

На правах рукописи

Небаба Степан Геннадьевич

**Технология подготовки изображений лиц к распознаванию
личности в видеопотоке в режиме реального времени
на основе компенсации ракурса и трекинга лиц**

05.13.17 – Теоретические основы информатики

**Автореферат
диссертации на соискание ученой степени
кандидата технических наук**

Томск – 2017

Работа выполнена в ФГБОУ ВПО Национальный исследовательский Томский политехнический университет.

**Научный
руководитель:**

Захарова Алёна Александровна
доктор технических наук

Официальные
оппоненты:

Калайда Владимир Тимофеевич
доктор технических наук, профессор кафедры оптико-электронных систем и дистанционного зондирования ФГАОУ ВО «Национальный исследовательский Томский государственный университет»

Князь Владимир Александрович
кандидат технических наук, заместитель начальника подразделения ФГУП «Государственный научно-исследовательский институт авиационных систем»

Ведущая
организация:

Институт прикладной математики
им. М.В. Келдыша
Российской академии наук, г. Москва

Защита состоится «19» октября 2017 г. в 15 час. 15 мин. на заседании диссертационного совета Д 212.268.05, созданного на базе Томского государственного университета систем управления и радиоэлектроники по адресу: 634050, г. Томск, пр. Ленина, 40, ауд. 201.

С диссертацией можно ознакомиться в научной библиотеке ТУСУРа по адресу: 634045, г. Томск, ул. Красноармейская, 146 и на сайте ТУСУРа:

<https://tusur.ru/urls/hacsl66d>

Автореферат разослан «__» _____ 2017 г.

Ученый секретарь
диссертационного Совета



Костюченко
Евгений Юрьевич

Общая характеристика работы

Актуальность работы. Задача распознавания лиц является одной из первых практических задач, которая послужила стимулом для развития теории распознавания объектов. В последнее время в различных сферах деятельности, связанных с обработкой информации, возрастает потребность быстрой и правильной идентификации личности человека в видеопотоке в режиме реального времени.

В общем случае задача идентификации личности в видеопотоке может быть разделена на несколько этапов: поиск области лица на изображении; слежение за положением обнаруженного лица на последовательности кадров (трекинг); предварительная обработка изображения лица с целью компенсации условий съемки; сравнение изображения лица с изображениями в базе изображений.

Задача поиска лиц на изображении на сегодняшний день успешно решена и применяется во многих технических устройствах, например, в фототехнике. Трекинг является необязательным этапом, но позволяет соотносить изображения лиц людей на разных кадрах и делать распознавание по лучшему изображению из последовательности кадров, косвенно повышая точность идентификации.

Задача распознавания личности является более сложной, и на данный момент существующие алгоритмы приближаются к точности распознавания лиц человеком только в строго заданных условиях, либо требуют вычислительные ресурсы, недоступные при обработке видеопотока в режиме реального времени.

Развитие теории и практики распознавания объектов во многом было обусловлено достижениями зарубежных ученых: Дж. Филиппса, Г. Векслера, Т. Канадэ, М. Турка, П. Виолы, М. Джонса, Б. Лукаса и других. Весомый вклад в область анализа изображений внесли российские ученые: Вапник В.Н., Журавлев Ю.И., Сойфер В.А., Загоруйко Н.Г., Местецкий Л.М., Алпатов Б.А., Горелик А.Л., Желтов С.Ю., Сергеев В.В., Васин Ю.Г., Кориков А.М., Катаев М.Ю., Калайда В.Т., Конушин А.С., Спицын В.Г., Князь В.А., Галактионов В.А. и другие.

Число ошибок идентификации личности у всех современных систем распознавания лиц в режиме реального времени имеет явную зависимость от качества изображений лиц, механических помех, ракурса съемки идентифицируемого лица, условий освещения.

Для решения этих проблем существует два пути: расширение базы эталонных изображений изображениями с различными усло-

виями съемки (что не всегда возможно сделать), а также моделирование условий съемки на эталонных изображениях при их сравнении с анализируемыми. Примером второго подхода является синтез трехмерных (3D) моделей лиц и сравнение рельефа и текстуры получаемых объектов (Paysan P., Knothe R., Amberg B., Romdhani S., Vetter T., 2009). Данный подход предоставляет широкие возможности, но требует создания 3D модели каждого лица в базе данных, что является ресурсоемкой и нетривиальной задачей.

Частая потребность в быстром некооперативном распознавании личности затрудняет применение как классических двумерных (2D) методов распознавания, так и методов, использующих 3D модели.

Существующие быстрые 2D методы и алгоритмы достигают хорошей точности только при соблюдении жестких требований по ракурсу изображений лица и условиям съемки, что обусловлено механизмами кластеризации и машинного обучения.

Учитывая вышеперечисленные проблемы методов идентификации личности в режиме реального времени по растровому изображению лица, можно сделать вывод о необходимости исследования существующих методов и алгоритмов предварительной обработки поступающих изображений как единой технологии.

Предварительная обработка может проводиться по нескольким критериям, таким как контрастность, резкость, освещение, ракурс. Технологию предварительной обработки можно представить в виде комплексного алгоритма для задачи обработки видеопотока, в том числе и в режиме реального времени, который должен учитывать все критерии и обладать достаточным быстродействием.

Проблема предварительной обработки изображений чаще всего рассматривается как набор частных задач по улучшению качества, компенсации освещения и ракурса съемки, но не как сложная система взаимозависимых алгоритмов.

Необходимость оптимизации всех шагов обработки и анализа изображений лиц под конкретную задачу распознавания, будь то фотографии или видеопоток, создают предпосылки для создания комплексной системы алгоритмов.

Цель работы: повышение точности и скорости распознавания личности в видеопотоке в режиме реального времени посредством технологии подготовки изображений, обеспечивающей компенсацию ракурса и трекинг лиц.

Для этого необходимо решить следующие задачи:

1. Провести обзор существующих методов анализа изображений: оценки схожести, контрастности и резкости, нормирование освещения, и выбрать оптимальные.

2. Исследовать возможность применения 3D моделей лица человека в распознавании лиц по изображению лица в режиме реального времени и разработать алгоритм синтезирования моделей лиц, подходящий для этих условий.

3. Разработать новый алгоритм, снижающий влияние ракурса на точность распознавания личности по изображению лица путем корректировки ракурса с помощью синтезированных моделей и применимый для задач реального времени.

4. Предложить и реализовать новый алгоритм трекинга лица в видеопотоке, обеспечивающий эффективное слежение за лицом и обладающий быстродействием, допускающим его применение в системах распознавания личности в режиме реального времени.

5. Разработать новую комплексную технологию автоматического анализа и обработки изображений лиц с целью их подготовки к распознаванию личности в видеопотоке в режиме реального времени и представить ее в виде комплексного алгоритма.

6. Реализовать разработанные алгоритмы в виде программной системы и провести вычислительные эксперименты с целью анализа их качества и эффективности.

Объектом исследования в данной работе является технология распознавания лиц по изображению лица для задач реального времени.

Предмет исследования – процесс предварительной обработки и анализа изображений лиц, используемых в алгоритмах распознавания, а также влияние этого процесса на точность и скорость получения конечного результата распознавания.

Методы исследований. В работе использованы методы теории алгоритмов, теории моделирования, методы планирования численных экспериментов, а также методы теории вероятностей и математической статистики для количественной обработки данных экспериментов.

Научная новизна полученных результатов заключается в следующем:

1. Предложена новая комплексная технология подготовки изображений лиц, обеспечивающая функционирование программных систем распознавания личности в видеопотоке в режиме реаль-

ного времени в автоматическом режиме, отличительными особенностями технологии являются авторские алгоритмы компенсации ракурса и трекинга лиц (пп. 5, 7, 14 паспорта специальности).

2. Предложена оригинальная модификация метода формирования деформируемой модели лица (Deformable face model, DFM), использующая принципы восстановления ландшафтных поверхностей методом интерполяции Шепарда и синтезирующая 3D модель лица человека в автоматическом режиме (пп. 3, 5, 7 паспорта специальности).

3. Разработан новый алгоритм формирования набора изображений лиц из одного фронтального изображения лица, основанный на применении предложенной модификации метода DFM и синтезировании проекций этой модели на плоскость (пп. 5, 7 паспорта специальности).

4. Разработан новый алгоритм трекинга лиц в видеопотоке, комбинирующий методы background subtraction и Виолы-Джонса, повышающий скорость поиска и трекинга лиц по сравнению с алгоритмом, использующим только метод Виолы-Джонса (пп. 5, 7 паспорта специальности).

Теоретическая значимость результатов диссертационного исследования состоит в том, что разработаны, исследованы и апробированы новые методы и алгоритмы для обработки и анализа изображений лиц в видеопотоке с целью их подготовки к распознаванию личности программными системами реального времени.

Практическая ценность и реализация результатов работы. Разработанные в диссертации алгоритмические и программные средства предназначены для решения задач предварительного анализа и обработки изображений лиц в видеопотоке. Они могут найти применение при создании отечественных программных систем поиска людей и верификации личности, например, при создании мобильных систем идентификации на удаленных объектах, функционирующих без доступа к сетевой инфраструктуре и серверам анализа данных в режиме реального времени. Программная система, использующая трекинг, может использоваться для анализа массовых моделей поведения, таких как исполнение инструкций на предприятиях со строгим технологическим процессом, планов учений, либо эвакуаций. Результаты исследования методов обработки изображений, полученные в ходе данной работы, и реализованное программное обеспечение технологии могут использоваться в процессе обу-

чения студентов специальностей, связанных с компьютерным зрением, трехмерным моделированием и обработкой изображений.

Степень достоверности результатов проведенных исследований подтверждается результатами тестирования, апробации и внедрения разработанных алгоритмов и технологии подготовки изображений лиц на собственных и общедоступных базах изображений. Полученные результаты согласуются с результатами, полученными другими авторами.

Основные положения, выносимые на защиту:

1. Новая комплексная технология подготовки изображений лиц к распознаванию в видеопотоке в режиме реального времени, внедренная в составе программной системы распознавания личности по изображению лица, основанная на методе классификации *k* Nearest Neighbours, позволяет повысить точность распознавания на 6-9 %.

2. Предложенная оригинальная модификация метода DFM, использующая метод интерполяции Шепарда, позволяет автоматически синтезировать рельеф лица по единственному плоскому изображению.

3. Разработанный новый алгоритм формирования набора изображений лиц из одного изображения для применения в распознавании личности, основанный на предложенной оригинальной модификации метода DFM, позволяет применять методы классификации лиц в условиях ограниченного набора исходных изображений для задач распознавания в режиме реального времени и повышает точность идентификации на 0,5-4%.

4. Предложенный новый алгоритм трекинга лиц, комбинирующий информацию, полученную методами background subtraction и Виолы-Джонса, работающий в режиме реального времени и снижающий вычислительную нагрузку на программную систему распознавания на 20 и более процентов в сравнении с базовым алгоритмом трекинга, в основе которого метод Виолы-Джонса.

Апробация работы. Результаты диссертационной работы обсуждались и докладывались на следующих симпозиумах, семинарах и конференциях: XII Международная научно-практическая конференция студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии» (Томск, 2014), 26-ая Международная научная конференция «GraphiCon 2016».

На основе анализа и сравнения методов обработки изображений, проведенного в ходе данной работы, разработан модуль учеб-

ной дисциплины по основам обработки изображений и распознаванию образов для студентов специальности Б3.В11 «Интеллектуальные и информационные системы», бакалавриат, направления 09.03.03 «Прикладная информатика» кафедры Программной Инженерии ИК ТПУ.

Тестовый стенд в виде программно-аппаратного комплекса, включающий в себя IP-камеру и персональный компьютер с разработанной программной системой, работает в Кибернетическом центре ИК ТПУ. На полученных данных были реализованы и протестированы функции обработки изображений и распознавания лиц.

Аналогичный программно-аппаратный комплекс прошел апробацию и внедрен в ООО Научно-производственная компания «Техника дела», где используется для формирования входного набора синтезированных изображений лиц в системе распознавания и сравнительного анализа методов обработки цифровых изображений.

Получены акт о внедрении разработанной программной системы и справка об использовании результатов.

Публикации. Основные результаты диссертационного исследования изложены в 11 печатных работах, в том числе в 5 статьях из списка рецензируемых журналов, рекомендованных ВАК РФ, в 2 статьях в журналах, индексируемых Scopus, в 1 статье в журнале, индексируемом Web of Science.

На алгоритмы, реализующие предложенные в ходе работы методы обработки изображений, было получено свидетельство о государственной регистрации программы для ЭВМ, получен акт о внедрении результатов диссертационных исследований.

Личный вклад. Постановка задач исследования по теме диссертации выполнена автором совместно с научным руководителем, д.т.н. А.А. Захаровой. Обзор существующих методов и алгоритмов анализа и обработки изображений выполнен совместно с С.Ю. Андреевым и М.А. Макаровым. Модификация метода DFM разработана совместно с А.А. Захаровой и С.Ю. Андреевым. Алгоритмы формирования базы изображений лиц и трекинга разработаны совместно с А.А. Захаровой. Новая комплексная технология подготовки изображений лиц к распознаванию, оптимизированная для использования в режиме реального времени, и результаты экспериментов, представленные в диссертации, получены лично автором.

Объем и структура работы. Диссертация состоит из введения, четырех глав, заключения, списка использованных источников из

121 наименования и 4 приложений. Объем основного текста диссертации составляет 154 страницы машинописного текста, иллюстрированного 53 рисунками и 19 таблицами.

Содержание работы

Во введении обосновывается актуальность работы в данном научном направлении, формулируются цель и задачи исследования, приводится краткое содержание работы по главам.

В первой главе проанализированы цели и задачи исследования, проведен обзор и анализ литературных источников по теме исследования, методов и алгоритмов, связанных с трекингом, оценками контрастности, резкости, коррекцией освещения и ракурса. Показана возможность создания на основе этих методов и алгоритмов новой технологии обработки изображений лиц, оптимизированной для функционирования в составе программного обеспечения (ПО) распознавания личности в режиме реального времени.

На основе анализа существующих методов трекинга объектов в видеопотоке, сформулированы требования к трекингу лиц системами распознавания в видеопотоке в режиме реального времени.

В результате обзора методов и алгоритмов расстановки особых точек лица, метода оценки угла поворота POSIT, а также существующих подходов к формированию 3D модели лица человека, сделан вывод о необходимости модификации существующих методов для создания автоматизированного и быстродействующего алгоритма контроля ракурса лиц на изображениях.

Во второй главе обоснован выбор методов и алгоритмов, используемых в составе создаваемой технологии и выбор средств разработки.

Описана *авторская модификация метода синтеза 3D модели лица с использованием интерполяции Шепарда* в вариантах использования обобщенной и деформируемой моделей лиц.

Интерполяция каждой точки 3D модели по близости к эталонным моделям является более точным решением в сравнении с выбором одной из большого числа этих 3D моделей лиц, в случае, если искомое лицо не находится в базе моделей. Для вычисления модели применен модифицированный метод интерполяции Шепарда, выбранный за его быстродействие в 3D пространстве и простоту автоматизации: этот алгоритм имеет всего два регулируемых параметра интерполяции.

Это позволяет получить уникальную модель анализируемого лица, которая будет точнее любой из эталонных моделей, а увеличение числа этих моделей аналогично ведет к улучшению точности предсказания деформируемой 3D модели лица. Предложенная модификация метода описывается следующим алгоритмом:

А). Поиск соответствия между найденными особыми точками и известными 3D моделями.

Для того чтобы найти соответствие, необходимо получить фронтальную проекцию множества точек 3D моделей лиц $P_i(x, y, z)$. То есть, найти такое множество точек $P_i(x^*, y^*)$, при котором координаты центров глаз будут находиться на строго горизонтальной линии (координата x^*), а срединные точки рта и переносицы – на строго вертикальной линии (координата y^*) (ГОСТ Р ИСО/МЭК 19794-5-2006).

Так как в выбранном для предложенного алгоритма наборе моделей координаты x и y соответствуют координатам x^* и y^* фронтальной проекции, то итоговая проекция получается отбрасыванием третьей координаты z . Из этой проекции можно сформировать фронтальное изображение лица по облаку точек $P_i(x, y)$ и текстуре 3D модели.

Далее с помощью алгоритма ASM вычисляется набор особых точек $E_j(x, y)$ для изображения, полученного из 3D модели, и изображения анализируемого лица.

Б). Нормирование особых точек относительно масштаба моделей (например, расстояние между центрами глаз) и центра координат (например, кончик носа).

Для нормирования плоского изображения необходимо определить как минимум две точки, относительно которых будет проводиться масштабирование остальных координат и сведение центров координат. В общем случае координаты изображения лица соответствуют пикселям этого изображения и по определению лежат в плоскости целых положительных чисел, а координаты проекции 3D модели могут быть дробными и отрицательными числами (зависит от формата хранения модели).

В рассматриваемой задаче удобнее всего при вычислении коэффициента масштабирования опереться на координаты особых точек, вычисленные алгоритмом ASM. Достаточно взять две устойчивых точки, координаты центров глаз для изображения проекции модели, $E_l(x_l, y_l)$ и $E_r(x_r, y_r)$, а также для изображения анализируемого

лица, $E_l(x_l, y_l)$ и $E_r(x_r, y_r)$. Расстояния между этими точками находятся по формулам

$$D = \sqrt{(x_l - x_r)^2 + (y_l - y_r)^2} \quad \text{и} \quad D' = \sqrt{(x_l - x'_r)^2 + (y_l - y'_r)^2}.$$

В результате коэффициент масштабирования для изображения анализируемого лица определяется как соотношение $K = D/D'$.

В качестве точки центра координат взята точка-кончик носа, $E_n(x_n, y_n)$. Для совмещения центров координат необходимо найти расстояние между координатами точки $E_n(x_n, y_n)$ изображения проекции модели и координатами точки $E'_n(x'_n, y'_n)$ изображения анализируемого лица:

$$D_x = x_n - x'_n, \quad D_y = y_n - y'_n.$$

После вычитания этих расстояний из каждой точки изображения лица и последующего умножения на коэффициент масштабирования K изображение лица и проекции модели становятся нормированными относительно одной и той же системы координат.

В). Вычисление расстояний для каждой особой точки изображения лица и соответствующих особых точек всех 3D моделей.

Так как полученные координаты особых точек для изображений нормированы по точкам центров глаз E_l и E_r , а также центра носа E_n , то расстояния между остальными точками массивов E и E' характеризуют различия изображения анализируемого лица и 3D модели, использованной при формировании фронтальной проекции:

$$F_i = \sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2}.$$

Г). Вычисление координаты глубины для особых точек изображения лица интерполяцией значений в особых точках моделей.

После вычисления расстояний для проекции каждой из 3D моделей появится возможность применить формулы интерполяции Шепарда, задав радиус R . Таким образом, для каждой особой точки изображения лица находится координата глубины z , позволяющая судить о рельефе анализируемого лица.

Д). Определение координаты глубины для всех остальных точек искомой модели постепенным вычислением новых точек между известными особыми точками до тех пор, пока число точек модели не достигнет числа точек эталонных моделей.

Для получения более подробной 3D модели лица необходимо расширить массив до некоторого предела $Q \leq N$, где N – число точек, описывающих исходные 3D модели.

В этом случае к описанной ранее процедуре дробления треугольников изображения лица добавляется вычисление координаты z для новой точки. Необходимые веса w_i новой точки вычисляются по средневзвешенному значению весов нескольких ближайших особых точек. Этим достигается гладкость и непрерывность результирующей 3D модели.

Предложен *новый авторский алгоритм формирования набора изображений лиц для методов распознавания с использованием 3D моделей лиц*. Индивидуальная 3D модель лица может быть использована для алгоритма построения изображения лица, повернутого на заданный угол. Для решения задачи получения набора изображений в различных ракурсах необходимо произвести следующие операции:

- Разбиение полученной модели лица на треугольники. В результате этой операции будет получен набор треугольников, описывающий область лица на изображении.
- Дробление треугольников по выбранному критерию (тесселяция), позволяющему снизить размеры используемых треугольников и приблизить используемую треугольную сетку к равномерной. В качестве критерия разбиения можно брать максимальную площадь треугольника либо максимальную длину стороны.
- Каждый из треугольников с помощью операций вращения относительно заданного центра и аффинного преобразования смещается и деформируется в соответствии с заданной величиной угла.

В результате выполнения алгоритмов из одного изображения лица удается получить набор изображений, который может использоваться в алгоритмах распознавания личности.

Предложен *новый алгоритм трекинга лиц в видеопотоке, основанный на методе вычитания фона background subtraction и детекторе лиц Виолы-Джонса*, подходящий для применения в режиме реального времени. Он оперирует только данными о положении обнаруженных лиц в кадре и изменениями фона видео.

Видеопоток можно представить в виде последовательности кадров (изображений)

$$V_s = [f_1, f_2 \dots f_N],$$

где N – общее число кадров, f_i – текущий кадр. Любое лицо, найденное в кадре, описывается как набор $F_{Cij} = \{If_{ij}, Rf_{ij}\}$, где If_{ij} – изображение лица, Rf_{ij} – его положение в кадре, i – номер кадра, j – номер лица в кадре.

При этом Rf можно представить как набор из двух точек

$$Rf = \{p_1=(x_1; y_1), p_2=(x_2; y_2)\},$$

где p_1 – верхняя левая точка прямоугольной области, p_2 – нижняя правая точка, x и y – их координаты.

Тогда трек лица k в видеопотоке – это последовательность лиц F_{Cik} , соответствующая одному человеку: $Tr_k = [F_{C1k}, F_{C2k} \dots F_{CNk}]$.

Основная задача в трекинге состоит в сопоставлении положений целевого объекта на последовательности кадров. В случае слежения за положением лица предполагается, что объекты в следующем кадре меняют свою позицию незначительно: найденное изображение лица принадлежит треку Tr_k , если точка-центр прямоугольника Rf , $p_c = (\frac{x_1+x_2}{2}; \frac{y_1+y_2}{2})$, лица F_{Cij} принадлежит области прямоугольника Rf лица $F_{C_{i-1,j}}$, найденного на предыдущем кадре и принадлежащего треку Tr_k .

Данный метод является простым с вычислительной точки зрения и для систем распознавания может применяться по умолчанию, несмотря на низкую точность, из-за необходимости использования детектора Виолы-Джонса для выделения областей лиц. Для снижения вычислительной сложности детектора лиц предлагается предварительный анализ кадров видео на предмет наличия областей, потенциально содержащих лица, методом вычитания фона и обнаружения движения background subtraction.

До запуска детектора добавляется операция вычисления разницы между кадрами: $f_T=f_i-f_0$, $f_M=f_i-f_{i-1}$. Вычисляются бинарные изображения f_T и f_M с порогом, зависящим от качества видеопотока. Далее скользящим окном со сторонами, равными половине стороны минимального окна детектора, производится оценка сумм белых пикселей. Те области изображения, в которых сумма белых пикселей превышает выбранный порог, отмечаются как подходящие для дальнейшего анализа. В этих областях находятся подвижные объекты, в том числе лица людей.

Такой подход позволяет избежать вычисления вейвлетов Хаара для всего кадра целиком, значительно снижая вычислительную сложность трекинга, основанного на детекторе Виолы-Джонса.

Предложен авторский алгоритм, описывающий новую комплексную технологию подготовки изображений лиц к распознаванию в видеопотоке в режиме реального времени следующей последовательностью действий (рис. 1):

(1). Формирование базы лиц с различными ракурсами с помощью индивидуальной 3D модели, полученной с помощью модификации метода DFM (авторский алгоритм).

(2). Поиск лиц на изображении (алгоритм Виолы-Джонса).

(3). Слежение за обнаруженными лицами в последовательности кадров (авторский алгоритм трекинга).

(4). Оценка контрастности и резкости изображений, отбрасывание зашумленных неинформативных кадров.

(5). Компенсация недостатков освещения сцены (алгоритмы Difference of Gaussians или Log-Gabor).

(6). Построение и сравнение биометрических векторов изображений лиц (алгоритм k Nearest Neighbors (kNN), либо любой аналогичный классификатор в т.ч. основанный на нейронных сетях).

Для всех предложенных алгоритмов представлены схемы.

В качестве критерия оценки эффективности предложенной технологии и включенных в нее алгоритмов были взяты ключевые

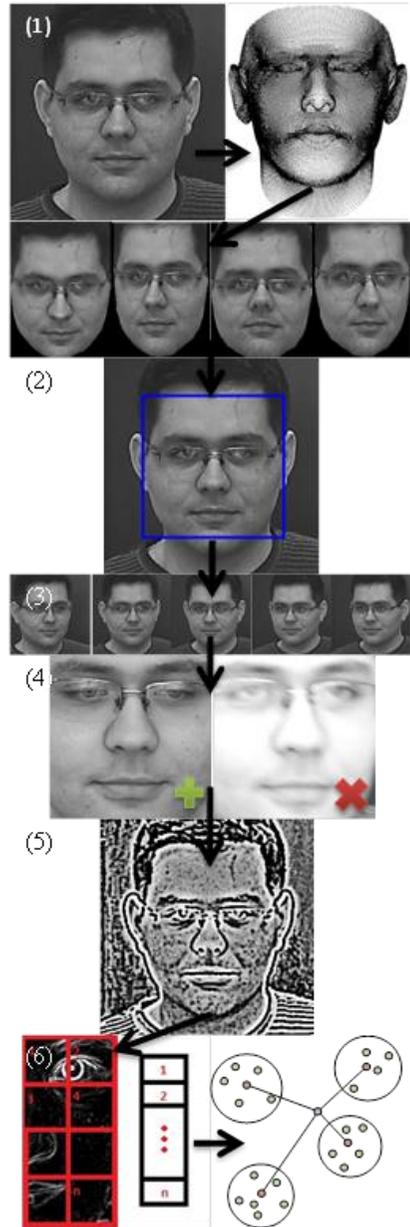


Рис. 1. Визуальное представление технологии

оценки качества результирующей системы распознавания личности по изображению лица: ошибка 1-го рода False Accept Rate (FAR), ошибка 2-го рода False Reject Rate (FRR), F_1 -мера, а также быстродействие. Для алгоритма трекинга в качестве критерия эффективности определена также длина получаемого трека в кадрах и скорость работы.

Разработанные алгоритмы были внедрены в программно-аппаратный комплекс, реализующий захват видеопотока с цифровой камеры, обработку кадров, распознавание личности по изображению лица и трекинг объектов. Программный комплекс был протестирован на базах изображений лиц и на видеофайлах.

Созданный программный комплекс состоит из нескольких модулей, осуществляющих основные функции и алгоритмы, применяемые при распознавании. Каждый из модулей имеет пользовательский интерфейс, реализованный с помощью QT.

В третьей главе описаны результаты тестирования алгоритмов трекинга и коррекции ракурса, а также результаты применения комплексной технологии предварительной обработки к решению задачи идентификации личности человека по изображению лица в видеопотоке.

Коррекция ракурса. Для проверки эффективности использования алгоритма автоматического синтезирования 3D моделей лиц при формировании набора изображений лиц проведено его тестирование в составе системы распознавания, которая была реализована в ходе исследования. Система основана на методах вычисления биометрических векторов стандартным отклонением (STD), снижения размерности векторов Principal Component Analysis (PCA) и классификаторе k Nearest Neighbors (kNN). Полученные векторы признаков сравнивались по метрике Евклида.

При сравнении результатов распознавания в качестве оценки эффективности алгоритма часто выбирается характеристика Equal Error Rate (EER). EER – равный уровень ошибок FAR и FRR, где FAR – вероятность ложного обнаружения, т.е. вероятность того, что система по ошибке признает подлинность пользователя, не зарегистрированного в системе, а FRR – вероятность пропуска цели, т.е. вероятность того, что система не признает подлинность зарегистрированного в ней пользователя.

Однако, такая оценка не всегда отражает реальную эффективность, так как она зависит не только от минимальных или опти-

мальных значений оценок FAR и FRR, но и от динамики их изменения, области применения алгоритмов. В системах, контролирующих вход людей на объекты ограниченного допуска, вероятность пропуска цели не такая критичная характеристика, как вероятность ложного обнаружения. В проведенных экспериментах оценкой эффективности системы идентификации личности принят уровень FRR при заданном уровне FAR=1% от выборки. В случае использования системы распознавания для поиска в кадре конкретного разыскиваемого лица ошибка FRR становится более значимой, и в таких задачах имеет смысл минимизировать именно её.

Базы изображений лиц для тестирования. Было использовано 3 базы изображений лиц: Caltech faces, FERET и собственная база.

Первое тестирование проводилось на собственной базе изображений. База состоит из 21 тысячи изображений лиц 12 человек в различных ракурсах и условиях съемки.

Второе тестирование проводилось на базе изображений лиц Caltech Faces. Эта база содержит изображения 26 человек, снятых во фронтальном ракурсе, но с различным качеством, освещением и мимикой, общее число изображений 450.

На рисунке 2 представлен пример изображения базы Caltech Faces и повернутых изображений, полученных с помощью авторского алгоритма формирования базы изображений.



Рис. 2. Слева направо: изображение лица из базы Caltech Faces; лица, повернутые на 20° влево и вправо по горизонтальной оси

Третье тестирование проводилось на базе изображений лиц FERET 2001 года. Эта база содержит изображения 996 человек, снятых в различных ракурсах (повороты на 0, 45 и 90 градусов по горизонтали относительно оптической оси камеры), с различным качеством, освещением и мимикой. Такая база не соответствует реальным задачам распознавания, но служит демонстрацией предела эффектив-

ности для алгоритмов компенсации различных ракурсов изображений лиц. Общее число изображений лиц, распознанных детектором Виолы-Джонса в реализации библиотеки openCV – 5335.

Результаты тестирования алгоритма распознавания с алгоритмом формирования наборов изображений. Полученные при тестировании результаты объединены в виде таблицы уровня ошибки FRR при фиксированном уровне FAR, а также в виде графика, показывающего соотношение уровней ошибок FAR и FRR для различных способов формирования набора входных изображений для системы распознавания личности.

На рисунке 3 и в таблице 1 представлены результаты тестирования системы распознавания с различными наборами изображений.

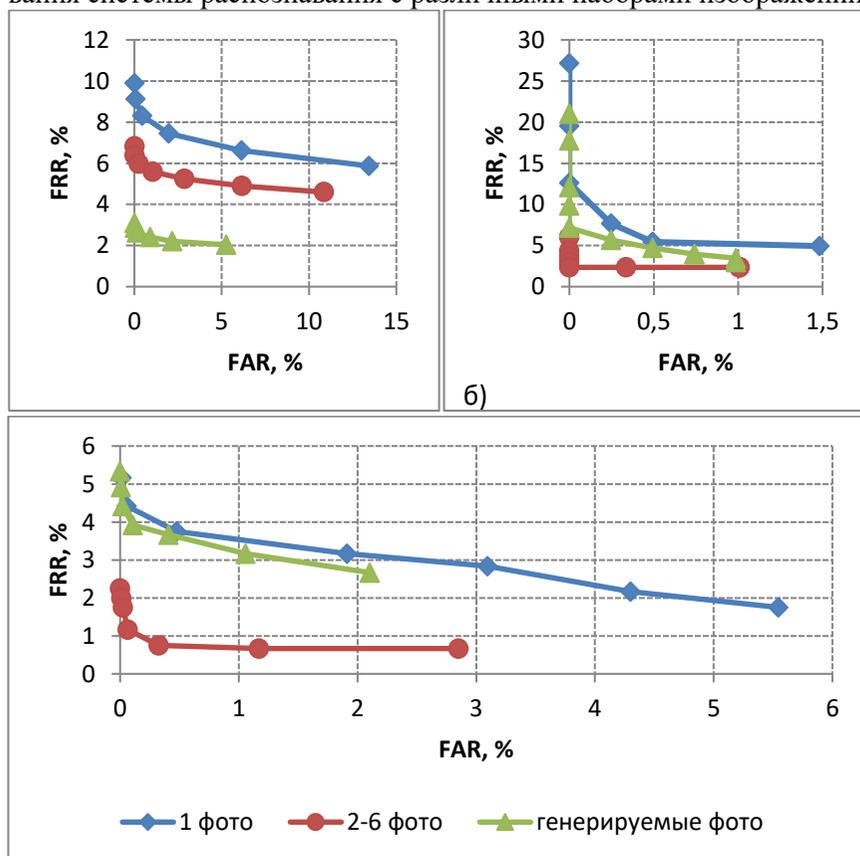


Рис. 3. График отношения уровня ошибок FRR к FAR
 а) видеофайлы; б) база Caltech Faces; в) база FERET

Таблица 1. Результаты тестирования системы распознавания личности на базах изображений лиц

База исходных изображений	Собственная	Caltech faces	FERET
Способ формирования базы	FRR, % (при FAR=1%)		
Одно изображение на человека	8,07	5,19	3,56
15 изображений на человека	4,82	2,35	0,88
15 сгенерированных алгоритмом изображений на человека	2,18	3,46	3,17

Как можно видеть из таблицы 1 и графиков на рисунке 3, результаты тестирования по выбранной метрике подтверждают эффективность алгоритма расширения набора исходных изображений: на собственной базе точность распознавания повысилась по сравнению с набором изображений из самой базы. На базах Caltech Faces и FERET тестирование показало, что набор синтезированных изображений повышает точность распознавания по сравнению с базой одиночных изображений.

Это позволяет утверждать, что использование синтетических изображений в разных ракурсах, полученных с помощью предложенной модификации DFM, повышает точность распознавания личности по изображению лица на 0,5-4% по сравнению с ограниченным набором изображений.

Трекинг. Для сравнительного тестирования взяты два варианта трекинга: трекинг на основе метода Виолы-Джонса и трекинг предложенным алгоритмом, использующим информацию о вычитании фона и движении в кадре.

Тестирование проводилось на собственной базе из 263 видеофайлов на следующей конфигурации компьютера: процессор Intel Core i7-3770 3.4 GHz, 16 Гб RAM, Windows 7 64-bit.

Разрешение видеофайлов 1280*720, на каждом видео присутствует один человек, совершающий разнообразные движения и поворачивающийся на различные углы. Всего детектором обнаружено 21735 изображений лиц.

Тестирование алгоритма, использующего метод Виолы-Джонса, дало следующие результаты: удалось составить 1846 треков, среднее время обработки одного кадра 16,75 мс.

Алгоритм, использующий предварительный анализ фона и движения, позволил снизить среднее время обработки одного кадра до 2,36 мс. Среднее время обработки кадров, содержащих движущиеся

объекты, снизилось до 12,14 мс, на 27%. Общее число треков – 1834, уменьшение произошло за счет исключения части ложных срабатываний детектора из последующей обработки.

Средняя длина трека в обоих случаях составляет 11 кадров, что позволяет применять предложенный алгоритм для вычисления усредненной оценки распознавания личности в треке.

Предложенный алгоритм применим для стационарной съемки и статического фона, на динамическом фоне этот алгоритм может привести к падению производительности системы.

Время обработки каждого кадра зависит от параметров детектора области лица, в тестовых примерах составляет порядка 10-15 мс. Предложенный алгоритм ускоренного трекинга объектов на статичном фоне не мешает применять иные методы повышения точности слежения за объектами на видео, в том числе и те, которые обеспечивают более высокую точность слежения при существенном снижении скорости (Девятков В.В., Алфимцев А.Н., 2011).

При съемке в реальных условиях возможно снижение средней длины трека, но общая тенденция сохранится, и трекинг как элемент предварительной обработки изображений является состоятельным и эффективным методом, позволяя усреднять оценку распознавания по треку.

Технология подготовки изображений лиц. Для оценки эффективности предложенных алгоритмов компенсации ракурса и трекинга лиц было проведено сравнение результатов, полученных с помощью предложенной комплексной технологии подготовки изображений, включающей авторские алгоритмы и известные методы повышения качества изображений, с мировыми аналогами систем распознавания на базе FERET (по информации, полученной на сайте Ayonix.com). В таблице 2 показано сравнение точности и скорости 4 доступных для сравнения алгоритмов.

Таблица 2. Сравнение результатов распознавания существующих систем и предложенной технологии

Критерий	Germany Company	Japan Company	Ayonix v4.0 3D Release	Технология с DFM
ЕЕR, %	97,9	97,9	99,6	97,4
Скорость поиска, мс	870	1200	468	152

Из таблицы 2 очевидно, что метод классификации kNN с применением предложенной технологии предварительной подготовки изображений уступает в точности ведущим мировым алгоритмам распознавания лиц, но заметно быстрее производит поиск по базе изображений, что подтверждает эффективность применения технологии для задач распознавания лиц в режиме реального времени.

В четвертой главе представлены результаты апробации и внедрения предложенного комплекса алгоритмов, а также результаты тестирования этой новой технологии на реальных данных.

На программную реализацию оригинальных авторских алгоритмов коррекции ракурса получено свидетельство о регистрации программы для ЭВМ.

На основе анализа и сравнения методов обработки изображений составлен учебный модуль по основам обработки изображений и распознаванию образов для студентов специальности БЗ.В11 «Интеллектуальные и информационные системы», бакалавриат, направления 09.03.03 «Прикладная информатика» кафедры Программной Инженерии ИК ТПУ.

С помощью IP-камеры, установленной на входе Института кибернетики Томского политехнического университета (ИК ТПУ) на видео фиксируются лица сотрудников и студентов, заходящих в корпус. Этот видеопоток обрабатывается разработанным программным комплексом непрерывно в режиме реального времени. Также на двух видеозаписях общей длительностью около 8 часов, полученных в этих условиях, был проведен тестовый контроль.

Эти видеозаписи сделаны в сложных условиях: разнообразные ракурсы лиц, освещение меняется в зависимости от времени суток.

На видеозаписях алгоритмом Виолы-Джонса обнаружено 16464 лиц. Технология подготовки лиц к распознаванию протестирована следующим на базе изображений лиц сотрудников института Кибернетики численностью 40 человек, распознавание лиц на кадрах проведено последовательно с каждым из алгоритмов обработки изображений классификатором kNN.

Среди всех найденных лиц число лиц людей, содержащихся в базе сотрудников ИК – 1193. Ввиду большого числа неизвестных лиц и сложных условий съемки принят фиксированный уровень FAR=10%.

В ходе тестирования технологии проведена оценка влияния авторского метода компенсации ракурса (DFM), в сравнении с мето-

дами компенсации освещения Difference of Gaussians (DoG) и Log-Gabor, на точность и скорость распознавания методом классификации изображений kNN при числе соседей $k=1$. Также подобрана комбинация алгоритмов обработки изображений для заданных условий (DFM + DoG + фильтр контрастности и резкости), минимизирующая уровень ошибки FRR и время обработки одного кадра. Результаты этого эксперимента представлены в таблице 3.

Таблица 3. Результаты тестирования компонентов новой технологии в составе системы распознавания личности

Алгоритм обработки	Число верных опознаний	FRR, %	Время вычисления, мс
kNN, $k=1$	15	62,54	16
kNN + трекинг	15	62,54	10
kNN + DFM	38	55,84	16
kNN + DoG	72	6,98	17
kNN + DoG + LogGabor	19	60,37	31
kNN + фильтр контрастности и резкости	67	14,38	16
Комбинация алгоритмов	96	2,94	11

Как можно видеть из таблицы 3, сочетание авторских алгоритмов компенсации ракурса и трекинга лиц в составе комплексной технологии позволяет добиться существенного улучшения результатов распознавания по соотношению ошибок FAR/FRR и повышения скорости обработки кадров, что свидетельствует об эффективности предложенной технологии обработки изображений для задач распознавания в режиме реального времени.

Программно-аппаратный комплекс, аналогичный установленному в ИК, был внедрен ООО Научно-производственная компания «Техника дела». Получен акт внедрения.

Результаты и выводы

В ходе выполнения диссертационной работы были получены следующие основные научные и практические результаты.

1. Выбраны наиболее подходящие с точки зрения эффективности и быстродействия методы анализа и обработки изображений в рамках технологии подготовки изображений лиц к распознаванию

личности в режиме реального времени по результатам оценки различных существующих методов и алгоритмов.

2. Произведен анализ существующих алгоритмов расстановки особых точек лица, оценки угла поворота, а также синтеза 3D моделей. Обоснована необходимость модификации существующих алгоритмов с целью повышения их скорости работы и автоматизации.

3. Разработан новый алгоритм трекинга лиц в видеопотоке, комбинирующий информацию, полученную методом background subtraction и методом Виолы-Джонса, и вычисляющий трек лица в видеопотоке со статичным фоном на 20 и более процентов быстрее алгоритма трекинга, основанного на детекторе Виолы-Джонса.

4. Разработана оригинальная модификация метода синтеза DFМ, использующая принципы восстановления ландшафтных поверхностей методом интерполяции Шепарда, позволяющая автоматически синтезировать рельеф лица по единственному плоскому изображению.

5. Разработан новый алгоритм формирования базы изображений лиц из одного фронтального изображения, основанный на предложенной модификации метода DFМ и позволяющий применять методы классификации лиц в условиях ограниченного набора исходных изображений, повышая точность программных систем распознавания личности, использующих методы классификации.

6. Предложена новая комплексная технология подготовки изображений лиц к распознаванию личности, основанная на авторских методах компенсации ракурса и трекинга лиц и рассчитанная на использование в программных системах распознавания в видеопотоке в режиме реального времени.

7. Протестированы авторские методы и алгоритмы в составе новой комплексной технологии обработки изображений лиц для программной системы распознавания личности на различных базах изображений лиц, как собственных, так и находящихся в открытом доступе.

8. Проведена апробация рассмотренных алгоритмов и методов в рамках комплексной технологии обработки изображений лиц, подтверждающая их быстроедействие и демонстрирующая повышение точности программных систем распознавания личности по изображению лица в видеопотоке в режиме реального времени.

9. Разработанное алгоритмическое и программное обеспечение внедрено в научно-образовательной и промышленной организациях,

а результаты проведенной работы - в учебный процесс, что подтверждено соответствующими документами.

Публикации по теме диссертации

Статьи в журналах, которые включены в перечень российских рецензируемых научных журналов и изданий для опубликования основных научных результатов диссертаций:

1. Небаба С.Г. Автоматическая оценка ракурса лица в кадре и приведение изображения к нулевым углам поворота / С.Г. Небаба, А.А. Захарова, С.Ю. Андреев // Научная визуализация, 7 (4), 2015.

2. Небаба С.Г. Алгоритм формирования индивидуальной трехмерной модели лица человека из растрового изображения лица и набора частных трехмерных моделей лица / С.Г. Небаба, А.А. Захарова, С.Ю. Андреев // Информационное общество, №5, 2015.

3. Савицкий Ю.В. Анализ методов извлечения вектора признаков из изображения лица человека для задачи идентификации / Ю.В. Савицкий, С.Г. Небаба, В.Г. Спицын, С.Ю. Андреев, М.А. Макаров // Научная визуализация, 8 (2), 2016.

4. Небаба С. Г. Тестирование технологии подготовки изображений лиц к распознаванию личности в видеопотоке в режиме реального времени // Современная наука: актуальные проблемы теории и практики. Серия «Естественные и технические науки», 2017. №3-4. С. 73-77.

5. Небаба С.Г. Алгоритм построения деформируемых 3D моделей человеческого лица и обоснование его применимости в системах распознавания личности / С.Г. Небаба, А.А. Захарова // Труды СПИИРАН. 2017. Вып. 52. С. 157-179.

6. Небаба С.Г. Трекинг изображений лиц в видеопотоке применительно к алгоритму распознавания лиц / С.Г. Небаба, А.А. Захарова // Кибернетика и программирование (принято к печати)

Scopus u Web of Science:

7. Nebaba S.G. Methods of Automatic Face Angle Recognition for Life Support and Safety Systems / S.G. Nebaba, A. A. Zakharova, T.V. Sidorenko, V.R. Viitman // The European Proceedings of Social & Behavioural Sciences EpSBS (January 2017), vol. XIX, pp. 735-744.

Другие публикации:

8. Небаба С.Г. Подготовка изображений лиц в видеопотоке к распознаванию и фильтрация неинформативных изображений / С.Г. Небаба, С.Ю. Андреев, М.А. Макаров, А.А. Захарова // Сборник трудов XII Международной научно-практической конференции сту-

дентов, аспирантов и молодых учёных «Молодёжь и современные информационные технологии». – Томск, 2014.

9. Небаба С. Г. Подготовка изображений лиц к распознаванию и фильтрация неинформативных изображений / С.Г. Небаба, С.Ю. Андреев, М.А. Макаров // Проблемы информатики. 2014. № 4. С. 53-60.

10. Nebaba S.G. Preparation of the Face Images in a Video Stream for Recognition and Filtering of Non-Informative Images / S.G. Nebaba, S.Y. Andreev, M.A. Makarov // Applied Mechanics and Materials, Sensors, Measurement, Intelligent Materials and Technologies III, 2015, Vols. 738-739, P. 648-655.

11. Небаба С.Г. Применение алгоритма формирования индивидуальной трехмерной модели человеческого лица в системе распознавания личности по изображению лица / С.Г. Небаба, А.А. Захарова // Труды 26-ой Международной научной конференции «GraphiCon2016», С. 310-313.

12. Небаба С.Г. Технология подготовки изображений лиц к распознаванию личности в видеопотоке в режиме реального времени // Материалы XX Юбилейной Международной конференции по вычислительной механике и современным прикладным программным системам (ВМСППС'2017), 24-31 мая 2017 г., Алушта. – М.: Изд. МАИ, 2017. – 816 с. С. 600-602.

Свидетельства о регистрации программ для ЭВМ:

13. Свидетельство о регистрации программы для ЭВМ №2017611032 «Создание набора изображений лиц для алгоритмов распознавания». Авторы: Небаба С.Г., Захарова А.А. Дата регистрации в Реестре программ для ЭВМ 19 января 2017 г.