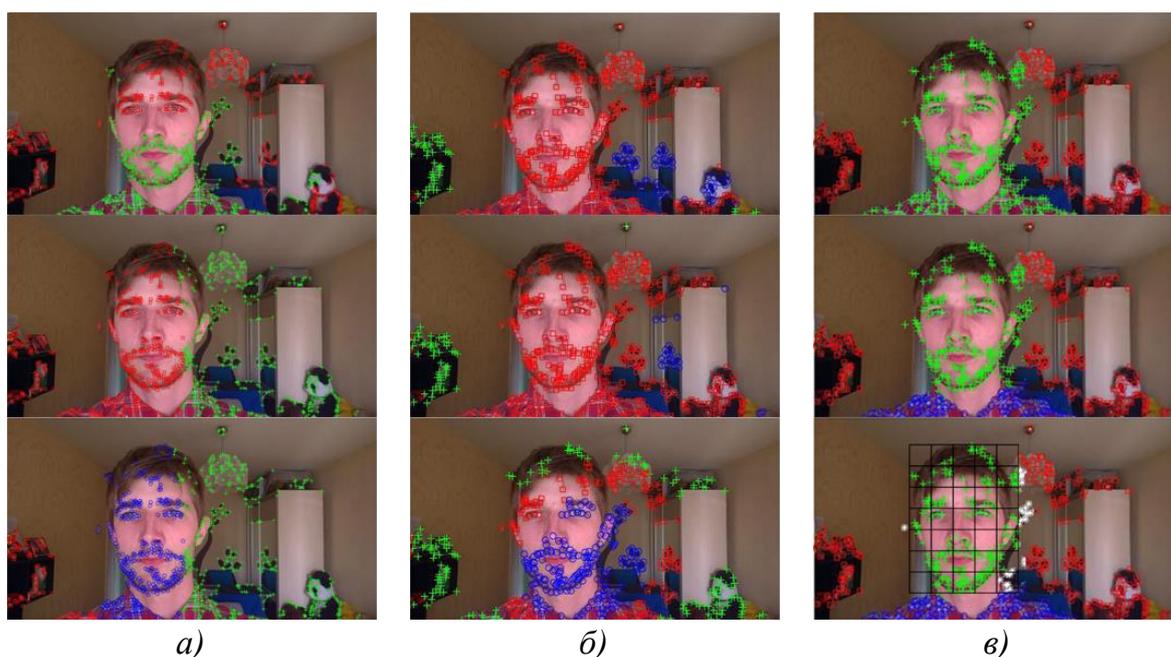


Рис. 2.13. Результаты кластеризации особенностей (по строкам сверху вниз: $\sigma = 1$, $\sigma = 100$, $\sigma = 200$): а – алгоритм нормализованных разрезов на графах, б – алгоритм кластеризации на основе случайного обхода графа без наложения сетки пропорций, в) алгоритм кластеризации особенностей на основе спектральной теории графов и ограничений

Блок-схема алгоритма представлена на рис. 2.14.



Рис. 2.14. Блок-схема алгоритма кластеризации особенностей на основе спектральной теории графов и ограничений для выделения лица человека на изображениях

2.4. Разработка алгоритма нахождения соответствий на изображениях с использованием тепловых ядер на графах

Нахождение соответствий на изображениях является важной задачей при отслеживании головы человека на видеопоследовательности [15]. Одним из недостатков существующих методов является то, что они определяют соответствия между всеми найденными особенностями и не учитывают их возможного отсутствия на сопоставляемых изображениях из-за изменения ракурса, освещения, шумов, перекрытий. Таким образом, в сопоставлении могут участвовать разные особенности, что повышает вероятность появления «выбросов». Для повышения устойчивости при решении поставленной задачи предлагается использовать структурный подход на основе графов.

Отслеживание должно осуществляться на всей видеопоследовательности. Этот процесс необходим, чтобы иметь возможность получить текущее положение и ориентацию головы человека в любой момент времени. Предполагается, что обработка видеопоследовательности будет осуществляться с частотой от 25 до 60 кадров в секунду. Следовательно, движение головы между кадрами будет небольшим.

На этапе инициализации выделены особые точки, описывающие голову человека на изображении и образующие граф. Положение головы в процессе отслеживания изменяется. Вместе с этим может определяться разный набор особых точек. Для устранения этой проблемы предлагается отслеживать не все обнаруженные особые точки, а только самые стабильные [4, 18].

Определение самых стабильных точек основано на термодинамическом уравнении. Термодинамическое уравнение широко применяется в физике и химии при изучении воздействия температуры на состояние газов [50]. Оно позволяет рассматривать граф в виде физической системы, внутри которой с течением времени изменяется распределение температуры нагретых газов. Так как этот процесс происходит равномерно, то в каждый момент времени можно определить самые горячие области. Они остаются нагретыми дольше всех остальных областей системы.

Термодинамическое уравнение позволяет оценить проходящую по ребрам графа информацию и рассматривать граф как функцию, зависящую от спектральных характеристик и параметра времени. Начальным

состоянием являются собственные векторы и собственные значения, которые вычисляются на основе спектральной декомпозиции матрицы графа.

Решением термодинамического уравнения является тепловое ядро. Тепловые ядра в зависимости от времени наблюдения за собственным пространством могут предоставить различную информацию о графе: при малых значениях – о связности и топологии, а при больших – о глобальной структуре графа [50].

Получив тепловые ядра графа в определенные моменты времени, можно определить «горячие» вершины. «Горячие» вершины – вершины графа, через которые на протяжении некоторого времени проходит наибольшее количество информации. Их взаимное расположение инвариантно к аффинным преобразованиям, которые может претерпеть объект [50]. Количество «горячих» вершин зависит от времени наблюдения за собственным пространством графа. Если рассмотреть граф автомобильных дорог, то самой «горячей» вершиной в нём будет перекресток, соединяющий две части города. В начальный момент времени количество автомобилей, прошедших рассматриваемый перекресток, не превышает аналогичные показатели на других перекрестках. Через некоторое время по этому перекрестку будет проходить намного больший трафик. Таким образом, по отношению к другим такой перекресток будет являться «горячим» [50].

Для анализа проходящей через особые точки информации предлагается использовать граф Делоне. Он строится согласно одноименной триангуляции. Данный метод триангуляции позволяет построить минимальное покрывающее дерево в связном неориентированном графе [74].

Чтобы устранить необходимость сопоставления большого числа полученных «горячих» вершин на изображениях по принципу «многие-ко-многим», что занимает значительное время, предлагается осуществить построение дерева. Уровни этого дерева будут образовываться путем объединения соседних областей. На самом нижнем уровне каждый лист дерева описывается одной особой точкой, а порог близости стремится к нулю. К корню дерева принадлежат все особые точки области головы на изображении. Порог близости в таком случае выбирается так, чтобы охватить все изображе-

ние головы. Промежуточные уровни дерева строятся путем постепенного увеличения этого порога.

Таким образом, нахождение соответствий будет осуществляться между особыми точками подграфов. Это должно в значительной степени увеличить быстродействие самого процесса нахождения соответствий, а также повысить точность.

На вход алгоритма отслеживания поступает текущий кадр видеопоследовательности и изображение области головы, полученное на этапе инициализации.

Выделяются особенности изображения с помощью метода SURF. Проводится триангуляция Делоне для построения графа $G=(V,E)$, где каждая вершина соответствует выделенной особенности, а некоторые пары вершин соединены ребрами согласно триангуляции. В данном случае используется невзвешенный ненаправленный граф, а матрица смежности имеет следующий вид:

$$A(u,v) = \begin{cases} 1, & \text{если } u, v \in E; \\ 0, & \text{в другом случае.} \end{cases} \quad (2.29)$$

По матрице смежности A вычисляется диагональная матрица степеней графа D . Затем находится нормализованная матрица Лапласа. Для невзвешенного графа она определяется следующим образом:

$$L_n = \begin{cases} 1, & \text{если } u = v \text{ и } d_v \neq 0; \\ -\frac{1}{\sqrt{d_u d_v}}, & \text{если } A(u,v) = 1; \\ 0, & \text{в другом случае.} \end{cases} \quad (2.30)$$

где d_u, d_v – степени вершин u и v соответственно.

Для вычисления спектра нормализованной матрицы Лапласа необходимо провести декомпозицию для получения собственных значений и собственных векторов $L_n = \Phi \Lambda \Phi^T$.

Распределение информации по графу выражается с помощью термодинамического уравнения, которое имеет вид:

$$\frac{\partial H_t}{\partial t} = -L_n H_t, \quad (2.31)$$

где t – время изменения состояния графа;

H_t – тепловое ядро.

Время влияет на получаемую информацию о графе. Если $t \rightarrow 0$, то можно получить только локальную информацию о связности и то-

пологии графа. Если $t \rightarrow \infty$, то будет известно о глобальной структуре графа [50]. Таким образом, значение t необходимо подбирать экспериментально.

Тепловое ядро является решением выражения (2.31), рассчитывается с помощью нормализованной матрицы Лапласа и представляет собой квадратную матрицу $|V| \times |V|$. Тепловое ядро, или матрица теплоты, в терминах собственного пространства для вершин u и v графа G имеет следующий вид:

$$H_t(u, v) = \sum_{i=1}^{|V|} e^{-\lambda_i t} \phi_i(u) \phi_i(v), \quad (2.32)$$

где λ_i – собственные значения матрицы Лапласа;

ϕ_i – собственные векторы матрицы Лапласа.

Зная тепловое ядро для графа, можно определить «горячие» вершины, которые будут сопоставляться. Они определяются по значению матрицы теплоты. Если граф имеет m вершин, то матрица теплоты H_t будет иметь m^2 элементов. Чтобы определить «горячие» вершины, из матрицы теплоты выбираются $(m+1)$ первых максимальных элементов.

«Горячими» вершинами являются те, которые удовлетворяют следующему условию:

$$n_h = \{i / H_t(i, i) \geq \varepsilon\}, \quad (2.33)$$

где ε – порог, который равен $m+1$ -ому максимальному элементу матрицы H_t .

«Горячие» вершины являются самыми стабильными из исходного набора. При увеличении времени наблюдения за собственным пространством их количество сокращается. При малом количестве итераций «горячих» вершин остается слишком много, что увеличивает трудоемкость. Напротив, при увеличении времени наблюдения «горячих» вершин становится мало, что мешает точно классифицировать отслеживаемый объект.

Опираясь на данные расчетов, было установлено, что оптимальной для поставленной задачи, является величина $t = 11$ (рис. 2.15).

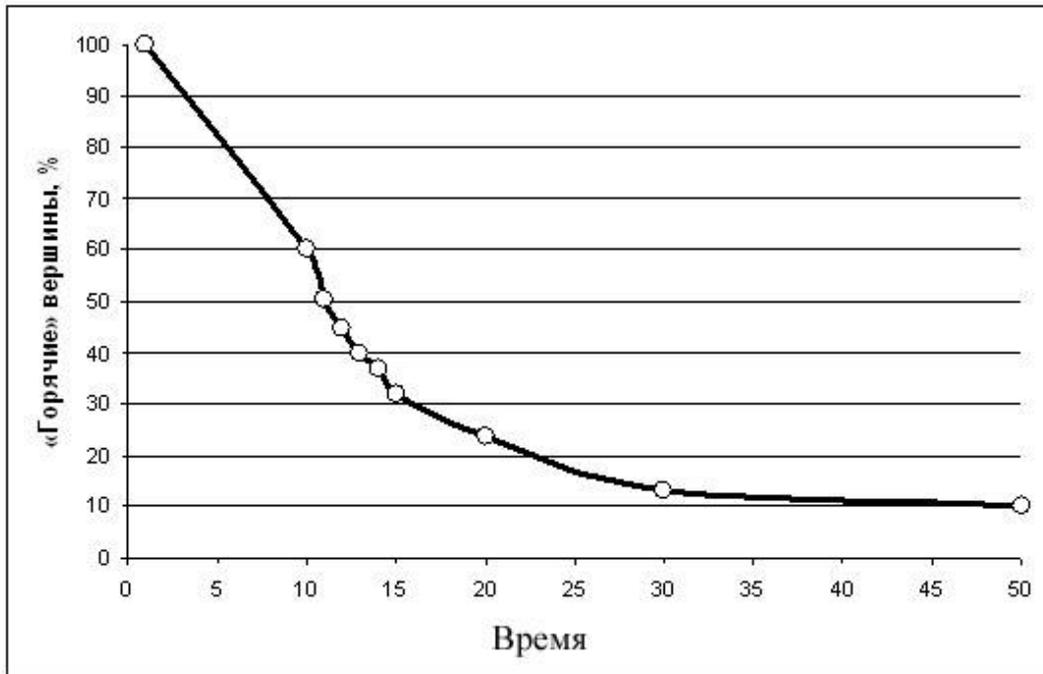


Рис. 2.15. Зависимость количества «горячих» вершин от количества итераций

При малом значении параметра t все вершины графа будут «горячими». Это противоречит самой идее отслеживания только самых стабильных вершин. При большом значении параметра t количество «горячих» вершин будет недостаточным для осуществления процесса сопоставления, так как будут отсутствовать особые точки многих областей, характерных именно для данного изображения. При установке параметра $t = 11$ произойдет заметное сокращение общего числа рассматриваемых вершин (до 50 % от начального количества). Это упростит дальнейшую работу с массивом данных. При этом особенности, точно описывающие изображение головы человека, будут сохранены.

Таким образом, для имеющихся особых точек изображения I алгоритм нахождения «горячих вершин» n_t состоит из следующих этапов.

Шаг 1. Выделяются особенности изображения с помощью метода SURF.

Шаг 2. Строится граф Делоне $G = (V, E)$.

Шаг 3. Находится нормализованная матрица Лапласа (2.30) и вычисляются ее собственные значения $\lambda_1, \lambda_2, \dots, \lambda_{|V|}$ и собственные векторы $\phi_1, \phi_2, \dots, \phi_{|V|}$.

Шаг 4. Рассчитывается матрица теплоты H_t (2.32).

Шаг 5. По матрице теплоты вычисляются «горячие» вершины n_h (2.33).

Обычно при сопоставлении графов или наборов точек используется принцип «многие-ко-многим», согласно которому каждая точка одного набора сравнивается с каждой точкой из второго. Этот процесс требует слишком много ресурсов и подвержен накоплению ошибок.

В разрабатываемом подходе предлагается представить полученные «горячие» вершины в виде иерархической структуры. На основе набора этих вершин строится дерево. Оно строится по иерархическому принципу, где наибольшее обобщение наблюдается на верхнем уровне, а максимальное уточнение – на самом нижнем. Выделение точек в кластеры представляет собой итеративный процесс с постепенным увеличением допустимого порога близости.

Точки u и v считаются соседними, если верно следующее условие:

$$d(u, v) \leq \mu, \quad (2.34)$$

где $d(u, v) = \sqrt{(u_x - v_x)^2 + (u_y - v_y)^2}$ – евклидово расстояние между точками;

μ – минимальный порог близости.

В результате получим уровни изображения, или сегменты, где самый нижний состоит из одиночных особенностей, представленных «горячими» вершинами. На верхних уровнях эти особенности будут объединяться в сегменты. На самом верхнем уровне иерархии все изображение описывается одним сегментом (рис. 2.16). Необходимо установить связь между регионами одного уровня, которые на плоскости являются соседними.

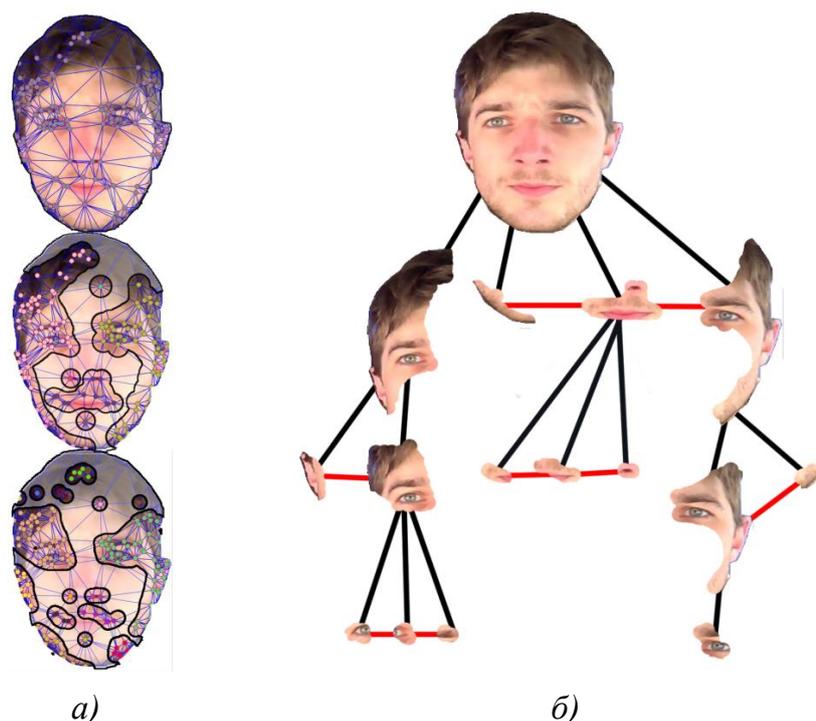


Рис. 2.16. Построение дерева «горячих» вершин:
 а – выделенные сегменты; б – дерево изображения

Таким образом, чтобы из имеющихся «горячих» вершин построить дерево, необходимо выполнить следующий алгоритм.

Шаг 1. Имеется N_H «горячих» точек. Устанавливается минимальный порог близости $\mu = 0$, чтобы на нижнем уровне дерева каждая особенность принадлежала одному сегменту. На данном этапе количество сегментов $N_S = N_H$. Шаг 2. Увеличивается порог μ . Рассматриваются все особенности в поисках соседей. Проверяется условие (2.34). Если точки являются соседними, то они принадлежат одному сегменту. В противном случае они относятся к разным сегментам.

Шаг 3. Если количество сегментов $N_S = 1$, то построение дерева считается завершенным, иначе повторяется шаг 2.

В результате построенное дерево состоит только из «горячих» вершин, которые являются наиболее стабильными особенностями изображения.

Используя алгоритмы нахождения «горячих» вершин и сегментации полученного набора данных, можно описать весь алгоритм сопоставления двух изображений. Сопоставление будет осуществляться по уровням дерева. Уровни зависят от порога близости μ . Для двух изображений I_1 и I_2 , которые необходимо сопоставить, алгоритм будет выглядеть следующим образом.

Шаг 1. Для изображений выполняется алгоритм нахождения «горячих» вершин.

Шаг 2. Полученные «горячие» вершины с помощью алгоритма построения дерева преобразуются в иерархическую структуру.

Шаг 3. Находятся соответствия между особенностями на определенном уровне на основе SURF-дескриптора.

Блок-схема алгоритма показана на рис. 2.17.

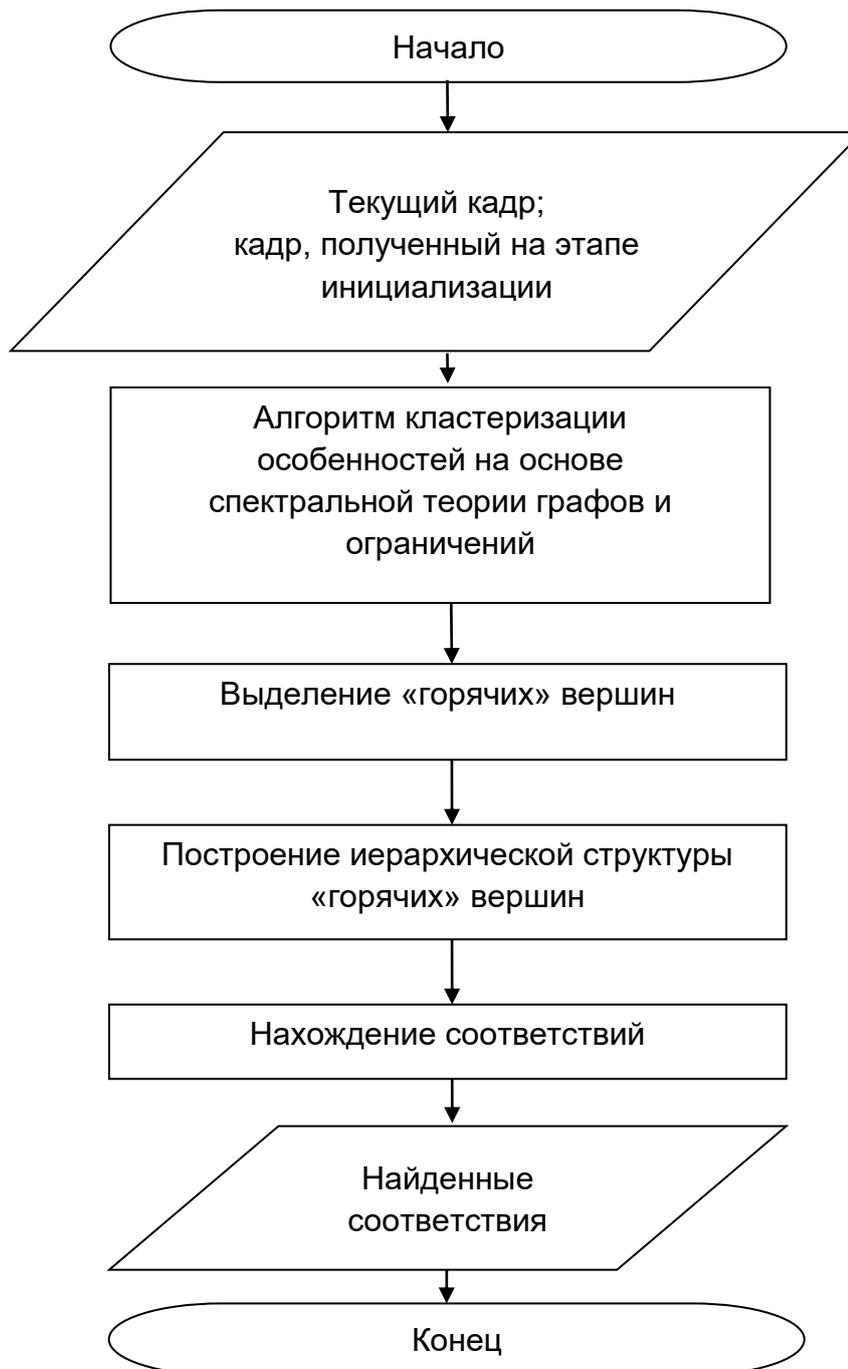


Рис. 2.17. Блок-схема алгоритма нахождения соответствий на изображениях с использованием тепловых ядер на графах

На рисунке 2.18 показаны соответствия, найденные с помощью метода SURF и разработанного алгоритма. Темными линиями показаны ложные соответствия. В разработанном алгоритме за счет выделения наиболее стабильных вершин в сопоставлении участвует меньшее количество особенностей. Так как эти особенности являются наиболее стабильными на изображении, то при использовании разработанного алгоритма уменьшается количество «выбросов».

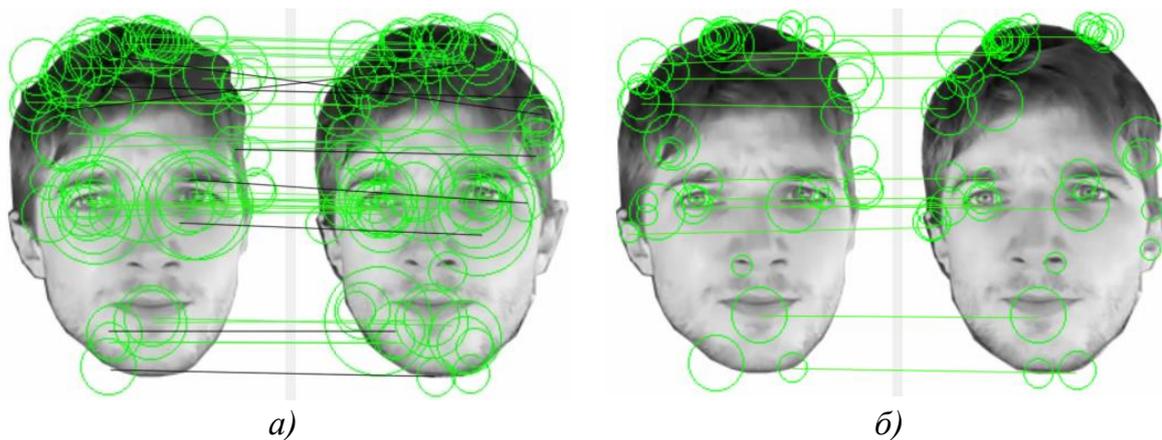


Рис. 2.18. Пример нахождения соответствий:

а – нахождение соответствий на основе метода SURF;

б – нахождение соответствий на основе разработанного алгоритма

2.5. Разработка алгоритма обнаружения головы человека на изображениях на основе вложения графа особенностей в векторное пространство

При анализе видеоизображений в задаче определения положения и ориентации головы человека в пространстве могут возникнуть ситуации, когда лицо человека пропадает из поля зрения. Это случается из-за резкого изменения скорости движения объекта, взаимного перекрытия в сцене, а также выхода объекта за рамки обзора области. Чтобы не приостанавливать работу системы и не запускать процесс инициализации повторно необходимо в автоматическом режиме производить поиск области головы на изображениях.

На данном этапе поиска области головы использование кластеризации на основе спектральной теории графов может быть недостаточным, так как необходимо обнаружить именно тот объект, который отслеживался ранее.

Для решения этой проблемы необходимо провести поиск по текущим кадрам видеопоследовательности. Для этого может использоваться изображение-эталон, полученное на этапе инициализации. Для сравнения изображения-кандидата с эталоном предлагается использовать спектральную теорию графов. Для решения поставленной задачи выполняется вложение графов, построенных по сравниваемым изображениям, в векторное пространство [8]. В этом случае графы сравниваемых изображений преобразуются в вектор числовых характеристик, на основе которого выполняется сравнение [157].

Под вложением графов понимается проецирование координат вершин графа в векторное пространство. В рассматриваемом случае для осуществления вложения предлагается использовать преобразование Юнга-Хаусхолдера [91].

При вложении графа в векторное пространство используется дифференциальная геометрия. Для описания связи между точками используются составные кривые. Если в обычной геометрии окружность характеризуется значением радиуса, то в дифференциальной геометрии используется значение ее кривизны. Кривизна является собирательным названием некоторого ряда характеристик, описывающих отклонение кривой от прямой, поверхности от плоскости. [157].

Каждый вложенный граф представляет собой матрицу координат, в которой каждый столбец является вектором координат соответствующей вершины графа. Любая пара таких векторов образует составную кривую, характеризующуюся параметрами кривизны. Кривизна состоит из ряда характеристик, описывающих отклонение кривой от прямой.

Таким образом, каждый граф можно описать матрицей параметров кривизны. С помощью метрики Хаусдорфа рассчитывается подобие между этими матрицами. Таким образом, можно сравнить графы [129].

Имея матрицу подобия, с помощью метода многомерного шкалирования MDS (Multi Dimensional Scaling) можно определить, насколько соответствуют друг другу объект-кандидат и объект-эталон. Метод MDS (MultiDimensional Scaling) используется для уменьшения размерности набора данных [62].

На вход алгоритма поступают текущее изображение и изображение-эталон, полученное на этапе инициализации. Следует отме-

тять, что при поворотах головы некоторые особенности на различных кадрах видеопоследовательности будут пропадать. Для снижения влияния этого процесса на надежность обнаружения предлагается отслеживать центры масс сегментов особенностей. Обычно такими областями являются характерные части головы человека: глаза, нос, уши, рот. При повороте головы набор особых точек в таких областях может измениться, но сама область будет присутствовать на изображении.

Проводится кластеризация особых точек. В ходе данного процесса все особые точки, выделенные на лице, группируются на основе метода связных компонент. Если в каком-либо кластере оказывается мало особенностей, то он не рассматривается далее. Таким образом, формируются сегменты, соответствующие наиболее характерным областям лица (рис. 2.19).

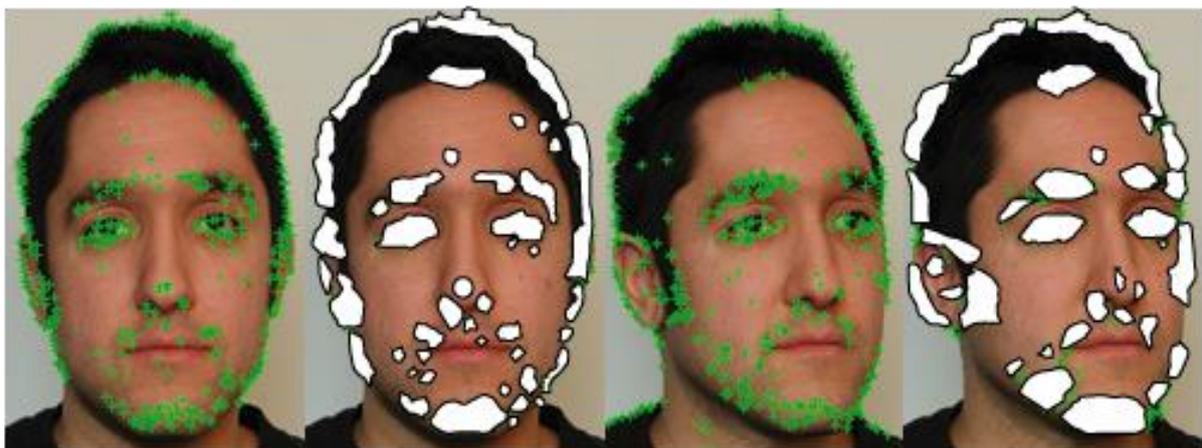


Рис. 2.19. Кластеризация особенностей, выделенных на изображении лица человека детектором SURF

Эксперименты проводились на изображениях, входящих в базы данных Массачусетского технологического института (MIT-CVCL Face Recognition Database). Эксперименты показали, что наиболее характерные кластеры остаются на изображениях при поворотах головы.

После кластеризации осуществляется построение графа. В качестве вершин выбираются центры масс полученных сегментов. Граф строится на основе триангуляции Делоне.

Нормализованная матрица Лапласа рассчитывается на основе выражения (2.30). На основе декомпозиции нормализованной матрицы Лапласа вычисляются спектральные характеристики графа.

Вложение графа основано на решении термодинамического уравнения (2.31). Тепловое ядро является решением этого уравнения и вычисляется с помощью собственных значений и собственных векторов (2.32).

При проецировании графа в векторное пространство с помощью спектральных характеристик используются значения ядра теплоты. Выполняется декомпозиция Юнга-Хаусхолдера, которая выглядит следующим образом:

$$H_t = Y^T Y, \quad (2.35)$$

где $Y = (y_1 \dots y_u \dots y_{|V|})$ – матрица координат размером $|V| \times |V|$, в которой каждый столбец является вектором координат соответствующей вершины.

Раскрывая это выражение, получаем $Y = e^{-\frac{\Lambda t}{2}} \Phi^T$. Следовательно, для вершины u координатный вектор рассчитывается следующим образом:

$$y_u = \left(e^{-\frac{\lambda_1 t}{2}} \phi_1(u), e^{-\frac{\lambda_2 t}{2}} \phi_2(u), \dots, e^{-\frac{\lambda_{|V|} t}{2}} \phi_{|V|}(u) \right)^T. \quad (2.36)$$

При вложении графа в векторное пространство используется дифференциальная геометрия, в которой для описания связи между элементами матрицы координат необходимо использовать составные кривые. По теореме Гаусса-Боне часть такой кривой может быть аппроксимирована дугой окружности [157]. Такие кривые характеризуются параметрами кривизны, значения которой могут быть получены по значениям кратчайшего пути по окружности (длина дуги) и квадратичному Евклидовому расстоянию (длина хорды) [157]. Параметры кривой изображены на рис. 2.20.

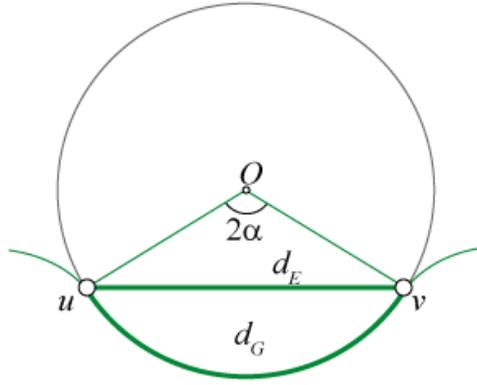


Рис. 2.20. Представление кривой с помощью дуги окружности

Квадратичное евклидово расстояние в этом случае имеет следующий вид:

$$d_E(u, v)^2 = (y_u - y_v)^T (y_u - y_v) = \sum_{i=1}^{|V|} e^{-\lambda_i t} [\phi_i(u) - \phi_i(v)]^2. \quad (2.37)$$

Составные кривые характеризуются параметрами кривизны. Для точек u и v значение кривизны рассчитывается следующим образом:

$$k(u, v) = \frac{1}{R(u, v)}, \quad (2.38)$$

где $R(u, v)$ – радиус окружности.

Далее радиус окружности будем обозначать R , кратчайший путь по окружности – d_G , евклидово расстояние – d_E .

Представим кривую дугой окружности с радиусом R . В таком случае длина дуги, то есть кратчайший путь по окружности между точками u и v , вычисляется следующим образом:

$$d_G = 2\alpha R, \quad (2.39)$$

где α – угол дуги.

При вложении графов кратчайшее расстояние характеризуется весом ребра. Следовательно, для невзвешенного графа можно принять $d_G = 1$. В таком случае угол α вычисляется следующим образом:

$$\alpha = \frac{1}{2R}. \quad (2.40)$$

Евклидово расстояние равно длине хорды в рассматриваемой окружности между точками u и v :

$$d_E = 2R \sin(\alpha). \quad (2.41)$$

В дальнейших расчетах необходимо избавиться от вычисления функции синуса неизвестного угла α . Для этого предлагается разложить функцию $\sin(x)$ в ряд Маклорена [60], то есть в ряд Тейлора при $x = 0$:

$$d_E = 2R \left(\alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \dots \right). \quad (2.42)$$

Принимается, что в расчетах используется точность до двух членов последовательности ряда Маклорена. Таким образом, имеем следующее:

$$d_E = 1 - \frac{1}{24R^2}. \quad (2.43)$$

Решая это уравнение в поисках R , получим следующее значение кривизны для точек u и v :

$$k(u, v) = \sqrt{24(1 - d_E)}. \quad (2.44)$$

Таким образом, для графа заполняется матрица кривизны.

Пусть имеются два графа $G_1 = (V_1, E_1, k_1)$ и $G_2 = (V_2, E_2, k_2)$, где V_1, V_2 – набор вершин, E_1, E_2 – набор ребер, k_1, k_2 – матрицы кривизны. Таким образом, расстояния между графами можно описать с помощью метрики Хаусдорфа:

$$HD(G_1, G_2) = \max_{i \in V_1} \max_{j \in V_1} \min_{I \in V_2} \min_{J \in V_2} \|k_2(I, J) - k_1(i, j)\|. \quad (2.45)$$

С помощью метрики Хаусдорфа вычисляется матрица расстояний для графов объекта-кандидата и объекта-эталоны. Для визуализации результатов используется метод MDS (MultiDimensional Scaling).

Таким образом, алгоритм обнаружения головы человека на изображениях на основе вложения графа особенностей в векторное пространство состоит из следующих шагов:

Шаг 1. Выделяются особенности изображения-кандидата.

Шаг 2. Проводится кластеризация особенностей.

Шаг 3. На основе центров масс кластеров строится граф Делоне. Рассчитывается нормализованная матрица Лапласа (2.30), вычисляются выражения (2.29 – 2.32).

Шаг 4. С помощью преобразования Юнга-Хаусхолдера происходит проецирование координат вершин графа в векторное пространство (2.36).

Шаг 5. Рассчитываются значения кривых, связывающих точки. Строится матрица кривизны на основе выражения 2.44.

Шаг 6. Вычисляется матрица расстояний между характеристиками графов изображения-эталона и текущего изображения (2.45).

Шаг 7. Для визуализации данных применяется метод многомерного шкалирования, позволяющий отобразить степень близости изображения-кандидата к какому-либо объекту из базы эталонов.

Примеры построения графов на основе изображений приведены на рис. 2.21. Вложение графов этих изображений в векторное пространство с помощью метода многомерного шкалирования представлено на рис. 2.22.

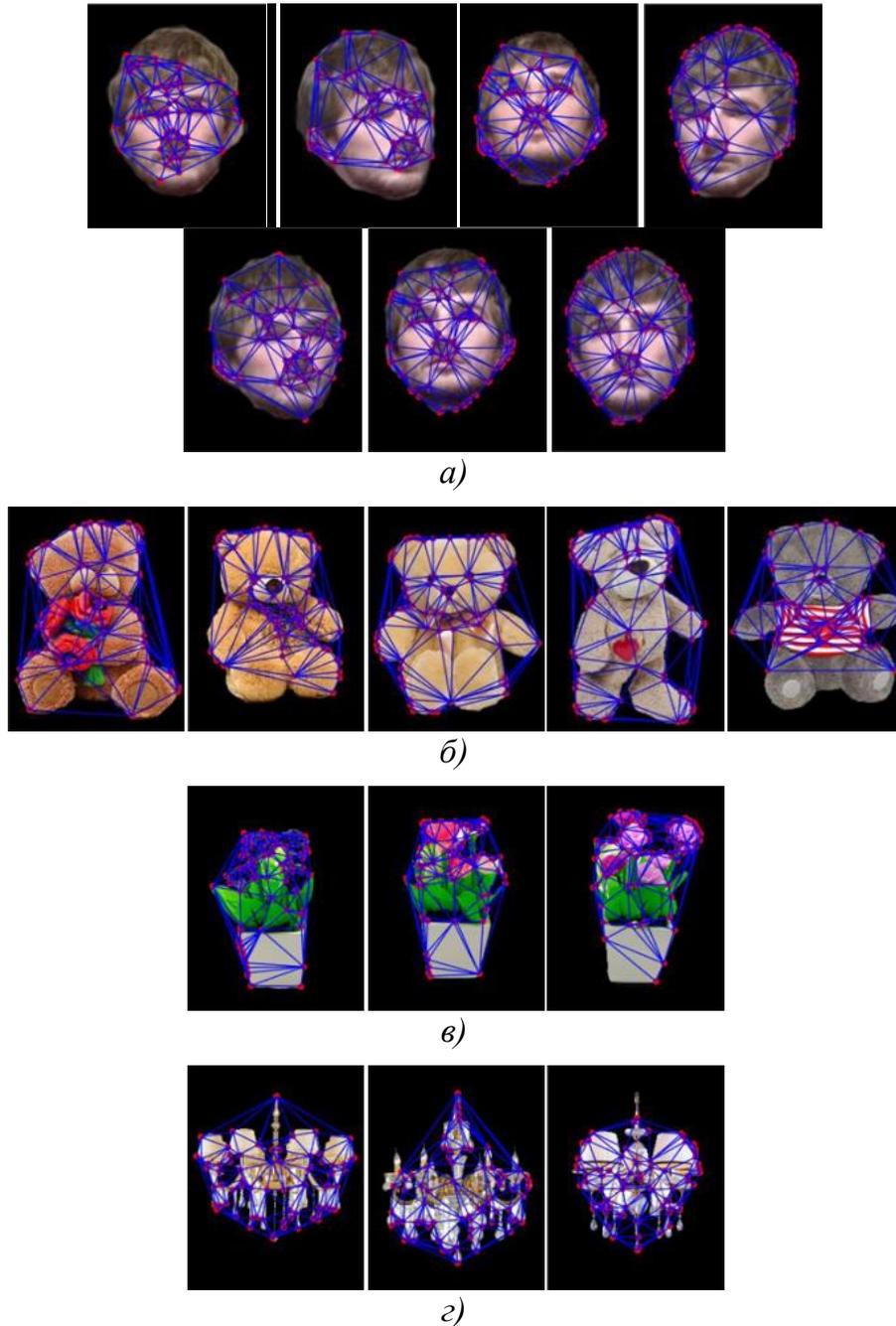


Рис. 2.21. Построение графа на основе изображений

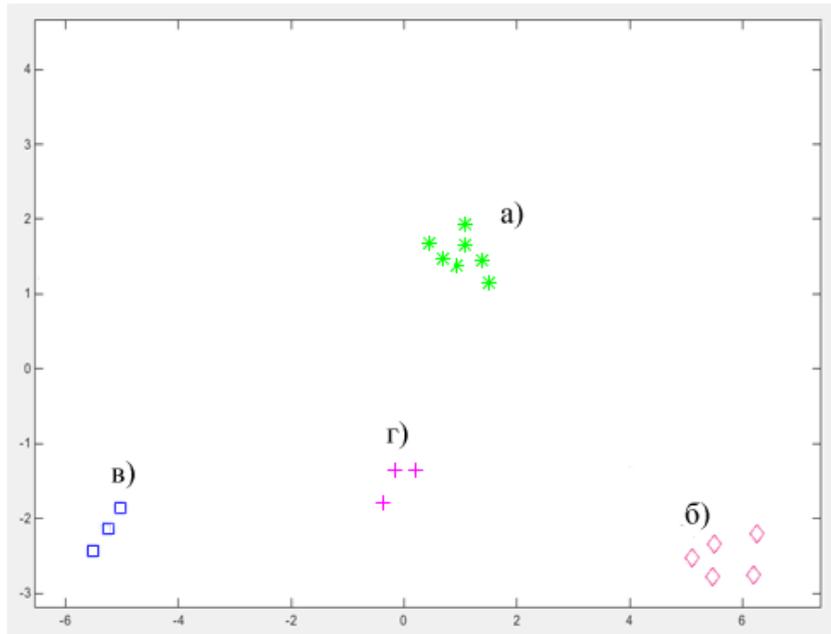


Рис. 2.22. Вложение графов изображений, представленных на рис. 2.21, в векторное пространство с помощью метода многомерного шкалирования

Глава 3. ИССЛЕДОВАНИЕ РАЗРАБОТАННЫХ МОДЕЛИ И АЛГОРИТМОВ ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ И ОРИЕНТАЦИИ ГОЛОВЫ ЧЕЛОВЕКА ПО СТЕРЕОИЗОБРАЖЕНИЯМ С ИСПОЛЬЗОВАНИЕМ ГРАФОВ

3.1. Исследование алгоритма кластеризации особенностей на основе спектральной теории графов и ограничений для выделения лица человека на изображениях

При исследовании алгоритма кластеризации особенностей на основе спектральной теории графов и ограничений для выделения лица человека на изображениях проводится сравнение результатов его работы с наиболее известными подходами по точности и скорости обработки [6, 55].

Для сравнительного анализа были выбраны следующие методы:

- метод нормализованного разреза (Normalized cuts – Ncut) [142];
- метод k-средних (k-means) [149];
- метод сильной связи подграфов (Highly Connected Subgraphs – HCS) [87].

Тестирование выполнялось на следующих наборах данных:

- тестовые изображения, состоящие из произвольного набора точек (рис. 3.1);
- тестовые изображения, состоящие из набора точек. Форма набор точек близка к изображению головы (рис. 3.7, а);
- изображения головы человека (рис. 3.8).

При сравнении разработанного алгоритма кластеризации особенностей на основе спектральной теории графов и ограничений для выделения лица человека на изображениях с методами нормализованного разреза, k-средних, сильной связи подграфов использовались тестовые изображения, представленные на рис. 3.1.

Следует отметить, что ограничения в этом случае не использовались, так как форма изображенных объектов далека от формы головы человека.

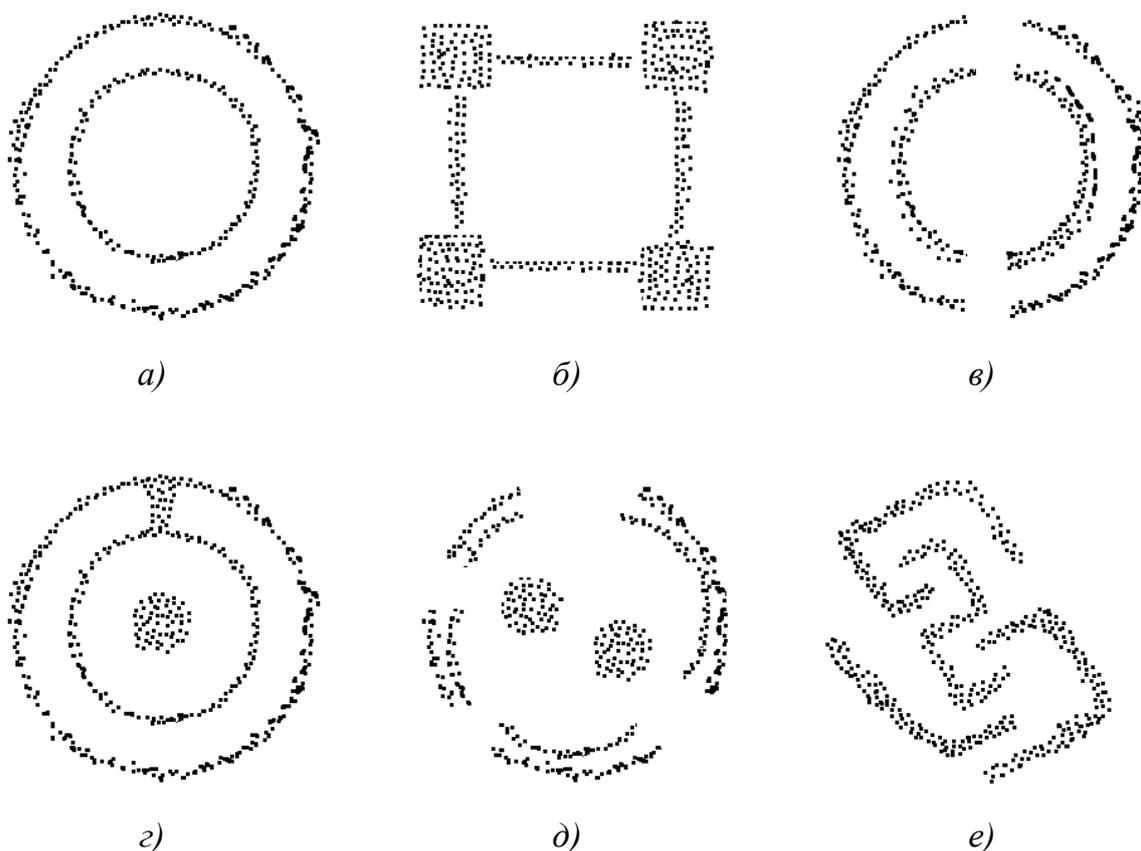


Рис. 3.1. Тестовые изображения, синтезированные для исследования алгоритма кластеризации особенностей: а – объект № 1; б – объект № 2; в – объект № 3; г – объект № 4; д – объект № 5; е – объект № 6

Метод нормализованного разреза хорошо работает в случаях с изображениями, на которых четко прослеживается кластерная структура (рис. 3.2.).

Метод k -средних наиболее целесообразно использовать в тех случаях, когда на изображении можно выделить четкие центры масс особенностей (рис. 3.3).

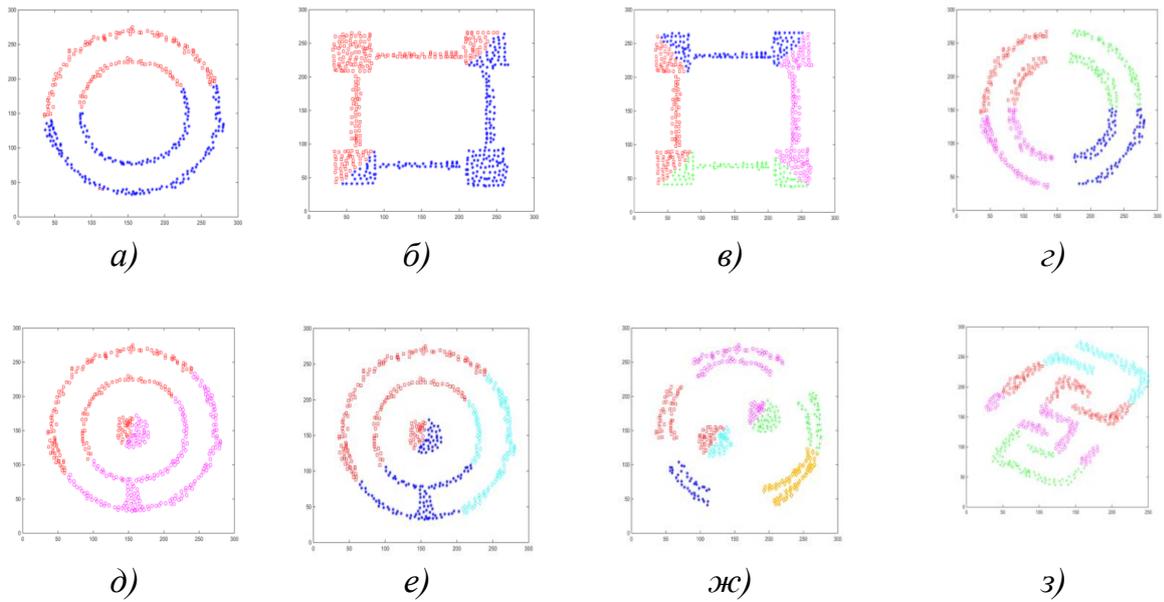


Рис. 3.2. Кластеризация особенностей тестовых изображений с помощью метода нормализованного разреза: а – Тест № 1, 2 кластера; б – Тест № 2, 2 кластера; в – Тест № 3, 3 кластера; г – Тест № 4, 4 кластера; д – Тест № 5, 2 кластера; е – Тест № 6, 3 кластера; ж – Тест № 7, 6 кластера; з – Тест № 8, 4 кластера

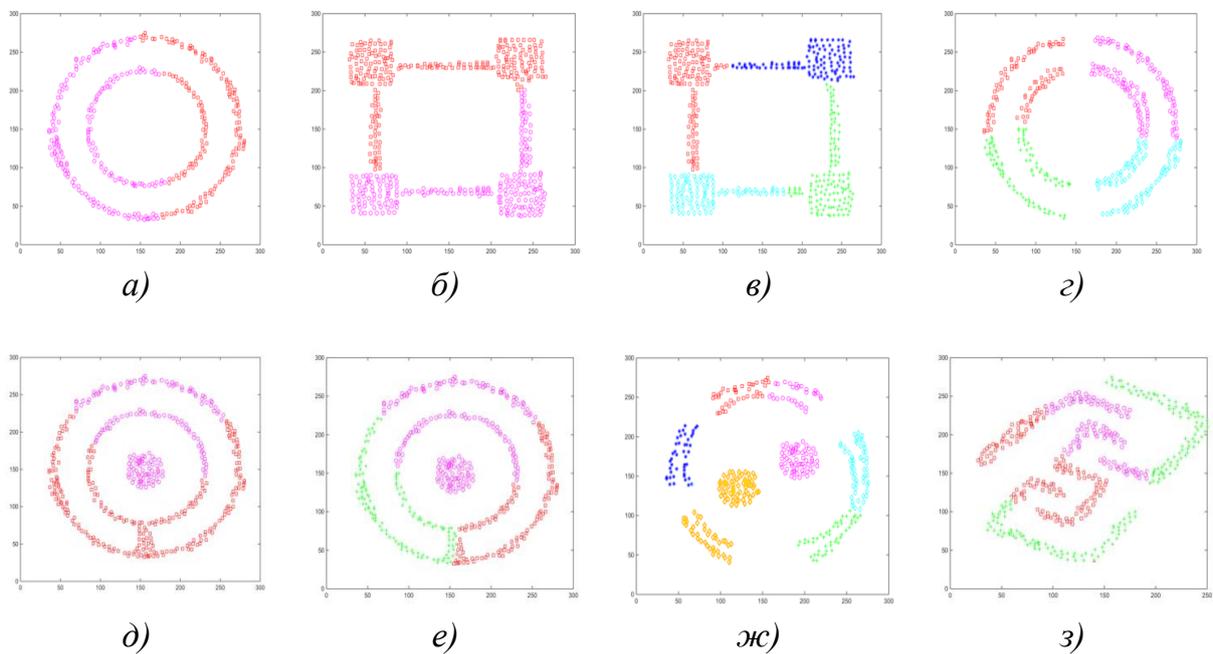


Рис. 3.3. Кластеризация особенностей тестовых изображений с помощью метода k-средних: а – Тест № 1, 2 кластера; б – Тест № 2, 2 кластера; в – Тест № 3, 3 кластера; г – Тест № 4, 4 кластера; д – Тест № 5, 2 кластера; е – Тест № 6, 3 кластера; ж – Тест № 7, 6 кластера; з – Тест № 8, 4 кластера

Метод сильной связи подграфов хорошо выделяет необходимое число кластеров в ситуациях, когда отдельные группы точек образуют единую структуру (рис. 3.4). Однако данный подход теряет в точности при возникновении шумов [87].

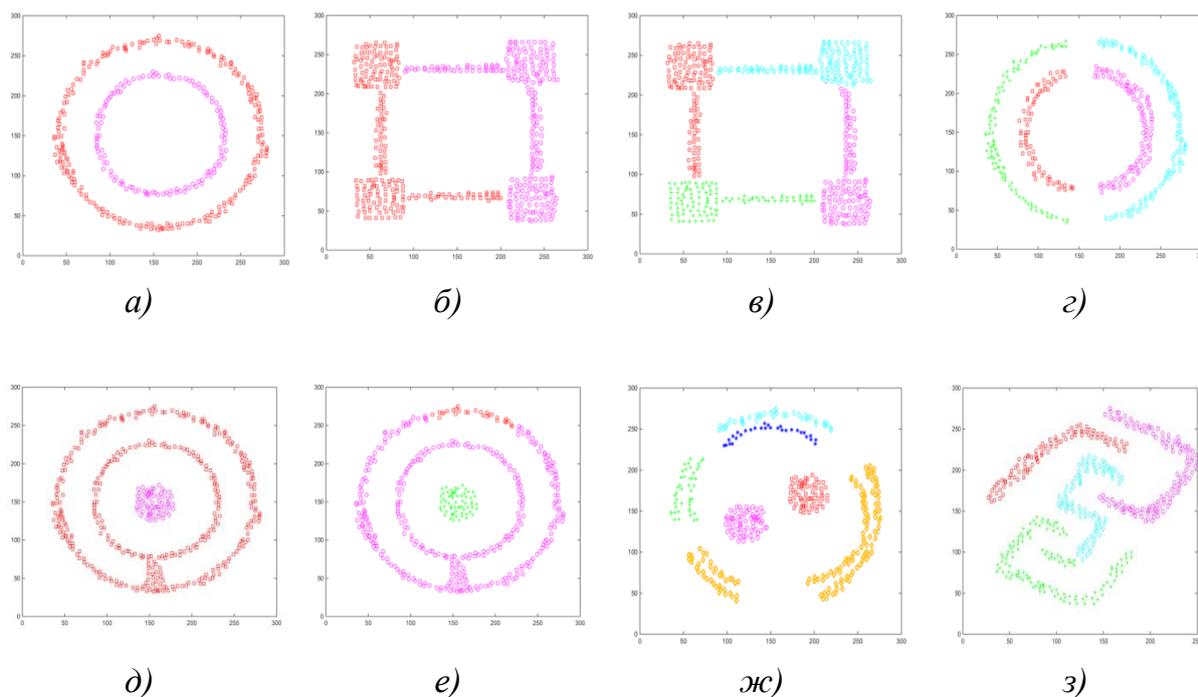


Рис. 3.4. Кластеризация особенностей тестовых изображений с помощью метода HCS: а – Тест № 1, 2 кластера; б – Тест № 2, 2 кластера; в – Тест № 3, 3 кластера; г – Тест № 4, 4 кластера; д – Тест № 5, 2 кластера; е – Тест № 6, 3 кластера; ж – Тест № 7, 6 кластеров; з – Тест № 8, 4 кластера

Разработанный алгоритм в большинстве случаев с тестовыми изображениями показывает хорошие результаты (рис. 3.5).

По диаграмме на рис. 3.6 видно, что точность кластеризации при использовании разработанного алгоритма без учета ограничений составляет почти во всех случаях более 90 %.

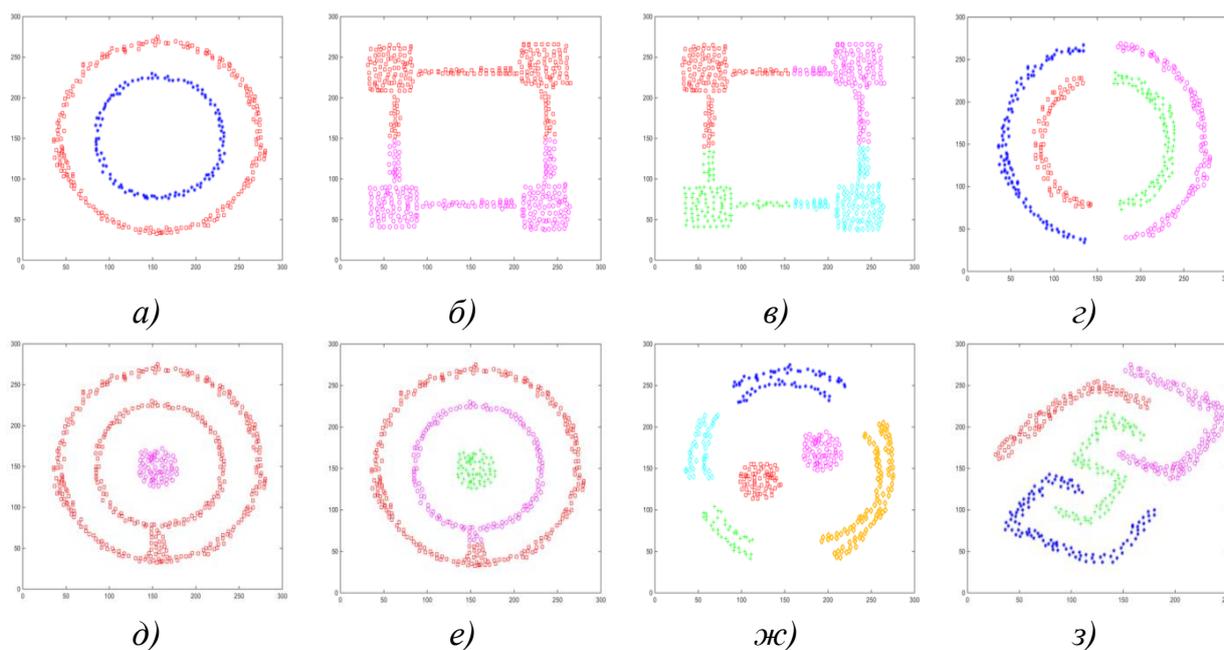


Рис. 3.5. Кластеризация особенностей тестовых изображений с помощью разработанного алгоритма: а – Тест № 1, 2 кластера; б – Тест № 2, 2 кластера; в – Тест № 3, 3 кластера; г – Тест № 4, 4 кластера; д – Тест № 5, 2 кластера; е – Тест № 6, 3 кластера; ж – Тест № 7, 6 кластеров; з – Тест № 8, 4 кластера

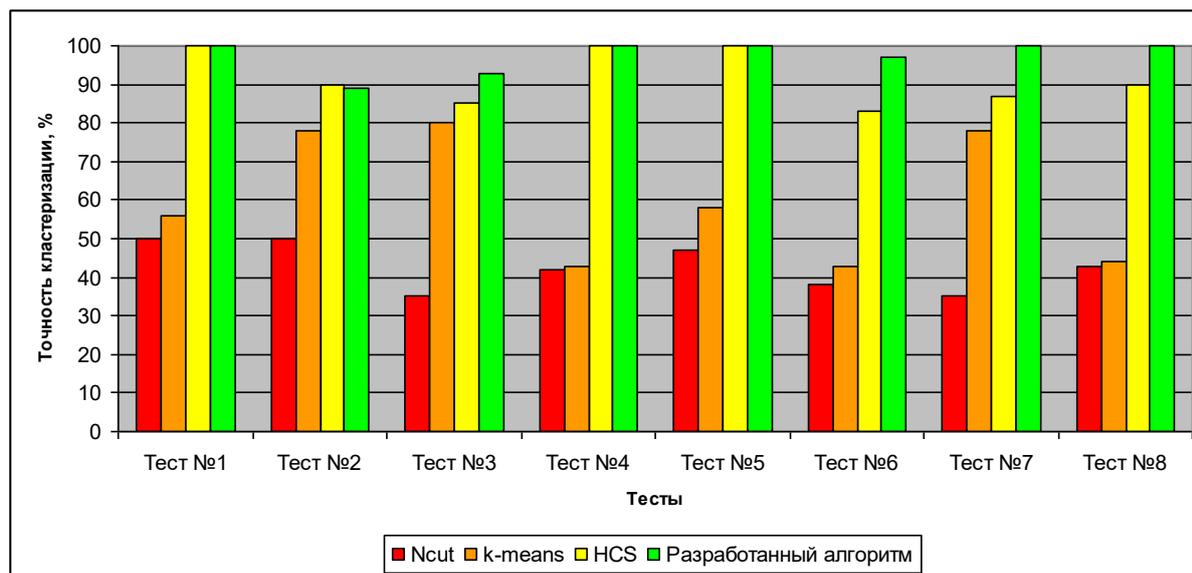


Рис. 3.6. Диаграммы, иллюстрирующие точность отнесения точек к кластерам на тестовых изображениях

Ниже приведены результаты кластеризации сгенерированного изображения, содержащего различные группы точек. Часть из них по форме подобна изображению головы человека (рис. 3.7, а). При кластеризации без наложения условий выделяются все отдельно расположенные точки (рис. 3.7, б).

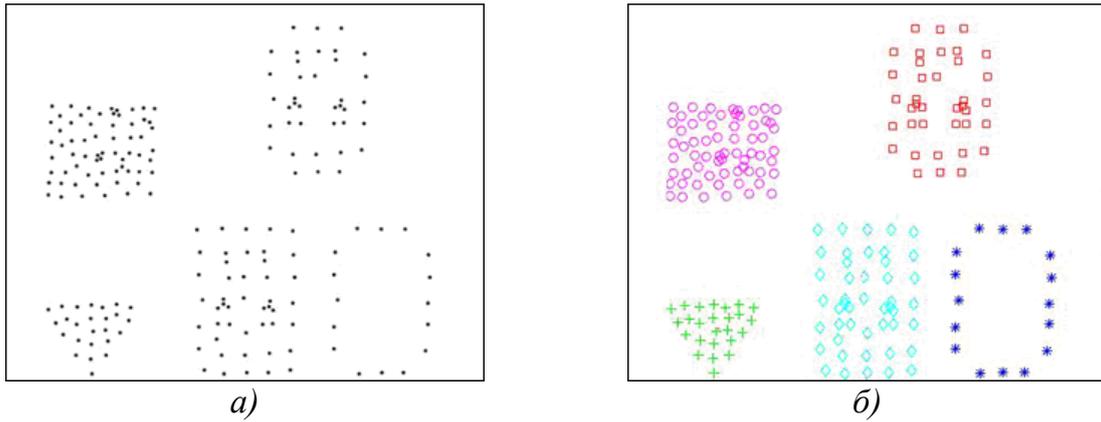


Рис. 3.7. Результаты разработанного алгоритма кластеризации особенностей на основе спектральной теории графов без использования ограничений: а – тестовые изображения, б – кластеризация без ограничений

При использовании тех или иных ограничивающих условий выделяются различные объекты (рис. 3.8).

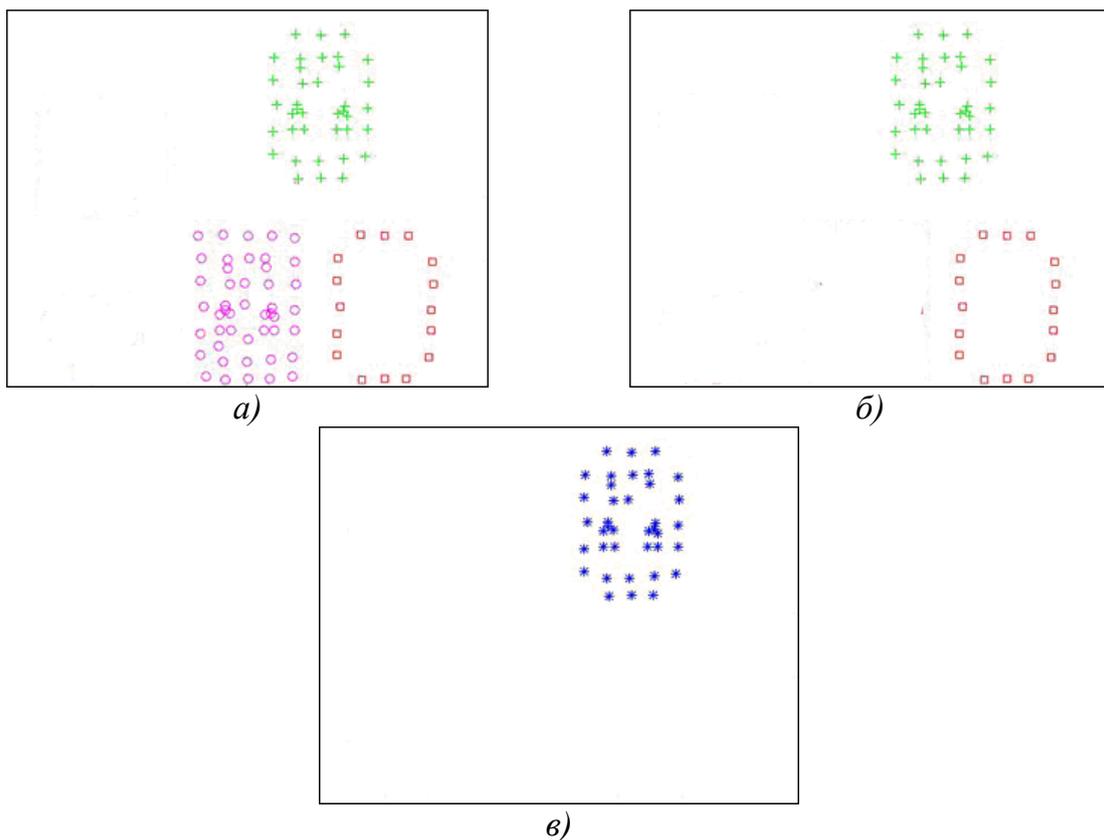


Рис. 3.8. Результаты разработанного алгоритма кластеризации особенностей на основе спектральной теории графов с использованием ограничений: а – применение условия, ограничивающего размер кластера на основе пропорций лица, выделено 3 объекта; б – применение условия вхождения особенностей в угловые ячейки сетки и условия вхождения особенностей в контур лица, выделено 2 объекта; в – применение всех условий, выделен 1 объект

При тестировании алгоритмов на реальных изображениях были использованы сцены, содержащие разное количество лиц. Количество кластеров, которые должны быть выделены, заранее не устанавливалось. Кластеризация осуществлялась до тех пор, пока не было установлено соответствие кластера сетке пропорций. При использовании не всего набора ограничивающих условий выделяются объекты, не являющиеся изображением лица человека. При использовании всех условий лица людей выделяются правильно (рис. 3.9 – 3.11).

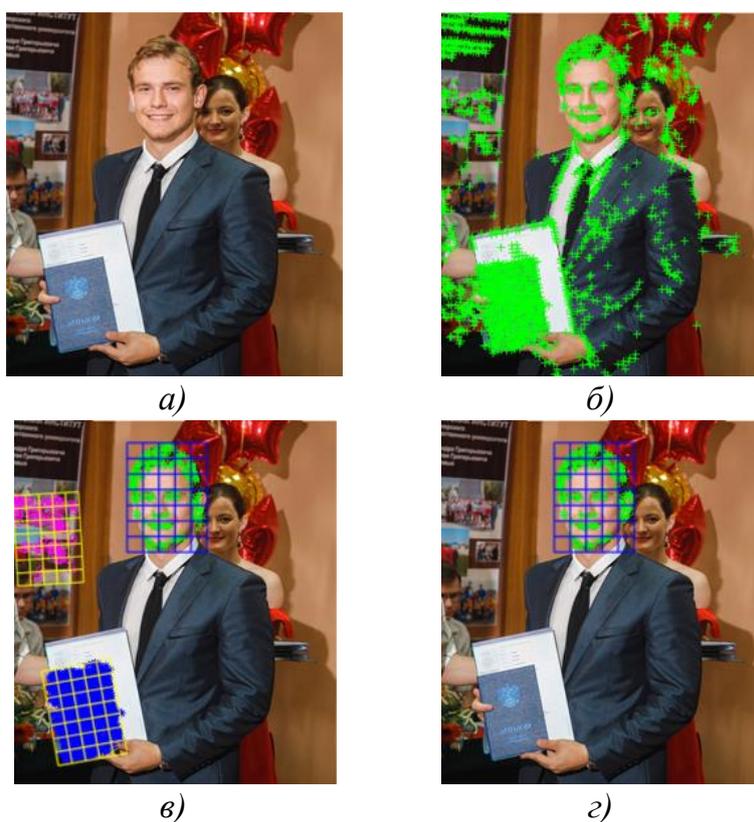


Рис. 3.9. Результаты разработанного алгоритма кластеризации особенностей на реальных изображениях (на изображении присутствует 1 человек): а – исходное изображение; б – выделенные особые точки сцены; в – кластеризация с неполными условиями (выделено 3 объекта, из которых 2 не являются лицами); г – кластеризация с полными условиями (выделено 1 лицо)

На рис. 3.11 присутствует 5 человек. Только с использованием полного набора ограничивающих условий алгоритм смог выделить 5 лиц.

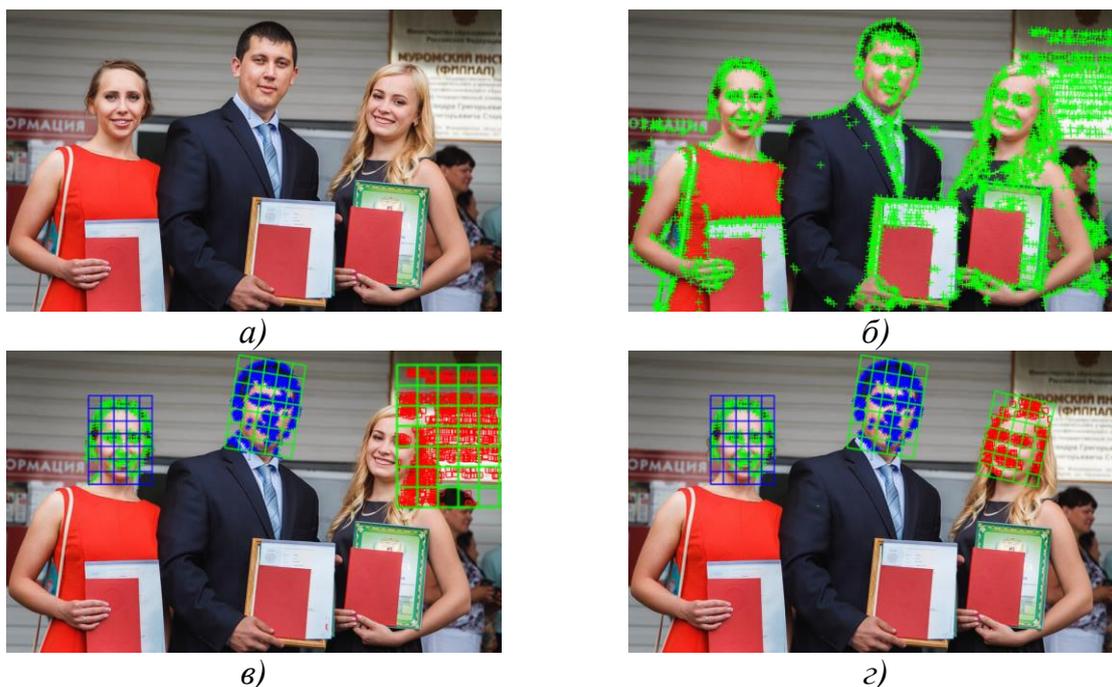


Рис. 3.10. Результаты разработанного алгоритма кластеризации особенностей на реальных изображениях (на изображении присутствуют 3 человека): а – исходное изображение; б – выделенные особые точки сцены; в – кластеризация с неполными условиями (выделено 3 объекта, из которых 1 не является лицом); г – кластеризация с полными условиями (выделено 3 лица)

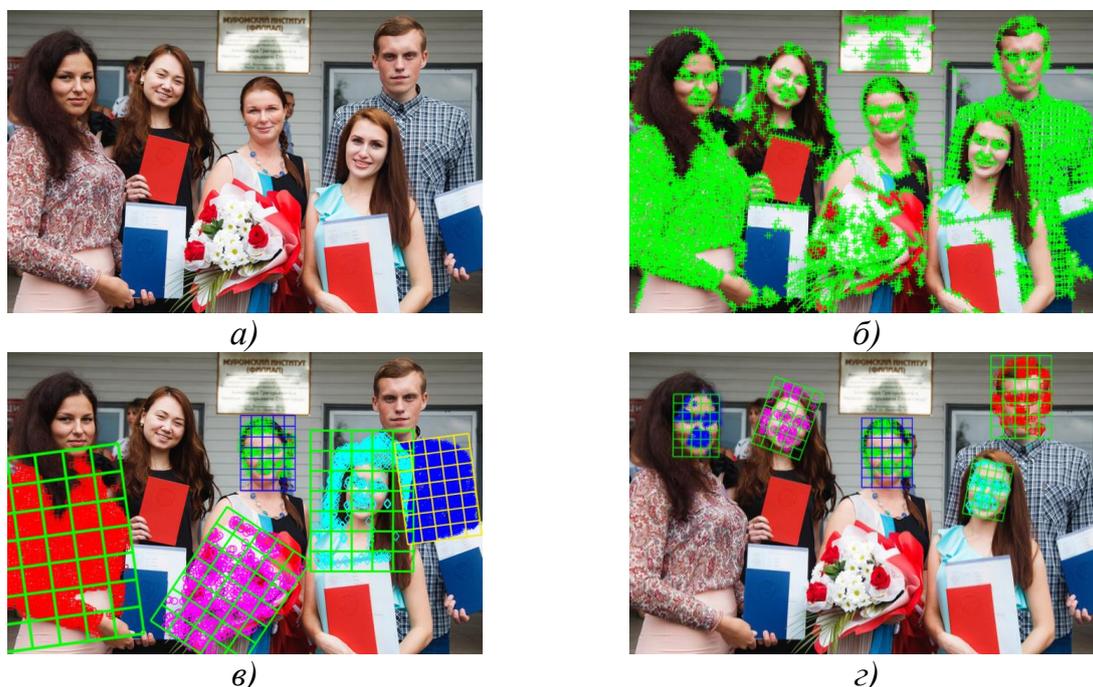


Рис. 3.11. Результаты разработанного алгоритма кластеризации особенностей на реальных изображениях (на изображении присутствуют 5 человек): а – исходное изображение; б – выделенные особые точки сцены; в – кластеризация с неполными условиями (выделено 5 объектов, из которых 4 не являются лицами); г – кластеризация с полными условиями (выделено 5 лиц)

Сложность разработанного алгоритма $O(n^2)$. Это объясняется необходимостью два раза рассчитывать собственные значения: на основе матрицы Лапласа и матрицы времени обхода. Однако дополнительные расчеты позволили добиться большей точности. Разработанный алгоритм осуществляет кластеризацию особенностей с использованием ограничений, что повышает в некоторых случаях точность выделения лиц на 27 % (рис. 3.12). Если изображение состоит из нескольких областей особенностей, то на каждой итерации будет происходить выделение одного кластера. Это происходит из-за того, что при расчете времени обхода учитываются все связи между всеми вершинами графа. Таким образом, применение априорной информации позволяет точно выделять кластеры особенностей, которые принадлежат именно голове человека на изображении.

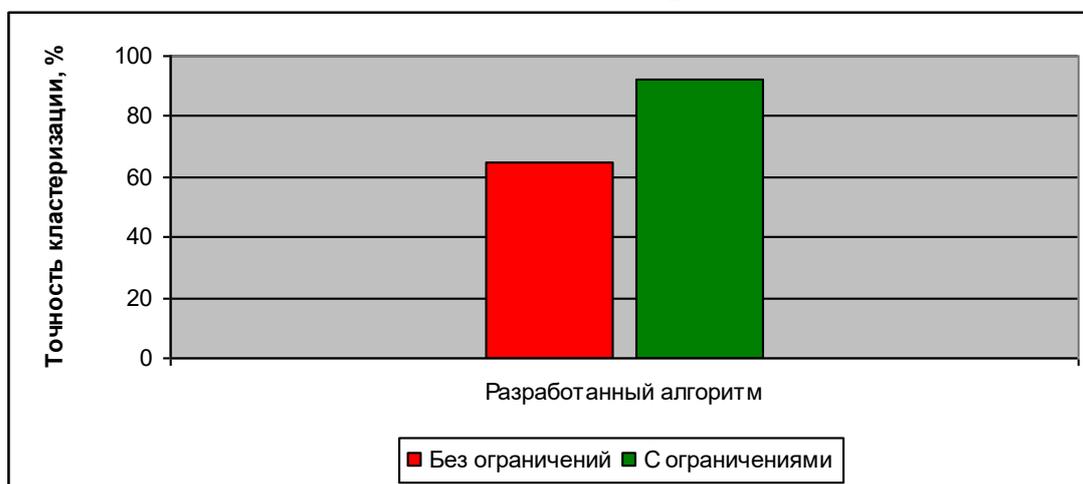


Рис. 3.12. Точность алгоритма кластеризации особенностей изображений лиц на основе спектральной теории графов без ограничений и с ограничениями

3.2. Исследование алгоритма нахождения соответствий на изображениях с использованием тепловых ядер на графах

Тестирование разработанного алгоритма осуществляется с использованием двух наборов изображений:

- изображений произвольных объектов;
- изображений лиц людей.

Сравнение работы алгоритма будет производиться с методом SURF, так как он считается некоторым стандартом для задач нахождения соответствий.

В качестве изображений произвольных объектов были использованы следующие наборы данных (рис. 3.13, 3.14):

- Carnegie Mellon University (CMU) Image Database;
- Columbia University Image Library (COIL-100).



Рис. 3.13. Тестовые изображения, используемые при исследовании алгоритма нахождения соответствий с использованием тепловых ядер на графах

Изображения лиц людей взяты из базы данных Массачусетского технологического института MIT-CBCL Face Recognition Database (рис. 3.14).



Рис. 3.14. Изображения лиц людей из базы Массачусетского технологического института (MIT-CBCL Face Recognition Database)

В ходе тестирования определялась точность нахождения соответствий при изменении угла поворота объектов в пространстве. На

рис. 3.15 и 3.16 приведены результаты нахождения соответствий с помощью метода SURF и разработанного алгоритма. Метод SURF выделяет больше особенностей и находит больше соответствий. В то же время количество выделенных особенностей у разработанного алгоритма меньше. Причиной этому служит то, что выделяются только самые стабильные особенности, расположение которых не меняется при повороте объекта [4, 54]. На основе проведенных исследований можно сказать, что с увеличением угла поворота количество правильных соответствий у алгоритма SURF уменьшается (рис. 3.15).

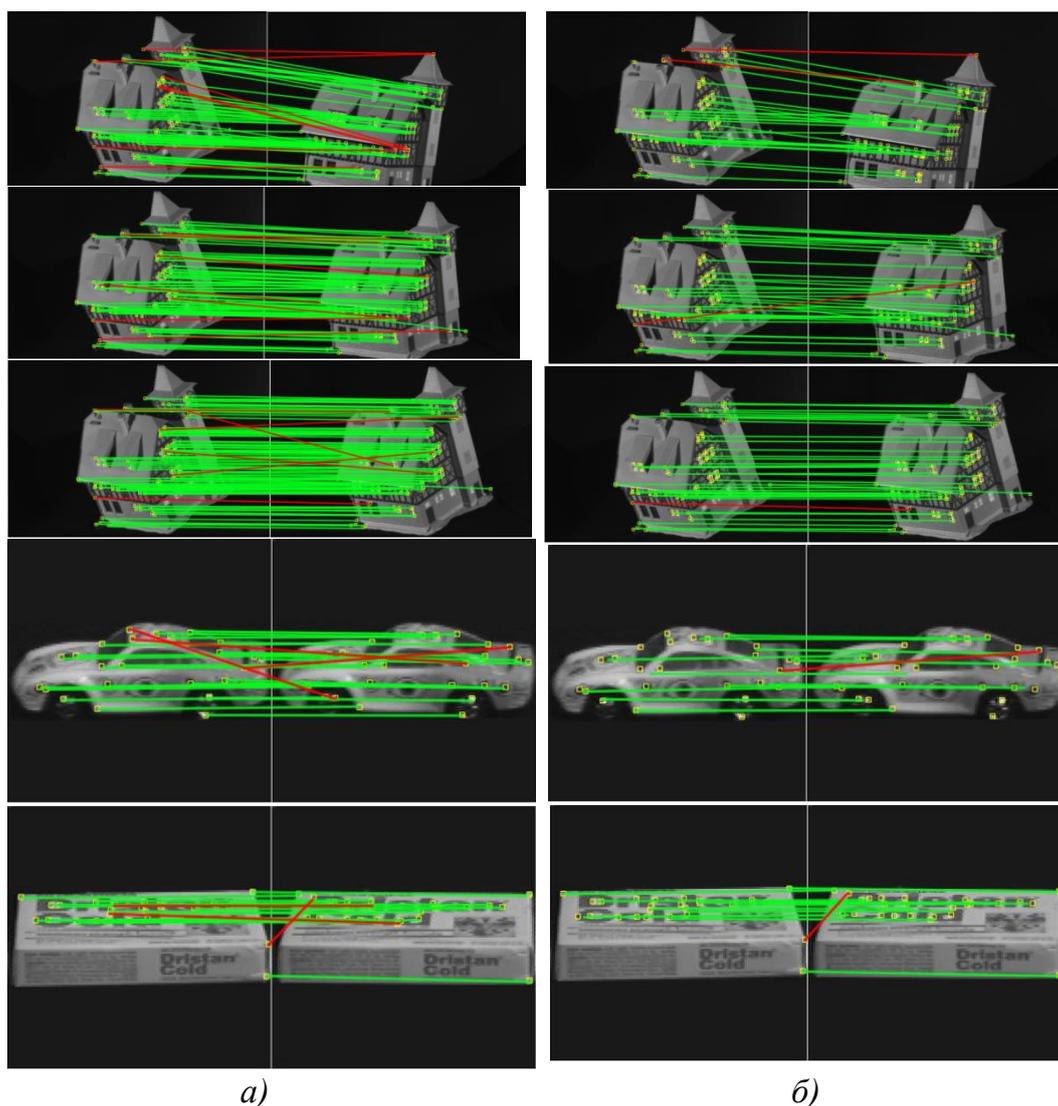


Рис. 3.15. Исследование алгоритма нахождения соответствий на тестовых изображениях с использованием тепловых ядер на графах:
 а – нахождение соответствий с использованием метода SURF;
 б – нахождение соответствий с использованием разработанного алгоритма

При тестировании изображений с лицами людей наблюдаются следующие особенности (рис. 3.16). Количество найденных верных соответствий при использовании алгоритма SURF заметно уменьшается при повороте головы на 15° и более. В то же время разработанный алгоритм более устойчив к изменению ракурса [4].

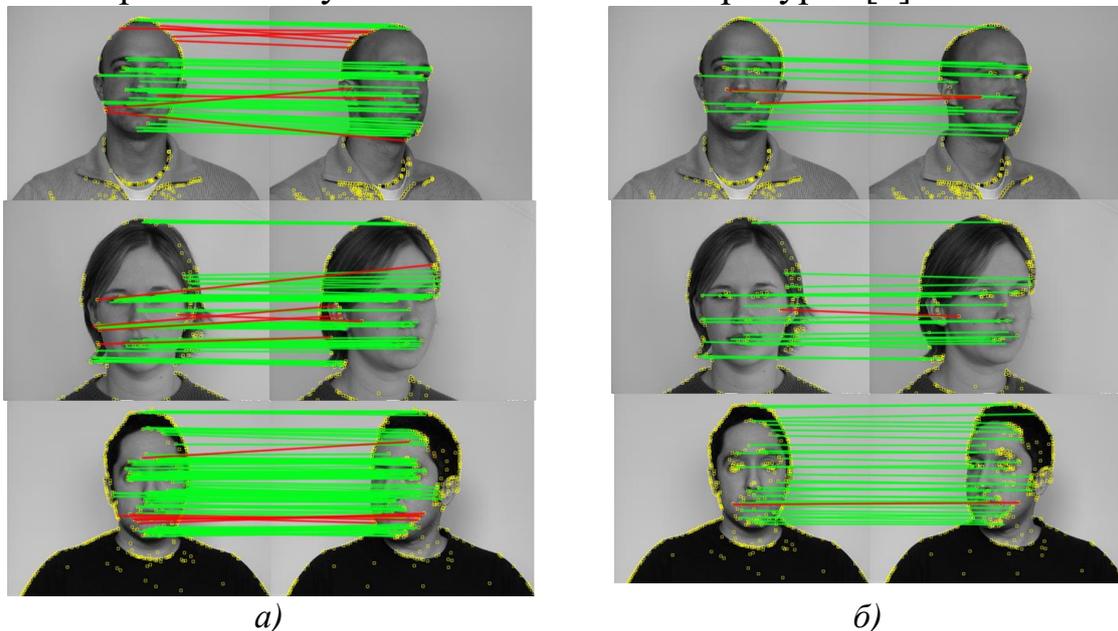


Рис. 3.16. Исследование алгоритма нахождения соответствий на изображениях лиц с использованием тепловых ядер на графах:
 а – нахождение соответствий с использованием метода SURF;
 б – нахождение соответствий с использованием разработанного алгоритма

При повороте объекта на угол до 50° количество ложных соответствий для разработанного алгоритма не превышает 10 %. Для алгоритма SURF это значение может достигать 50 % (рис. 3.17).

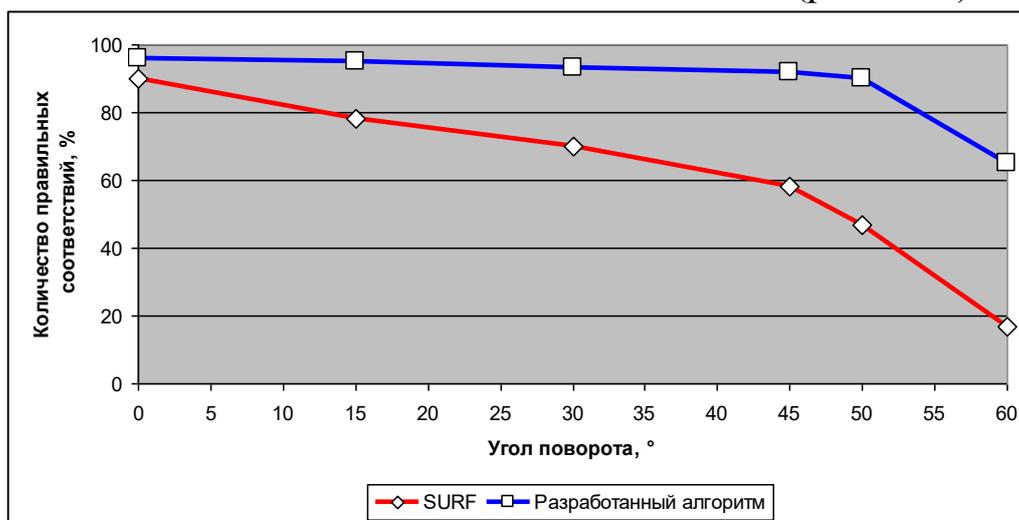


Рис. 3.17. Зависимость количества правильных соответствий от угла поворота объектов

3.3. Исследование алгоритма обнаружения головы человека на изображениях на основе вложения графа особенностей в векторное пространство

В ходе исследования были использованы следующие базы данных изображений:

- Carnegie Mellon University (CMU) Image Database;
- база данных Массачусетского технологического института MIT-CVCL Face Recognition Database.

Было проведено исследование разработанного алгоритма на предмет правильного отнесения объектов к определенным группам на основе вложения графов в векторное пространство. Также выполнялось сравнение с алгоритмом вложения, в основе которого лежит использование всех особых точек сцены [80].

В ходе тестирования использовались изображения из базы данных CMU. На используемых изображениях меняется угол поворота объекта в пространстве (рис. 3.18).

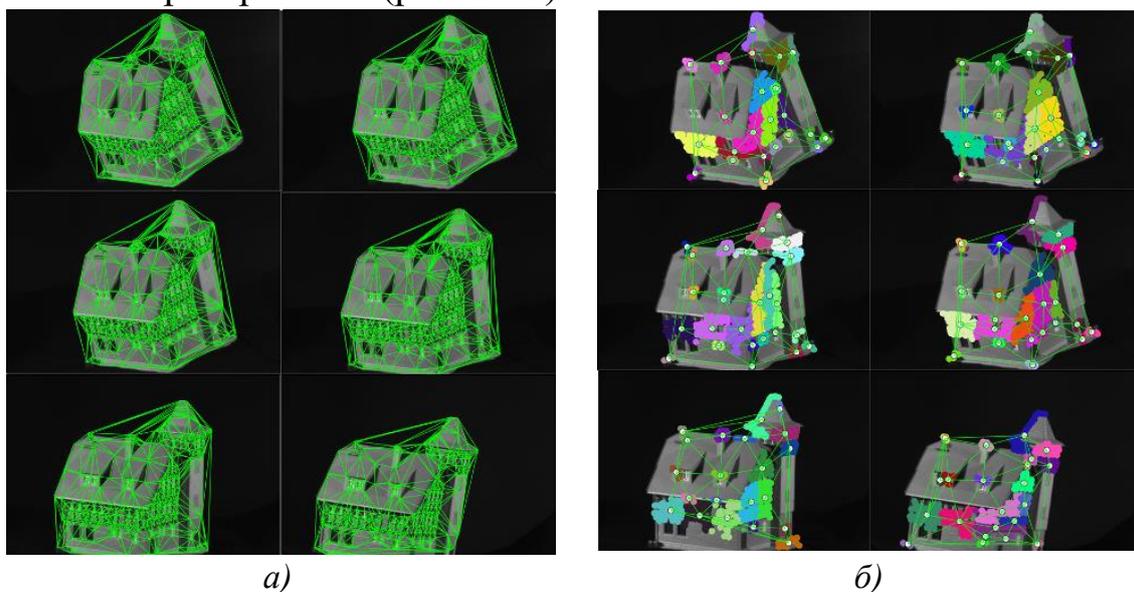


Рис. 3.18. Построение графа объекта по тестовым изображениям:
а – вершинами графа являются все особенности изображения;
б – вершинами графа являются центры масс сегментов особенностей

Были получены следующие результаты вложения графов с использованием метода многомерного шкалирования (рис. 3.19). Несмотря на то, что используются изображения одного объекта, в ситуациях, когда вершинами графа являются все особые точки изображе-

ния, алгоритм выделяет две разные группы. Напротив, разработанный алгоритм, рассматривая в качестве вершин графа центры масс сегментов особенностей, отнес объекты к одной группе [8, 19].

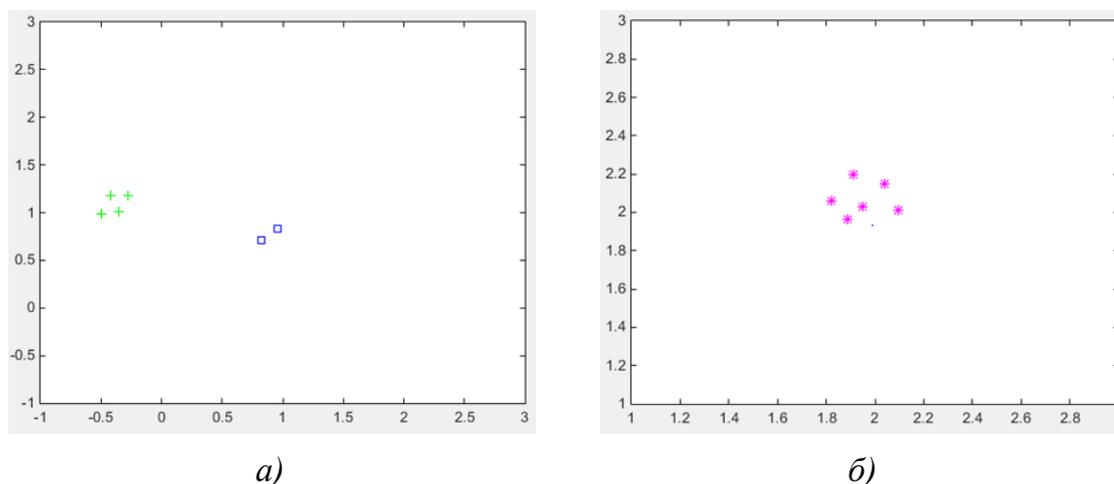


Рис. 3.19. Результаты вложения графов, изображенных на рисунке 3.18, в векторное пространство: а – вершинами графа являются все особые точки изображения; б – вершинами графа являются центры масс сегментов особенностей

При проецировании графов, построенных по изображениям лица, наблюдаются аналогичные результаты. На рисунке 3.20 показано построение графов по изображениям лиц. На рисунке 3.21 показаны результаты вложения графов изображений лиц, представленные с помощью метода многомерного шкалирования. Разработанный алгоритм вложения разделил на группы изображения лиц разных людей.

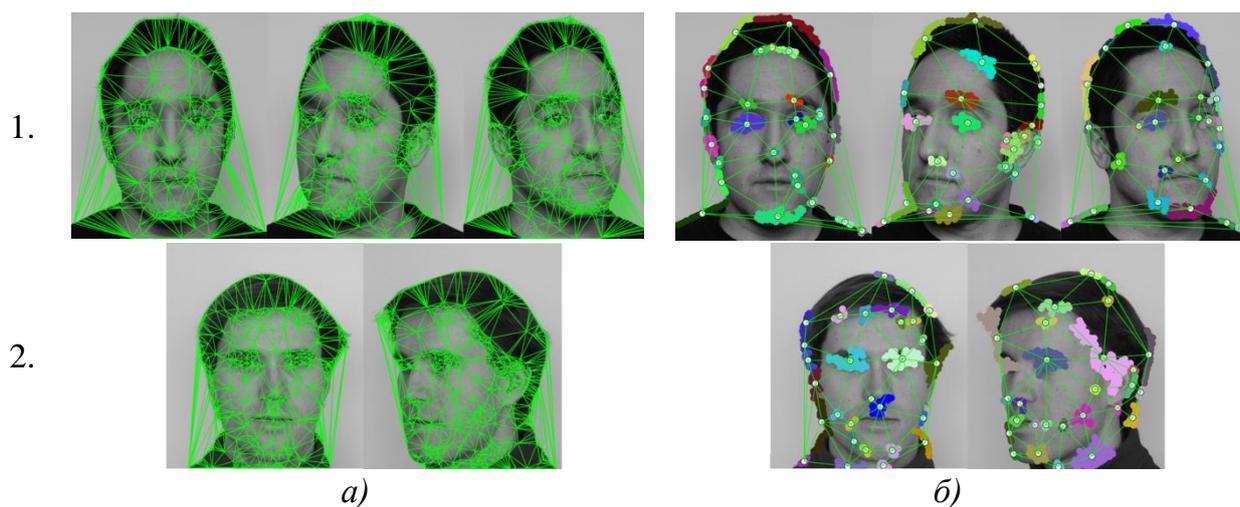


Рис. 3.20. Построение графов по изображениям лиц: а – вершинами графа являются все особенности изображения; б – вершинами графа являются центры масс сегментов особенностей

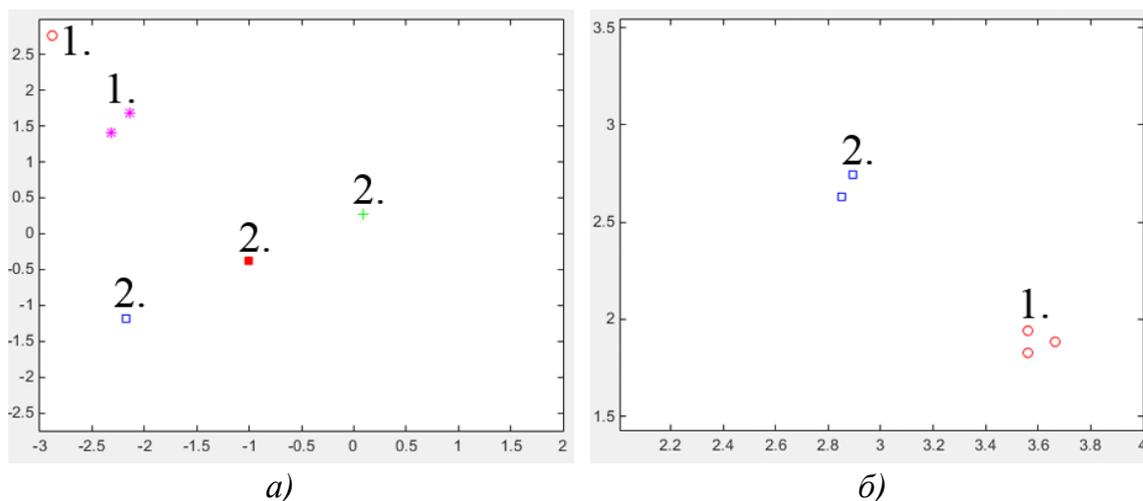


Рис. 3.21. Результаты вложения графов, изображенных на рисунке 3.20, в векторное пространство: а – вершинами графа являются все особенности изображения; б – вершинами графа являются центры масс сегментов особенностей

Основная функция алгоритма – обнаружение головы человека на изображении. В данном случае сравнение работоспособности осуществлялось с методом Виолы-Джонса [151], так как он является некоторым стандартом в области компьютерного зрения и широко применяется в задачах выделения лиц на изображениях. В основе его работы лежит использование каскадов фильтров.

В задаче определения положения и ориентации головы человека необходимость обнаружить ранее отслеживаемую голову может возникнуть в любой момент. В таких случаях голова человека может быть повернута относительно любой оси глобальных координат.

Было выполнено тестирование разработанного алгоритма на изображениях лиц из базы данных MIT-CVCL. В данном случае диапазон кивка, наклона и поворота составлял от 0 до 90°. При тестировании алгоритма для каждого лица из базы данных была составлена выборка изображений, на которых голова человека имела различную ориентацию в пространстве. Были получены следующие результаты (рис. 3.22). Главной особенностью разработанного алгоритма является то, что при любом значении угла поворота в плоскости XOY правильное обнаружение лица было не ниже 90%. Это связано с тем, что спектральные характеристики графа не зависят от маркировки его вершин [8].

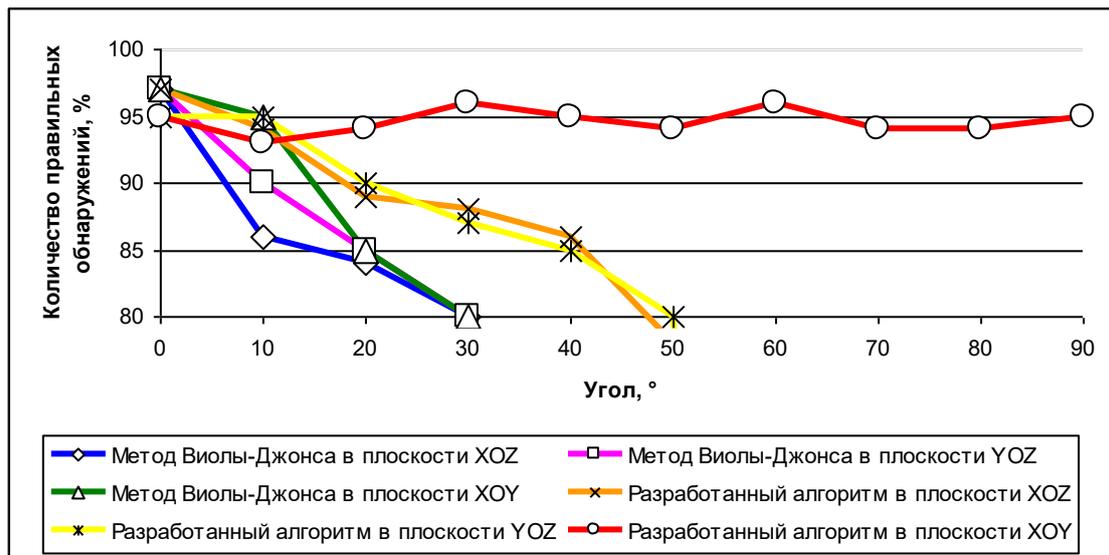


Рис. 3.22. Графики зависимости количества правильных обнаружений головы от угла поворота в плоскостях XOY , XOZ и YOZ для метода Виолы-Джонса и разработанного алгоритма

При использовании метода Виолы-Джонса с увеличением углов поворота, кивка и наклона снижается точность правильного обнаружения лица человека. Напротив, разработанный алгоритм обладает большей точностью. В случаях вращения объекта в плоскости XOY точность обнаружения остается неизменной.

Достоинством подхода является полная инвариантность к повороту изображения на плоскости, так как спектральные характеристики графа не зависят от маркировки его вершин. Также вложение графов в векторное пространство позволяет осуществить приближенное сравнение структур. Сравнение векторов осуществляется намного быстрее и точнее, чем нахождение структурных соответствий.

3.4. Исследование модели определения положения и ориентации головы человека с использованием стереоизображений

В ходе исследования подход на основе разработанной модели определения положения и ориентации головы сравнивался с известными методами (табл. 3.1) [10, 19]:

- методы шаблонов (алгоритм сопоставления шаблонов) [113];
- методы массивов детекторов (алгоритм быстрого обнаружения лица человека с помощью каскада детекторов) [162];

- методы нелинейной регрессии (алгоритм определения положения и ориентации головы с помощью регрессии и вероятностного подхода) [76];
- методы вложенных многообразий (алгоритм определения положения и ориентации головы с использованием результатов вложения многообразий) [147];
- гибкие модели (алгоритм определения положения и ориентации головы с помощью активной модели внешнего вида и стереосистемы) [68];
- геометрические методы (алгоритм определения положения и ориентации головы с помощью выделения лицевых особенностей и расчета отношений между ними) [88];
- гибридные методы: статические шаблоны и отслеживание [118];
- гибридные методы: динамические шаблоны и отслеживание [53];
- гибридные методы: вероятностная модель и отслеживание [67].

Таблица 3.1. *Исследование модели определения положения и ориентации головы человека с использованием стереоизображений*

Метод	Диапазон (поворот, кивок, наклон)	Ошибка (поворот, кивок, наклон)	Автоматическая инициализация	Обработка перекрытий
Методы шаблонов (Moore, 2011)	$\pm 80^\circ$; $\pm 25^\circ$; $\pm 10^\circ$.	4,1°; 2,3°; 2,4°.	Нет	Нет
Методы массивов детекторов (Zhang, 2007)	$\pm 90^\circ$; $\pm 30^\circ$; $\pm 30^\circ$.	9°; 8°; 9°.	Нет	Нет
Методы нелинейной регрессии (Drouard, 2015)	$\pm 90^\circ$; $\pm 90^\circ$; 0.	8,4°; 8,5°; 0.	Нет	Нет
Методы вложенных многообразий (Sundararajan, 2015)	$\pm 50^\circ$; $\pm 45^\circ$; $\pm 20^\circ$.	4,9°; 5,9°; 4,7°.	Да	Да

Окончание табл. 3.1

Метод	Диапазон (поворот, кивок, наклон)	Ошибка (поворот, кивок, наклон)	Автоматическая инициализация	Обработка перекрытий
Гибкие модели (Chen, 2014)	$\pm 55^\circ$; $\pm 45^\circ$; $\pm 45^\circ$.	$3,8^\circ$; $3,2^\circ$; $1,4^\circ$.	Нет	Нет
Геометрические методы (Natem, 2015)	$\pm 45^\circ$; $\pm 45^\circ$; $\pm 45^\circ$.	$2,65^\circ$; $2,56^\circ$; $3,01^\circ$.	Да	Нет
Гибридные мето- ды: статические шаб- лоны и отслежива- ние (Murphy-Chutorian, 2008)	$\pm 40^\circ$; $\pm 40^\circ$; $\pm 20^\circ$.	$3,2^\circ$; $1,6^\circ$; $1,3^\circ$.	Да	Нет
Гибридные мето- ды: динамические шаблоны и отсле- живание (Baltrusaitis, 2011)	$\pm 55^\circ$; $\pm 40^\circ$; $\pm 25^\circ$.	$3,39^\circ$; $4,67^\circ$; $2,38^\circ$.	Да	Нет
Гибридные мето- ды: вероятностная мо- дель и отслежива- ние (Cabrerera, 2015)	$\pm 40^\circ$; $\pm 30^\circ$; $\pm 20^\circ$.	$1,5^\circ$; $1,7^\circ$; $0,5^\circ$.	Да	Нет
Разработанный подход	$\pm 50^\circ$; $\pm 50^\circ$; $\pm 25^\circ$.	$3,1^\circ$; $2,75^\circ$; $0,8^\circ$.	Да	Да

Сравнение осуществлялось по следующим параметрам:

- диапазон определяемых положений головы;
- максимальное значение ошибки;
- наличие автоматической инициализации;
- обработка ситуаций взаимного перекрытия объектов в сцене.

Из табл. 3.1 видно, что большинство рассматриваемых подходов обладает сходным диапазоном определяемых углов поворота, кивка и

наклона. В то же время при использовании методов массивов детекторов максимальная ошибка составляет 9° . Максимальная ошибка при использовании разработанной модели не превышает $3,1^\circ$. Предложенный набор алгоритмов позволяет осуществлять автоматическую инициализацию. Также возможен поиск отслеживаемого объекта в случае его потери из области видимости камеры, при возникновении взаимных перекрытий в сцене. Производительность совокупности разработанных алгоритмов составляет 25 кадров в секунду, что сопоставимо с существующими аналогами (рис. 3.23). В то же время предложенная модель обладает большей точностью и набором функциональных возможностей по сравнению с аналогами [5, 7, 11].

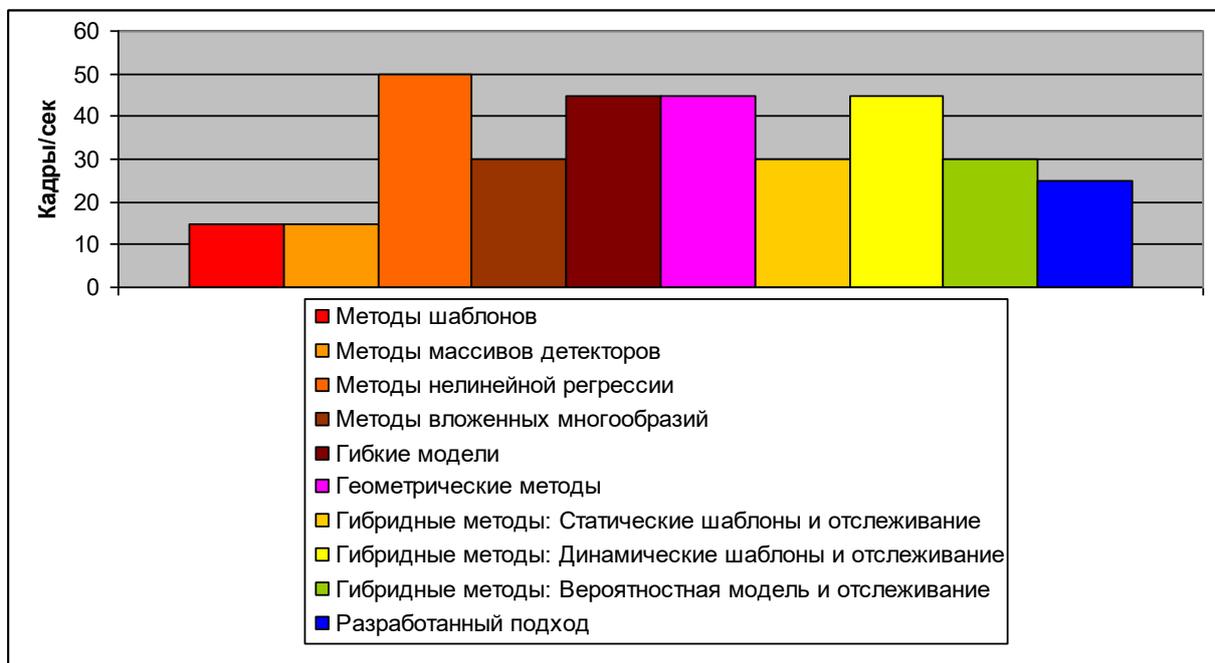


Рис. 3.23. Сравнение производительности разработанного подхода определения положения и ориентации головы человека с аналогами

Глава 4. ПРИМЕНЕНИЕ РАЗРАБОТАННЫХ МОДЕЛИ И АЛГОРИТМОВ В ТРЕНАЖЕРАХ ТРАНСПОРТНЫХ СРЕДСТВ

4.1. Структура системы определения положения и ориентации головы человека по стереоизображениям

Разработанная система реализована в среде программирования Visual Studio Community 2015 с использованием языка программирования C# [11, 12]. При разработке системы использовались как стандартные функции среды, так и сторонние библиотеки:

- библиотека OpenCV (работа с видеоизображениями);
- библиотека ALGLIB (использование расширенных математических функций);
- интерфейс WebCamLib (работа с видеокамерой);
- функции калибровки камеры Camera Calibration Tools.

Для сохранения настроек программы использовалась технология записи в «ini» файлы.

Разработанную систему можно представить в виде следующей структуры (рис. 4.1).

Модуль стереосистемы. Используется для согласованной работы видеокамер, входящих в состав стереоустановки. Включает в себя алгоритм калибровки.

Модуль инициализации. Используется для осуществления процесса начального поиска головы пользователя на изображении и вычисления дескриптора для последующего обнаружения. Содержит в себе алгоритм спектральной кластеризации особенностей изображений с ограничениями и алгоритм вычисления структурного дескриптора.

Модуль отслеживания. Осуществляет слежение за дескриптором головы пользователя между кадрами видеопоследовательности. Содержит в себе алгоритм нахождения соответствий на изображениях с использованием тепловых ядер на графах.

Модуль вычисления углов ориентации головы. С помощью разработанной модели вычисляются текущие значения углов ориентации и трехмерные координаты головы в пространстве.

Модуль обнаружения головы. Осуществляет поиск объекта, если он потерян из области обзора стереоустановки. Содержит алгоритм обнаружения головы на основе структурного дескриптора.

Модуль передачи информации. Представляет собой интерфейс между системой определения положения и ориентации головы человека и транспортным тренажером.

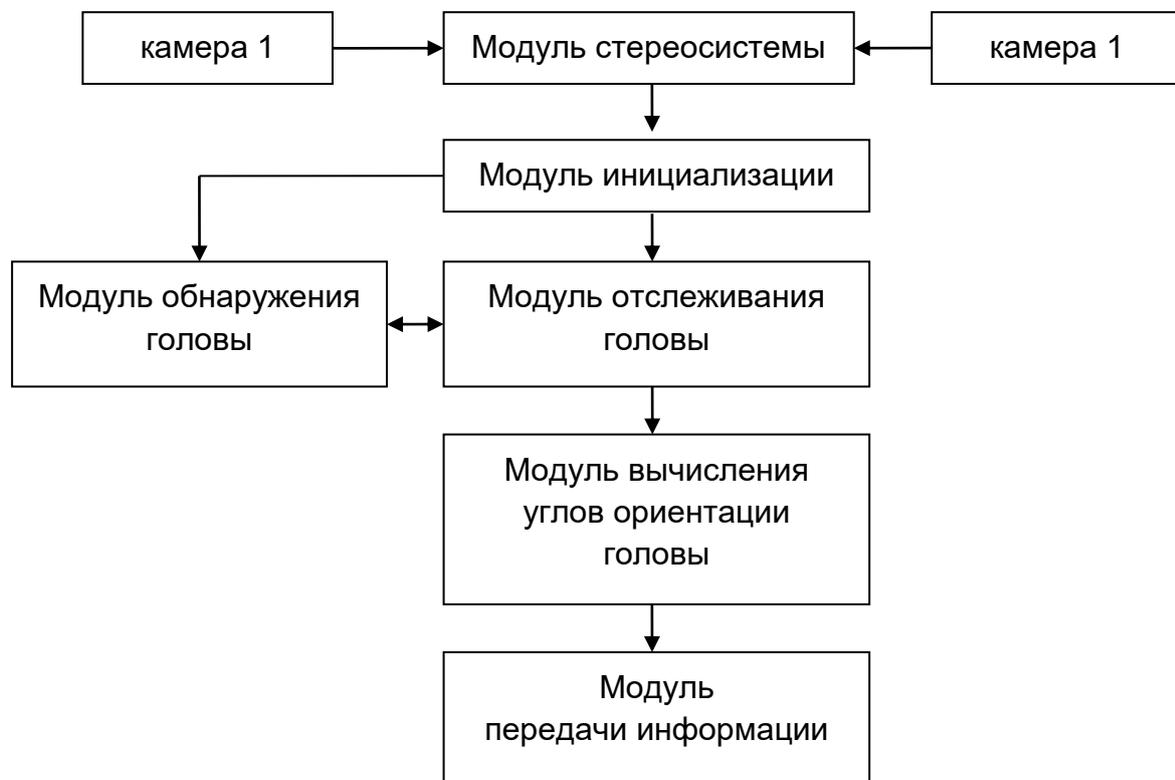


Рис. 4.1. Структура разработанной системы определения положения и ориентации головы человека по стереоизображениям с использованием графов

В состав аппаратного обеспечения системы входят две веб-камеры Logitech C300, с матрицей 1.3 Мп. Они крепятся на специальном кронштейне, который устанавливается на мониторе (рис. 4.2). Каждая камера соединяется с компьютером с помощью кабеля USB 2.0.



Рис. 4.2. Фиксация стереосистемы на мониторе тренажера

4.2. Методика настройки разработанной системы определения положения и ориентации головы человека по стереоизображениям

Перед работой система определения положения и ориентации головы устанавливается на компьютер. Для корректной работы системы необходимо, чтобы были выполнены минимальные системные требования: операционная система: Windows 7; библиотека Microsoft.Net Framework 3.5; процессор Intel Core i5-4200M @ 2.50Гц; оперативная память (ОЗУ) 2ГБ; свободное место на диске 500МБ.

Если необходимая библиотека Microsoft .Net Framework будет отсутствовать на момент установки, то будет предложено скачать программное обеспечение с официального сайта Microsoft. Во время установки предлагается выбрать камеры, которые будут объединены в стереосистему. Также это можно будет сделать в настройках приложения. Главная форма программы изображена на рисунке 4.3.

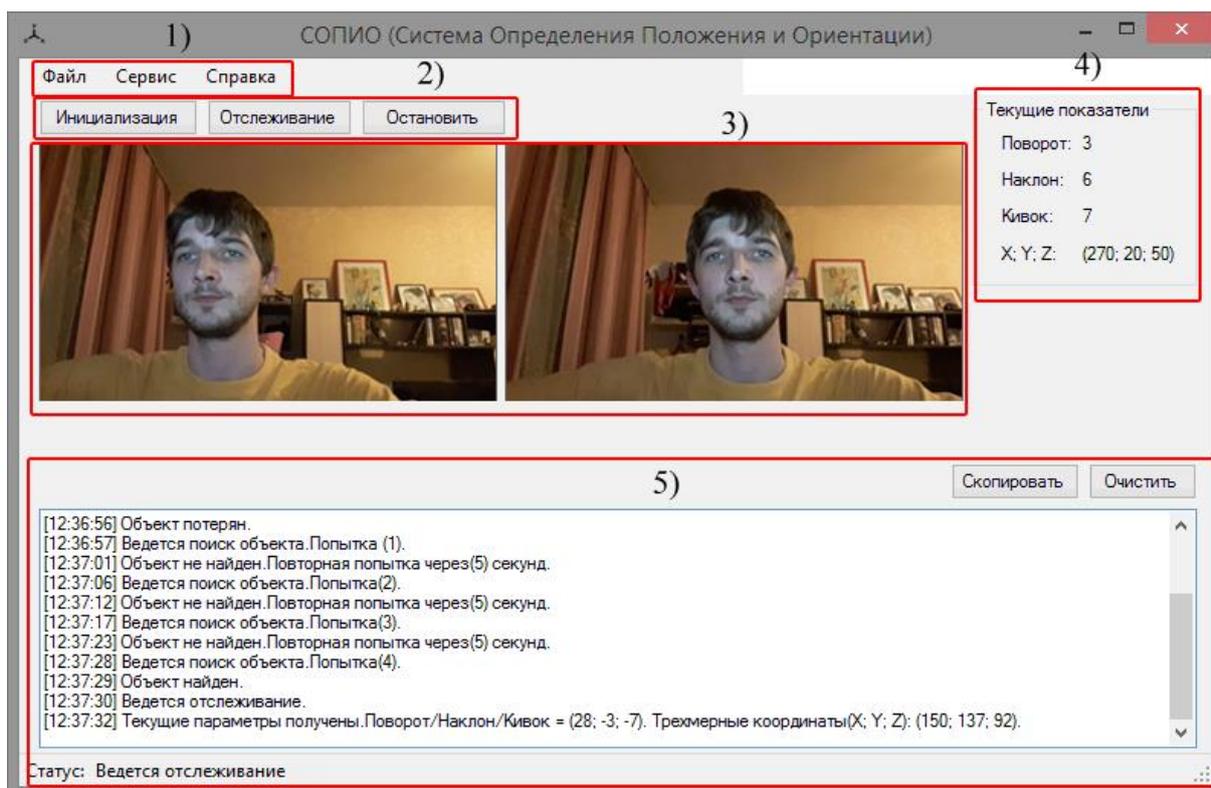


Рис. 4.3. Главная форма программы: 1) верхнее меню; 2) кнопки запуска модулей программы; 3) изображения с камер; 4) текущие показатели; 5) отображение информации о работе системы

На главной форме располагаются следующие элементы:

1. Верхнее меню;
2. Кнопки запуска: «Инициализация», «Отслеживание», «Остановить»;
3. Изображения, получаемые с двух камер стереоустановки;
4. Текущие показатели параметров положения и ориентации головы;
5. Текстовое поле отображения и записи информации обо всех происходящих в системе событиях с кнопками «Скопировать» и «Очистить».

Верхнее меню состоит из нескольких разделов (рис. 4.4):

1. «Файл». В этом меню находятся элементы: «Новый проект»; «Загрузить проект»; «Сохранить проект»; «Выход».
2. «Сервис». Данное меню состоит из элементов: «Калибровка камеры»; «Параметры».
3. «Справка». В данном меню можно просмотреть предоставленную о системе справочную информацию.

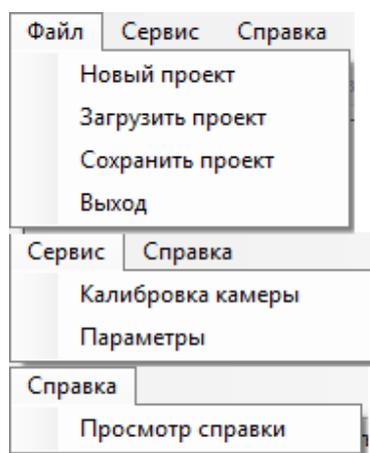


Рис. 4.4. Содержание верхнего меню и его подпунктов

«Новый проект». Предоставляется выбор директории, где будут располагаться все файлы проекта. В файлы проекта входят файл настроек инициализации, файл параметров калибровки стереоустановки, файлы видеопоследовательности.

«Загрузить проект». Позволяет загрузить ранее созданный проект.

«Сохранить проект». Сохраняет текущие файлы видеопоследовательности, которые по умолчанию располагались в папке для вре-

менных файлов, а также файлы настройки инициализации и калибровки стереоустановки.

«Выход». Закрывает приложение и очищает директорию для временных файлов.

«Калибровка камеры». Калибровка стереоустановки необходима для успешной работы системы. На форме расположены области, в которых будут показаны изображения для каждой из камер с подписями в виде имен файлов (рис. 4.5). Для каждой камеры должно быть три изображения.

Кнопка «Загрузить» позволяет выбрать файлы изображений, на основе которых будет проведена калибровка.

Нажатие на кнопку «Калибровать» запускает калибровку, если загружены все шесть изображений. Параметр «Ширина ячейки (мм)» содержит значения по умолчанию, если величины заданы в соответствующем пункте меню настроек. В случае успешной калибровки пользователь увидит надпись «Параметры камеры получены».

Полученные параметры калибровки можно сохранить в отдельный файл с помощью кнопки «Сохранить файл калибровки», который можно использовать в других проектах с той же стереосистемой. Для этого можно воспользоваться кнопкой «Загрузить файл калибровки».

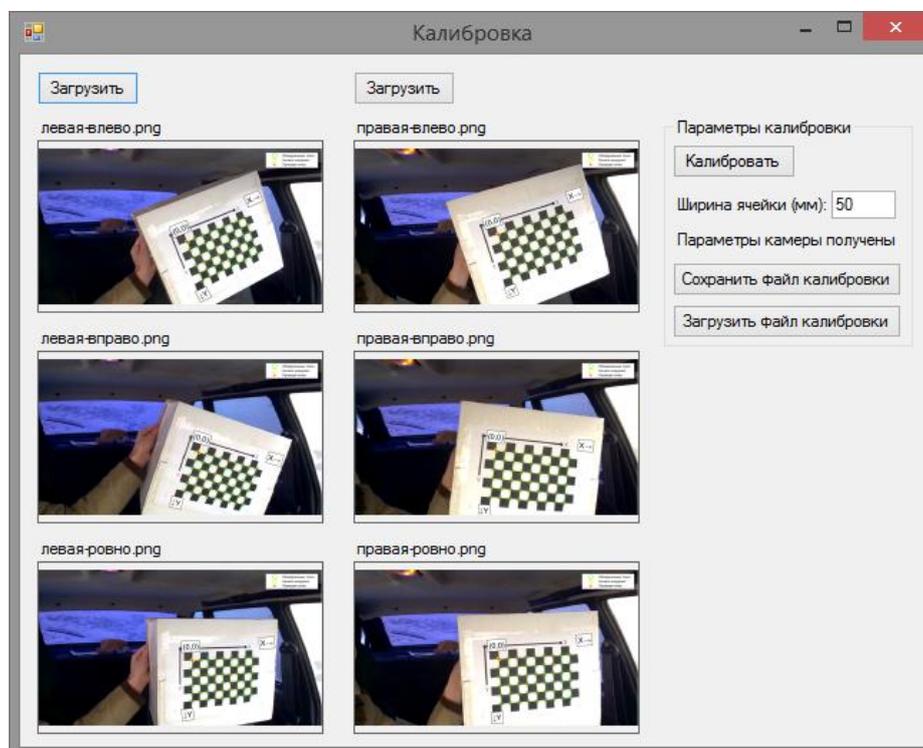


Рис. 4.5. Диалоговое окно калибровки стереосистемы

Для запуска процесса калибровки необходимо подготовить калибровочный шаблон, на изображении которого можно выделить особые точки (рис. 4.6). Размеры шаблона: ширина 27 см, высота 20 см.

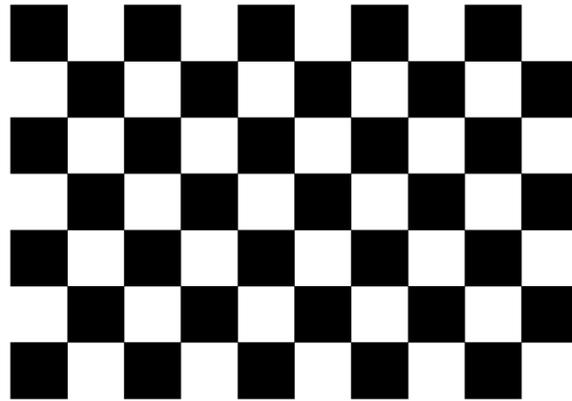


Рис. 4.6. Калибровочный шаблон

Затем шаблон необходимо разместить в области видимости стереосистемы и сделать по три снимка для каждой из камер. Полученные изображения необходимо загрузить в программу. Для загрузки трех изображений в окне «Открытие» необходимо выбрать все три файла для одной камеры (кнопка «Shift» или «Ctrl») (рис. 4.7).

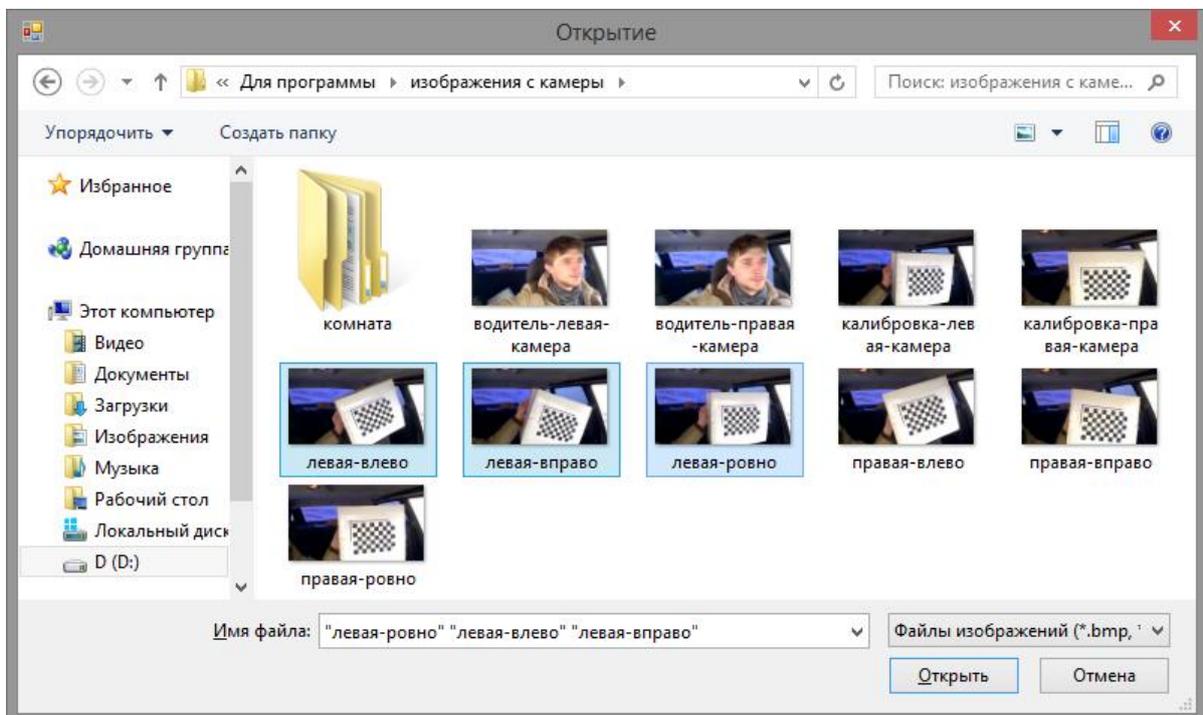


Рис. 4.7. Диалоговое окно открытия файла

После выполняется процесс калибровки, при котором будут вычислены внутренние и внешние параметры стереосистемы. С помощью этих параметров станет возможным определять трехмерные координаты отслеживаемой головы пользователя. Пример отображения стенда с результатами калибровки приведен на рис. 4.8.

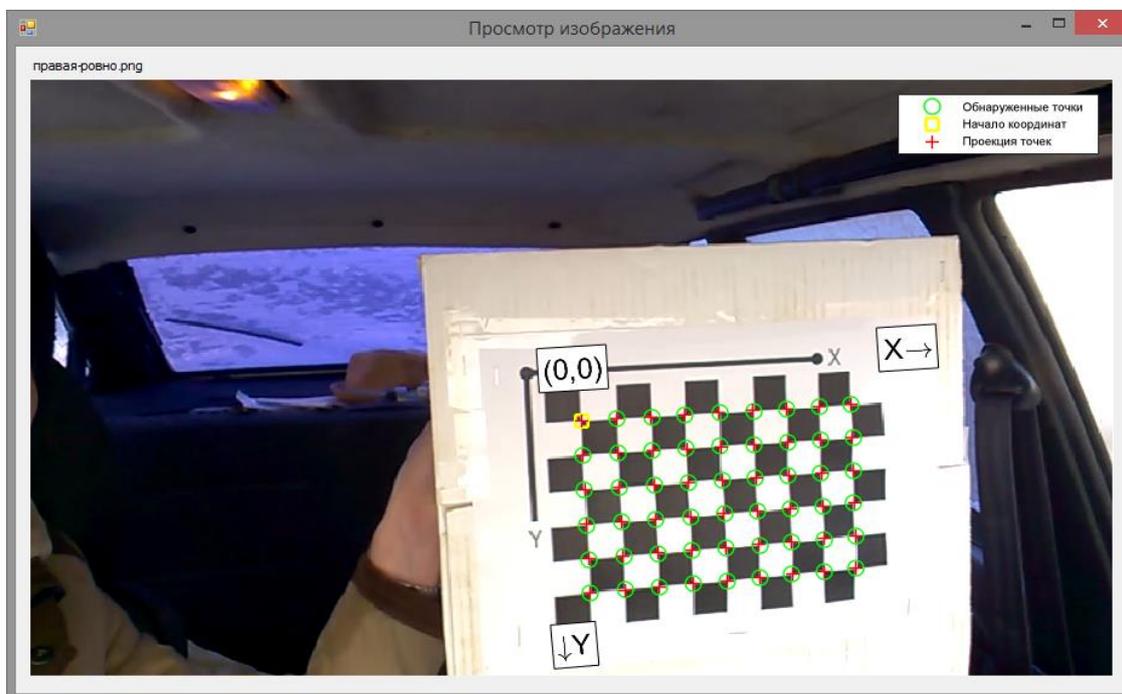


Рис. 4.8. Пример отображения калибровочного стенда

На форме «Параметры» располагаются все настройки системы (рис. 4.9):

- настройки стереосистемы. Включают в себя выбор левой и правой камер стереосистемы. Отображаются все доступные в данный момент камеры в операционной системе. Также можно выбрать разрешение, яркость и контрастность получаемых видеоизображений;

- настройки калибровки позволяют указать ширину ячейки калибровочного шаблона;

- настройка записи и сохранения видеопоследовательности. Указывается директория по умолчанию, в которую будут записываться временные файлы видеопоследовательности. Параметр «Продолжительность (мин.)» позволяет установить время последовательной записи;

- кнопка «По умолчанию» вернет предустановленные параметры.

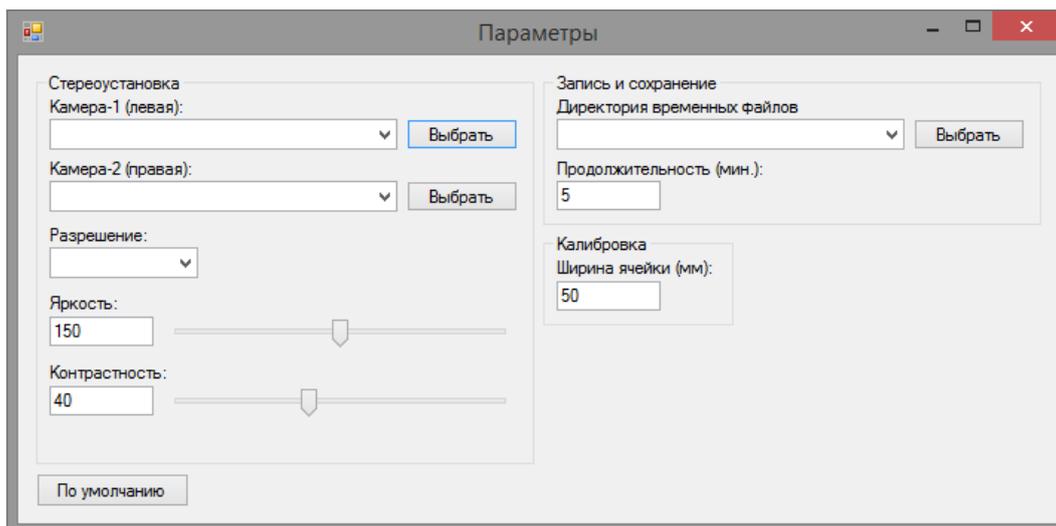


Рис. 4.9. Диалоговое окно установки параметров системы

«Справка». Содержит в себе описание всех функциональных возможностей программы, а также подсказки о настройке и работе всей системы.

«Инициализация». Запуск режима инициализации необходим перед непосредственной работой системы (рис. 4.10). На протяжении данного процесса от пользователя требуется находиться в неподвижном состоянии. Относительно стереосистемы голова повернута анфас. После инициализации система формирует численное описание дескриптора области головы на изображении. Также определяется начальное положение головы, относительно которого в дальнейшем будут вычисляться текущие показатели. Численное представление дескриптора определяется один раз для каждой камеры.

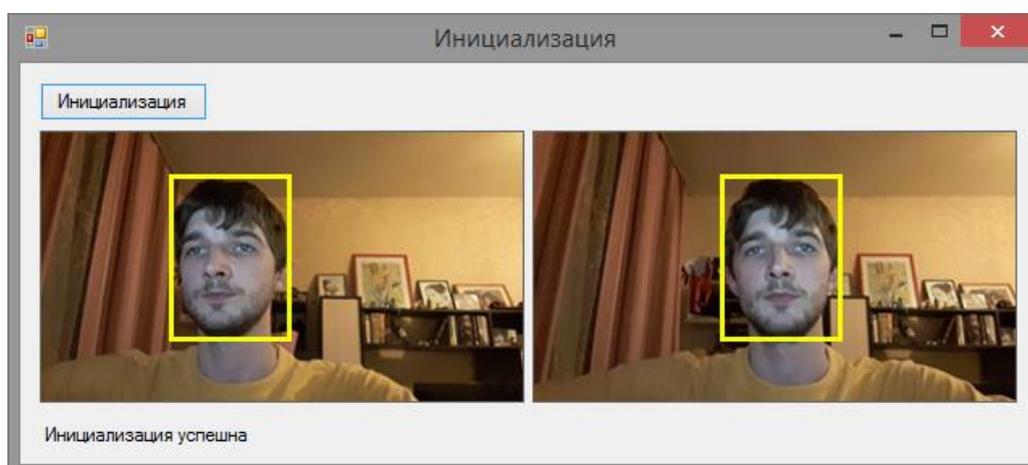


Рис. 4.10. Диалоговое окно инициализации

«Отслеживание». После процесса инициализации возможно отслеживание головы пользователя на видеопоследовательности (рис. 4.11). На данном этапе производится сопоставление изображений, чтобы не потерять отслеживаемый объект. В программе на изображениях, получаемых с камер стереоустановки, отмечается область, которая представляет собой голову пользователя. Во время работы видеопоследовательность сохраняется в указанный на этапе настройки каталог, чтобы в дальнейшем можно было загрузить ее в программу и провести анализ. Продолжительность видеопоследовательностей указывается в меню «Параметры».

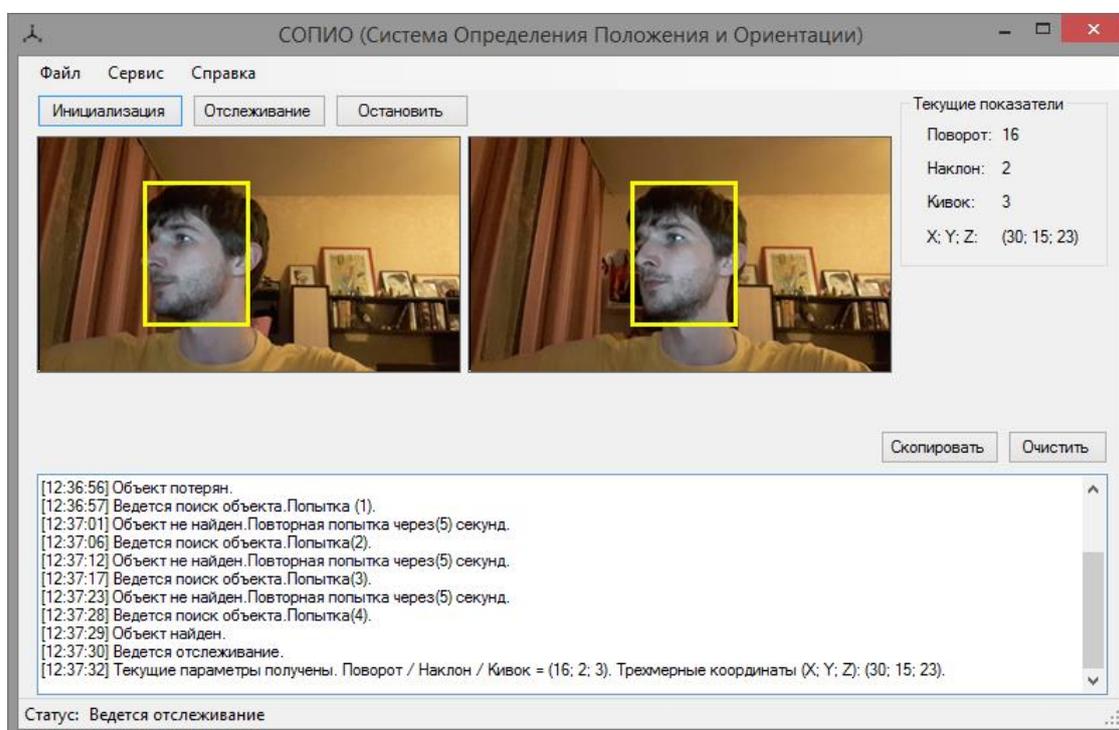


Рис. 4.11. Отслеживание головы пользователя по видеопоследовательности

«Текущие показатели». Отображается вся информация по текущим значениям углов поворота, наклона, кивка. Углы рассчитываются относительно начального положения головы, которое определено на этапе инициализации.

«Журнал событий». Отображается информация о текущих процессах, происходящих в системе. Над текстовым полем располагаются кнопки «Скопировать» и «Очистить». При нажатии на кнопку «Скопировать» осуществляется копирование всей информации о событиях, произошедших в системе, в буфер обмена. При нажатии на кнопку «Очистить» происходит очистка текстового поля.

При функционировании системы могут произойти следующие события (рис. 4.12):

– «Не обнаружены камеры». Этот статус означает, что при установке не были выбраны камеры по умолчанию. Для решения проблемы необходимо перейти в пункт «Параметры» и выбрать две камеры для стереоустановки;

– «Камеры не откалиброваны». Перед использованием системы необходимо провести калибровку камер;

– «Объект не инициализирован». Перед началом отслеживания головы пользователя необходимо пройти этап инициализации;

– «Ведется отслеживание». Камеры откалиброваны, объект инициализирован, осуществляется отслеживание объекта;

– «Текущие параметры получены». Во время отслеживания головы пользователя система получает текущие значения углов кивка, наклона, поворота и трехмерных координат головы в пространстве;

– «Объект потерян». Пользователь покинул область видимости или произошло перекрытие каким-либо объектом. На данном этапе запустится режим автоматического поиска головы пользователя;

– «Ведется поиск объекта». Выполняется алгоритм обнаружения головы на изображении;

– «Объект не найден». В процессе поиска не был обнаружен ни один объект, соответствующий изображению головы. Поиск будет автоматически повторяться каждые 5 секунд до тех пор, пока голова не будет найдена;

– «Объект найден». В результате поиска голова была обнаружена. Система продолжает работу в режиме отслеживания.

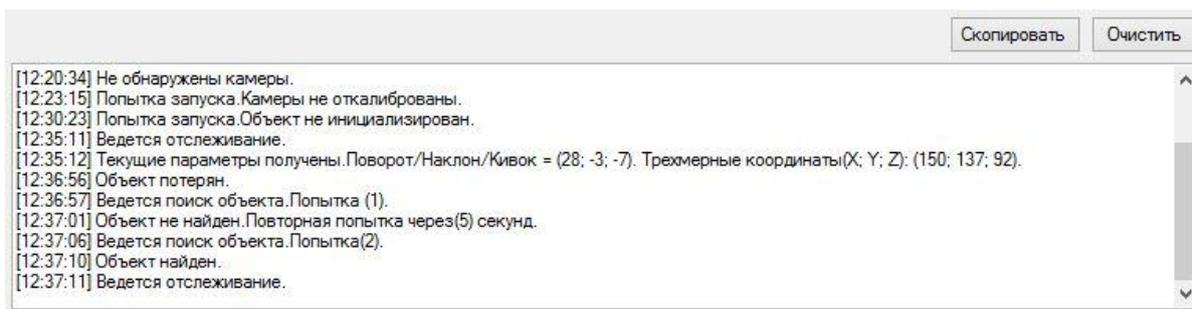


Рис. 4.12. Диалоговое окно отображения происходящих событий

4.3. Внедрение системы определения положения и ориентации головы человека по стереоизображениям

Разработанная система определения положения и ориентации головы человека по стереоизображениям используется в тренажере механика-водителя. Акт внедрения разработанных алгоритмов на «АО» Муромское СКБ приведен в приложении В.

Основное назначение системы – определение положения и ориентации головы человека. Полученная информация помогает более реалистично отображать виртуальное пространство за счет визуализации изменений, возникающих на экране тренажера при поворотах и смещениях головы обучаемого. Программное обеспечение позволяет:

- выделить голову пользователя на изображении на этапе инициализации;
- отследить перемещение головы пользователя на видеопоследовательности;
- произвести обнаружение головы в случае потери объекта в области видимости видеокамеры;
- определить координаты головы в пространстве;
- определить значения углов поворота, наклона и кивка головы.

Система была интегрирована в тренажер механика-водителя в виде одного из модулей (рис. 4.13). В целом тренажер механика-водителя состоит из нескольких компонентов. Условно состав модулей можно разделить на следующие подгруппы: модули кабины механика-водителя; модули рабочего места инструктора; программные

модули, обеспечивающие взаимодействие первой и второй группы модулей, система определения положения и ориентации головы человека.



Рис. 4.13. Система определения положения и ориентации головы человека по стереоизображениям в структуре тренажера механика-водителя

Кабина механика-водителя содержит следующие модули:

- модуль органов управления транспортным средством;
- модуль имитации слуховых и акселерационных ощущений;
- модуль средств отображения внешней визуальной обстановки.

Модуль средств отображения внешней визуальной обстановки осуществляет визуализацию виртуального пространства в соответствии с расположением транспортного средства.

Система определения положения и ориентации вычисляет значения углов поворота и смещений головы и передает их в модуль средств отображения внешней визуальной обстановки. В соответствии с вычисленными значениями изменяется отображение внешнего пространства в кабине тренажера.

Модуль системы объективного контроля позволяет передавать управляющие команды от инструктора к водителю.

Перед началом работы тренажера необходимо провести процесс калибровки стереоустановки с помощью соответствующего калибровочного шаблона.

Перед работой обучаемый размещается в кабине тренажера. Происходит ознакомление с элементами управления и отображения информации. Инструктор запускает процесс инициализации и передает соответствующие команды пользователю. Когда инициализация проведена, инструктор получает уведомление. После инициализации осуществляется процесс обучения на тренажере. Инструктор выбирает необходимые упражнения и заданный маршрут. Водитель выполняет поставленные перед ним задачи по управлению транспортным средством.

ЗАКЛЮЧЕНИЕ

В заключении сформулированы основные результаты, полученные в ходе работы:

1. Обоснована актуальность создания системы определения положения и ориентации головы человека для тренажеров транспортных средств, заключающаяся в необходимости изменять обзор внешней визуальной обстановки в зависимости от перемещений головы.

2. Проведены обзор и анализ методов и систем определения положения и ориентации объектов, показывающие необходимость разработки новых модели и алгоритмов, позволяющих определять текущую позицию головы человека на основе изображений для тренажеров транспортных средств.

3. Разработана и исследована модель определения положения и ориентации головы человека на основе стереоизображений, которая отличается предварительной трехмерной реконструкцией объекта, кластеризацией особенностей лица с использованием сетки пропорций, нахождением соответствий на изображениях с использованием тепловых ядер на графах, обнаружением лица с использованием структурного дескриптора и позволяет отслеживать кивок и поворот головы до 50° , наклон до 25° , что превосходит известные подходы.

4. Разработан и исследован алгоритм кластеризации особенностей на основе спектральной теории графов и ограничений для выделения лица человека на изображениях, отличающийся использованием сетки пропорций и позволяющий повысить точность выделения объектов до 27 % по сравнению подходами без ограничений.

5. Разработан и исследован алгоритм нахождения соответствий на изображениях с использованием тепловых ядер на графах, отличающийся использованием иерархических структур, построенных по наиболее стабильным областям изображения, и позволяющий уменьшать количество «выбросов» до 10% при угле поворота объекта в пространстве до 50° , что превосходит метод SURF.

6. Разработан и исследован алгоритм обнаружения головы человека на изображениях на основе вложения графов в векторное пространство, отличающийся использованием центров масс кластеров особенностей изображения в качестве вершин графа и обладающий инвариантностью к повороту изображения на плоскости, а также способностью обнаруживать объекты с углом поворота в пространстве до 50° , что превосходит известные подходы.

7. Результаты проведенных исследований апробированы на предприятии АО «Муромское СКБ». Реализована система определения положения и ориентации головы человека на основе стереоизображений, позволяющая повысить эффект присутствия обучаемого в синтезируемой виртуальной среде транспортного тренажера за счет изменения поля зрения.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Алешин, В.П. Особенности реализации виртуального окружения для тренировки сложных режимов пилотирования в учебно-тренажерных комплексах [Текст] / В.П. Алешин, В.О. Афанасьев, А.С. Клименко, С.В. Клименко, В.Н. Пугач, С.И. Ротков, А.Д. Сандлер, В.Ф. Уразметов // Приволжский научный журнал. – 2014. – №. 3 (31). – С. 91 – 99.

2. Алпатов, Б.А. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление [Текст] / Б.А. Алпатов, П.В. Бабаян, О.Е. Балашов, А.И. Степашкин. – М.: Радиотехника, 2008. – 176 с.

3. Бабенко, В.С. Виртуальная реальность [Текст]: Толковый словарь терминов / В.С. Бабенко. – СПб.: ГУАП, 2006. – 87 с.

4. Баринов, А.Е. Алгоритм нахождения соответствий на изображениях с использованием тепловых ядер на графах [Текст] / А.Е. Баринов // Научно-технический вестник Поволжья. – 2016. – № 4 – С. 66-68.

5. Баринов, А.Е. Алгоритм определения положения и ориентации трехмерных объектов по изображениям с использованием спектральной теории графов [Текст] / А.Е. Баринов, А.А. Захаров // Наука и образование в развитии промышленной, социальной и экономической сфер регионов России. VII Всероссийские научные Зворыкинские чтения: сб. тез. докл. Всероссийской межвузовской научной конференции. – Муром: Изд.-полиграфический центр МИ ВлГУ. – 2015. – С. 138-139.

6. Баринов, А.Е. Алгоритм спектральной кластеризации с ограничениями для выделения лица человека на изображениях [Текст] / А.Е. Баринов, А.А. Захаров, А.Л. Жизняков // Динамика систем, механизмов и машин. – 2016. – Т. 2. (1). – С. 222-228.

7. Баринов, А.Е. Алгоритмы определения положения и ориентации головы человека на основе видеоизображений с использованием спектральной теории графов [Текст] / А.Е. Баринов, А.А. Захаров // XXII Международная научно-техническая конференция «Информационные системы и технологии» (ИСТ–2016). – Н. Новгород: Нижегородский государственный технический университет имени Р.Е. Алексеева. – 2016. – С. 398.

8. Баринов, А.Е. Обнаружение головы человека на изображениях на основе вложения графов в векторное пространство [Текст] / А.Е. Баринов // Современные наукоемкие технологии. – 2016. – №. 12 (1). – С. 29-34.

9. Баринов, А.Е. Разработка алгоритма определения положения и ориентации 3D-объектов на основе извлечения градиентных особенностей изображения [Текст] / А.Е. Баринов, А.А. Захаров // Ползуновский вестник. – 2014. – № 2. – С. 50-54.

10. Баринов, А.Е. Разработка модели определения положения и ориентации головы человека по видеоизображениям [Текст] / Баринов А.Е., Захаров А.А. // Наука и образование в развитии промышленной, социальной и экономической сфер регионов России. VIII Всероссийские научные Зворыкинские чтения: сб. тез. докл. Всероссийской межвузовской научной конференции. – Муром: Изд.-полиграфический центр МИ ВлГУ. – 2017. – С. 269-270.

11. Баринов, А.Е. Разработка системы определения положения и ориентации головы человека на основе видеоизображений для тренажеров транспортных средств [Текст] / А.Е. Баринов, А.А. Захаров // Наука и образование в развитии промышленной, социальной и экономической сфер регионов России. VIII Всероссийские научные Зворыкинские чтения: сб. тез. докл. Всероссийской межвузовской научной конференции. – Муром: Изд.-полиграфический центр МИ ВлГУ. – 2016. – С. 225-226.

12. Баринов, А.Е. Система определения положения и ориентации трехмерных объектов по изображениям с использованием спектральной теории графов [Текст] / А.Е. Баринов, А.А. Захаров. – Свидетельство о государственной регистрации программы для ЭВМ. – № 2015619541; заявл. 20.07.2015; дата регистрации 04.09.2015.

13. Ваганов, В.И. Вождение автотранспортных средств [Текст] / В.И. Ваганов, А.А. Рывкин. – М.: Транспорт, 1990. – 224 с.

14. Вежневцев, В. Система управления курсором «мышь» при помощи движений головы пользователя [Текст] / В. Вежневцев, Р. Шоргин, А. Вежневцев // Труды конференции Graphicon-2006. – 2006. – С. 455-458.

15. Визильтер, Ю.В. Обработка и анализ изображений в задачах машинного зрения [Текст]: Курс лекций и практических занятий / Ю.В. Визильтер, С.Ю. Желтов, А.В. Бондаренко, М.В. Ососков, А.В. Моржин. – М.: Физматкнига, 2010. – 672 с.

16. Гридин, В.Н. Система регистрации лиц людей по изображениям камер видеонаблюдения [Текст] / В.Н. Гридин, В.С. Титов, М.И. Труфанов // Методы и средства систем обработки информации. – Курск: Юго-западный государственный университет. – 2010. – С. 90-93.

17. Желтов, С.Ю. Автоматическое выделение на цветных цифровых изображениях лица человека и его характерных черт [Текст] / С.Ю. Желтов, Ю.В. Визильтер, М.В. Ососков, С.П. Каратеев, И.В. Бекетова // Вестник компьютерных и информационных технологий. – 2005. – № 10. – С. 2-7.

18. Захаров, А.А. Использование тепловых ядер на графах для нахождения соответствий на стереоизображениях [Текст] / А.А. Захаров, А.Е. Баринов, А.Ю. Тужилкин // Наука и образование в развитии промышленной, социальной и экономической сфер регионов России. VIII Всероссийские научные Зворыкинские чтения: сб. тез. докл. Всероссийской межвузовской научной конференции. – Муром: Изд.-полиграфический центр МИ ВлГУ. – 2017. – С. 269-270.

19. Захаров, А.А. Разработка моделей и алгоритмов определения положения и ориентации головы человека по видеоизображениям с использованием графов [Текст] / А.А. Захаров, А.Е. Баринов, А.Л. Жизняков // Оптико-электронные приборы и устройства в системах распознавания образов, обработки изображений и символической информации. Распознавание – 2017: сб. материалов XII Междунар. науч.-техн. конф. – Курск: Юго-западный государственный университет. – 2017. – С. 167-169.

20. Каминский, А. Мастер вождения автомобиля [Текст] / А. Каминский. – М.: «Манн, Иванов и Фербер», 2014. – 228 с.

21. Катаев, М.Ю. Восстановление углов поворота головы человека по изображениям [Текст] / М.Ю. Катаев, Н.В. Ковалев, А.А. Грибоедов // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2012. – Т. 2. – №. 1. – С. 238-242.

22. Катаев, М.Ю. Методика оценки угла поворота головы человека по анализу изображения [Текст] / М.Ю. Катаев, Н.В. Ковалев, Н.Г. Катаева // Информатика и системы управления. – 2013. – №. 1 (35). – С. 62-68.

23. Колесниченко, Д. 150 ситуаций на дороге, советы опытного инструктора по вождению [Текст] / Д. Колесниченко. – М.: АСТ, 2009. – 302 с.

24. Крестинин, И.А. Метод особых точек в задачах поиска лиц на графических изображениях [Текст] / И.А. Крестинин, О.С. Середин // Известия Тульского государственного университета. Технические науки. – 2008. – № 3. – С. 218-227.

25. Кривцов, О.А. Отслеживание положения головы человека в пространстве на основе анализа видеокадров [Текст] / О.А. Кривцов, А.М. Корилов // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2008. – №. 2(18), ч. 2. – С. 32-39.

26. Кудряшов, П.П. Быстрый поиск человеческих лиц на изображениях [Текст] / П.П. Кудряшов, С.А. Фоменков // Вестник компьютерных и информационных технологий. – 2007. – № 7. – С. 14-17.

27. Кумов, В.С. Разработка и исследование метода оценки ракурса по координатам контрольных точек 2D изображения лица [Текст] / В.С. Кумов, А.В. Самородов // Наука и образование: научное издание МГТУ им. Н.Э. Баумана. – 2016. – № 1. – С. 78-89.

28. Куприянов, В.В. Лицо человека [Текст] / В.В. Куприянов, Г.В. Стовичек. – М.: Медицина, 1988. – 272 с.

29. Лукьяница, А.А. Цифровая обработка видеоизображений [Текст] / А.А. Лукьяница, А.Г. Шишкин. – М.: «Ай-Эс-Эс Пресс», 2009. – 518 с.

30. Масанов, А.Н. Применение диагностики тренажера в целях обучения [Текст] / А.Н. Масанов, Р.А. Симаков // Математические и технические средства обработки данных и знаний. – Ташкент: НПО "Кибернетика" АН РУз. – 1999. – С. 127-131.

31. Масанов, А.Н. Формирование требований к компьютерному имитатору визуальной обстановки применительно для водителя наземного транспорта [Текст] / А.Н. Масанов // Математические и технические средства обработки данных и знаний. – Ташкент: НПО "Кибернетика" АН РУз. – 1999. – С. 122-126.

32. Небаба, С.Г. Автоматическая оценка ракурса лица в кадре и приведение изображения к нулевым углам поворота [Текст] / С.Г. Небаба, А.А. Захарова, С.Ю. Андреев // Научная визуализация. – 2015. – Т. 7 (4). – С. 101-108.

33. Небаба, С.Г. Алгоритм построения деформируемых 3D моделей лица и обоснование его применимости в системах распознавания личности [Текст] / С.Г. Небаба, А.А. Захарова // Труды СПИ-ИРАН. – 2017. – № 3 (52). – С. 157-179.

34. Потапов, А. С. Распознавание образов и машинное восприятие: общий подход на основе принципа минимальной длины описания [Текст] / А.С. Потапов. – СПб.: Политехника, 2007. – 548 с.

35. Приоров, А.Л. Разработка и анализ алгоритма выделения лиц на изображениях [Текст] / И.С. Игнатов, М.Н. Голубев, В.В. Хрящев // Проектирование и технология электронных средств. – 2008. – № 2. – С. 58-62.

36. Прэтт, У. Цифровая обработка изображений [Текст]: Пер. с англ. / У. Прэтт. – М.: Мир, 1982. – Кн. 2. – 480 с.

37. Сойфер, В.А. Методы компьютерной обработки изображений [Текст] / В.А. Сойфер. – М.: ФИЗМАТЛИТ, 2003. – 784 с.

38. Трухин, А.В. Анализ существующих в РФ тренажерно-обучающих систем [Текст] / А.В. Трухин // Открытое и дистанционное образование. – 2008. – №. 1. – С. 32-39.

39. Туголукова, М.А. Техническое обеспечение создания виртуальной и дополненной реальности [Текст] / М.А. Туголукова, Е.П. Борисова // Сборник материалов Всероссийской научно-методической конференции «Виртуальная и дополненная реальность-2016: состояние и перспективы». – М.: ГПБОУ МГОК. – 2016. – С. 326-331.

40. Фаворская, М.Н. Модели локализации лиц на изображениях [Текст] / М.Н. Фаворская, А.И. Пахирка // Системы управления и информационные технологии. – 2008. – Т. 33, № 3.3. – С. 404–408.

41. Форсайт, Д. Компьютерное зрение. Современный подход [Текст]: Пер. с англ. / Д. Форсайт, Ж. Понс. – М.: Издательский дом «Вильямс», 2004. – 928с.

42. Фу, К. Структурные методы в распознавании образов [Текст] / К. Фу. – М.: Мир, 1977. – 317 с.

43. Хорн, Б.К.П. Зрение роботов [Текст]: Пер. с англ. / Б.К.П. Хорн. – М.: Мир, 1989. – 487 с.
44. Цветкович, Д. Спектры графов. Теория и применение [Текст] / Д. Цветкович, М. Дуб, Х. Захс. – Киев: Наукова думка, 1984. – 384 с.
45. Яне, Б. Цифровая обработка изображений [Текст] / Б. Яне. – М.: Техносфера, 2007. – 584 с.
46. Ан, К.Н. 3D head tracking and pose-robust 2D texture map-based face recognition using a simple ellipsoid model [Текст] / К.Н. Ан, М.Ж. Chung // Intelligent robots and systems. – 2008. – P. 307-312.
47. Artas, M. Incremental PCA for on-line visual learning and recognition [Текст] / М. Artas, М. Jogan, А. Leonardis // International conference and pattern recognition. – 2002. – Vol. 16. – P. 781-784.
48. Ва, S. A probabilistic framework for joint head tracking and pose estimation [Текст] / S. Ва, J.-М. Odobez // Proceedings of the 17th international conference on pattern recognition. – 2004. – Vol. 4. – P. 264-267.
49. Ва, S. From camera head pose to 3d global room head pose using multiple camera views [Текст] / S. Ва, J.-М. Odobez // Proceedings international workshop classification of events, activities and relationships. – 2007.
50. Bai, X. Characterising graphs using the heat kernel [Текст] / X. Bai, R.C. Wilson, E.R. Hancock // British machine vision conference. – 2005. – P. 315-324.
51. Balasubramanian, V. Biased manifold embedding: supervised isomap for person-independent head pose estimation [Текст] / V. Balasubramanian, S. Panchanathan // Computer vision and computer graphics. Theory and applications. – 2009. – Vol. 21. – P. 177-188.
52. Balasubramanian, V.N. Biased manifold embedding: a framework for person-independent head pose estimation [Текст] / V.N. Balasubramanian, J. Ye, S. Panchanathan // IEEE Conference on computer vision and pattern recognition. – 2007.
53. Baltrusaitis, T. Real-time inference of mental states from facial expressions and upper body gestures [Текст] / T. Baltrusaitis, D. McDuff, N. Banda, M. Mahmoud, R. Kaliouby, P. Robinson, R. Picard // Proceedings of the IEEE international conference on automatic face gesture recognition and workshops. – 2011. – P. 909-914.

54. Barinov, A.E. An algorithm for determining the position and orientation of 3-D objects from images using spectral graph theory [Текст] / A.E. Barinov, A.A. Zakharov // Applied mechanics and materials. – 2015. – Vol. 770. – P. 585-591.

55. Barinov, A.E. Clustering using a random walk on graph for head pose estimation [Текст] / A.E. Barinov, A.A. Zakharov // 2015 International conference on mechanical engineering, automation and control systems (MEACS). – 2015.

56. Barinov, A.E. Development of the synthesis algorithm of tree-dimensional objects from video images [Текст] / A.E. Barinov, A.A. Zakharov // 11-th international conference on pattern recognition and image analysis: new information technologies (PRIA-11-2013). – Samara: IPSI RAS. – 2013. – Vol. 1. – P. 167-169.

57. Barinov, A.E. Graph matching algorithm in position and orientation estimation tasks [Текст] / A.E. Barinov, A.A. Zakharov // Crimean conference “Microwave & telecommunication technology”. – 2014. – P. 411-412.

58. Bay, H. SURF: Speeded up robust features [Текст] / H. Bay, A. Ess, T. Tuytelaars, L. Van Gool // Computer vision and image understanding. – 2008. – Vol. 110. – №. 3. – P. 346-359.

59. Belkin, M. Laplacian eigenmaps for dimensionality reduction and data representation [Текст] / M. Belkin, P. Niyogi // Neural computation. – 2003. – Vol. 15. – № 6. – P. 1373-1396.

60. Beyer, W.H. CRC standard mathematical tables [Текст] / W.H. Beyer. – Boca Raton: CRC Press, 1987. – P. 299-300.

61. Beymer, D.J. Face recognition under varying pose [Текст] / D.J. Beymer // Computer vision and pattern recognition. – 1994. – P. 756-761.

62. Borg, I. Modern multidimensional scalling: theory and applications [Текст] / I. Borg, P. Groenen. – New York: Springer-Verlag, 2005. – 355 p.

63. Bradski, G. Learning OpenCV: Computer Vision with the OpenCV Library [Текст] / G. Bradski, A. Kaehler. – Sebastopol: O’Reilly Media, 2008. – 556 p.

64. Breitenstein, M.D. Head pose estimation from passive stereo images [Текст] / M.D. Breitenstein, J. Jensen, C. Hoilund, T.B. Moeslund, L. Van Gool // Scandinavian conference on image analysis. – 2009. – P. 219-228.

65. Brown, L. Comparative study of coarse head pose estimation [Текст] / L. Brown, Y. L. Tian // Proceedings workshop motion and video computing. – 2002. – P. 125-130.

66. Burger, P. Self initializing head pose estimation with a 2D monocular USB camera [Текст] / P. Burger, M. Rothbacher // Technischer bericht. – München: Technische universität, 2013.

67. Cabrera, C.R. Fast head pose estimation for human-computer interaction [Текст] / C.R. Cabrera, M. García-Montero, R. López-Sastre, T. Tuytelaars // Iberian conference on pattern recognition and image analysis (IbPRIA). – 2015.

68. Chen, Y. A method of head pose estimation based on active shape model and stereo vision [Текст] / Y. Chen, M. Fu, Y. Yang, W. Song // Control conference. – 2014.

69. Chung, F.R.K. Discrete green's functions [Текст] / F.R.K. Chung, S.T. Yau // Journal of combinatorial theory. – 2000. – №. 91. – P. 191-214.

70. Cootes, T.F. Active appearance models [Текст] / T. F. Cootes, G. J. Edwards, C. J. Taylor // IEEE transactions on pattern analysis and machine intelligence. – 2001. – Vol. 23. – P. 681-685.

71. Cootes, T.F. Groupwise construction of appearance models using piece-wise affine deformations [Текст] / T.F. Cootes, C.J. Twining, V. Petrovic, R. Schestowitz, C.J. Taylor // British machine vision conference. – 2005. – P. 879-888.

72. Cootes, T.F. View-based active appearance models [Текст] / T. F. Cootes, G.V. Wheeler, K.N. Walker, C.J. Taylor // Image and vision computing. – 2002. – Vol. 20. – №. 9-10. – P. 657-664.

73. Dahmane, A. Learning symmetrical model for head pose estimation [Текст] / A. Dahmane, S. Larabi, C. Djeraba, I. Bilasco // Proceedings of the international conference on pattern recognition. – 2012. – P. 3614-3617.

74. Delaunay, B. Sur la sphere vide [Текст] / B. Delaunay // Izvestia akademii nauk SSSR, otdelenie matematicheskikh i estestvennykh nauk. – 1934. – P. 793-800.

75. Dong, Y. A robust and efficient face-tracking kernel for driver inattention monitoring system [Текст] / Y. Dong, Z. Hu, K. Uchimura, N. Murayama // IEEE intelligent vehicles symposium. – 2010. – P. 929–934.

76. Drouard, V. Head pose estimation via probabilistic high-dimensional regression [Текст] / V. Drouard, S. Ba, G. Evangelidis, A. Deleforge, R. Horaud // Proceedings IEEE International conference on image processing. – 2015.

77. Edenborough, N. Driver state monitor from Delphi [Текст] / N. Edenborough, R. Hammoud, A. Harbach, A. Ingold, B. Kisacanin, P. Malawey, T. Newman, G. Scharenbroch, S. Skiver, M. Smith, A. Wilhelm, G. Witt, E. Yoder, H. Zhang // Proceedings IEEE computer society conference on computer vision and pattern recognition. – 2005. – Vol. 2. – P. 1206–1207.

78. Fernandez, A. Driver distraction using visual-based sensors and algorithms [Текст] / A. Fernandez, R. Usamentiaga, J.L. Carus, R. Casado // Sensors. – 2016. – № 16(11): 1805.

79. Fu, Y. Graph embedded analysis for head pose estimation [Текст] / Y. Fu, T.S. Huang // Proceedings of the 7th international conference on automatic face and gesture recognition, FGR'06. – 2006. – P. 3-8.

80. Ghawalby, H. Heat kernel embeddings, differential geometry and graph structure [Текст] / H. Ghawalby, E.R. Hancock // Axioms. – 2015. – Vol. 4. – P. 275-293.

81. Gonzalez, R.C. Digital image processing [Текст] / R.C. Gonzalez, R.E. Woods. – New Jersey: Prentice Hall, 2002. – 793 p.

82. Gorbatshevich, V. Face pose recognition based on monocular digital imagery and stereo-based estimation of its precision [Текст] / V. Gorbatshevich, Y. Vizilter, V. Knyaz, S. Zheltov // ISPRS annals of the photogrammetry, remote sensing and spatial information sciences. – 2014. – Vol. 5. – P. 257-263.

83. Gourier, N. Estimating face orientation from robust detection of salient facial structures [Текст] / N. Gourier, D. Hall, J. Crowley // Proceedings pointing 2004 workshop: visual observation of deictic gestures. – 2004. – P. 17-25.

84. Gourier, N. Head pose estimation on low resolution images [Текст] / N. Gourier, J. Maisonnasse, D. Hall, J. Crowley // Multimodal technologies for perception of humans: proceedings first international workshop classification of events, activities and relationships. – 2007. – P. 270-280.

85. Gurbuz, S. Model free head pose estimation using stereovision [Текст] / S. Gurbuz, E. Oztop, N. Inoue // Pattern recognition. – 2012. – Vol. 45. – P. 33-42.

86. Hartley, R. Multiple view geometry in computer vision [Текст] / R. Hartley, A. Zisserman. – Cambridge university press, 2003. – 655 p.

87. Hartuv, E. A clustering algorithm based on graph connectivity [Текст] / E. Hartuv, R. Shamir // Information processing letters. – 2000. – Vol. 76. (4-6). – P. 175-181.

88. Hatem, H. Head pose estimation based on detecting facial features [Текст] / H. Hatem, Z. Beiji, R. Majeed, J. Waleed, M. Lutf // International journal of multimedia and ubiquitous engineering. – 2015. – Vol. 10. – №. 3. – P. 311-322.

89. Heinzmann, J. 3-D facial pose and gaze point estimation using a robust real-time tracking paradigm [Текст] / J. Heinzmann, A. Zelinsky // Proceedings international conference on automatic face and gesture recognition. – 1998. – P. 142-147.

90. Horprasert, T. Computing 3-D head orientation from a monocular image sequence [Текст] / T. Horprasert, Y. Yacoob, L. Davis // Proceedings international conference on automatic face and gesture recognition. – 1996. – P. 242-247.

91. Householder, A.S. Unitary triangulation of a nonsymmetric matrix [Текст] / A.S. Householder // Journal of the ACM. – 1958. – Vol. 5. – P. 339-342.

92. Hu, N. Head pose estimation by non-linear embedding and mapping [Текст] / N. Hu, W. Huang, S. Ranganath // Proceedings IEEE international conference on image processing. – 2005. – Vol. 2. – P. 342-345.

93. Hu, Y. Estimating face pose by facial asymmetry and geometry [Текст] / Y. Hu, L. Chen, Y. Zhou, H. Zhang // Proceedings international conference on automatic face and gesture recognition. – 2004. – P. 651-656.

94. Jebara, T. Parametrized structure from motion for 3D adaptive feedback tracking of faces [Текст] / T. Jebara, A. Pentland // Proceedings IEEE conference computer vision and pattern recognition. – 1997. – P. 144-150.

95. Jimenez, P. Face tracking and pose estimation with automatic 3D model construction [Текст] / P. Jimenez, J. Nuevo, L.M. Bergasa // IET computer vision. – 2009.

96. Jimenez-Pinto J. Optical flow and driver's kinematics analysis for state of alert sensing [Текст] / J. Jimenez-Pinto, M. Torres-Torriti // Sensors. – 2013. – Vol. 13. – P. 4225-4257.

97. Jones, M. Fast multi-view face detection [Текст] / M. Jones, P. Viola. – Mitsubishi electric research laboratories, technical report 096. – 2003. – 11 p.

98. Kaczmarek, P. EKF-based method for kinematic configuration estimation of finger-like structure using low grade multi-IMU system [Текст] / P. Kaczmarek, J. Tomczynski, T. Mankowski // IEEE international conference on multisensor fusion and integration for intelligent systems (MFI-2016). – 2016.

99. Kaufmann, A. Capturing full body motion [Текст] / A. Kaufmann // Distributed systems seminar. – 2013.

100. Kok, M. An optimization-based approach to human body motion capture using inertial sensors [Текст] / M. Kok, J. D. Hol, T. B. Schon // IFAC proceedings volumes. – 2014. – Vol. 47 (3). – P. 79-85.

101. Kousidis, S. MINT.tools: tools and adaptors supporting acquisition, annotation and analysis of multimodal corpora [Текст] / S. Kousidis, T. Pfeiffer, D. Schlangen // Proceedings of Interspeech. – 2013.

102. Krinidis, M. 3-D head pose estimation in monocular video sequences using deformable surfaces and radial basis functions [Текст] / M. Krinidis, N. Nikolaidis, I. Pitas // IEEE transactions on circuits and systems for video technology. – 2009. – Vol. 19. – №. 2. – P. 261-272.

103. Kruger, N. Determination of face position and pose with a learned representation based on labeled graphs [Текст] / N. Kruger, M. Potzsch // Image and vision computing. – 1997. – Vol. 15. – №. 8. – P. 665-673.

104. Lades, M. Distortion invariant object recognition in the dynamic link architecture [Текст] / M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. Malsburg, R.P. Wurtz, W. Konen // IEEE transactions on computers. – 1993. – Vol. 42. – P. 300-311.

105. Li, Y. Support vector machine based multi-view face detection and recognition [Текст] / Y. Li, S. Gong, J. Sherrah, H Liddell // Image and vision computing. – 2004. – Vol. 22. – №. 5. – P. 413-427.
106. Liao, W.K. Robust, real-time 3D face tracking from a monocular view [Текст] / W.K. Liao, D.Fidaleo, G.Medioni // EURASIP journal on image and video processing. – 2010. – P. 1-15.
107. Luxburg, U. Hitting and commute times in large random neighborhood graphs [Текст] / U. Luxburg, A. Radl, M. Hein // Journal of machine learning research. – 2014. – № 15. – P. 1751-1798.
108. Ma, Y. Sparse Bayesian regression for head pose estimation [Текст] / Y. Ma, Y. Konishi, K. Kinoshita, S. Lao, M. Kawade // Proceedings IEEE international conference on pattern recognition. – 2006. – P. 507-510.
109. McKenna, S. Real-time face pose estimation [Текст] / S. McKenna, S. Gong // Real-time imaging. – 1998. – Vol. 4. – №. 5. – P. 333-347.
110. Melia, M. A random walks view of spectral segmentation [Текст] / M. Melia, J. Shi // NIPS. – 2000. – P. 873-879.
111. Menache, A. Understanding motion capture for computer animation / A. Menache // The Morgan Kaufmann series in computer graphics. – 2011. – 254 p.
112. Moon, H. Estimating facial pose from a sparse representation / H. Moon, M. Miller // Proceedings IEEE international conference on image processing. – 2004. – P. 75-78.
113. Moore, S. Local binary patterns for multi-view facial expression recognition [Текст] / S. Moore, R. Bowden // Computer vision image understanding. – 2011. – Vol. 115. – P. 541-558.
114. Morency, L.P. Adaptive view-based appearance models [Текст] / L.P. Morency, A. Rahimi, T. Darrell // Proceedings IEEE conference computer vision and pattern recognition. – 2003. – P. 803-810.
115. Morency, L.P. Pose estimation using 3D view-based eigenspaces [Текст] / L.P. Morency, P. Sundberg, T. Darrell // Proceedings IEEE international workshop analysis and modeling of faces and gestures. – 2003. – P. 45-52.

116. Munoz, E. Fast 6D pose from a single RGB image using cascades forests templates [Текст] / E. Munoz, Y. Konishi, C. Beltran, V. Murino, A. Del Bue // 2016 IEEE/RSJ international conference on intelligent robots and systems (IROS). – 2016.

117. Murphy-Chutorian, E. Head pose estimation in computer vision: A survey [Текст] / E. Murphy-Chutorian, M. Manubhai Trivedi // IEEE transactions on pattern analysis and machine intelligence. – 2009. – Vol. 31. – №. 4. – P. 607-626.

118. Murphy-Chutorian, E. HyHOPE: Hybrid head orientation and position estimation for vision-based driver head tracking [Текст] / E. Murphy-Chutorian, M.M. Trivedi // Intelligent vehicles symposium. – 2008. – P. 512-517.

119. Newman, R. Real-time stereo tracking for head pose and gaze estimation [Текст] / R. Newman, Y. Matsumoto, S. Rougeaux, A. Zelinsky // Proceedings IEEE international conference automatic face and gesture recognition. – 2000. – P. 122-128.

120. Niese, R. Facial expression recognition based on geometric and optical flow features in colour image sequences [Текст] / R. Niese, A. Al-Hamadi, A. Farag, H. Neumann, B. Michaelis // IET computer vision. – 2012. – Vol. 6. – P. 79-89.

121. Niyogi, S. Example-based head tracking [Текст] / S. Niyogi, W.T. Freeman // Proceedings of the 2nd international conference on automatic face and gesture recognition. – 1996. – P. 374-375.

122. Ohayon, S. Robust 3D head tracking using camera pose estimation [Текст] / S. Ohayon, E. Rivlin // Proceedings 18th international conference on pattern recognition. – 2006. – P. 1063-1066.

123. Oka, K. Head pose estimation system based on particle filtering with adaptive diffusion control [Текст] / K. Oka, Y. Sato, Y. Nakanishi, H. Koike // Proceedings IAPR conference machine vision applications. – 2005. – P. 586-589.

124. Osuna, E. Training support vector machines: an application to face detection [Текст] / E. Osuna, R. Freund, F. Girosi // Proceedings IEEE conference computer vision and pattern recognition. – 1997. – P. 130-136.

125. Pol, D. Head pose estimation for a domestic robot [Текст] / D. Pol, R.H. Cuijpers, J.F. Juola // Proceedings of the 6th international conference on human robot interaction (HRI-2011). – 2011.

126. Pustka, D. Optical outside-in tracking using unmodified mobile phones [Текст] / D. Pustka, J.P. Hulss, J. Willneff, F. Pankratz, M. Huber, G. Klinker // IEEE international symposium on mixed and augmented reality (ISMAR). – 2012.

127. Qiu, H. Image segmentation using commute times [Текст] / H. Qiu, E.R. Hancock // British machine vision conference. – 2005. – P. 929-938.

128. Raytchev, B. Head pose estimation by nonlinear manifold learning [Текст] / B. Raytchev, I. Yoda, K. Sakaue // Proceedings international conference on pattern recognition. – 2004. – P. 462-466.

129. Rockafellar, R. Variational analysis [Текст] / R. Rockafellar, W. Tyrrel, J.B. Roger. – Springer-Verlag, 2005. – 117 p.

130. Rowley, H.A. Rotation invariant neural network-based face detection [Текст] / H.A. Rowley, S. Baluja, T. Kanade // Proceedings IEEE conference on computer vision and pattern recognition. – 1998. – P. 38-44.

131. Rudovic, O. Coupled Gaussian processes for pose-invariant facial expression recognition [Текст] / O. Rudovic, S. Member, M. Pantic, I. Patras, S. Member // IEEE transactions on pattern analysis and machine intelligence. – 2013. – Vol. 35. – P. 1357-1369.

132. Saeed, A. Frame-based facial expression recognition using geometrical features [Текст] / A. Saeed, A. Al-Hamadi, R. Niese, M. Elzobi // Advances in human-computer interaction. – 2014. – Vol. 4. – P. 1-13.

133. Sarkar, S. Quantitative measures of change based on feature organization: eigenvalues and eigenvectors [Текст] / S. Sarkar, K.L. Boyer // Computer vision and pattern recognition. – 1996. – P. 478-483.

134. Schodl, A. Head tracking using a textured polygonal model [Текст] / A. Schodl, A. Haro, I. Essa // Proceedings workshop perceptual user interfaces. – 1998.

135. Scott, G. Feature grouping by relocalisation of eigenvectors of the proximity matrix [Текст] / G. Scott, H. Longuet-Higgins // British machine vision conference. – 1990. – P. 103-108.

136. Seemann, E. Head pose estimation using stereo vision for human-robot interaction [Текст] / E. Seemann, K. Nickel, R. Stiefelhagen // Proceedings international conference on face and gesture recognition. – 2004. – P. 626-631.
137. Seo, K. Face pose estimation system by combining hybrid ICA-SVM learning and re-registration [Текст] / K. Seo // Asian Conference on Computer Vision. – 2004.
138. Sharma, S. Template matching approach for face recognition system [Текст] / S. Sharma // International journal of signal processing systems. – 2013. – Vol. 1. – №. 2. – P. 284-289.
139. Sheerman-Chase, T. Online learning of robust facial feature trackers [Текст] / T. Sheerman-Chase, E.J. Ong, R. Bowden // International computer vision workshops. – 2009. – P. 1386-1392.
140. Sherrah, J. Fusion of perceptual cues for robust tracking of head pose and position [Текст] / J. Sherrah, S. Gong // Pattern recognition. – 2001. – Vol.34. – №. 8. – P. 1565–1572.
141. Sherrah, J. Understanding pose discrimination in similarity space [Текст] / J. Sherrah, S. Gong, E. Ong // British machine vision conference. – 1999. – P. 523-532.
142. Shi, J. Normalized cuts and image segmentation [Текст] / J. Shi, J. Malik // IEEE transactions on pattern analysis and machine intelligence. – 2000. – №. 8 (22). – P. 888-905.
143. Silva, I.M. A method for creating interactive, user-resembling avatars [Текст] / I.M. Silva, R.C. Moiola // PeerJ Computer Science. – 2017.
144. Srinivasan, S. Head pose estimation using view based eigenspaces [Текст] / S. Srinivasan, K. Boyer // Proceedings international conference pattern recognition. – 2002. – P. 302-305.
145. Stengel, M. An affordable solution for binocular eye tracking and calibration in head-mounted displays [Текст] / M. Stengel, S. Grogorick, M. Eisemann, E. Eisemann, M. Magnor // Proceedings of the 23rd ACM international conference on Multimedia, 2015. – P. 15-24.
146. Stiefelhagen, R. Modeling focus of attention for meeting indexing based on multiple cues [Текст] / R. Stiefelhagen, J. Yang, A. Waibel // IEEE transactions neural networks. – 2002. – Vol. 13. – №. 4. – P. 928-938.

147. Sundararajan, K. Head pose estimation in the wild using approximate view manifolds [Текст] / K. Sundararajan, D.L. Woodard // IEEE conference on computer vision and pattern recognition workshops. – 2015. – P. 50-58.

148. Tobon, R. The MOCAP book: a practical guide to the art of motion capture [Текст] / R. Tobon. – Foris force, 2010. – 258 p.

149. Vassilvitskii, A. K-means++: The advantages of careful seeding [Текст] / A. Vassilvitskii, D. Vassilvitskii, S. Vassilvitskii // Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms. – 2007. – P. 1027-1035.

150. Vezhnevets, V. Method for localization of human faces in color-based face detectors and trackers [Текст] / V. Vezhnevets // Third international conference on digital information processing and control in extreme situations. – 2002. – P. 51-56.

151. Viola, P. Rapid object detection using a boosted cascade of simple features [Текст] / P. Viola, M. Jones // Proceedings IEEE conference computer vision and pattern recognition. – 2001. – P. 511-518.

152. Viola, P. Robust real time object detection [Текст] / P. Viola, M. Jones // IEEE ICCV workshop on statistical and computational theories of vision. – 2001.

153. Wang, J.G. EM enhancement of 3D head pose estimated by point at infinity [Текст] / J.G. Wang, E. Sung // Image and vision computing. – 2007. – Vol. 25. – № 12. – P. 1864-1874.

154. Wilson, H. Perception of head orientation [Текст] / H. Wilson, F. Wilkinson, L. Lin, M. Castillo // Vision research. – 2000. – Vol. 40. – №. 5. – P. 459-472.

155. Wu, J. A two-stage head pose estimation framework and evaluation [Текст] / J. Wu, M. Trivedi // Pattern recognition. – 2008. – Vol. 41. – №. 3. – P. 1138-1158.

156. Wu, Y. Wide-range, person- and illumination in sensitive head orientation estimation [Текст] / Y. Wu, K. Toyama // Proceedings IEEE international conference automatic face and gesture recognition. – 2000. – P. 183-188.

157. Xiao, B. Heat kernel, Riemannian manifolds and graph embedding [Текст] / B. Xiao, E.R. Hancock // LNCS. – 2004. – P. 198-206.

158. Xiao, J. Real-time combined 2D+3D active appearance models [Текст] / J. Xiao, S. Baker, I. Matthews, T. Kanade // IEEE conference on computer vision and pattern recognition. – 2004. – Vol. 2. – P. 535-542.
159. Yang, R. Model-based head pose tracking with stereovision [Текст] / R. Yang, Z. Zhang // Proceedings international conference automatic face and gesture recognition. – 2002. – P. 242-247.
160. Zakharov, A.A. An algorithm for 3D-object reconstruction from video using stereo correspondences [Текст] / A.A. Zakharov, A.E. Barinov // Pattern recognition and image analysis. – 2015. – Vol. 25. – P. 117-121.
161. Zakharov, A.A. Recognition of human pose from images based on graph spectra [Текст] / A.A. Zakharov, A.E. Barinov, A.L. Zhyznyakov // The international archives of the photogrammetry, remote sensing and spatial information sciences. – 2014. – Vol. 5. – P. 9-12.
162. Zhang, Z. Head pose estimation in seminar room using multi view face detectors [Текст] / Z. Zhang, Y. Hu, M. Liu, T. Huang // In multimodal technologies for perception of humans, international workshop classification of events activities and relationships, CLEAR. – 2007. – Vol. 4122.
163. Zhao, G. Large head movement tracking using sift-based registration [Текст] / G. Zhao, L. Chen, J. Song, G. Chen // Proceedings of the 15th international conference on multimedia, MULTIMEDIA'07. – 2007. – P. 807-810.
164. Zhao, L. Real-time head orientation estimation using neural networks [Текст] / L. Zhao, G. Pingali, I. Carlbom // Proceedings international conference on image processing. – 2002. – P. 297-300.
165. Zhu, X. Face detection, pose estimation, and landmark localization in the wild [Текст] / X. Zhu, D. Ramanan // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2012. – P. 2879-2886.

Научное издание

ЗАХАРОВ Алексей Александрович
БАРИНОВ Алексей Евгеньевич
ЖИЗНЯКОВ Аркадий Львович

МОДЕЛИ И АЛГОРИТМЫ ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ
И ОРИЕНТАЦИИ ГОЛОВЫ ЧЕЛОВЕКА ПО ВИДЕОИЗОБРАЖЕНИЯМ
С ИСПОЛЬЗОВАНИЕМ ГРАФОВ

Монография

Издается в авторской редакции

Подписано в печать 23.08.19.
Формат 60x84/16. Усл. печ. л. 7,21. Тираж 50 экз.

Заказ

Издательство

Владимирского государственного университета
имени Александра Григорьевича и Николая Григорьевича Столетовых.
600000, Владимир, ул. Горького, 87.