

ОБЗОР МЕТОДОВ РАСПОЗНАВАНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ СУБЪЕКТА

Дубровский П. С., Севрюк В. Э., Захарьев В. А.

Кафедра систем управления, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {victoria.sevryuk, joikervik}@gmail.com, zahariev@bsuir.by

В докладе рассмотрены основные области применения и особенности процесса автоматического распознавания эмоций. Представлен обзор актуальных методов распознавания эмоционального состояния на основе анализа речевого сигнала и изображений лица, направлений для дальнейших исследований.

ВВЕДЕНИЕ

Распознавание эмоционального состояния человека является актуальным направлением исследований, результаты которого могут быть использованы во многих сферах человеческой деятельности таких как медицина, психология, маркетинг, обеспечение безопасности. Существует восемь основных биполярных эмоций: радость и печаль, гнев и страх, доверие и отвращение, удивление и ожидание. Первичные эмоции могут быть выражены в разной интенсивности и могут быть смешаны друг с другом, образуя различные эмоциональные состояния. Методы распознавания эмоций используют различные типы входных данных, т.е. выражение лица, речь, жесты и языка тела, а также физические сигналы, такие как электрокардиограммы (ЭКГ), электромиографии (ЭМГ), электродермальная активность, температура кожи, гальваническое сопротивление, пульс и дыхание.

I. РАСПОЗНАВАНИЕ ЭМОЦИЙ НА ОСНОВЕ РЕЧЕВОГО СИГНАЛА

Эмоциональная окраска речевого сигнала (РС) относится к невербальному типу информации, оценивая которую можно делать вывод о психоэмоциональном состоянии диктора, определять его отношение к объекту описания, оценивать намерения говорящего, прогнозировать наступление целевых ситуаций или проявление определенного поведения индивида.

Анализ и распознавание эмоциональной окраски речи диктора начинается с извлечения вектора признаков для описания тех или иных параметров речевого сигнала. Можно выделить четыре общих класса признаков: просодические, спектральные, энергетические и темпоритмические [2].

Просодические признаки отражают такие перцептивные параметры речи как ударение, тон, интонацию. К просодическим признакам также относятся: частота основного тона (ЧОТ), скорость речи, джиттер. ЧОТ является одним из главных признаков, поскольку при изменении эмоционального состояния существенно изменя-

ются мгновенные значения и профиль контура ЧОТ (Рис. 1).

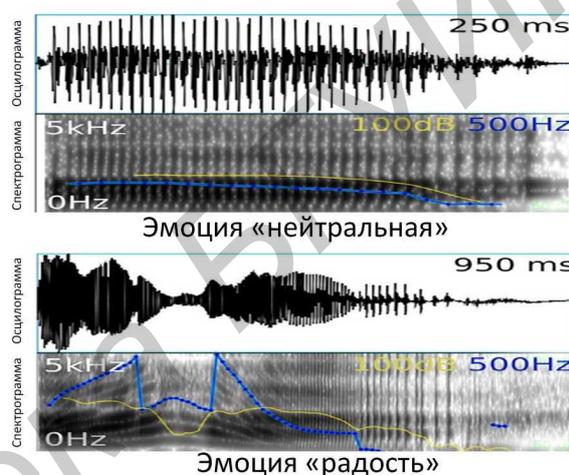


Рис. 1 – Изменение параметров РС в зависимости от типа эмоции

В качестве энергетических признаков выступает кратковременная энергия, а также ее различные вариации. Одним из часто используемых признаков является шиммер, который позволяет оценивать изменения пиковых значений амплитуд при переходе от одного кратковременного фрейма к другому или амплитудную модуляцию речевого сигнал. Темпо-ритмические параметры речи отражают скорость произнесения, временные параметры отдельных звуков или отдельных слов [3].

На основе выделяемого набора информативных признаков строится классификатор, который обучается на предварительно подготовленном наборе тренировочных данных. Классификация эмоциональных состояний производится в соответствии либо с задачами построения анализатора (оценки удовлетворенности, уровня стресса, усталости и т. п.), либо с выбранной моделью описания (набор базовых эмоций, непрерывная классификация и т. п.). Как правило, с ростом числа возможных вариантов классификации, точность распознавания эмоциональных состояний значительно снижается. Соответственно, количество классов, используемых для обучения выбирается небольшим [2].

Наиболее популярными техниками классификации являются следующие: поиск ближайших соседей, метод опорных векторов, скрытые марковские модели, модель смеси нормальных распределений, модели на основе нечеткой логики, байесовские классификаторы максимума вероятности. Самой важной проблемой в распознавании эмоций по голосу является извлечение дискриминационных признаков, которые эффективно характеризуют различное эмоциональное состояние[3].

II. РАСПОЗНАВАНИЕ ЭМОЦИЙ НА ОСНОВЕ ИЗОБРАЖЕНИЯ

Обнаружение присутствия человеческого лица на изображении является сложной задачей из-за возможности различных вариаций лица. Различные размеры, углы, освещение и позы человеческого лица на изображении могут вызывать эти сложности. Кроме того, наличие очков, бороды, волос и макияжа имеет значительный эффект в лицевом облике. Распознавание эмоций на основе изображения состоит из следующих этапов[4]: захват и отслеживание лица, определение особых точек и формирование характеристического вектора признаков, классификация в пространстве признаков (Рис. 2).



Рис. 2 – Распознавание на основе изображений лица

Современные методы распознавания используют три характерных признака для распознавания: уровень серого, движение и частоты. Разные эмоции приводят к разному уровню серого на цифровом изображении лица пользователя. Однако анализ этого характерного признака требует серьезной предварительной цифровой обработки изображения для нормализации входного изображения. Характерный признак: движение, использует информацию об изменении положения определённых точек на лице. Характерный признак: частоты, использует разницу между различными цифровыми изображениями лица в частотной области [5].

С другой стороны, методы распознавания можно разделить на три области: целостные и локальные распознавания, экстрагирования деформации и движения, геометрические характеристики и характеристика внеш-

ности. Методы целостных распознаваний анализируют эмоции человеческого лица в целом, чтобы затем найти разницу между разными изображениями[5]: Principal Component Analysis (PCA), Independent Component Analysis (ICA), Fisher's Linear Discriminants (FLD), Local Feature Analysis (LFA), Fisher Actions, Hidden Markov Models, и кластерный анализ. Методы локальных распознаваний анализируют отдельные части лица, например: рот, брови и глаза. Типичные методы: Facial Actions Code System (FACS), Local PCA, Вейвлеты и Нейронные сети. Методы экстрагирования деформации и движений для распознавания берут за основу изменения органов на лице, когда меняются различные выражения: Active Shape Model (ASM), Point Distribution Model (PDM). Методы геометрических характеристиках основаны на форме и положении различных частей лица человека, чтобы извлечь характерные векторы, которые представляют геометрические характеристики лица.

III. ЗАКЛЮЧЕНИЕ

В докладе рассмотрены основные методы автоматического распознавания эмоций с использованием различных видов информации о субъекте, в том числе при помощи мимики и речевых сигналов. Собрана база данных изображений и речевых сигналов для последующего анализа.

Необходимо отметить, что основную задачу представляет разработка алгоритма оценки эмоционального состояния человека по кадрам видеопотока. Эта проблема является наиболее перспективным направлением исследований, так как выражение эмоций человеком всегда требует определенного времени, а значит, простой анализ статического изображения не может давать точный результат. Для более точной оценки эмоционального состояния человека, наряду с анализом лицевой мимики следует также анализировать речь, интонацию и движения. Полноценный анализ состояния человека невозможен без комплексной системы определения эмоций. Создание подобной системы является целью дальнейшей работы.

IV. СПИСОК ЛИТЕРАТУРЫ

1. R. El Kaliouby, P. Robinson, Mind reading machines: Automated inference of cognitive mental states from video. – 2004. – Vol. 1. – P. 682-688
2. D. Kaminska, A. Pelikant, Recognition of human emotion from a speech signal based on Plutchik's model. – 2012. – P. 165-171
3. R. Plutchik, Emotion: a psychoevolutionary synthesis. – 2000.
4. K. Crammer, Y. Singer, On the algorithmic implementation of multi-class SVMs. JMLR. 2. . – 2001. – P. 265-292
5. Sujun Z. Facial Expression Recognition Algorithm Based on Active Shape Model and Gabor Wavelet. – 2010. – P. 40-45.