

Работа выполнена на кафедре “Системы обработки информации и управления” Московского государственного технического университета имени Н.Э. Баумана

Семенова Яна Сергеевна

Научный
руководитель:

Кандидат технических наук,
доцент Филиппович Андрей Юрьевич

СИСТЕМА РАСПОЗНАВАНИЯ МИМИКИ ЧЕЛОВЕКА

Ведущая
организация:

Факультет «Информатики и система
управления» Московского государствен-
ного технического университета имени
Н.Э. Баумана

АВТОРЕФЕРАТ

диссертации магистра
техники и технологии

Код (ОКСО)

Наименование направления

230100

Информатика и вычислительная техника

Код

Магистерская программа

68

Квалификация:
Магистр техники и технологии

Защита магистерской диссертации состоится «24» июня 2010 г. в 9.00 на заседании диссертационного совета при Московском государственном техническом университете имени Н.Э. Баумана по адресу: 105005, Москва, 2-я Бауманская д.5.

С магистерской диссертацией можно ознакомиться в библиотеке Московского государственного технического университета имени Н.Э. Баумана.

Общая характеристика работы

Актуальность. Задолго до возникновения речевого общения, взаимодействие между людьми происходило на доступном невербальном уровне. Мимика стала не только рефлексорной, но и социально значимой. Изучение эмоциональных картин позволяет получить представление о сознательном и бессознательном отношении человека к окружающему миру.

С математической точки зрения, описание поверхности лица человека по сложности не отличается от детализации, например, поверхности земли. Расположение мимических линий относительно друг друга определяется силой экспрессии и может быть не только словесно описано, но и измерено количественно.

Существует множество визуальных методик классификации и трактования эмоций, но все они сложны и трудоемки для изучения, требуют специальной квалификации. Эксперт, как правило, владеет только одной методикой, а результаты его заключений носят субъективный характер.

Ни одна из визуальных методик не стандартизирована и не описывает всего многообразия мимических проявлений. Затруднен сравнительный анализ результатов исследований эмоций, поскольку в каждой лаборатории ученые выбирают разные методики, которые отличаются по видам и параметрам измерения.

Если экспериментальные данные получать посредством фото или видеосъемки и воспользоваться возможностью современных высокоскоростных процессоров обрабатывать большие объемы цифровых изображений, можно осуществить актуальную задачу автоматизации процесса описания и классификации эмоций.

Цель. Целью настоящего исследования является разработка системы распознавания мимики человека для количественного измерения силы экспрессии, определяющей яркость и выразительность мимической картины.

Задачи. Для реализации поставленной цели в работе решаются следующие задачи:

1. Проанализировать литературные источники, посвященные изучению мимики человека, изучить основные принципы существующих классификаторов мимических проявлений.

2. В описанных мимических картинах выделить значимые элементы мимики, которые могут быть индикаторами мимических проявлений.

3. Сопоставить функционал существующих программ по распознаванию мимики, методы, лежащие в их основе, определить их слабые и сильные стороны для формирования собственного метода распознавания мимики.

4. Смоделировать схему исследования для практической реализации метода распознавания мимики человека, отличного от существующих аналогов.

5. Написать пилотную версию программы, модифицирование которой можно было проводить итерационно, по разработанной схеме исследования.

6. В завершении всего комплекса исследований написать итоговый прототип программы распознавания мимики.

Задача интерпретации полученных результатов, т.е. собственно оценка эмоционального состояния человека, выходит за рамки данной работы, поскольку носит психологический или психиатрический характер и может быть решена только соответствующими специалистами (например, психиатрами или клиническими психологами).

Методы исследования. Результаты проведенных и представленных в магистерской диссертации исследований получены с использованием теорий вероятности, системного анализа, методов главных компонент, скрытых Марковских моделей, алгоритма обратного распространения ошибки для обучения искусственных нейронных сетей.

Научная новизна. В диссертации обоснованы и выносятся на защиту следующие основные положения, полученные по результатам проведенных исследований:

1. Разработанная система автоматизированного анализа цифрового панорамного изображения, содержащего лицо человека, позволяющая объективизировать полученные данные.
2. Алгоритм формирования банка цифровых эталонных изображений и классификатора эмоций человека.
3. Результаты экспериментальных исследований, полученных с помощью разработанного программного комплекса обработки цифровых изображений лица.

Обоснованность и достоверность научных положений, рекомендаций и выводов. Обоснованность научных положений, рекомендаций и выводов определяется корректным использованием методов исследования. Достоверность положений и выводов магистерской диссертации подтверждается результатами экспериментальных данных, полученных при использовании программного комплекса. Предложенные определения и классификации апробированы на конференциях и в научных публикациях.

Практическая ценность работы. Практическая ценность работы состоит в создании автоматизированной системы распознавания мимики человека, предназначенной для применения в научных исследованиях, экспертизах и коммуникационных приложениях. Система предоставляет возможность объективизации данных исследований мимики человека, количественного измерения эмоций, служит инструментом в работе экспертов, психиатров, клинических психологов, перед которыми стоит задача анализа мимики человека с целью составления описательных карт. Разработанный программный комплекс осуществляет последовательную обработку цифровых данных в соответствии с этапами анализа мимики человека. Модульная конфигурация системы позволяет встраивать ее в различные приложения, где необходимо задействовать возможности анализатора.

Апробация работы. Основные положения магистерской диссертации докладывались и получили одобрение на Научной школе "Компьютерная графика и математическое моделирование (Visual Computing)" в 2009 г., на Научно-методическом семинаре НОК CLAIM №4/10, а также на Научной межвузовской конференции преподавателей, аспирантов, молодых ученых и специалистов «Печатные средства информации в современном обществе» посвященной 80-летию Московского государственного университета печати в 2010 г. Результаты внедрения продемонстрировали эффективность применения разработанной в рамках работы системы распознавания.

Основное содержание работы

Магистерская диссертационная работа состоит из введения, четырех глав, заключения, списка использованной литературы из 56 наименований и приложений. Общий объем магистерской диссертации составляет 106 страниц машинописного текста.

Во введении описывается структура работы, формулируются цели и задачи работы, указаны предпосылки, которые делают работу актуальной в настоящее время.

В первой главе вводятся понятия мимики и мимических картин, экспрессии, мимиогенной зоны, мимического паттерна, рассматривается трактование различных видов эмоциональных реакций человека в современных психологических исследованиях и приводится пример описания основных мимических картин.

В работе проанализированы современные подходы к выделению значимых мимических черт, подлежащих анализу, и выполнено сравнение двух основных подходов к распознаванию мимики: визуального и автоматизированного. Проанализированы методики визуального анализа, основанные на описательных методах. Отмечены основные принципы проведения успешного эксперимента по распознаванию мимики. Описываются особенности физиологии восприятия человека.

Исследования основываются на работах таких зарубежных авторов как Экман П., Фризен У., Bartlett M., Bernhaupt R., Duda J.R, Fu K.S., Gonzalez R.C, Hazlett R., Kaiser S., Kuilenburg H., Lyons, M.J., Mahike S., Mandryk, R.L., а также работах отечественных авторов: Барабанщиков В.А., Барцев С.И., Гилев С.Е., Охонин В.А., Витт Н.В., Ермолаева-Томина Л.Б., Галушкин А.И., Зинченко Е.А., Лепешенков К., Чепин Е., Попова И.А., Бакеев Д.А., Ильина И.В. и др.

Проведен сравнительный анализ существующих методов визуального описания мимики человека. Указанные методы сгруппированы по типам действия, показано различие визуальных методик по видам единиц измерения, интенсивности и времени измеряемого действия. По результатам анализа составлена сравнительная таблица методов измерения мимических движений в разрезе различий по видам и типам мимических движений (Табл. 1).

Метод	Тип действия	Интенсивность действия	Время действия
Ekman & Friesen (1976, 1978)	Измерение всех мышечных движений; 44 единиц движений	Четыре действия по три контрольных точки интенсивности	Старт-стоп (начало, максимум, смещение)
Frois-Wittmann (1930)	28 описаний	Не описывает	Не описывает
Fulcher (1942)	отсутствие/наличие 16 мышечных движений	Рейтинг по объему движений в каждой области лица	Не описывает
Ermiane & Gergerian (1978)	Измерение всех видимых движений; 27 мышечных движений	Каждое действие оценивается по трех-бальной шкале	Не описывает
Landis (1924)	22 описания	Каждое действие измеряется по четырех-бальной шкале	Не описывает

Таблица 1. Сравнение анатомических методов

Выполнена постановка задач, требующих решения в рамках магистерской диссертации.

Во второй главе исследуются вопросы проведения необходимых мероприятий по предварительной обработке цифровых изображений, поскольку очевидна зависимость результатов точности распознавания мимики от качества исходных данных. Введены понятия обнаружения и локализации лица на панорамном изображении, рассмотрено описание вариантов процессов распознавания образов и приведен пример описания в виде графа. Даются рекомендации по установке параметров исходного цифрового изображения.

Предложенный процесс предварительной обработки цифровых изображений состоит из пяти основных этапов.

Этап 1. Кадрирование. В связи с тем, что на вход программы может подаваться не только фотография, но и потоковое видео, возникает необходимость выделения и выбора кадров, подлежащих распознаванию. Пользователю предоставлена возможность установки временного промежутка между захватами соседних кадров в потоковом видео. Эффективность распознавания во многом будет зависеть от правильного выбора и подготовки кадра. Проводится анализ кадра на предмет пригодности как параметра функции распознавания.

Этап 2. Локализация лица на панорамном изображении. На втором этапе обработки осуществляется поиск лица на анализируемом кадре. Детектирование лица человека на панорамном изображении осуществляется по алгоритму Viola-Jones. Изображение сканируется рамкой заданного размера и строится пирамида копий. Построенная пирамида анализируется заранее обученными каскадами Хаара, и на изображении находятся области, отвечающие заданным признакам. По крайним точкам описанной области происходит построение прямоугольника, в котором предполагается расположение найденного лица человека.

Проводится проверка горизонта по уровню глаз и при необходимости осуществляется поворот во фронтальной плоскости. Поворот в горизонтальной плоскости не производится, поскольку мышцы на лице человека могут работать не симметрично и восстановление

невидимой части лица по аналогии с видимой не является информативным при анализе двумерного изображения.

Выделенный прямоугольник, содержащий лицо человека, сохраняется в отдельный кадр, который поступает на дальнейшую обработку. В некоторых случаях возможна частичное удаление фона, что улучшает качество распознавания.

Этап 3. Пропорциональное масштабирование кадра. На основании сравнения высоты и ширины выделенного кадра, считается коэффициент масштабирования. Итоговый масштаб выбирается с учетом требований классификатора, необходимых при сравнении тестового изображения с эталонами. В результате масштабирования может произойти потеря качества цифрового изображения, результатом чего станет неверная типизация мимической картины.

Этап 4. Бинаризация кадра. На данном этапе выбирается порог яркости изображения и все пиксели сравниваются относительно выбранного порога. Цветное цифровое изображение преобразуется в черно-белое. Если кадр возвращается с последующего этапа как не валидный, он преобразовывается в изображение с градациями серого и снова поступает на проверку. Таким образом делается попытка увеличения информативности кадра.

Этап 5. Проверка валидности кадра. На последнем этапе проверяются все параметры кадра на предмет соответствия текущему классификатору. В случае неудачной проверки в потоковом видео выбирается следующий кадр, а при проверке фотографии пользователю предлагаются варианты по улучшению качества фотографии.

Необходимо отметить, что проверка валидности анализируемого исходного цифрового изображения происходит на всех этапах обработки. При отбраковывании кадра делаются попытки улучшения его параметров и он снова подается на вход отказавшейся от него функции. Данный механизм увеличивает время обработки кадра, но в тоже время снижает потерю информативности исследуемых цифровых объектов.

В третьей главе описывается модель бизнес-процесса разработки классификатора и составления банка эталонных изображений мимических картин. Представлен алгоритм применения классификатора в процессе распознавания мимики. Показаны возможности пользователя по изменению настроек процесса распознавания.

Проведен анализ существующих систем распознавания мимики человека путем группирования по методам, лежащим в основе их классификации (Табл. 2).

Т и п	Метод		Предмет изучения	Ограничения	Примеры программ
	Объект	Модель			
1	Плотность потока, опорные точки	Скрытые модели Маркова	Фронтальное положение головы	Только основные мимические картины	mplab.ucsd.edu
2	Опорные точки, коэффициенты Колмогорова	Нейронные сети	Незначительные повороты	Основные и комбинированные мимические картины	www.seestorm.com/technologies/cv/
3	Опорные точки	Комбинированные методы	Любые видимые лицевые изменения	Движения лица в реальном времени	www.vicarvision.nl www.agisoft.ru/projects/emotions.html

Таблица 2. Классификация автоматизированных методов распознавания эмоций

Большинство существующих программ распознавания носят коммерческий характер, поэтому информация об использованных в них методах и алгоритмах закрыта, что затрудняет процесс проведения классификации. Приведенная попытка не претендует на полноту и точность, поскольку программы постоянно совершенствуются, в них появляются новые функции, недоступные на момент проведения текущего исследования.

Задача распознавания мимики является частным случаем задачи распознавания образов. На предварительном этапе исследовались три метода, которые чаще всего применяются в современных исследованиях классификации объектов: метод главных компонент, скрытые модели Маркова, алгоритм обратного распространения ошибки.

Каждый метод был имплементирован в отдельном классе. Тестирование проводилось независимо друг от друга на одной и той же обучающей и тестовой выборке. После сопоставления результатов для итоговой реализации была выбрана искусственная нейронная сеть, обученная по алгоритму обратного распространения ошибки (Рис.1).

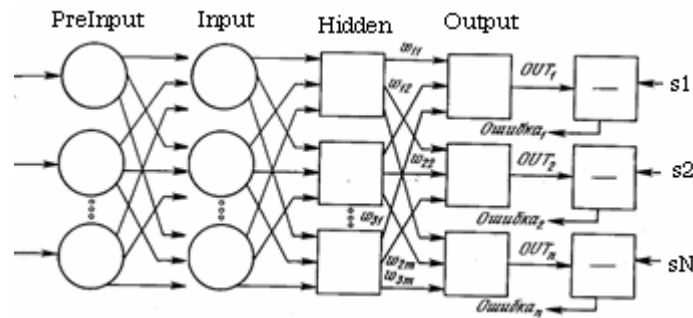


Рис. 1. Схема трехслойной сети.

Стрелками оказан механизм обратного распространения ошибки.

В качестве сжимающей или активационной функции выбрана сигмоида вида $F(x) = 1/(1 + e^{-x})$, которую можно представить:

$$OUT = \frac{1}{1 + e^{-NET}},$$

где NET – сумма произведений сигнала на вес;

OUT – сумма, модифицированная сигмоидальной функцией.

Производная сигмоидальной функции имеет вид:

$$\frac{\partial OUT}{\partial NET} = OUT(1 - OUT).$$

В обучающую выборку разработанной сети вошли семь эталонов, мимические картины на которых значительно отличаются по силе экспрессии (Рис.2). Таким образом, нейронная сеть получила семь выходов. Результирующая сумма всех выходов равна 1.



Рис. 2. Обучающая выборка [Paul Ekman, Wallace V. Friesen, Joseph C. Hager]

В тестовую выборку вошли цифровые изображения из той же базы эмоций, составленной Экманом, что обеспечило качество экспертной оценки. Результаты распознавания произвольно выбранных изображений давали ожидаемое низкое соответствие, поскольку исходная обучающая выборка включает минимальный набор возможных мимических представлений.

После проведения нескольких серий экспериментов по подбору параметров сети, было отобрано несколько различных конфигураций сетей, которые показали более 90% правильных распознаваний. Оценка точности распознавания проводилась путем сравнения значения мимической картины на максимально соответствующем эталоне со значением тестового эталона. Выбранные сети предоставлены пользователю в качестве вариантов выбора классификатора.

В результате проведенной автоматизированной классификации, пользователю предоставляются ранжированные варианты процентного соответствия тестируемого кадра каждому из эталонов. Результаты визуально представляются в виде гистограммы. На основании сформированных вариантов пользователь принимает окончательное решение по оценке мимической картины.

Оценка пользователя заносится в файл отчета и учитывается при анализе точности работы классификатора.

В четвертой главе рассматриваются вопросы реализации и практического применения системы анализа мимики человека. Описывается программный комплекс, созданный для поддержки этапов выполнения исследования, включающий расширенную часть настройки классификатора. Рабочее название созданного проекта «Mimic Analyzer».

Система имеет модульную архитектуру: модуль подготовки классификатора, модуль распознавания и модуль представления результатов (Рис. 3). Сохранение результатов работы каждого модуля организовано в виде файловой системы, работа с которой может осуществляться независимо от самого приложения.



Рис. 3. Архитектура системы

При реализации системы распознавания мимики человека на основе информации, полученной на этапе анализа исходных данных, разработан алгоритм построения графа-диалога для получения необходимых результатов распознавания.

Точность распознавания зависит от величины ошибки, заданной при обучении классификатора. Предоставление подробных настроек параметров нейронной сети позволяет проводить тонкую настройку инструмента распознавания.

Полный граф диалога взаимодействия с пользователем обра-

зуется динамически в зависимости от входных данных и выбранных пользователем параметров проведения процесса распознавания. На базе сформированного графа производится ситуационный диалог с пользователем. Далее выполняется поиск похожей картины в базе эталонов классификатора. Цель работы системы можно считать достигнутой, в случае нахождения хотя бы одного варианта соответствия мимической картины, превышающего заданную пороговую величину.

Система представлена не только как windows-приложение. Разработанный веб-сервис системы позволяет проводить анализ мимических картин в он-лайн режиме. Данный сервис является площадкой для постоянного тестирования обновлений в программной реализации, а также является источником пополнения банка данных.

На веб-страничке разработанной системы по адресу [http://philippovich.ru/Persons/Semenova_Yana/mimic.htm] можно ознакомиться с последними изменениями в программе.

В приложениях рассмотрены примеры применения системы распознавания мимики в виде руководства пользователя, приведен список сокращений и реферативный обзор основной литературы, а также список использованных источников.

Основные результаты работы

1. Проанализированы основные описательные методы измерения мимических картин. Проведен анализ существующих подходов к классификации мимики, по результатам которого составлена сравнительная таблица.
2. В описанных мимических картинах выделены значимые элементы мимики, которые являются индикаторами мимических проявлений.
3. Сопоставлен функционал существующих программ по распознаванию мимики, рассмотрены использованные в них методы организации процесса распознавания мимики.
4. Предложены алгоритмы формирования банка эталонных изображений.
5. Разработан алгоритм задания параметров для последующего обучения классификатора и сохранения результатов.
6. Разработан программный комплекс, основу которого составляет экспериментальная автоматизированная система распознавания мимики человека, которая осуществляет обработку пользовательских эталонов, позволяет получать количественные результаты при распознавании мимических картин.

Публикации

1. Семенова, Я.С. Распознавание мимики человека с использованием растровых изображений / Я.С. Семенова // Научная школа для молодых ученых «Компьютерная графика и математическое моделирование (Visual Computing)»: тезисы и доклады. М. – 2009 г. – С. 179-183.
2. Семенова, Я.С. Система распознавания мимики человека / Я.С. Семенова // Сборник учебно-методических работ и статей аспирантов и студентов Научной межвузовской конференции преподавателей, аспирантов, молодых ученых и специалистов «Печатные средства информации в современном обществе» посвященной 80-летию Московского государственного университета печати. – 2010 – в печати.