

КУРГАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Кулешов Сергей Викторович

ПРИМЕНЕНИЕ СКРЫТЫХ МАРКОВСКИХ МОДЕЛЕЙ  
ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ

**РЕФЕРАТ**

Курган 2003 г.

Работа выполнена в Курганском Государственном университете.

Научный руководитель: кандидат физико-математических наук  
доцент, Симахин В. А.

В данной работе содержится обзор биометрических систем идентификации человека основанных на распознавании лиц, описано их назначение и область применения.

Подробно описывается выбор проектных решений и математических методов решения задачи. Содержится описание стадии проектирования и реализации системы на основе описанных алгоритмов, производится анализ, и сравнение эффективности различных подходов, которые были реализованы в данной системе.

Стр. 64, Рис. 22, Табл.2, Библиогр.: 14 назв.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Автоматические охранные системы с "фейс-контролем" и мониторингом несанкционированной активности, игрушки-роботы и автономные космические зонды, анализ документов и изображений, видеобазы данных, самообучающиеся компьютерные системы, управление компьютером с помощью жестов, интерфейсы прикладных программ. Это лишь некоторые примеры использования технологий "компьютерного зрения" - одного из самых перспективных исследовательских направлений, которое открывает новые методы взаимодействия с компьютером без использования клавиатуры или мыши. Чтобы проложить дорогу к широкому потребителю, нужны дружелюбные и персонализированные способы общения с компьютером. Это значит, что интерфейсы нового поколения должны идентифицировать окружающую человека обстановку и, как минимум, узнавать его самого.

Итак, необходимо научить компьютер идентифицировать личность человека. Какую технологию для этого выбрать? Существует большое разнообразие методов идентификации, причем многие из них получили широкое коммерческое применение и используются уже немало лет. На сегодняшний день в основе наиболее распространенных технологий верификации и идентификации лежит использование паролей и персональных идентификаторов (personal identification number — PIN) или документов типа паспорта, водительских прав. Однако такие системы слишком уязвимы и могут легко пострадать от подделки, воровства и просто человеческой забывчивости. Поэтому все больший интерес вызывают методы биометрической идентификации, позволяющие определить личность человека по его физиологическим характеристикам путем распознавания по образцам. Классический пример биометрии – анализ отпечатков пальцев, а к новейшим технологиям относятся распознавание сетчатки и радужной оболочки глаза.

Применение этих биометрических методов оправдано в банковской сфере и некоторых предприятиях для ограничения доступа к секретной информации, однако у них есть один серьезный недостаток. Такие технологии слишком навязчивы и в физиологическом, и в социальном смысле. Они требуют от человека чтобы он занял определенное положение относительно чувствительного элемента и простоял так несколько секунд. И эта процедура вряд ли изменится, поскольку технология требует самого тщательного пространственного анализа. Надо иметь в виду, что, общаясь между собой, люди не опознают друг друга путем сканирования сетчатки, поэтому подобные методы идентификации кажутся им не вполне естественными.

Интерфейс типа «остановись и продекларируй себя» нужен приложениям с высокими требованиями к безопасности (некоторая задержка заставит пользователя осознать важность проблемы). Однако для магазина, в котором узнают своих постоянных клиентов, для информационного киоска, который вас

«помнит», или здания, которое «знает», кто в нем живет, нужно совсем другое. Для интеллектуальных сред нового поколения лучше всего подходят технологии распознавания лиц и голоса. Они ненавязчивы (распознавание происходит на расстоянии, не задерживая и не отвлекая человека), они, как правило, пассивны (не требуют специального уровня освещенности), они не ограничивают пользователя в свободе перемещений, к тому же потребляют мало энергии и недороги. Но самое важное, по-видимому, то, что люди обычно узнают друг друга по лицам и голосам, значит, не будут испытывать неудобств с системой, основанной на аналогичных способах распознавания. Также новый стимул разработке систем распознавания человека по лицу дало обострение ситуации с международным терроризмом. Установка подобных систем в местах массового скопления людей (аэропортах, крупных торговых центрах) должна способствовать раннему выявлению лиц, находящихся в розыске.

# 1 АНАЛИТИЧЕСКИЙ ОБЗОР

## 1.1 Описание предметной области

Не так давно проблема распознавания лиц считалась одной из сложнейших задач искусственного интеллекта и компьютерного зрения. Однако целая череда успешных реализаций прошлого десятилетия показала, что этот метод идентификации личности не только технически осуществим, но и экономически выгоден.

Когда разрешимость проблемы распознавания лиц стала очевидна, мечты об интеллектуальных средах породили мощную волну интереса к этой задаче как со стороны исследователей, так и со стороны инвесторов. Появилось несколько преуспевающих коммерческих предприятий. Сегодня ряд компаний представляют на рынке программные системы распознавания лиц, которые обеспечивают распознавание высокой точности и имеют базы данных с более чем тысячами изображений людей.

Для того чтобы достичь своих первых успехов, разработчикам потребовалось четко определить и реализовать методы распознавания по образцу и хорошо разобраться в довольно непростом процессе генерации изображений. Кроме того, исследователи обнаружили, что им поможет знание присущих людям регулярных свойств. Например, в основе того или иного цвета человеческой кожи лежит одномерное многообразие цветового пространства, а цветовые оттенки зависят, прежде всего, от концентрации меланина. Существует ограниченное число вариантов геометрии человеческого лица, и если человек смотрит на камеру, его лицо, по существу, двумерно. Сегодня идет работа над тем, чтобы ослабить некоторые ограничения существующих алгоритмов распознавания лиц и добиться их устойчивой работы в случае изменения в освещенности, старения или смены выражения лица. Кроме того, исследователи заняты решением проблемы восприятия тех или иных перемен внешнего облика человека, таких как, скажем, волосы на лице, очки или макияж. На этот счет существует уже несколько частных решений. Но область исследований настолько широка, что технологии компьютерного зрения и распознавания образов просто обречены на вечную доработку и совершенствование. Неудивительно, что в этом направлении самостоятельно работают многочисленные исследовательские центры по всему миру, а узкоспециализированные разработки применяются в продуктах различных компаний. Например, технологии распознавания образов в роботе Sony и зонде NASA созданы "с нуля" и закрыты для конкурентов.

На рис.1 представлена формальная модель биометрической системы идентификации человека на основе распознавания лиц. Система должна на основе фотографии, используя БД определить личность человека (дать идентификатор, однозначно указывающий на конкретную личность) и показатель уверенности. При

невозможности распознавания (отсутствие описания соответствующего человека в БД) или неудовлетворительном качестве фотографии система выдает сообщение о невозможности идентификации данного человека.

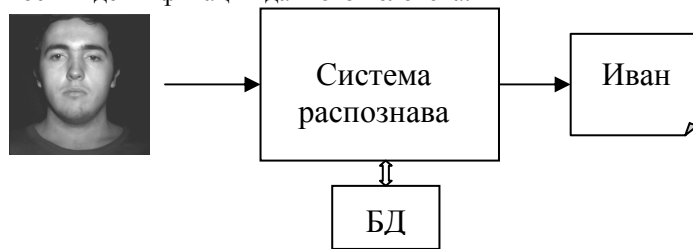


Рис. 1 Формальная модель биометрической системы.

## 1.2 Обзор существующих методов решения

Распознавание лиц имеет столь же давнюю историю, что и проблема компьютерного зрения (computer vision). Эта задача имеет серьезную практическую перспективу и вызывает большой исследовательский интерес. Несмотря на то, что другие методы идентификации (отпечатки пальцев или сканирование радужной оболочки), возможно, более точны, распознавание лиц всегда было в центре внимания ученых в силу своей «неагрессивности» по отношению к пользователям. Этот метод опознания личности для человека естественен и реализуется на интуитивном уровне.

Все системы распознавания лиц можно разделить на 2 класса. К первому классу относятся системы, которые проверяют соответствие предложенной фотографии конкретному человеку. В этих системах вначале необходимо набрать свой код, фамилию или вставить идентификационную карту, а система лишь сравнивает внешность человека с заявленной внешностью. Системы второго класса не требуют от человека называть себя. Они сами выбирают кандидатов и проверяют соответствие лица человека и лиц кандидатов.

При всём многообразии различных алгоритмов и методов распознавания изображений, типичный метод распознавания состоит из трёх компонент, рис. 2

1. преобразование исходного изображения в начальное представление (может включать в себя как предобработку, так и математические преобразования, например вычисление главных компонент);
2. выделение ключевых характеристик (например берётся первые  $n$  главных компонент или коэффициентов дискретного косинусного преобразования);

3. механизм классификации (моделирования): кластерная модель, метрика, нейронная сеть и т.п.

Кроме этого, построение метода распознавания опирается на априорную информацию о предметной области (в данном случае – характеристики лица человека), и корректируется экспериментальной информацией, появляющейся по ходу разработки метода.

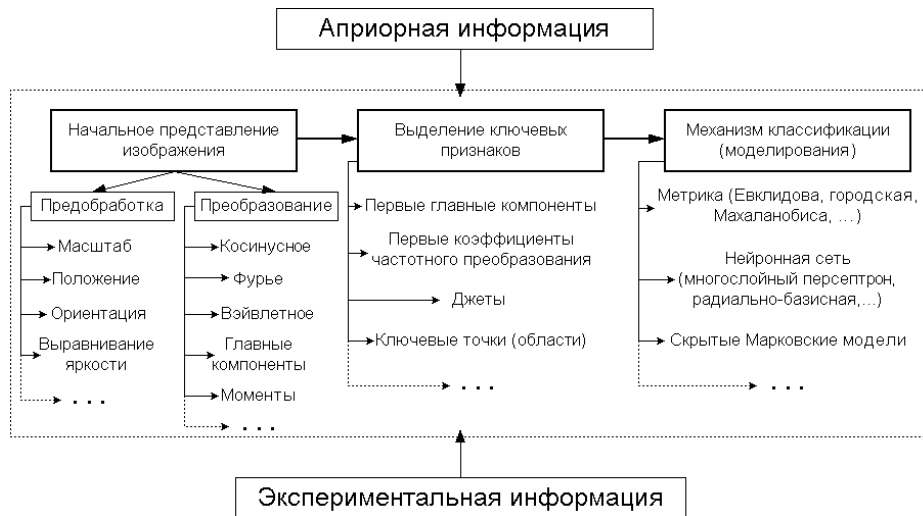


Рис. 2 Структура методов распознавания изображений.

### Методы, основанные на геометрических характеристиках лица

Один из самых первых методов – это анализ геометрических характеристик лица. Изначально применялся в криминалистике и был там детально разработан. Потом появились компьютерные реализации этого метода. Суть его заключается в выделении набора ключевых точек (или областей) лица и последующем выделении набора признаков. Каждый признак является либо расстоянием между ключевыми точками, либо отношением таких расстояний. В отличие от метода сравнения эластичных графов, здесь расстояния выбираются не как дуги графов. Наборы наиболее информативных признаков выделяются экспериментально.

Ключевыми точками могут быть уголки глаз, губ, кончик носа, центр глаза и т.п., рис.3. В качестве ключевых областей могут быть прямоугольные области, включающие в себя: глаза, нос, рот.

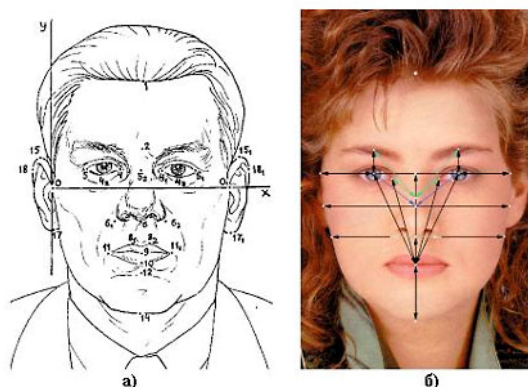


Рис. 3 Идентификационные точки и расстояния: а) используемые при криминалистической фотоэкспертизе; б) наиболее часто применяемые при построении автоматизированных систем идентификации.

В процессе распознавания сравниваются признаки неизвестного лица, с признаками, хранящимися в базе.

Задача нахождения ключевых точек приближается к трудоёмкости непосредственно распознавания, и правильное нахождение ключевых точек на изображении во многом определяет успех распознавания.

Поэтому изображение лица человека должно быть без помех, мешающих процессу поиска ключевых точек. К таким помехам относят очки, бороды, украшения, элементы причёски и макияжа. Освещение желательно равномерное и одинаковое для всех изображений.

Кроме того, изображение лица должно иметь фронтальный ракурс, возможно с небольшими отклонениями. Выражение лица должно быть нейтральным. Это связано с тем, что в большинстве методов нет модели учёта таких изменений.

Таким образом, данный метод предъявляет строгие требования к условиям съёмки, нуждается в надёжном механизме нахождения ключевых точек для общего случая. Кроме того, требуется применение более совершенных методов классификации или построения модели изменений.

В общем случае этот метод не самый оптимальный, однако для некоторых специфических задач перспективен. К таким задачам можно отнести документный контроль, когда требуется сравнить изображение лица, полученного в текущий момент с фотографией в документе. При этом других изображений этого человека не имеется, и следовательно механизмы классификации, основанные на анализе тренировочного набора, недоступны.



### Метод сравнения эталонов

Сравнение эталонов (*Template Matching*) заключается в выделении областей лица на изображении (рис. 4), и последующем сравнении этих областей для двух различных изображений. Каждая совпавшая область увеличивает меру сходства изображений. Это так же один из исторически первых методов распознавания человека по изображению лица. Для сравнения областей используются простейшие алгоритмы, типа попиксельного сравнения.



Рис. 4 Сравнимые области-эталон лица.

Недостаток этого метода заключается в том, что он требует много ресурсов как для хранения участков, так и для их сравнения. В виду того, что используется простейший алгоритм сравнения, изображения должны быть сняты в строго установленных условиях: не допускается заметных изменений ракурса, освещения, эмоционального выражения и пр.

### Методы оптического потока

Алгоритмы оптического потока используются в основном для анализа движения. Используя два или более последовательных кадра изображения, можно рассчитать двумерное векторное поле, называемое **оптическим потоком** (*Optical Flow*), которое отражает актуальное или наиболее вероятное смещение точек изображения от кадра к кадру.

Оптический поток рассчитывался для двух произвольных изображений лица, для получения меры соответствия изображений. Эти два изображения считались последовательными кадрами. Затем вычислялось векторное поле, наилучшим образом отображающее одно изображение в другое, в смысле минимизации расстояния между изображениями и с учётом геометрических ограничений, таких как относительное расположение точек изображения. Алгоритм находил наиболее соответствующие блоки. Поиск осуществлялся иерархически, начиная с больших блоков, и затем разбивая их на меньшие блоки. Таким образом, строилась пирамида соответствия изображений.

Используя векторное поле оптического потока, строились различные меры соответствия, с помощью которых в базе находилось изображение, ближайшее к неизвестному.

База данных представляла собой 76 изображений 36-и человек с различиями в ракурсе и выражении лица. Распознавание осуществлялось несколькими способами.

В первом способе блоки 8x8 неизвестного изображения заменялись на наиболее соответствующие блоки сравниваемого изображения (рис. 5, 6). Затем вычислялось Евклидово расстояние между неизвестным и полученным изображением. Было достигнуто 92% точности распознавания. Учитывая то, что в базе находилось только одно изображение нужного человека и по два на всех остальных, это хороший результат.



Рис. 5 Отображение неизвестного изображения на известное, один и тот же человек.

Слева направо: неизвестное изображение, изображение из базы данных, неизвестное изображение, в котором блоки заменены блоками из известного изображения.

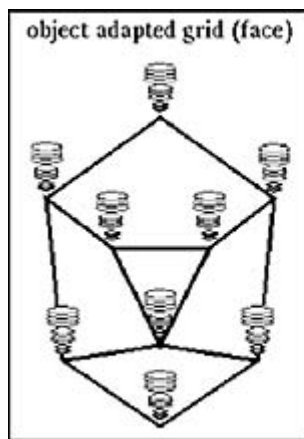


Рис. 6 Отображение неизвестного изображения на изображение из базы.  
Изображения разных людей.

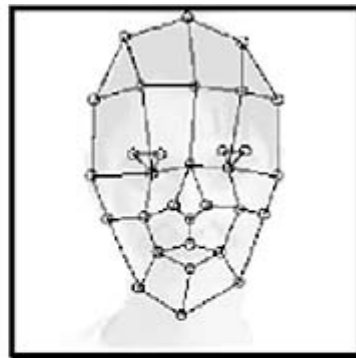
Как оказалось, даже при отображении изображения одного человека на изображение другого, результирующее изображение оказывается визуально очень похожим на исходное (рис. 6). Точность оказалась хуже, вследствие больших искажений у блоков, не представляющих важности (таких как волосы например). Использование центральной части лица дало немного лучший результат.

К недостаткам этого метода в первую очередь относится его вычислительная трудоёмкость. На специализированном параллельном компьютере сравнение двух изображений занимало больше минуты. Метод неспособен извлекать компактный набор характеристик для хранения и поиска в базе.

### Другие методы



(a)



(b)

Чтобы получить общее представление лица, так называемую адаптированную к объекту решетку (object-adapted grid), система по 70 моделям лица создает обобщенный граф лица (face bunch graph) (a). Затем (b) сравнивает заданное изображение с обобщенным графом и ищет отправные точки. С помощью гибкого сравнения создается граф изображения, и затем для опознания лица полученное изображение сравнивается с лицами из базы данных.

Рис. 7 Система Университета Южной Калифорнии распознает лица путем гибкого сравнения с графом (elastic graph matching).

К 1993 году было заявлено о существовании нескольких алгоритмов, правильно решающих задачу в среде с минимальным числом ограничений. Чтобы выявить их реальный потенциал, агентство DARPA и Исследовательская лаборатория армии США разработали программу Feret (face recognition technology). Цель этой программы – оценить эффективность предложенных алгоритмов и стимулировать развитие лучших из них.

В ходе реализации программы Feret несколько алгоритмов продемонстрировали самый высокий уровень точности распознавания для больших баз данных (1200 человек) в наиболее сложных для распознавания условиях тестирования. Один из них - алгоритм, разработанный в Университете Южной Калифорнии (рис. 7).

В тестовые базы данных Feret заносились лица в различных положениях, разного размера и по-разному освещенные, похожие на фотографии для водительских удостоверений или полицейского архива. Работа всех алгоритмов была близка к идеалу в том случае, если база данных включала не более 200 лиц, а изображения получались при сходных условиях. Интересно, что даже простое сравнение корреляций иногда удавалось провести с той же данных ровно из 200 элементов. Из этого со всей очевидностью следует, что признать точностью на базе новый алгоритм достойным рассмотрения можно только в том случае, если он протестирован на базе данных из по крайней мере 200 человек и продемонстрировал эффективность распознавания не ниже 95% для изображений типа фотографий для удостоверений.

В более масштабных тестах программы Feret (с числом изображений не менее 1196) эффективность всех алгоритмов была примерно одинаковой. В этой связи трудно или даже невозможно провести четкие различия между ними (особенно если согласовать даты тестирования). Для фронтальных изображений, сделанных в один и тот же день, приемлемая точность распознавания, как правило, составляет 95%. Для изображений, сделанных разными аппаратами и при разном освещении, точность, как правило, падает до 80%. Для изображений, сделанных с разницей в год, точность распознавания составляет примерно 50%. При этом стоит

заметить, что даже 50 процентов – это более чем приемлемая точность работы системы.

Другим направлением распознавания является применение нейронных сетей [3]. На рис. 8 приведен пример архитектуры многослойного персептрона. В нем число входов сети равно количеству пикселей входного изображения. Число выходов соответствует количеству человек (классов). Для обучения каждому новому представлению приходилось тщательно подбирать вручную параметры обучения и другие характеристики сети, что, однако, не гарантировало сходимости. Распознавание на тестовой выборке показало точность при различном делении обучающей выборки и других изменениях параметров от 90 до 98%, в среднем 94%. Главным образом это зависит от качества обучающей выборки, т.е. от того, каким образом выборка будет случайно разделена на обучающие и тестовые изображения.

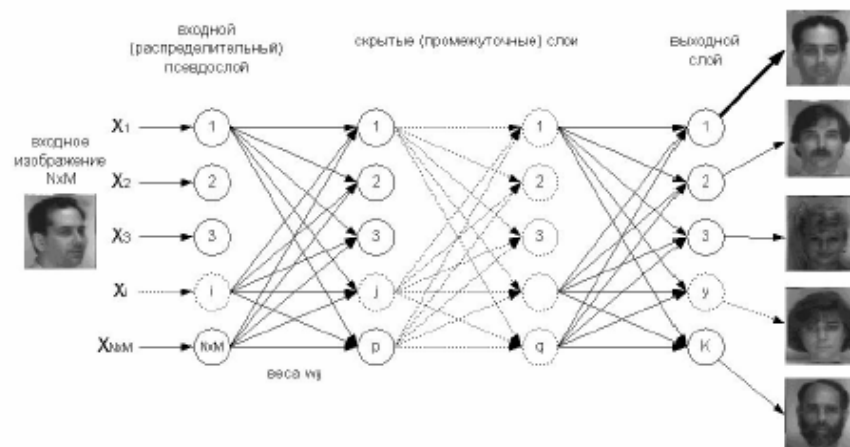


Рис. 8 Архитектура многослойного персептрона.

Также перспективными являются методы с восстановлением 3-х мерного изображения [4,9,10]. Под трехмерной моделью лица понимается совокупность поточечной карты трехмерного рельефа лица и набора отдельных интегральных характеристик, как то: расположение глаз, высота носа и другие. Эти характеристики вычисляются исходя из карты рельефа, таким образом являющейся основой для построения трехмерной модели лица, то есть совместить картинку с двух камер и получить третье изображение с заданными параметрами с учетом расстояния между камерой и различными частями объекта. Это называется

эпиполярным совмещением изображений, то есть математическим смещением точки наблюдения, соответствующим параллельному расположению лучей зрения обеих камер. Разработаны два типа алгоритмов совмещения: быстрый (20 кадров в секунду), который уже сегодня работает в реальном времени на Pentium 4, и медленный, но более точный (1-4 кадров в секунду). Второй алгоритм разработан для реального режима времени персональных компьютеров недалекого будущего.

### **1.3 Обзор существующих систем**

Сейчас имеется несколько коммерческих продуктов для распознавания лиц. Дать окончательную оценку сложно, однако, похоже, три системы выбиваются в лидеры – это разработки компаний Visionic, Viisage и Miros.

В основе приложения FaceIt компании Visionic лежит алгоритм анализа локальных признаков, разработанный в Университете Рокфеллера. Одна коммерческая компания в Великобритании интегрировала FaceIt в телевизионную антикриминальную систему под названием Mandrake. Эта система ищет преступников по видеоданным, которые поступают с 144 камер, объединенных в замкнутую сеть. Когда устанавливается идентичность, система сообщает об этом офицеру безопасности.

Еще один лидер рынка систем распознавания лиц, компания Viisage, использует алгоритм, разработанный в МТИ. Коммерческие компании и властные структуры во многих американских штатах и в ряде других стран используют систему компании Viisage вместе с идентификационными удостоверениями, например, водительскими правами.

В системе распознавания лиц TrueFace компании Miros используется технология нейронных сетей, а сама система применяется в комплексе выдачи наличных денег корпорации Mr.Payroll и установлена в казино и других увеселительных заведениях многих штатов.

Также существует система «Видео Око» (Россия) -для работы системы применяются ч/б камеры высокого разрешения, время принятия решения 1.5 сек для Pentium 200. Регистрируемые события : движение в кадре, изменение фона, лицо в кадре, лицо распознано. Метод работы не описан.

## 2 ЦЕЛЬ И ЗАДАЧИ ИССЛЕДОВАНИЯ

### 2.1 Постановка задачи

Необходимо исследовать метод распознавания изображений на базе метода скрытых Марковских моделей и спроектировать биометрическую систему идентификации человека на основе распознавания лиц.

### 2.2 Требования

Система идентификации человека на основе распознавания лиц должна на входе которой подается фотография человека или нескольких человек, которые периодически выделяются из видеопотока, получаемого с камеры или вводятся из файла с фотографией. На выходе система выдает список кандидатов, похожих на предлагаемого человека. Для каждого кандидата система выводит показатель уверенности, показывающую степень схожести предлагаемой фотографии и известных системе людей. Список кандидатов ранжируется по степени уверенности.

Система должна хранить в БД некоторую информацию необходимую для проведения идентификации человека в виде набора признаков. Процесс идентификации производится в несколько этапов:

- на вход системы подается фотография или кадр видео потока.
- производится определение есть ли на фотографии лицо человека
- выделяется предположительное место расположения лица
- определяются границы лица
- происходит масштабирование полученной части фотографии
- производится идентификация
- система сообщает смогла ли она идентифицировать человека на фотографии

В работе системы выделяются два основных этапа которые требуют значительных вычислительных затрат: анализ сцены для выделения области лица, распознавание лица на выделенном фрагменте фотографии

При этом потребуется исследовать различные алгоритмы анализа сцены и распознавания образа и подобрать оптимальное сочетание алгоритмов для получения максимальной точности распознавания. Главным объектом исследования

сравнение является методов [1,2] и [5] и алгоритмы предварительной обработки лиц для возможности применения к ним этого метода.

Система может делать ошибки следующего рода:

- Не опознан известный системе человек
- Неизвестный системе человек опознан как какой-либо известный
- Известный системе человек опознан как другой известный



## 3 ПРОЕКТИРОВАНИЕ СИСТЕМЫ

### 3.1 Описание математических методов

Основная последовательность операций при распознавании лиц приведена на рис. 9.



Рис. 9 Основная последовательность операций при распознавании лиц.

#### 3.1.1 Подготовка изображения

Этап подготовки изображения включает в себя непосредственный ввод изображения с камеры и предварительную обработку.

Ввод производится через программный интерфейс (API) устройства ввода.

Предварительная обработка состоит из нормализации по яркости и контрастности. Эта обработка напрямую не относится к распознаванию, а

корректирует разброс параметров камеры и освещения. Как правило, параметры данной обработки однократно задаются при установке системы, но, при необходимости, могут изменяться, адаптируя систему к внешним условиям. Пример использования нормализации приведен на рис. 10



Рис. 10 Пример использования алгоритма нормализации.

### 3.1.2 Выделение лица из общей сцены

Для выделения лица возможно использование метода анализа сцены, использующего тот факт, что объект распознавания - человеческое лицо имеет весьма специфическую природу [7]. Используя плоскую модель человеческого лица (рис. 11) для задания областей интереса при помощи корреляционного метода проверяется наличие в рассматриваемой области глаз, носа, рта, причем сама область для проверки выбирается путем отыскания связанных областей эллиптической формы [6].

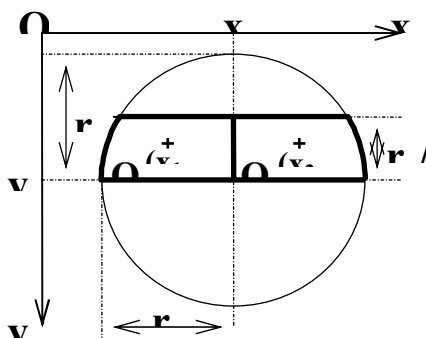


Рис. 11 Плоская модель лица человека.

Также один из перспективных методов, используемых для выделения области лица, основан на восстановлении карты высот на изображении [8,9]. Применив выделения связанных областей на карте высот можно определить набор областей интереса, одной из которых окажется лицо. Достоинством этого метода является устойчивость к ошибкам, возникающим из-за неравномерного освещения лица. Пример карты высот, полученный методом [9] приведен на рис. 12

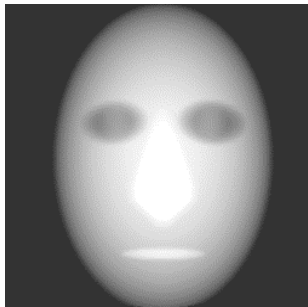


Рис. 12 Карта высот лица человека.

Для использования в данной системе был выбран метод идентификации объектов, основанный на преобразовании Хаара.

Основные идеи дискретного преобразования Хаара состоят в следующем. Пусть некоторый сигнал задан дискретной последовательностью отсчетов  $S_n$ . Выберем два последовательных отсчета,  $a = S_k$  и  $b = S_{k+1}$ , и перейдем к их среднему и разности:

$$p = (a + b)/2,$$

$$q = b - a$$

Полезность такого представления состоит в том, если в сигнале присутствуют значительные корреляции между последовательными отсчетами, то величина разности  $q$  мала (в пределе  $a=b$ ,  $q=0$ ), и может быть представлена в ЭВМ меньшим числом бит. При этом не происходит потери информации, поскольку имеется обратное преобразование:

$$a = p - q/2,$$

$$b = p + q/2$$

Если выполнить описанное преобразование для всех последовательных пар отсчетов сигнала  $S_n = \{S_{n,m} | m < 2^n\}$  длины  $2^n$ , то он распадается на два сигнала половинной длины:

$$S_{n-1,m} = (S_{n,2m} + S_{n,2m+1}) / 2$$

$$D_{n-1,m} = S_{n,2m+1} - S_{n,2m}$$

Вектор средних значений  $S_{n-1,m}$  можно рассматривать, как огрубленное (сжатое) представление исходного вектора  $S_{n,m}$ , а вектор разностей  $D_{n-1,m}$  - как детализирующую информацию, необходимую для перехода из сжатого представления к исходному. Далее, такое же преобразование можно применить к сжатому сигналу  $S_{n-1,m}$ , переходя к еще более компактному и огрубленному представлению. Рекурсивно выполняя преобразование  $n$  раз, мы получаем из исходного сигнала  $n$  его версий с огрублением на разных масштабах. Самое грубое представление  $S_{0,0}$  - это просто среднее от всего сигнала.

Описанная процедура и есть дискретное преобразование Хаара. Полное число коэффициентов разложения равно исходному числу отсчетов -  $2^n$ , однако информационная нагрузка их (и, соответственно, требуемая точность представления) различна. Обратное преобразование выполняется также рекуррентно, по обратным формулам. Полное число операций пропорционально  $n$  (что эффективнее, чем быстрое преобразование Фурье, требующее  $O(n \log n)$  операций). Имеются удобные схемы проведения вычислений (аналогичные методу прогонки), при которых все коэффициенты преобразования размещаются на месте хранения исходного сигнала, т.е. не требуется дополнительной памяти для ЭВМ или для специализированного процессора.

Необходимо отметить, что описанная схема вычислений эквивалентна разложению кусочно- постоянного аналога дискретного сигнала по базису  $e_{j,k}(x) = 2^{j/2} s(2^j x - k)$ , где  $s$  - вейвлет Хаара.

Рекуррентное преобразование Хаара имеет по меньшей мере два ценных практических свойства, определяющие перспективность его применения:

- При хранении и передаче больших объемов информации степень ее сжатия и подробность представления можно варьировать в широких пределах, зануляя (не передавая и не храня) поправочные коэффициенты  $d$ , начиная с требуемого уровня подробности.
- Передачу сигнала можно осуществлять последовательно, начиная с самых грубых уровней, несущих в предельно компактной форме основной объем информации, с дальнейшим последовательным уточнением, вплоть до полного восстановления. На нужном уровне подробности передачу данного фрагмента (кадра) можно прекратить, и перейти к следующему фрагменту. При поиске информации в базе данных можно увеличивать скорость доступа за счет снижения подробности.

Для выделения области интереса используется следующий алгоритм, работающий на основе преобразования Хаара. Способ выделения объектов

первоначально предложен П. Виолой. Сначала, классификатор (*cascade of boosted classifiers working with haar-like features*) обучается на нескольких сотнях видов образца конкретного объекта (лица человека), которые являются положительными примерами. Все изображения должны быть приведены к стандартному виду (по параметрам освещенности, положению объекта и размеру). Также классификатору должны быть предъявлены отрицательные образцы: произвольные изображения того же размера, но не содержащие лиц.

После предварительного обучения классификатора он может применяться к области интереса (того же самого размера, который был использован при обучении) в входном образе. Классификатор принимает решение "1" если область, с некоторой вероятностью, является требуемым объектом (лицом), и "0" в противном случае. Для поиска объекта на фотографии классификатор применяется к фрагменту изображения, перемещаемому по фотографии. Принцип работы классификатора позволяет его применение к объектам различного размера без их предварительного масштабирования. Для этого процедура сканирования должна быть проведена несколько раз с разными весовыми коэффициентами.

Классификатор состоит из различных более простых классификаторов (этапов), которые последовательно применяются к области интереса, пока на не котором этапе кандидат не отвергается или пока все не закончены этапы успешно. Основные классификаторы являются деревом решений, не менее, чем с 2 уровнями. Алгоритм использует следующие Хагг-подобные характеристики (рис. 13)

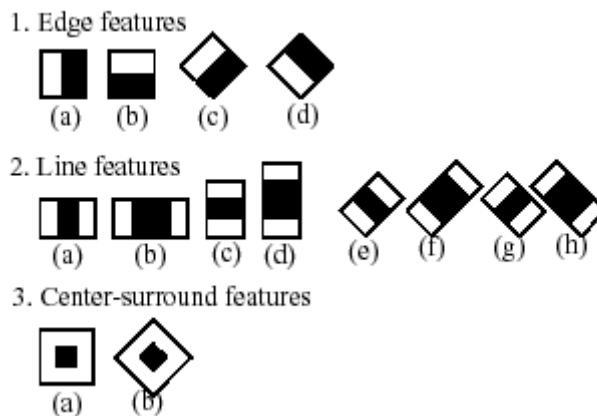


Рис. 13 Признаки в алгоритме выделения лица.

Характеристика использованная в конкретном классификаторе определяется своей формой (1a, 2b и т.п.), позицией в пределах области интереса и коэффициентом. Например, в случае третьей характеристики строки (2c) ответ

вычислен как различие между суммой пикселей образа под прямоугольником, покрывающим целую характеристику (включая две белых полосы и черную полосу в середине) и суммой пикселей образа под черной полосой умноженной на 3 (компенсация размера области).

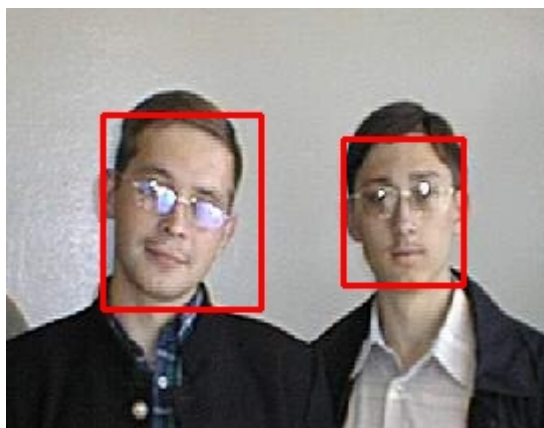


Рис. 14 Пример работы алгоритма выделения лиц.

Нормализация изображения состоит из масштабирования лица для приведения его к основному размеру (размеру фотографии на которых производилось обучение системы). В качестве алгоритма масштабирования была выбрана билинейная интерполяция. Данный алгоритм дает хорошее качество изображения при масштабировании и требует минимального количества ресурсов, что немаловажно при решении данной задачи.

### **3.1.3 Распознавание изображения**

В качестве методов распознавания были исследованы методы классификации геометрических изображений [1,2] и метод скрытых Марковских моделей (СММ) [5].

#### **3.1.3.1 Методы классификации геометрических изображений.**

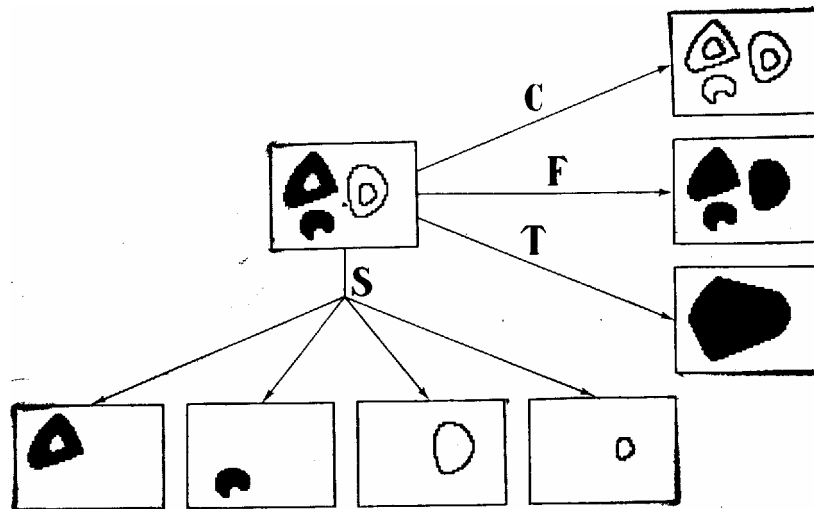
Методы классификации геометрических изображений моделируют способность человека находить принцип классификации (разделяющее правило) некоторых геометрических объектов по малому числу примеров. При этом необходимо, чтобы система находила тот же принцип классификации (среди большого числа априори возможных), который в этой же задаче находит человек.

Такое требование естественно привело к следующим принципам организации системы:

- 1) язык системы, описывающий разделяющие правила, должен иметь терминологическое сходство с языком, которым в этих же задачах пользуется человек;
- 2) поскольку на одном языке бывает возможно описать несколько разделяющих правил, удовлетворяющих данному материалу обучения, последовательность перебора разделяющих правил в системе должна быть также близкой к человеческой.

Метод рассчитан на решение задач, в которых человек для описания разделяющих правил применяет только «геометрические» термины: фигура, линия, часть фигуры, контур, площадь, длина, наклон, угол, подмножество и т. п. Входными объектами для программы являются плоские черно-белые картинки. Для обучения вводится набор картинок, разбитых на классы. Цель метода обучения классификации состоит в том, чтобы в результате обработки этих картинок сформулировать правило, отличающее один класс картинок от другого.

Метод состоит в применении к изображению последовательности операторов (рисующие, измерительные, решающие) и получения в результате их работы некоторого числового критерия. Пример работы рисующих операторов приведен на рис. 15.



*выделение контура C, заполнение контура P, натягивание оболочки T;  
разделение по связности S*

Рис. 15 Рисующие операторы в методе классификации геометрических изображений.

Перебирая в различные последовательности применения операторов можно получить различные критерии для данного изображения. Пример графа возможных преобразований изображений при использовании 4 операторов: выделение контура (C), заполнение контура (P), натягивание оболочки (T) и разделение по связности (S) приведен на рис 16.

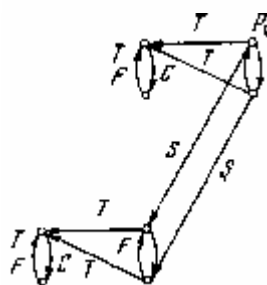


Рис. 16 Граф возможных преобразований изображений в методе классификации геометрических изображений.

После предъявления системе набора изображений разбитых на классы, производится формирование разделяющего правила путем формирования первичного ключа таблицы признаков, т.е. набора признаков наименьшей длины, который с требуемой степенью точности разделяет изображения на классы. Значения этих признаков (существенные признаки) заносятся в базу данных и в дальнейшем используются для распознавания.

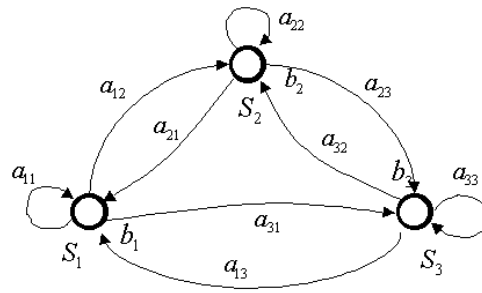
Исследования этого метода показали ограниченную возможность его применимости для распознавания лиц. Это связано с малым учетом априорной информации в методе о лицах человека. Так как человеческое лицо сложно с помощью формальных алгоритмов привести к строго геометрическому двухцветному изображению, то этот метод оказывается неэффективным в связи с тем, что он изначально построен на классификации геометрических объектов. Испытания показали большое количество ошибок при распознавании одного человека с разными эмоциональными выражениями: алгоритм распознавал их как разных людей. Это связано с «предвзвешенными» алгоритма [2]. Также имеются сложности построения предварительных фильтров для приведения фотографии к геометрическому виду.

### 3.1.3.2 *Скрытые Марковские модели*

Марковские модели являются мощным средством моделирования различных процессов и распознавания образов. По своей природе Марковские



модели позволяют учитывать непосредственно пространственно-временные характеристики сигналов, и поэтому получили широкое применение в распознавании речи, а в последнее время – изображений (в частности изображений лиц).



$$O = \{ Y, G, R, G, G, B, R, R, Y, B \}$$

$$S = \{ 2, 1, 1, 3, 2, 2, 2, 3, 3, 1 \}$$

Рис. 17 Схема Марковской модели на примере последовательности наблюдений  $O$  и последовательности состояний  $S$ .

Каждая модель  $\lambda = (A, B, \pi)$ , рис. 17, представляет собой набор  $N$  состояний  $S = \{S_1, S_2, \dots, S_N\}$  (на рисунке – вершины графа), между которыми возможны переходы (на рисунке - дуги). В каждый момент времени система находится в строго определённом состоянии. В наиболее распространённых Марковских моделях первого порядка полагается, что следующее состояние зависит только от текущего состояния.

При переходе в каждое состояние генерируется наблюдаемый символ, который соответствует физическому сигналу с выхода моделируемой системы.

Набор символов для каждого состояния  $V = \{v_1, v_2, \dots, v_M\}$ , количество символов  $M$ . Выход, генерируемый моделью, может быть так же непрерывным. Существуют так же модели, в которых набор символов для всех состояний одинаков. Символ в состоянии  $q_t = S_i$  в момент времени  $t$  генерируется с вероятностью  $b_j(k) = P[v_{k,t} | q_t = S_j]$ . Набор всех таких вероятностей составляет матрицу  $B = \{b_j(k)\}$ .

Матрица  $A = \{a_{ij}\}$  определяет вероятность перехода из одного состояния в другое состояние:  $a_{ij} = P[q_{t+1} = S_j | q_t = S_i]$ ,  $1 \leq i, j \leq N$ . Считается, что  $A$  не зависит от времени. Если из каждого состояния можно достичь любого другого за один переход, то все  $a_{ij} > 0$ , и модель называется эргодической.

Так же модель имеет вероятность начальных состояний  $\pi = \{\pi_i\}$ , где  $\pi_i = P[q_1 = S_i]$ .

Обычно в реальных процессах последовательность состояний является скрытой от наблюдения и остаётся неизвестной, а известен только выход системы, последовательность наблюдаемых символов  $O = O_1 O_2 \dots O_T$ , где каждое наблюдение  $O_t$  – символ из  $V$ , и  $T$  – число наблюдений в последовательности. Поэтому такие модели называют скрытыми Марковскими моделями (СММ, по английски – Hidden Markov Models, НММ).

Модель  $\lambda = (A, B, \pi)$  с настроенными параметрами может быть использована для генерирования последовательности наблюдений. Для этого случайно, в соответствии с начальными вероятностями  $\pi$  выбирается начальное состояние, затем на каждом шаге вероятность  $B$  используется для генерации наблюдаемого символа, а вероятность  $A$  – для выбора следующего состояния. Вероятность  $P$  генерирования моделью последовательности состояний  $O$ :

$$P(O | \lambda) = \prod_{t=1}^T b_{q_t}(O_t), \text{ где } Q = q_1 q_2 \dots q_T \text{ – последовательность состояний.}$$

Предполагается, что наблюдения статистически независимы.

В распознавании образов скрытые Марковские модели применяются следующим образом. Каждому классу  $i$  соответствует своя модель  $\lambda_i$ . Распознаваемый образ (речевой сигнал, изображение и т.д.) представляется в виде последовательности наблюдений  $O$ . Затем для каждой модели  $\lambda_i$  вычисляется вероятность того, что эта последовательность могла быть сгенерирована именно этой моделью. Модель  $\lambda_j$ , получившая наибольшую вероятность, считается наиболее подходящей, и образ относят к классу  $j$ .

В связи с этим появляются несколько вопросов, называемых тремя основными задачами скрытых Марковских моделей.

Имея последовательность наблюдений  $O = O_1 O_2 \dots O_T$  и настроенную модель  $\lambda = (A, B, \pi)$ , как оценить вероятность  $P(O | \lambda)$  генерации этой моделью данной последовательности наблюдений? Эта задача называется задачей распознавания.

Имея последовательность наблюдений  $O = O_1 O_2 \dots O_T$  и настроенную модель  $\lambda = (A, B, \pi)$ , как подобрать последовательность состояний  $Q = q_1 q_2 \dots q_T$ , чтобы она была оптимальной (в соответствии с некоторым критерием, аналитически эта задача неразрешима)? Другими словами это задача объяснения. Она нужна для последующей коррекции параметров модели.

Каким образом корректировать параметры модели  $\lambda$ , для того чтобы максимизировать  $P(O | \lambda)$ ? Т.е. как сделать так, чтобы модель больше соответствовала своему классу, одним из образов которого является данная последовательность наблюдений (или несколько различных последовательностей). Это задача обучения.

Первая задача имеет точное аналитическое решение, называемое процедурой прямого-обратного прохода. Последующие две задачи не имеют точного аналитического решения. Для решения второй задачи используется алгоритм Витерби, для третьей – алгоритм Баума-Вельча. Оба этих метода являются разновидностями градиентного спуска и решаются оптимизационными методами.

Для того, чтобы сократить вычисления, в распознавании речи используются линейные модели, рис. 18. В таких моделях каждое состояние имеет только одно последующее, так же переход возможен обратно в то же состояние. Такие модели учитывают временные характеристики речевого сигнала: определённый порядок следования участков сигнала, их взаимное расположение, возможность локальных растяжений или сжатий. Это позволяет их применять и в распознавании изображений.

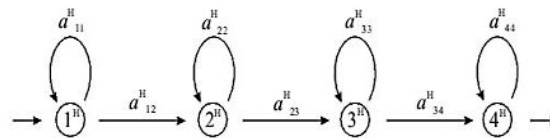


Рис. 18 Линейная Марковская модель.

Одна из первых работ, применяющая СММ для распознавания изображений лиц – это работа Ф.Самарии, которой предшествовали работы по распознаванию изображений других видов скрытыми Марковскими моделями. В этой работе распознавание осуществлялось как простейшими одномерными

линейными СММ, так и псевдодвумерными. Введение второго измерения позволило повысить точность распознавания с 85% до 95% на базе ORL.

Суть двумерных Марковских моделей заключается в том, что в отличие от одномерных линейных СММ, они позволяют моделировать искажения изображения и взаимное расположение участков не отдельно по горизонтали или вертикали, а в обоих направлениях одновременно. Для уменьшения вычислительной сложности применяются псевдодвумерные СММ (Pseudo-2D Hidden Markov Models, P2D-HMM). Такая модель состоит из нескольких линейных вертикальных моделей нижнего уровня, и одной линейной горизонтальной модели верхнего уровня, на вход которой поступают выходы моделей нижнего уровня, рис. 19. Каждое состояние модели верхнего уровня включает в себя последовательность состояний соответствующей модели нижнего уровня. Модели нижнего уровня не связаны между собой. Изначально в модели верхнего уровня были вертикальными. В последующих работах модели верхнего уровня были сделаны горизонтальными (как это и изображено на рисунке), в связи с тем, чтобы вертикальные модели нижнего уровня могли учесть факт того, что глаза могут находиться на разной высоте. Таким образом псевдодвумерная модель позволяет учесть локальные деформации и взаимное расположение участков изображений. Но в отличие от оптических потоков и других методов сопоставления деформациями, псевдодвумерная модель учитывает характер деформаций, а то какими именно могут быть возможные деформации, псевдодвумерные СММ усваивают в процессе обучения. Другими словами, участок, соответствующий глазу, никогда не будет сопоставлен например участку на месте рта, как это может быть в оптическом потоке.

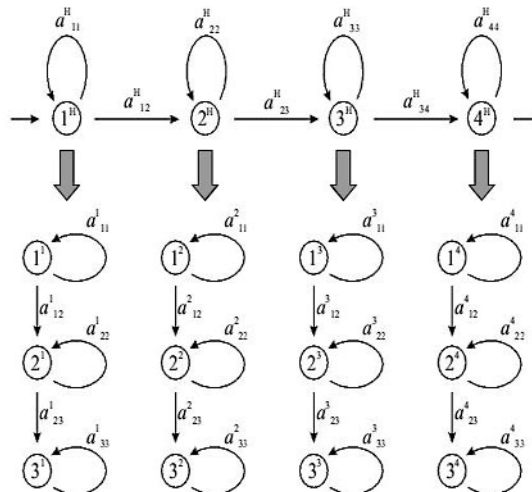


Рис. 19 Псевдодвумерная скрытая Марковская модель.

Наблюдениям, подаваемым на вход СММ, являлись квадратные участки изображений, рис. 20. Было обнаружено, что участки, извлекаемые с 75% перекрытием друг с другом, давали наилучшую точность распознавания [13].

В работах [13], продолжающих идею Самарии использовались дальнейшие улучшения способов начального представления изображения и алгоритмов тренировки.

Для каждого квадратного участка изображения  $16 \times 16$  вычислялось двумерное дискретное косинусное преобразование, и этот участок представлялся в виде набора первых 15-ти коэффициентов ( $u + v \leq 4$ , п. 2.7.2.). Это позволило повысить точность распознавания на 2%. Кроме того, такое представление позволяет более точно, чем при масштабировании представлять изображение, используя меньший объём информации.

Для увеличения тренировочного набора использовались так же зеркально отражённые по вертикали изображения [5]. Это позволило учесть более широкий диапазон ракурсов.

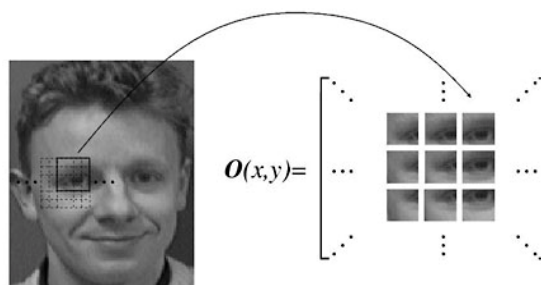


Рис. 20 Извлечение участков-образцов наблюдения.

Для СММ важное значение имеет начальная инициализация модели. В [13], в качестве начальной инициализации всех моделей использовались все изображения из тренировочного набора. Затем модель каждого класса донастраивалась на свои изображения.

Полезное свойство распознавания по коэффициентам дискретного косинусного преобразования заключается в том, что оно позволяет работать непосредственно со сжатыми изображениями, такими как JPEG и MPEG, в которые на сегодняшний день являются распространёнными форматами хранения изображений и видео.

Как можно видеть, правильный выбор метода классификации (моделирования), начального представления изображения и учёт его особенностей в

сочетании с применением априорных знаний о предметной области, дало успешный результат.

Недостатком СММ является то, что СММ не обладает различающей способностью. Т.е. алгоритм обучения только максимизирует отклик каждой модели на свои классы, но не минимизирует отклик на другие классы, и не выделяются ключевые признаки, отличающие один класс от другого. Для устранения этого недостатка можно использовать алгоритм ранжирования вероятностей, заключающийся в следующем. На обучающем наборе, каждая модель реагирует на изображения-примеры с некоторой вероятностью. Отсортированные таким образом модели образуют исходное ранжирование. Для неизвестного изображения модели так же ранжируются по вероятностям отклика на неизвестное изображение. Большая величина отклонения полученного ранжирования от исходного сигнализирует о том, что изображение принадлежит к неизвестному классу лица.

Таким образом, похожие классы могут оказаться слабо различимыми (как это и было в [13] с единственной ошибкой) и при увеличении объёма базы или использования в более широких условиях СММ может оказаться ненадёжными.

### **3.1.3.3      *Дискретное косинусное преобразование***

Высокая эффективность сжатия, которую дает этот алгоритм, основана на том факте, что в матрице частотных коэффициентов, образующейся из исходной матрицы после дискретного косинусного преобразования, низкочастотные компоненты расположены ближе к левому верхнему углу, а высокочастотные - внизу справа. Это важно потому, что большинство графических образов на экране компьютера состоит из низкочастотной информации, так что высокочастотные компоненты матрицы можно безболезненно выбросить. "Выбрасывание" выполняется путем округления частотных коэффициентов. После округления отличные от нуля значения низкочастотных компонент остаются, главным образом, в левом верхнем углу матрицы. Округленная матрица значений кодируется с учетом повторов нулей. В результате графический образ сжимается более чем на 90% , теряя очень немного в качестве изображения только на этапе округления.

Основным этапом работы алгоритма является дискретное косинусное преобразование (ДКП), представляющее собой разновидность преобразования Фурье. Оно позволяет переходить от пространственного представления изображения к его спектральному представлению и обратно.

Дискретное косинусное преобразование ДКП (discrete cosine transform - DCT) довольно схоже с ДПФ. Оно также является разделимым линейным преобразованием, т. е. двумерное преобразование эквивалентно двум одномерным,

произведенным последовательно по каждой из осей [14]. Двумерное ДКП осуществляется в соответствии со следующей формулой:

$$D(p,q) = a_p a_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} S(m,n) \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}$$

при  $0 \leq p \leq M-1, 0 \leq q \leq N-1$

$$a_p = \begin{cases} 1/\sqrt{M} & p = 0 \\ \sqrt{2/M} & 1 \leq p \leq M-1 \end{cases}, a_q = \begin{cases} 1/\sqrt{N} & q = 0 \\ \sqrt{2/N} & 1 \leq q \leq N-1 \end{cases}.$$

где S – матрица исходного изображения размера MxN,

D – матрица после применения ДКП,

M и N - соответственно количество строк и столбцов исходного изображения S.

## 4 РЕАЛИЗАЦИЯ СИСТЕМЫ

### 4.1 Выбор средств разработки

Данная система может быть реализована на базе различных аппаратных платформ. Главным требованием является наличие поддержки арифметики с плавающей точкой и производительность не менее 400 MIPS. Особенно перспективным является построение системы на базе встраиваемых компьютеров или микроконтроллеров (например производства Texas Instruments), что обеспечит системе необходимую мобильность. Система реализована на базе аппаратной платформы Intel PC, что связано с наличием большого инвентария существующих прикладных программных средств, позволяющих более эффективно пройти этап исследования. Выбор Intel PC среди настольных систем обоснован и тем, что под данную платформу существует большой модельный ряд камер, используемых для ввода изображения.

Выбор операционной системы также продиктован наличием инвентария и его доступностью в нашей стране. В этом качестве хорошо подходит любая ОС семейства MS Windows.

При выборе языка программирования учитывались следующие критерии:

- Возможность написания кода программы, который в дальнейшем позволял бы добиться максимальной производительности при ее (программы) работе (язык assembler не подходит из-за трудоемкости написания большого количества кода).
- Простая и эффективная работа с графическими изображениями.

Этим критериям отвечают C (Borland C Builder, Visual C), Pascal (Borland Delphi, Kylix), Java. Другие распространенные языки программирования мало пригодны для решения данной задачи.

Важным фактором при выборе языка служит переносимость кода как между ОС так и между аппаратными платформами. Этому критерию отвечают C и Java.

Учитывая все вышеперечисленные требования, для реализации системы выбран язык C, так как Java не обеспечит требуемого быстродействия (особенно для встраиваемых систем).

Так как данная работа носит исследовательский характер, и требует в большей степени простоты разработки, было решено в качестве средства разработки для решения описанной выше задачи использовать Borland C Builder.



## 4.2 Структура системы

Разрабатываемая система должна иметь модульную структуру для обеспечения возможности различных вариантов компоновки модулей, реализованных различными методами. Это снизит трудоемкость исследования и, при необходимости, упростит возможную процедуру внедрения системы в промышленную эксплуатацию. Общая структура системы приведена на рис. 21.

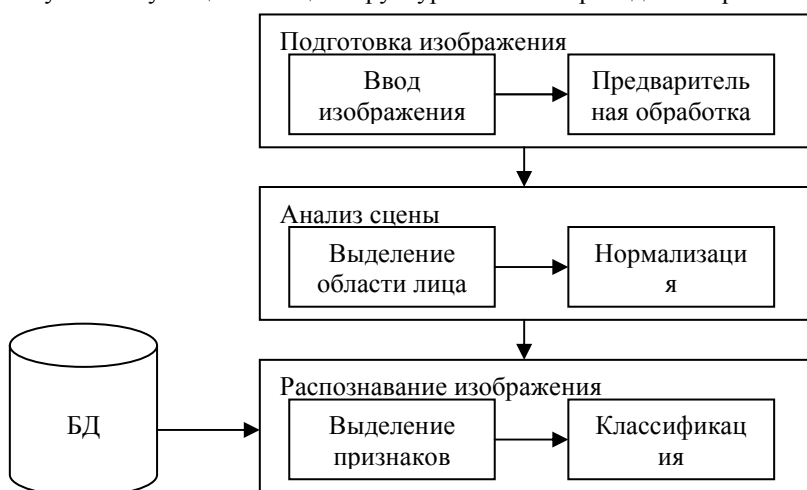


Рис. 21 Архитектура системы

Разработанная система представляет собой набор библиотек, выполняющих функции распознавания лиц. Для тестирования и демонстрации системы существует программная оболочка, работающая в двух режимах.

- режим обучения;
- режим распознавания.

В режиме обучения строится модель лица человека на основании тренировочных изображений, предоставляемых на вход модуля вместе с уникальным идентификатором данного человека. Полученная модель может быть занесена в БД и использована в дальнейшем для работы модуля распознавания.

В режиме распознавания обрабатывается введенное изображение и сравнивая его с моделями из БД моделей лиц подбирает лицо с наибольшей мерой сходства. Эти данные могут записываться в журнал сообщений или активировать некоторые действия.

Система реализована с использованием функций библиотеки Open Computer Vision версии 3.0 производства компании Intel [11]. Эта библиотека представляет собой набор функций в виде исходных текстов, предназначенных для реализации цифровой обработки изображений (IPL – image processing library). В системе используются средства ввода и представления изображений с помощью библиотек IPL, а также, функция преобразования Хаара и функция, реализующая алгоритм Витерби для построения скрытой Марковской модели.

Система выполнена в виде консольного приложения для платформы Win32, вызываемого из мастер-приложения. В зависимости от параметра запуска она работает с видеопотоком с устройства видеозахвата, установленного в системе по умолчанию или с файлом изображения в формате BMP, JPG. После запуска происходит считывание параметров моделей из базы лиц в оперативную память и создаются виртуальные каталоги элементов системы. Приложение создает 2 окна: в первом окне выводится обрабатываемое изображение (кадр видеопотока или фотография) на котором белыми прямоугольниками выделяются фрагменты, на которых, предположительно, находятся лица. Во втором окне выводится первый из найденных фрагментов с лицом, после этого производится распознавание. После распознавания система формирует системное сообщение, которое может быть принято мастер-приложением и обработано в соответствии с предназначением системы. Для настройки параметров устройства видеозахвата возможен вызов диалогов настройки формата и свойств камеры непосредственно во время работы системы при нажатии соответствующих клавиш.

#### **4.3 Структура базы лиц**

При распознавании лиц по каждому изображению строится скрытая Марковская модель, которые хранятся в виде значений вероятности перехода между состояниями и получения наблюдения в каждом состоянии. Эти модели хранятся в виде иерархической файловой структуры, изображенной на рис. 22. Текстовый формат хранения информации позволяет применять широкий набор инструментальных средств, что незаменимо при исследовательской работе. Иерархическая структура каталогов позволяет легко производить операции с добавлением и удалением объектов без применения системы. В дальнейшем, при развитии системы возможно применение этой иерархической модели, с использованием СУБД для сокращения времени поиска модели с заданными параметрами при большом количестве лиц.

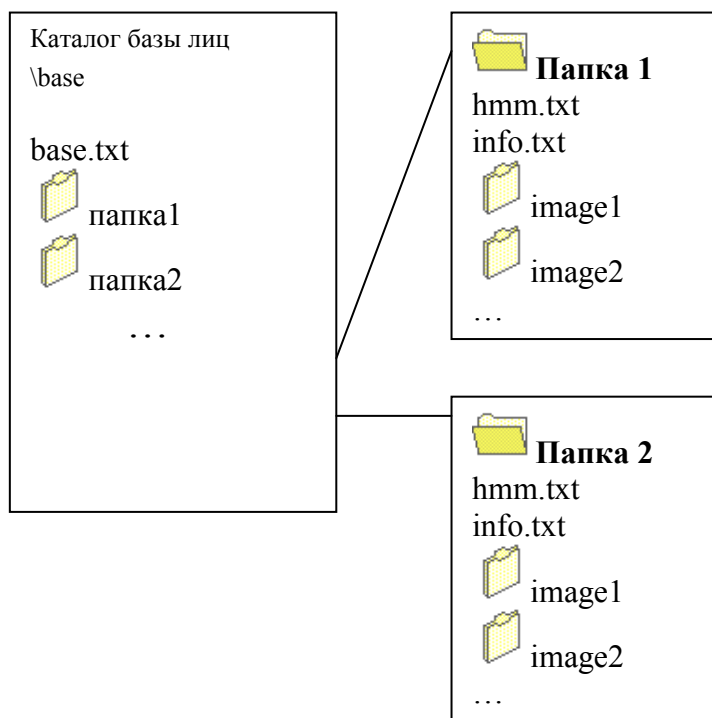


Рис. 22 Структура базы лиц.

Формат файла info.txt для содержательной о человеке приведен в табл. 4.2.

Таблица 4.2 Формат файла info.txt

Номер строки	Значение	Описание
1	PERSONAL INFO	Сигнатура файла
2	Фамилия, имя, отчество	Содержательный описатель человека
3	Имя файла изображения 1	Имя файла изображения
4	Имя файла изображения 2	Имя файла изображения

...	...	...
-----	-----	-----

Наличие файлов изображений требуется только при обучении для построения моделей, и после обучения системы не является обязательным.

Формат файлов hmm.txt для модели лица.

```
<BeginExternalHMM>
    данные о внешней модели
<EndExternalHMM>

<BeginEmbeddedHMM>
    <NumStates> количество_состояний

    <State> 1
    <Mean> значения_вероятности_наблюдений
    <Inverted_Deviation> значения_отклонений

    <State> 2
    ...
    <TransP> матрица_вероятностей_переходов
<EndEmbeddedHMM>

<BeginEmbeddedHMM>
    данные о следующей внутренней модели
    ...
<EndEmbeddedHMM>

...
```

## **5 ОЦЕНКА ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ**

Система тестировалась на стандартной базе фотографий людей ORL, применяемой для тестирования систем распознавания лиц. Эта база содержит фотографии 40 человек, по 10 фотографий на каждого человека с различным эмоциональным состоянием.

### **5.1 Подсистема выделения лица**

Данная подсистема достаточно эффективно работает и хорошо выделяет лицо (или набор лиц) из общей сцены при условии хорошего качества изображения и равномерного освещения лица. Примерно в 80% случаев система правильно выделила лицо человека, т.е. алгоритм выделения, который использует система, является достаточно эффективным. Для подсистемы выделения лица очень важно освещение лица человека и это необходимо учитывать при получении материалов для системы.

### **5.2 Подсистема распознавания лица**

Подсистема распознавания лица на основе скрытых Марковских моделей имеет хорошее качество распознавания на фотографиях, приведенных к стандартному виду. На базе ORL было достигнуто 90% точность распознавания. Полезное свойство распознавания по коэффициентам дискретного косинусного преобразования заключается в том, что оно позволяет работать непосредственно со сжатыми изображениями, такими как JPEG и MPEG, в которые на сегодняшний день являются распространёнными форматами хранения изображений и видео.

Недостатком СММ является то, что СММ не обладает различающей способностью. Т.е. алгоритм обучения только максимизирует отклик каждой модели на свои классы, но не минимизирует отклик на другие классы, и не выделяются ключевые признаки, отличающие один класс от другого.

При увеличении объёма базы или использования в более широких условиях СММ может оказаться ненадёжными. Требуется либо доработка данного алгоритма, либо применение другого подхода для решения задачи распознавания.

В целом, система показала достаточно хорошие результаты работы как на специально подготовленных изображениях, так и на реальных изображениях, полученных с камеры при условии хорошего качества изображения и равномерного освещения лица. Примерно в 80% случаев система правильно выделила и распознала лицо человека.



## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

### Книги:

1. Максимов В.В. Система, обучающаяся классификации геометрических изображений, Сборник "Моделирование обучения и поведения", М., "Наука", 1975, с.29-120
2. Бонгард М.М. Проблема узнавания, М., "Наука", 1967

### Статьи:

3. Брылюк Д., Старовойтов В. Разработка нейросетей, выбор их параметров, их оптимизация для распознавания видеоизображений людей. Минск: BSUIR, 2001.
4. Бутаков Е.А., Островский В.И., Фадеев И.Л. Обработка изображений на ЭВМ М.: Радио и связь, 1987
5. Ara V. Nefian, Monson H. Hayes III. Face recognition using an embedded hmm. Georgia Tech Lorraine.
6. Govinandaraju V., David B.Sher, Srihari R.K., Srihari S.N. Locating Human Faces In Newspaper Photographs. IEEE, 1989, CH2752-4, pp.549-554.

### Препринты:

7. Мурынин А.Б., Серебренников О.А., Кулаков В.В. Автоматическая система распознавания личности по характеристикам изображения и голоса, ВЦ РАН, 1997, 46с
8. Мурынин А.Б., Цурков В.И. Принципы оптимизации информативных признаков в задаче автоматического распознавания лица человека компьютерной системой. Сообщения по вычислительной математике, ВЦ РАН, 1997, 34 с.
9. Мурынин А.Б., Матвеев В.А. Идентификация объектов по стереоизображениям. Оптимизация алгоритмов восстановления поверхности, ВЦ РАН, 1997.
10. Зорин С.А., Матвеев И.А., Мурынин А.Б., Сеньков Р.В., Цурков В.И. Восстановление рельефа лица в задаче автоматической идентификации личности. Сообщения по прикладной математике, ВЦ РАН, 1997, 34 с.

### Нормативно-технические документы:

11. Open Source Computer Vision Library. Reference Manual.- Intel Corp, 2002

### Материалы, размещенные в сети Internet :

12. Д. Брилюк, В. Старовойтов. Распознавание человека по изображению лица и нейросетевые методы - <http://daily.sec.ru/dailypblshow.cfm?rid=5&pid=4425&pos=1&stp=5>

13. Eickeler S., Muller S., Rigoll G. High performance face recognition using Pseudo 2-D Hidden Markov Models // Gerhard-Mercator-University Duisburg, Germany, 1998
14. Обнаружение и локализация лица на изображении. - [http://cgm.graphicon.ru:8080/issue2/face\\_detection/index.html](http://cgm.graphicon.ru:8080/issue2/face_detection/index.html)