

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ М. В. ЛОМОНОСОВА



На правах рукописи

Чернышов Виктор Геннадьевич

БИОМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ЛИЧНОСТИ ПО ИЗОБРАЖЕНИЮ ВНЕШНЕЙ
СТОРОНЫ ЛАДОНИ НА БАЗЕ МОБИЛЬНОГО УСТРОЙСТВА

05.13.17 — Теоретические основы информатики

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Москва

2019

Работа выполнена на кафедре математических методов прогнозирования факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова.

Научный руководитель: **Местецкий Леонид Моисеевич**
доктор технических наук, профессор
Московского государственного университета
имени М. В. Ломоносова

Официальные оппоненты: **Миркин Борис Григорьевич**
доктор технических наук, профессор
Национального исследовательского
университета «Высшая школа экономики»

Вишняков Борис Ваисович
кандидат физико-математических наук,
начальник лаборатории ФГУП ГосНИИАС

Ведущая организация: Тульский государственный университет

Защита состоится «6» февраля 2020 г. в 14:00 часов на заседании диссертационного совета Д 002.073.05 при Федеральном исследовательском центре «Информатика и управление» РАН по адресу: 119333, г. Москва, ул. Вавилова, д. 44, корп. 2.

С диссертацией можно ознакомиться в библиотеке и на сайте ФИЦ ИУ РАН <http://frccsc.ru>.

Автореферат разослан «9» декабря 2019 г.

И. о. учёного секретаря
диссертационного совета Д 002.073.05
д. т. н.

Матвеев И. А.

1 Общая характеристика работы

Актуальность темы. Технология идентификации личности по изображению внешней стороны ладони, полученному с помощью мобильного устройства (телефона или планшета), представляется весьма привлекательной по двум причинам. Во-первых, повсеместное распространение таких устройств позволяет рассматривать эту технологию как не требующую специального оборудования и навыков — и потому как доступную и удобную для применения в учреждениях образования, здравоохранения и культуры. Во-вторых, получение и использование изображений внешней стороны ладони является социально приемлемым для большинства людей. Это отличает данный признак от других биометрических параметров: отпечатков пальцев, радужной оболочки и т. п.

Несмотря на очевидные достоинства такой технологии и наличие возможностей для её практического применения, на современном рынке подобные биометрические системы не представлены. Основная причина состоит в том, что существующие методы идентификации по изображению ладони ориентированы на использование в стационарных установках, ведущих съёмку в условиях выбранного фона, постоянной освещённости и фиксированного ракурса визирования. Но съёмка на мобильном устройстве ладони в свободном положении не может обеспечить такие условия. Предложенный Бакиной¹ подход к идентификации личности по форме ладони работает при строго ортогональном положении плоскости ладони относительно оси камеры, что трудно обеспечить при съёмке мобильным устройством «с руки». Вариация ракурса съёмки приводит к большим ошибкам распознавания.

Работа Жанга², посвящённая идентификации по складкам кожи на суставах пальцев, показывает, что данный признак является достаточно информативной индивидуальной характеристикой. Однако представленная технология основана на использовании специально-го стационарного устройства, обеспечивающего точное позиционирование согнутых пальцев в фиксированном положении. В то же время, в изученной научной литературе не встречаются аналогичные системы на основе **мобильных** устройств, обладающие приемлемой точностью и практической применимостью.

Таким образом, актуальность темы диссертации определяется востребованностью технологии идентификации личности по изображению внешней стороны ладони с помощью мобильного устройства и отсутствием подобных систем в свободном доступе.

¹Бакина И. Г., Местецкий Л. М. Многомодальная идентификация личности по форме ладони и голосу // Таврический вестник информатики и математики. — 2008. — № 1 — С. 59–65.

²Online finger-knuckle-print verification for personal authentication / Lin Zhang, Lei Zhang, David Zhang, Nailong Zhu // Pattern recognition. — 2010. — Vol. 43, no. 7. — Pp. 2560–2571.

Цель исследования. Целью исследования является создание метода идентификации человека по изображениям внешней стороны ладони, который может быть использован в системе распознавания на базе мобильного устройства. Разработка такого метода призвана подтвердить гипотезу о возможности идентификации личности с приемлемой точностью на основе изображений внешней стороны ладони, полученных с помощью мобильного устройства. Неотъемлемой частью работы является реализация соответствующего программного комплекса, обосновывающего эффективность метода и его частей.

Выбор в качестве модальности **внешней** стороны ладони, а также желаемые показатели точности и скорости работы метода идентификации во многом продиктованы требованиями к системе распознавания в целом:

точность — небольшой уровень ошибок (т.е. приемлемый для предметной области, в которой планируется внедрение);

скорость — распознавание должно осуществляться в режиме реального времени;

доступность — система должна состоять из распространённых и недорогих компонентов;

социальная приемлемость — система не должна вызывать негативную реакцию пользователей в силу социальных стереотипов или религиозных установок;

удобство — процесс идентификации должен быть интуитивно понятен для конечного пользователя и не вводить дополнительных ограничений (например, не требовать выравнивания ладони перед камерой), эксплуатация не должна быть сложна для владельца (например, минусом может являться сложный процесс разворачивания системы на новом месте);

гибкость — возможность работы в различных внешних условиях (например, при разном освещении и при съёмке «с руки»).

Описанная в работе мобильная система распознавания по внешней стороне ладони (по форме ладони и кожным складкам пальцев) разработана, в первую очередь, для идентификации в малых группах. Возможной областью её использования являются учреждения, где существует потребность в контроле посещаемости: вузы, школы, больницы, библиотеки.

В научной литературе не представлено систем идентификации по ладони в целом (и по внешней её стороне — в частности), одновременно удовлетворяющих всем вышеуказанным требованиям. Возможные причины этого кроются в сложности возникающих научно-технических задач, которые необходимо решить при создании такой системы. Более того,

создание подобных систем стало обоснованным в связи со значительным прогрессом характеристик мобильных устройств и снижением их цены, произошедшими за последние годы. Разработке и проектированию системы биометрической идентификации по внешней стороне ладони сопутствует ряд неопределённых факторов, которые сводятся к нескольким принципиальным вопросам.

1. Насколько уникальными являются признаки формы ладони и рисунка кожных складок и в какой мере их достаточно для идентификации личности?
2. Существует ли возможность надёжного извлечения этих признаков из изображений, полученных с помощью общедоступных мобильных камер (без использования специального оборудования) при свободном расположении ладони во время съёмки?
3. Возможно ли решить задачу извлечения признаков и провести идентификацию на их основе в реальном времени при работе на мобильном устройстве в различных условиях применения?
4. Каким образом построить систему распознавания и организовать процесс идентификации, чтобы получившийся программный комплекс оказался практически востребованным?

Получение ответов на эти вопросы составляет основное содержание исследования, проведенного в данной работе.

Методы исследования. В исследовании использованы методы вычислительной геометрии, теории графов, теории анализа и распознавания изображений, компьютерной графики, анализа данных и машинного обучения. Для сбора данных и проведения вычислительных экспериментов использовались методы разработки многозвенных программных систем — в том числе, мобильных и веб-приложений.

Предлагаемый подход к решению. В работе предлагаются и обосновываются следующие принципы построения метода идентификации по изображению внешней стороны ладони и основанной на нём системы распознавания на базе мобильного устройства.

- Применяется клиент-серверная структура программного комплекса с мобильным устройством в качестве клиента.
- Декомпозиция задачи идентификации на задачи детектирования и распознавания, их структурное распределение между клиентом и сервером:

- задача клиента — детектирование, т. е. отбор подходящих для распознавания изображений ладони и последующая отправка их на сервер;
 - задача сервера — распознавание, т. е. вычисление признаков, сравнение с базой эталонов и выдача результата.
- Для детектирования ладони и локализации областей кожных складок на пальцах используется непрерывное медиальное представление формы ладони.
 - Формирование признакового описания ладони осуществляется с помощью двумерных фильтров Габора и дескрипторов особых точек, применяемых к областям изображения внешней стороны ладони, содержащим кожные складки вторых фаланг пальцев.
 - Классификация изображений ладони происходит путём сравнения их признакового описания с базой эталонов — используется метод ближайшего соседа со специальной мерой сходства.

Научная новизна.

1. Впервые выполнено исследование, обосновывающее возможность идентификации личности по изображениям внешней стороны ладони со свободным положением руки в кадре в системе компьютерного зрения, работающей в реальном времени. Программная реализация на базе серийного мобильного устройства также выполнена впервые.
2. Новизна предложенного подхода к решению задачи состоит в декомпозиции процесса идентификации на задачи детектирования и распознавания ладони, и разработке системы с клиент-серверной архитектурой, выполняющей детектирование на стороне клиента и распознавание на стороне сервера.
3. Для реализации подхода предложены оригинальные методы обработки и распознавания изображений ладони человека: метод детектирования ладони в системе компьютерного зрения на основе непрерывного медиального представления; метод формирования признакового описания ладони на основе двумерных фильтров Габора и дескрипторов особых точек, применённых к областям кожных складок пальцев; метод определения сходства наблюдаемой ладони с эталонными изображениями на основе сформированных признаковых описаний.

Научные результаты, выносимые на защиту. На защиту выносятся следующие результаты данной работы:

1. Метод биометрической идентификации личности по изображениям внешней стороны ладони со свободным положением руки в кадре, включающий детектирование ладони по признакам формы объекта и классификацию по текстурным признакам.
2. Метод детектирования ладони на основе непрерывного медиального представления изображения.
3. Метод формирования признакового описания ладони по изображениям кожных складок пальцев на базе двумерных фильтров Габора.
4. Мера сходства изображений ладони, основанная на признаках, полученных с помощью двумерных фильтров Габора и SIFT-дескрипторов особых точек.
5. Обоснование работоспособности предложенных методов путём реализации программного комплекса на базе мобильного устройства. Экспериментально доказано, что рисунок кожных складок на внешней стороне средних фаланг пальцев является индивидуальным признаком личности человека, дающим возможность уверенно проводить идентификацию в малых группах.

Все результаты, выносимые на защиту, получены автором самостоятельно.

Научная значимость. Научный результат работы заключается в разработке метода идентификации по изображению внешней стороны ладони, а также в разработке подхода к созданию систем идентификации по ладони на базе представленного метода. Работа экспериментально доказывает, что изображения внешней стороны ладони достаточно для уверенной идентификации личности человека при помощи камеры мобильного устройства.

Ключевыми особенностями метода являются: разделение задачи идентификации на задачи детектирования и распознавания, использование непрерывного медиального представления формы ладони, применение методов машинного обучения. Научная значимость решения определяется также универсальностью и гибкостью предложенного подхода:

- сфера применимости подхода не ограничена смартфонами и планшетами — аналогичное программное обеспечение может быть реализовано для другой подходящей аппаратной платформы: персонального компьютера с внешней камерой или специализированного устройства с камерой и вычислительными ресурсами;
- разделение задачи идентификации на детектирование и распознавание не обязывает к физическому разнесению этих модулей между устройствами (клиентом и сервером) —

возможно чисто структурное деление в рамках одного универсального устройства с нужными характеристиками;

- для построения информативного описания ладони в работе использованы области кожных складок пальцев на её тыльной стороне, однако в рамках предложенного подхода могут быть исследованы возможности построения системы, работающей с другими текстурными особенностями — например, с рисунком на внутренней стороне ладони.

Практическая значимость. Основной недостаток наиболее эффективных систем биометрической идентификации — необходимость использования специального оборудования (датчики, сканеры) и сопутствующей инфраструктуры. Как следствие, подобные системы дороги не только при покупке, но и при последующей эксплуатации, что делает их недоступными при небольшом финансировании.

Практическим результатом данной работы является программное обеспечение, решающее задачу идентификации по тыльной стороне ладони в малых группах, не выходя за рамки возможностей современных мобильных устройств. Иными словами, любой человек, имеющий смартфон или планшет, получает возможность на его основе за несколько минут внедрить для своего коллектива систему идентификации по форме и текстуре ладони.

Достоверность результатов. Достоверность результатов работы подтверждается корректностью постановок рассматриваемых задач исследования, применением математически обоснованных методов их решения, программной реализацией и вычислительными экспериментами, проведенными на реальных базах изображений человеческих ладоней. Реализованная в рамках исследования система биометрической идентификации по внешней стороне ладони прошла апробацию в качестве инструмента контроля посещаемости студентов и учёта посетителей на фестивале науки.

Публикации и апробация. Материалы исследования опубликованы в 9 печатных работах, из них: 2 статьи в рецензируемых журналах из списка ВАК РФ [1, 2], 4 статьи в сборниках трудов конференций [3–6] и 3 тезиса докладов [7–9].

Результаты исследования докладывались на следующих конференциях и семинарах:

1. 16-ая всероссийская конференция «Математические методы распознавания образов» ММРО-16, Казань, 2013;
2. 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding (OGRW 2014), Germany, Koblenz, 2014 [4];

3. 10th International Conference on Computer Vision Theory and Applications (VISAPP 2015), Germany, Berlin, 2015 [5];
4. 6-я научно-техническая конференция «Техническое зрение в системах управления-2015», Москва, 2015 [7];
5. 17-я всероссийская конференция «Математические методы распознавания образов» ММРО-17, Светлогорск, 2015 [8];
6. 26-я международная конференция по компьютерной графике и машинному зрению (ГрафиКон 2016), Нижний Новгород, 2016 [6];
7. 18-я всероссийская конференция «Математические методы распознавания образов» ММРО-18, Таганрог, 2017 [9];
8. семинар «Морфологический анализ данных» под руководством проф. Ю. П. Пытьева, Москва, 13 ноября 2014 г.;
9. совместный с Хуавэй международный научный семинар «Интеллектуальная обработка изображений и видео», Москва, 22 ноября 2017 г.;
10. XIII Московский Фестиваль науки НАУКА 0+, Москва, 12-14 октября 2018 г.;
11. the 5th International Professor's Day on ICT Algorithm Design (ICTAD-2018), Russia, Moscow, November 28-29, 2018;
12. CMC MSU — Zhejiang international workshop, Russia, Moscow, December 14, 2018.

В рамках исследования разработана мобильная система идентификации личности по ладони, с помощью которой осуществлялся контроль посещаемости студентов двух групп ВМК МГУ. Также система применялась для учёта посетителей на XIII Московском фестивале науки Nauka 0+. Описания отдельных результатов работы включены в отчёты по проектам РФФИ 14-01-00716, 16-57-45054, 17-01-00917.

Обоснование специальности. Данная работа по своей тематике и направленности полученных результатов соответствует следующим пунктам паспорта специальности 05.13.17 — «Теоретические основы информатики»:

5) Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечениях, разработка и исследование методов и алгоритмов анализа текста, устной речи и изображений.

7) Разработка методов распознавания образов, фильтрации, распознавания и синтеза изображений, решающих правил. Моделирование формирования эмпирического знания.

Структура и объём. Работа состоит из введения, 4 глав, заключения и библиографии. Работа содержит 134 страницы основного текста, включая 60 иллюстраций и 12 таблиц. Перечень библиографических источников включает 71 наименование.

2 Содержание работы

Во **введении** обоснована актуальность темы диссертационной работы, сформулированы цель исследования и требования к системе распознавания, аргументирована научная новизна исследования, описан подход к решению, показана практическая значимость полученных результатов, представлены выносимые на защиту научные положения.

В **главе 2** приводится обзор имеющихся методов биометрической идентификации личности по ладони.

Выполнен обзор методов идентификации по форме ладони. Среди рассмотренных можно выделить методы на основе анализа контура, методы на основе спектра, метод на основе гибких объектов. Методы требовательны к качеству сегментации, при этом не дают высоких показателей точности (EER обычно превышает 2%) — сказывается как вариабельность формы ладони в рамках одного человека, так и относительно высокая похожесть формы ладоней разных людей. Тем не менее, это не исключает использование формы ладони в качестве дополнительной модальности в некоторых системах биометрической идентификации. Почти во всех изученных работах предполагается стационарный сценарий получения изображения ладони — с **фиксацией** на стекле сканера или на поверхности перед объективом неподвижно закреплённой камеры. Попытка переноса этого решения на мобильное устройство показала существенное снижение качества распознавания за счёт сильной изменчивости формы в зависимости от угла съёмки.

Описаны методы идентификации по кожным складкам пальцев. В уже упомянутой ранее статье Жанга, посвящённой модификации алгоритма Competitive Code ³, предлагается метод кодирования признаков описания на основе разноориентированных двумерных фильтров Габора — сходный подход использован автором при разработке собственного метода. Съёмка осуществляется с помощью специально разработанного **стационарного** сканера. Собранная таким образом база изображений пальцев используется для тестирования

³Kong Adams Wai-Kin, Zhang David. Competitive Coding Scheme for Palmprint Verification // Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04). — Vol. 1 of ICPR '04. — Pp. 520–523.

в большинстве релевантных работ. Рассмотрены методы генерации признакового описания, основанные на прочих преобразованиях: дискретном двумерном преобразовании Фурье, модифицированном конечном дискретном преобразовании Радона, на базе моногенного сигнала.

Показано, что **мобильные** системы идентификации на основе ладони (по рисунку ладони, форме ладони, кожным складкам пальцев внешней стороны ладони) слабо представлены в литературе и, в силу своих характеристик, не годятся для решения реальных прикладных задач. Среди них ещё менее распространены мобильные системы, работающие с кожными складками пальцев, — единственная известная автору работа ⁴ в этом направлении имеет невысокую точность распознавания. Более того, разработанное приложение ориентировано для персональной аутентификации, а не для идентификации в малых группах.

Основным выводом обзорной главы является то, что в научной литературе не представлена система биометрической идентификации человека по снимку внешней стороны ладони на базе мобильного устройства, удовлетворяющая требованиям к системе распознавания.

Глава 3 содержит описание предлагаемого метода идентификации по изображению внешней стороны ладони. В **разделе 3.1** изложены основные положения и определения применяемой в работе методологии машинного обучения и рассмотрен цикл исследования, применённый для построения обучаемого алгоритма. В контексте исследования биометрическая идентификация личности по изображению ладони сводится к задаче классификации.

- Каждому человеку взаимно-однозначно назначена метка класса $y \in Y$, где $Y = \{1, \dots, M\}$, M — число непересекающихся классов.
- Изображение ладони человека является объектом $x \in X$. Произвольный класс K_y множества объектов X сформирован из изображений ладони, принадлежащих человеку с меткой класса y .
- Метки классов y_i известны только на конечном подмножестве объектов $\{x_1, \dots, x_l\} \subset X$, при этом пары «объект-ответ» (x_i, y_i) называются прецедентами, а совокупность таких пар $X^l = (x_i, y_i)_{i=1}^l$ — обучающей выборкой.
- Искомый обучаемый по X^l алгоритм классификации $a(x): X \rightarrow Y$ должен давать ответ на вопрос «какому человеку принадлежит изображение ладони x ?».

В **разделе 3.2** описана предложенная автором концепция, лежащая в основе метода идентификации по изображению внешней стороны ладони. Её суть заключается в деком-

⁴Cheng KamYuen, Kumar Ajay. Contactless finger knuckle identification using smartphones //2012 BIOSIG - Proceedings of the International Conference of the Biometrics Special Interest Group. — 2012. — Pp. 1–6.

позиции задачи идентификации на задачи детектирования и распознавания. *Задача детектирования* — обнаружение объекта (в нашем случае ладони) на изображении и, в случае наличия, проверка «пригодности» изображения ладони для последующего распознавания. *Задача распознавания* — генерация признакового описания по изображению ладони, прошедшей проверки на «пригодность», и непосредственно классификация ладони путём сопоставления с базой признакового описания ладоней-эталонов.

При этом детектирование выполняется на мобильном устройстве, а распознавание на удалённом сервере. Задача распознавания является вычислительно трудной даже для персонального компьютера — как минимум потому, что одним из её этапов является «точная» сегментация. Это значит, что сервер не может обрабатывать большое число изображений в режиме реального времени, что вкупе с ограничениями сети предъявляет определённые требования к модулю детектирования. Алгоритм детектирования должен допускать к дальнейшей обработке как можно меньше «непригодных» изображений, т. е. иметь низкое значение FAR (иначе сервер не справится с большим количеством изображений). Вместе с тем, при соблюдении распознаваемым пользователем процедуры предъявления ладони должно поддерживаться невысокое значение FRR (это даёт приемлемое общее время распознавания).

В разделе 3.3 определяется, что является *корректной ладонью* в контексте задачи детектирования:

- ладонь вместе с частью запястья находится целиком в кадре;
- посторонние предметы не закрывают ладонь, т. е. отсутствуют часы, кольца, браслеты, длинные рукава и т. п.;
- все пальцы расставлены (разведены).

Некорректная ладонь — ладонь, которая не является корректной в терминах определения выше. *Корректными изображениями (кадрами)* будем называть те изображения (кадры), которые содержат корректную ладонь. *Некорректные изображения (кадры)*, в свою очередь, либо не содержат ладони вообще, либо содержат некорректную ладонь. Соответствующие проверки назовём проверками на *корректность*.

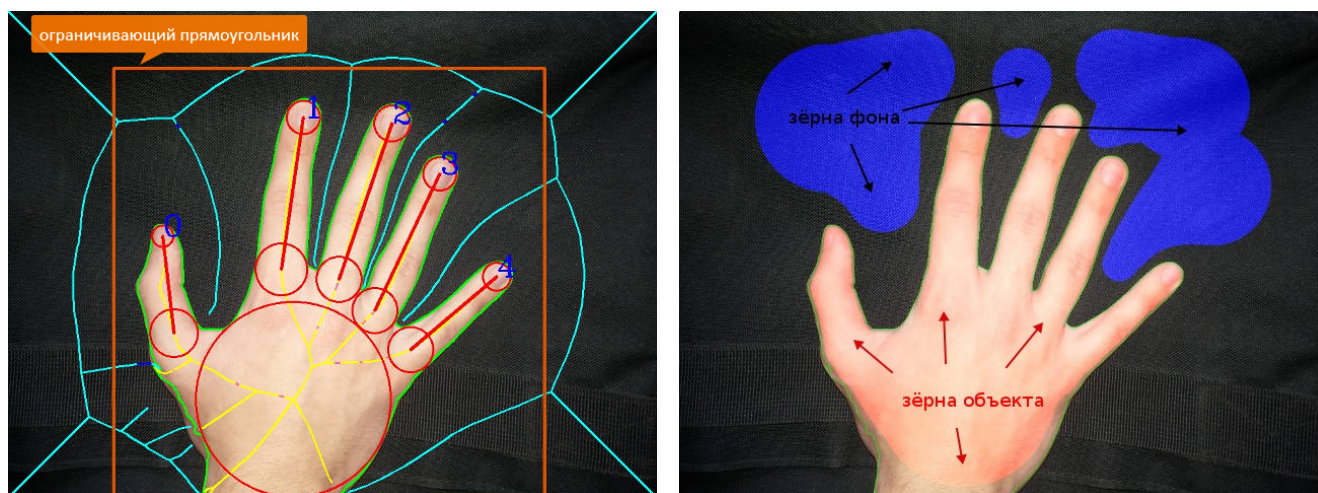
В разделе 3.4 предложен алгоритм детектирования, в котором можно выделить ряд последовательных этапов.

1. Обнаружение объекта-кандидата на роль ладони в кадре (задача «быстрой» сегментации, для решения которой выбрана бинаризация по Оцу).

2. Создание скелетного представления ладони на основе циркулярных графов⁵, построение семантической разметки (рис. 1а) и проверка ладони на корректность. Задачи семантической разметки: выделение оснований и концов пальцев, упорядочивание пальцев, поиск вершин внешнего скелета между пальцами. Предложенная процедура разметки [2] досрочно завершается (и ладонь считается некорректной), если на очередном шаге алгоритма ветви-кандидаты не удовлетворяют соответствующим критериям (например, базовые пороговые проверки расстояний, отсеивающие ветви-кандидаты, которые анатомически не могут являться пальцами в силу длины, расположения оснований и концов пальцев) или после действий остаётся менее 5 ветвей-кандидатов (например, после исключения побочных ветвей-кандидатов, соответствующих одному пальцу).
3. Определение расфокусированных кадров. Можно сформулировать основную идею так: если исходное изображение мало отличается от результата размытия по Гауссу, то кадр, вероятнее всего, расфокусирован. Так как фон с неоднородной текстурой может вносить существенные искажения в картину разности, то анализ проводится только для части кадра, лежащей внутри контура ладони и относящейся к пальцам.
4. В случае корректной ладони, находящейся в фокусе, — формирование данных для отправки в модуль распознавания. В частности, на базе семантической разметки формируются два множества пикселей с фиксированными метками (объекта или фона), которые не изменяются в процессе работы алгоритма сегментации (рис. 1b). Пиксели из указанных множеств также называют *зёрнами* (объекта или фона соответственно). Такую ладонь, пригодную для последующего распознавания, назовём *валидной*.

Таким образом, корректность ладони является необходимым (но не достаточным) условием её валидности, и только валидные ладони участвуют в последующем распознавании. Скелетное представление фигуры отражает её топологическую структуру, кроме того оно мало изменяется при деформациях ладони — это положительно влияет на стабильность работы метода детектирования. Более того, все этапы алгоритма детектирования допускают эффективную реализацию, что обеспечивает скорость обработки видеопотока на мобильном устройстве в режиме реального времени. Новизна предлагаемого автором метода заключается в успешном применении скелетного представления фигуры для решения задачи детектирования ладони.

⁵Местецкий Леонид Моисеевич. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. — М.: ФИЗМАТЛИТ, 2009. — С. 288.



(а) Скелетное представление ладони на основе циркулярных графов и семантическая разметка. (б) Зёрна объекта (красный) и фона (синий) для последующей «точной» сегментации.

Рис. 1: Использование скелетного представления фигуры в методе детектирования ладони.

Раздел 3.5 посвящён описанию метода распознавания. Для сегментации применяется алгоритм минимизации энергии на основе разрезов графов OneCut⁶ с использованием полученных ранее зёрен («точная» сегментация), продемонстрировавший оптимальное сочетание качества сегментации и скорости работы. Высокая эффективность во многом обеспечена предложенным способом получения «больших» зёрен объекта (ладони) и фона. Полученное бинарное изображение передается в модуль извлечения признаков формы, а также используется для выделения области интереса при извлечении текстурных признаков.

Генерация признаков формы описана в **разделе 3.5.2** и происходит следующим образом. По бинарному изображению (маске ладони) строится скелет, проводится семантическая разметка ладони и вычисляются признаки.

1. Спектр (распределение медиальной ширины)⁷.
2. Ломаная, последовательно соединяющая вершины пальцев и точки между ними. В качестве признака используются нормализованные длины ее звеньев (вектор длины 8).
3. Кривизны пальцев. Для каждого пальца вычисляется максимальное расстояние между точками скелета и осью пальца по обе стороны от последней (каждый палец описывается вектором длины 2).

⁶Grabcut in One Cut / Meng Tang, Lena Gorelick, Olga Veksler, Yuri Boykov // International Conference on Computer Vision. — 2013. — December.

⁷Местецкий Л. М. Медиальная ширина фигуры - дескриптор формы изображений //Машинное обучение и анализ данных. — 2014. — Т. 1, № 9. — С. 1291–1318.

Для сравнения спектров двух изображений выбрана EMD_{L1} -метрика, для сравнения векторов признаков ломаной и кривизны пальцев применена евклидова метрика. Результаты распознавания на основе признаков формы приводятся в **разделе 4.3.2** с экспериментами.

Для генерации текстурных признаков необходимо предварительно выделить области кожных складок пальцев. В **разделе 3.5.3** описан алгоритм, который на основе семантической разметки ладони локализует соответствующие участки изображения. Суть алгоритма в следующем. Изображение пальца с помощью перспективного преобразования приводится к прямоугольнику фиксированного размера (рис. 2). После предобработки, улучшающей контрастность и убирающей шум в однородных областях изображения, к полученному изображению применяется детектор границ Канни (Canny). Результатом его работы является *изображением границ* (рис. 2с) — бинарное изображение, ненулевые пиксели которого обозначают найденные границы. Для фиксированного размера окна поиска назовём *плотностью границ* количество ненулевых пикселей на изображении границ внутри данного окна. Поиск области складок проводится только в одном направлении — вдоль вертикали изображения границ. Точка, являющаяся центром окна с максимальной плотностью границ, выбирается в качестве центра искомой области кожных складок. По анатомическим причинам во время поиска центр окна не выходит за пределы верхней половины изображения пальца.

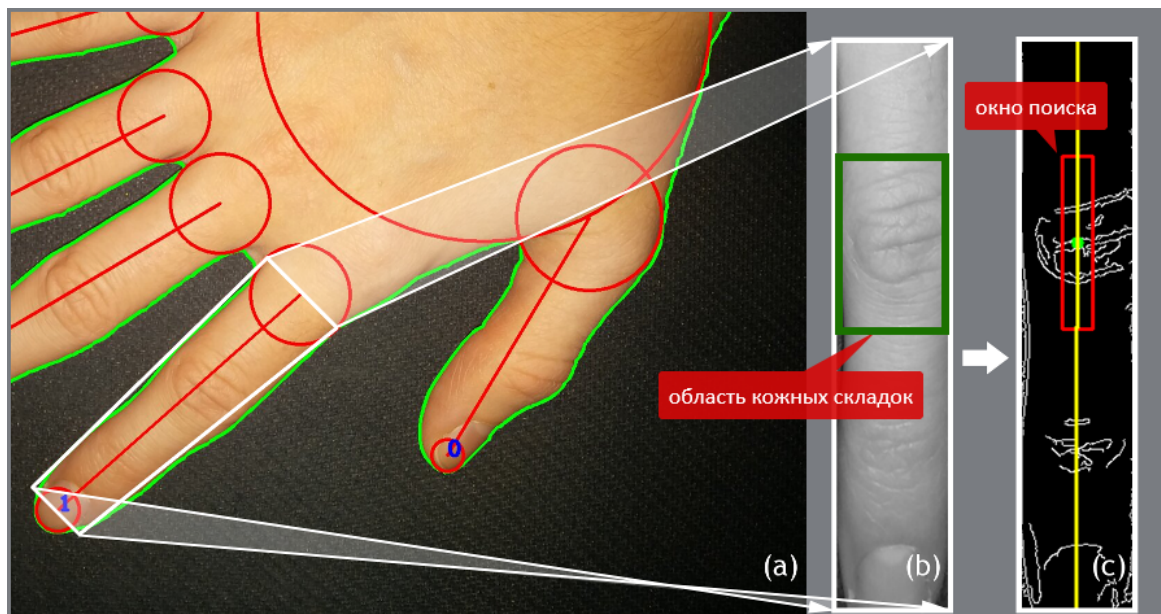


Рис. 2: Извлечение (а) и нормализация (b) области пальца. Локализация центра области складок пальца на изображении границ (с) с помощью окна поиска.

В **разделе 3.5.4** представлен метод извлечения текстурных признаков на основе фильтра Габора. В нём используется та же основная идея по извлечению локальной информации

об ориентации с помощью семейства двумерных фильтров Габора, что и в упомянутом ранее методе Competitive Code. Рассмотрим предлагаемый метод подробнее.

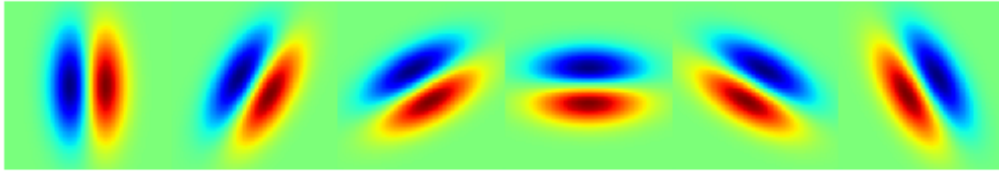
Фильтр Габора — линейный фильтр, импульсная характеристика которого имеет вид функции, представляющей собой произведение Гауссовой и гармонической функций:

$$G_R(x, y; \lambda, \theta, \varphi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \varphi\right),$$

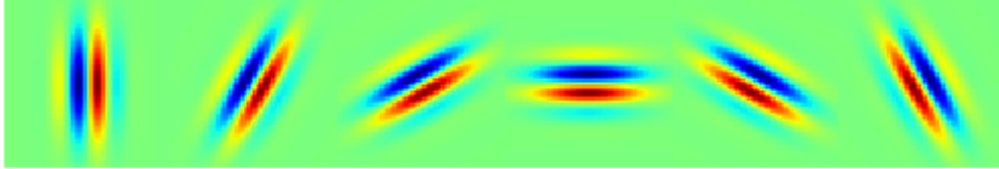
где

$$x' = x \cos \theta + y \sin \theta, y' = -x \sin \theta + y \cos \theta.$$

Согласно другой интерпретации, выше представлена действительная часть комплексной функции Габора. Далее под словами «фильтр Габора» и «функция Габора» мы будем понимать именно функцию G_R . Смысл параметров фильтра следующий: σ — стандартное отклонение Гауссова ядра, отвечает за амплитуду фильтра; λ определяет длину волны гармонической части фильтра; θ — угол поворота фильтра; γ определяет форму фильтра (чем ближе значение к единице, тем менее вытянутую форму имеет фильтр); φ — фаза фильтра.



(a) Первая эпоха.



(b) Вторая эпоха.

Рис. 3: Визуализация используемых двумерных фильтров Габора.

Структура двумерного фильтра Габора позволяет адаптировать его для извлечения информации о характере кожных складок на изображении: их толщине и направлении. Выполним свёртку изображения с несколькими фильтрами, имеющими разные ориентации. Значением признака $F(x, y)$ в данной точке (x, y) изображения I является номер фильтра с минимальным откликом:

$$F(x, y) = \arg \min_j \{I(x, y) * G_R(x, y; \lambda, \theta_j, \varphi, \sigma, \gamma)\},$$

где $I(x, y)$ — окрестность изображения с центром в точке (x, y) ; θ_j — угол поворота фильтра, $\theta_j = j\pi/J, j = 0, \dots, J-1$, J — количество различных ориентаций. При выборе

$J = 6$ (обусловлено адаптацией теории вейвлетов к физиологическим особенностям человека при прочих фиксированных параметрах получается семейство из 6 разноориентированных фильтров (рис. 3(а) или (б)). Для адаптации изложенного метода кодирования локальной информации об ориентации под поставленную задачу идентификации личности по изображению внешней стороны ладони автором внесены следующие принципиальные изменения.

1. Используется два семейства фильтров Габора с разными параметрами σ и λ , чтобы учесть текстурные особенности с разным характерным размером. Семейство из J фильтров, отличающихся между собой ориентацией (параметром θ_j), называется *эпохой* (рис. 3).
2. Вместо вектора признаков с длиной, равной количеству пикселей в области кожных складок пальца, используется значения $F(x, y)$ в вершинах равномерной прямоугольной сетки. Это позволяет существенно сократить размер вектора, и, как следствие, время извлечения и сравнения признаков, сохраняя приемлемый уровень точности.
3. Для каждого изображения ладони вычисляются текстурные признаки областей кожных складок каждого из трёх пальцев. Сопоставление изображений областей кожных складок выполняется для каждого из пальцев отдельно, затем полученные значения расстояний объединяются с помощью функции MIN. Это значительно повышает качество классификации в сравнении с работой только по одному пальцу.

При сравнении текстурных признаков (**раздел 3.5.5**) применяется расстояние, учитывающее круговую природу ориентации в каждой точке. Пусть есть два вектора $f^{(1)}$ и $f^{(2)}$ размерности N , содержащие Габор-признаки на базе одной из эпох для двух сравниваемых изображений кожных складок. Расстояние между ними будет равно:

$$GDist(f^{(1)}, f^{(2)}) = \frac{1}{3N} \sum_{j=1}^N \min\{d_j, 6 - d_j\}, \quad d_j = |f_j^{(1)} - f_j^{(2)}|.$$

Чтобы компенсировать ошибки поиска области кожных складок, при сопоставлении векторов признаков конкретной эпохи выполняются сдвиги вверх-вниз точек сетки, в которых берутся значения признаков. В диссертационной работе шаг сдвига равен 3 пикселя, используется 3 сдвига вверх и 3 вниз. В качестве результирующего расстояния по конкретной эпохе $GShiftDist(f^{(1)}, f^{(2)})$ берётся минимальное расстояние, полученное при попарных сдвигах векторов признаков относительно друг друга. Обозначим за $GShiftDist_i^k$ расстояние между парой изображений кожных складок пальца i на базе эпохи k . Тогда расстояние по Габор-признакам между изображениями кожных складок пальца i вычисляется следующим

образом:

$$GShiftDist_i = GShiftDist_i^1 + \alpha_g * GShiftDist_i^2, i \in \{1, 2, 3\},$$

где α_g — настраиваемый во время обучения параметр.

В разделе 3.5.6 представлен метод сравнения на основе дескрипторов SIFT (scale-invariant feature transform, размерность получаемого вектора признаков равна 128), инвариантных к масштабированию, изменению ориентации, изменению освещённости и частично инвариантных к аффинным искажениям. На сравниваемых изображениях строятся дескрипторы SIFT (рис. 4), которые используются для вычисления меры различия между пальцами описанным ниже способом.

Пусть найдено n_1 ключевых точек на первом изображении I_1 области кожных складок пальца и n_2 — на втором изображении I_2 . Обозначим $s_1 = \{(p_1^1, d_1^1), (p_2^1, d_2^1), \dots, (p_{n_1}^1, d_{n_1}^1)\}$ — множество пар ключевых точек и соответствующих дескрипторов для первого изображения, $s_2 = \{(p_1^2, d_1^2), (p_2^2, d_2^2), \dots, (p_{n_2}^2, d_{n_2}^2)\}$ — для второго изображения. Для каждого дескриптора d_i^1 ключевой точки p_i^1 на одном изображении вычисляется ближайший дескриптор на втором, в качестве метрики используется расстояние Евклида:

$$(p_{i*}^2, d_{i*}^2) = \arg \min_{(p^2, d^2) \in s_2} |d_i^1 - d^2|$$

Затем выполняется фильтрация полученных пар с целью отсеять ложные соответствия. Если после фильтрации осталось достаточное количество n_0 пар точек (дескрипторов), считается, что это изображения одного и того же пальца:

$$SDist(I_1, I_2) = \begin{cases} 0, & \text{если } n_0 \geq n_{pairs}; \\ 1, & \text{иначе.} \end{cases}$$

где n_{pairs} — пороговое значение, которое может выставляться экспертно или настраиваться во время обучения на конкретной базе ладоней. За $SDist_i$ обозначим расстояние между двумя изображениями кожных складок пальца i .

Раздел 3.5.7 объединяет результаты предыдущих разделов в единую функцию расстояния между изображениями ладоней. Итоговое расстояние между парой изображений кожных складок пальца i с использованием Габор- и SIFT-признаков вычисляется как линейная комбинация соответствующих расстояний:

$$Dist_i = GShiftDist_i + \alpha_s SDist_i,$$

где α_s — настраиваемый во время обучения параметр. Предложенное автором итоговое расстояние $Dist$ между изображениями ладоней получается с помощью функции MIN из попарных

расстояний между изображениями кожных складок каждого из трёх пальцев:

$$Dist = \min\{Dist_1, Dist_2, Dist_3\}.$$

В качестве классификатора используется метод ближайшего соседа (1NN) с функцией расстояния $Dist$. **Раздел 3.5.8** содержит выводы к **разделу 3.5**, посвящённому методу распознавания. В **разделе 3.6** изложены выводы к **главе 3**, посвящённой методу идентификации.

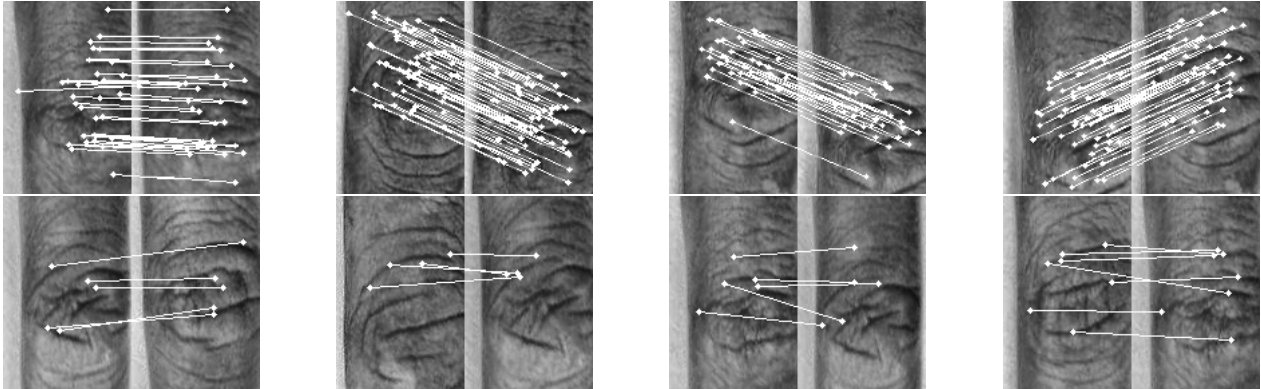


Рис. 4: Процедура сопоставления SIFT-дескрипторов для нормализованных изображений кожных складок пальцев. В верхнем ряду в поиске соответствия участвуют области кожных складок одного и того же человека, в нижнем ряду — разных людей.

В **главе 4** описано экспериментальное исследование методов детектирования и распознавания. В **разделе 4.1** приведено оборудование, использованное для сбора собственных баз изображений ладоней. В **разделе 4.2** содержится описание собственных (База-1 и База-2) и сторонних (Hand Geometric Points Detection Database, сокр. HGC и The Hong Kong Polytechnic University Contactless Hand Dorsal Images Database, сокр. PolyU HD) баз изображений ладоней, а также протоколов их сбора. В **разделе 4.3** представлены проведённые вычислительные эксперименты. **Раздел 4.3.1** содержит результаты тестирования метода детектирования на Базе-1 и HGC, которые подтверждают быстродействие и низкий уровень ошибок разработанного метода, а также его пригодность для работы с видеопотоком на мобильном устройстве. На базе HGC также проведено тестирование метода семантической разметки — полученные результаты сопоставимы с лучшими алгоритмами, участвовавшими в соответствующем состязании (2011 Hand Geometric Points Detection Competition).

В **разделе 4.3.2** изложены результаты тестирования метода распознавания на Базе-2 (съёмка «с руки») и PolyU HD (статичная съёмка) — табл. 1. Признаки формы продемонстрировали неудовлетворительные результаты на Базе-2. Причина, по-видимому, состоит в том, что признаки формы ладони сильно зависят от угла съёмки. Более того, форма ладони сама по себе обладает высокой внутриклассовой вариабельностью (это гибкий объ-

| Метод | Ошибка, % База-2 | AUC База-2 | EER, % База-2 | Ошибка, % PolyU HD | AUC PolyU HD | EER, % PolyU HD |
|------------------------------|---------------------|---------------|------------------|-----------------------|-----------------|--------------------|
| Кривизна | 21.0 | 0.9692 | 8.7 | 10.6 | 0.9750 | 7.7 |
| Ломаная | 22.1 | 0.9589 | 11.2 | 12.6 | 0.9703 | 7.9 |
| Спектр | 24.3 | 0.9407 | 15.1 | 13.4 | 0.9672 | 9.8 |
| Габор (палец 1) | 5.7 | 0.9923 | 4.0 | 11.1 | 0.9734 | 6.5 |
| Габор (палец 2) | 6.3 | 0.9877 | 3.6 | 12.3 | 0.9723 | 8.4 |
| Габор (палец 3) | 6.1 | 0.9955 | 3.7 | 11.4 | 0.9757 | 7.1 |
| Габор (три пальца) | 1.4 | 0.9996 | 1.1 | 1.8 | 0.9990 | 1.3 |
| Габор + SIFT (три пальца) | 0.9 | 0.9997 | 0.7 | 1.4 | 0.9997 | 0.8 |

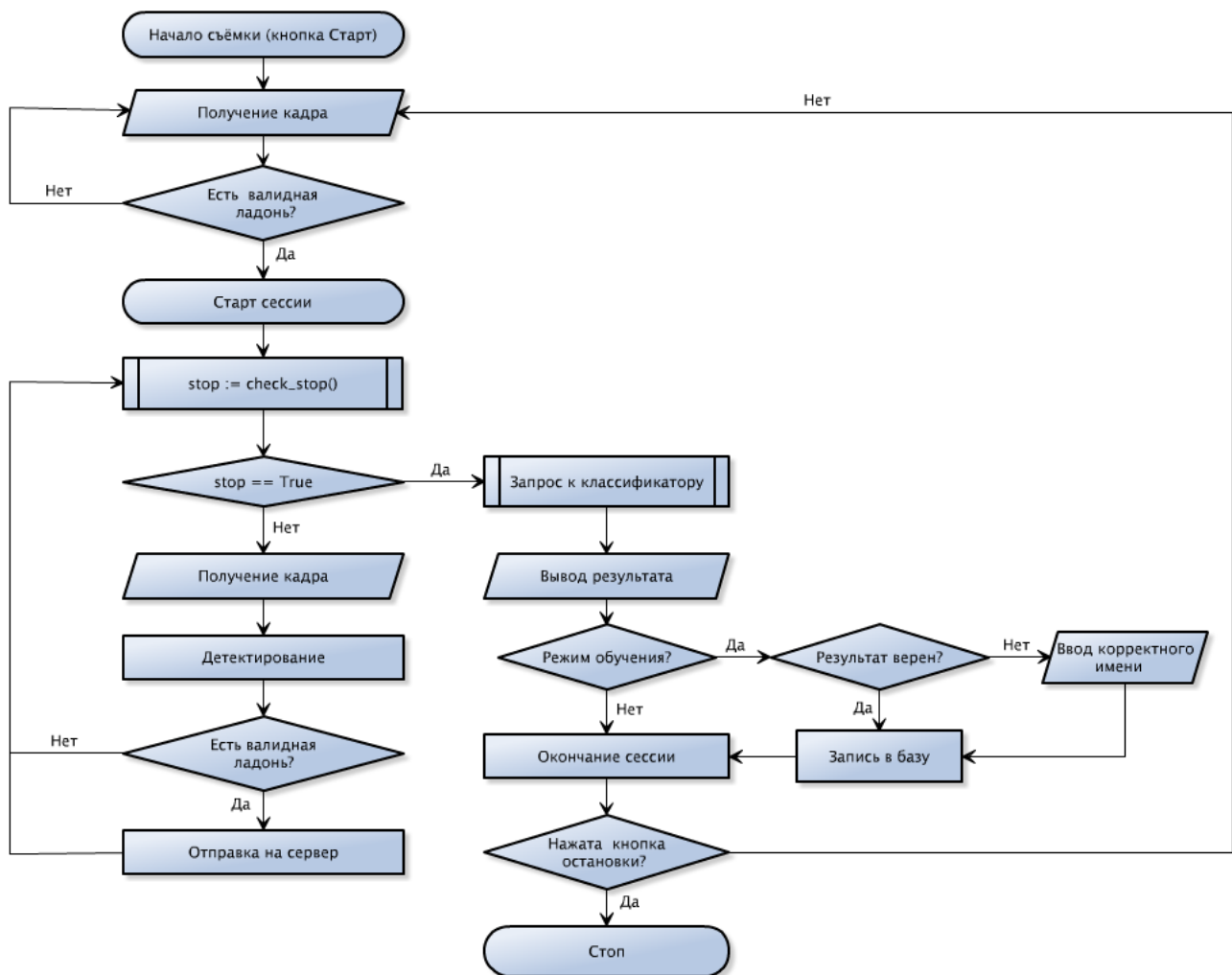
Таблица 1: Тестирование метода распознавания на Базе-2 и на базе PolyU HD при различных признаках. Столбцы «Ошибка, %...» содержат ошибку идентификации, усреднённую по $p = 10$ случайным разбиениям базы на обучающую и тестовую части в соотношении 2: 1 (оценка скользящего контроля). Столбцы «EER, %...» и «AUC, %...» — содержат характеристики ROC-кривой, построенной по разбиению базы, на котором достигается медианное значение EER среди всех разбиений.

ект) и невысокой межклассовой (несложно найти людей, чьи ладони визуально сильно похожи). Вследствие этого, признаки формы не используются в итоговом классификаторе. Тестирование подтвердило то, что расстояние между изображениями ладоней на основе объединения Габор-признаков трёх пальцев работает значительно лучше, чем на основе Габор-признаков любого из пальцев по отдельности. Опытным путём обосновано, что линейная комбинация Габор- и SIFT-признаков увеличивает точность классификации относительно использования только Габор-признаков.

Проведённые вычислительные эксперименты показали высокую информативность и устойчивость предложенного автором признакового описания на основе фильтров Габора и дескрипторов SIFT. Экспериментально обосновано использование двух семейств фильтров Габора для улучшения качества распознавания. Опытным путём доказано, что использование графовой («точной») сегментации значительно увеличивает точность метода распознавания. Достигаемые показатели точности метода распознавания ($EER \leq 0.8\%$) сравнимы с результатами передовых методов идентификации по кожным складкам внешней стороны пальцев, рассмотренных в обзоре. Количество людей (48 и 165 соответственно) в использованных базах позволяет утверждать, что разработанный метод подходит для идентифика-

ции в малых группах. Экспериментально доказано, что рисунок кожных складок пальцев, полученный из изображений внешней стороны ладони, является индивидуальным признаком личности человека, дающим возможность уверенной идентификации в малых группах при помощи мобильного устройства. **Главу 4** завершает **раздел 4.4** с выводами.

Рис. 5: Схема сессии. Здесь `check_stop()` — процедура проверки условий остановки сессии, она устанавливает флаг окончания сессии либо в положение `true` (сессия завершается) или `false` (сессия продолжается).



Глава 5 содержит описание разработанной в ходе исследования мобильной системы биометрической идентификации по внешней стороне ладони. В **разделе 5.1** изложена общая характеристика системы. Наряду с точностью и скоростью (обеспечиваются характеристиками метода идентификации), системы удовлетворяет требованиям доступности, гибкости и удобства, которые были предъявлены при постановке задачи исследования. В **разделе 5.2** содержатся определения и принципы работы с мобильным приложением. Основным при работе с системой является понятие *сессии*, схема которой приведена в разделе в виде набора блок-схем (рис. 5). Сессия — непрерывный и ограниченный по времени процесс съёмки

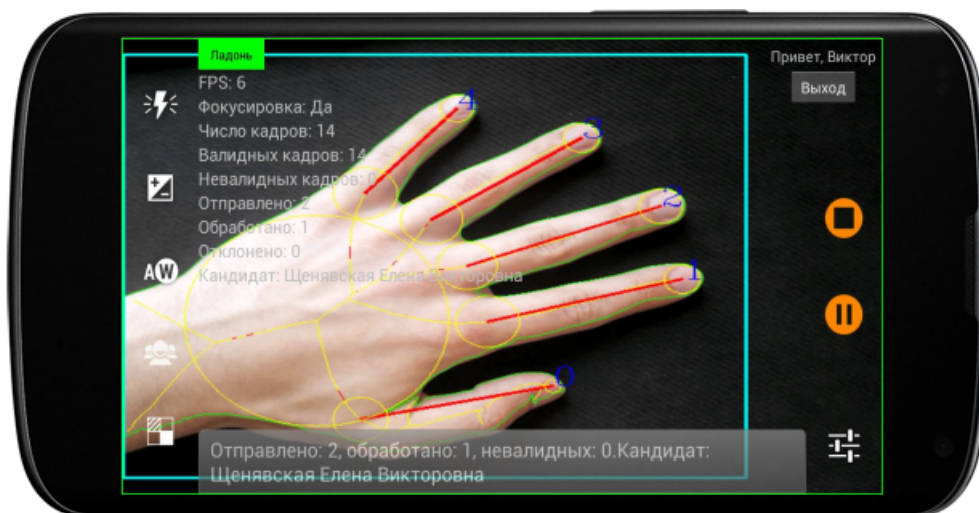


Рис. 6: Экран приложения во время сессии.

ладони одного человека (рис. 6). Вторым ключевым компонентом разработанной системы является сервер. К его задачам относятся классификация изображений ладони и хранение базы биометрических прецедентов. Общая схема работы сервера описана в **разделе 5.3**. Основными компонентами сервера являются: модуль распознавания, база данных, очередь задач, интерфейс взаимодействия с сервером. В **разделе 5.4** представлено описание веб-интерфейса. Через него администраторами осуществляется доступ ко всей биометрической информации, полученной при работе с системой и хранящейся на сервере. Ключевым и самым часто используемым компонентом веб-интерфейса является таблица посещений. В ней отражается статистика распознаваний пользователей по датам. По строкам расположены пользователи, по столбцам — даты распознаваний. В **разделе 5.5** приведены технические детали, касающиеся разработки кода. В силу того, что компоненты системы запускаются в существенно различных средах, для разработки приложения и сервера использовались несколько языков программирования (C++, Java и Python), а также активно использовалась кросс-компиляция кода. **Главу 5** завершает **раздел 5.6** с выводами.

В **заключении** сформулированы основные результаты диссертации.

Список публикаций

1. Chernyshov V. G., Mestetskii L. M. Mobile Computer Vision System for Hand-Based Identification // *Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications*. — 2015. — Vol. 25, no. 2. — Pp. 209–214.
2. Chernyshov V., Mestetskiy L. Real-time hand detection using continuous skeletons // *Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications*. — 2016.

— Vol. 26, no. 2. — Pp. 368–373.

3. *Chernyshov V., Mestetskiy L.* Mobile Machine Vision System for Palm-Based Identification // Proceedings of the 11th International Conference on Pattern Recognition and Image Analysis: New Information Technologies (PRIA-11-2013), Samara, Russia, 23-28 September, 2013. — Vol. 2. — IPSI RAS Samara, 2013. — Pp. 398–401.
4. *Chernyshov Victor.* Real-Time Hand Detection Using Continuous Skeletons // OGRW 2014 - Proceedings of the 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding, Koblenz, Germany, 1-5 December, 2014. — 2014. — Pp. 200–204.
5. *Chernyshov Victor.* Efficient Hand Detection on Client-server Recognition System // VISAPP 2015 - Proceedings of the 10th International Conference on Computer Vision Theory and Applications, Berlin, Germany, 11-14 March, 2015. — Vol. 2. — 2015. — Pp. 461–468.
6. *Чернышов В. Г.* Мобильная система контроля посещений по геометрии и текстуре ладони // Сборник трудов 26-ой международной конференции по компьютерной графике и зрению «ГрафиКон 2016», г. Нижний Новгород, Россия, 19-23 сентября 2016 г. — 2016. — С. 373–377.
7. *Чернышов В. Г.* Задачи детектирования и распознавания ладони в клиент-серверной системе биометрической идентификации // Сборник тезисов конференции «Техническое Зрение в Системах Управления-2015», г. Москва, Россия, 17-19 марта 2015 г.
8. *Чернышов В. Г., Местецкий Л. М.* Использование текстурных особенностей костяшек пальцев в клиент-серверной системе биометрической идентификации // Сборник тезисов 17-ой конференции «Математические Методы Распознавания Образов» (ММРО-17), г. Светлогорск, Россия, 19-25 сентября 2015 г.
9. *Чернышов В. Г.* Метод биометрической идентификации по изображению внешней стороны ладони // Сборник тезисов 18-ой конференции «Математические Методы Распознавания Образов» (ММРО-18), г. Таганрог, Россия, 9-13 октября 2017 г.