

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ИЗМЕРЕНИЯ ЭМОЦИЙ АБОНЕНТА МОБИЛЬНОГО ТЕЛЕФОНА

Михайлов В. О.

ФГБОУ ВПО «Пермский государственный национальный исследовательский университет», Пермь, Россия, (614990, Пермь, ул. Букирева, 15), e-mail: mr.wp@mail.ru

В статье описывается актуальность задачи распознавания эмоций абонента по голосу. Приводятся существующие алгоритмы, распознающие эмоции абонентов, и описаны сильные и слабые стороны этих алгоритмов. Предложен авторский алгоритм, позволяющий устранять недостатки описанных алгоритмов. Показано, что одним из достоинств авторского алгоритма, основанного на общей математической теории эмоциональных роботов с неабсолютной памятью, является возможность его использования для моделирования эмоций мобильных телефонов и вычисления их воспитания. В статье описано программное приложение, основанное на авторском алгоритме и определяющее численное значение эмоции, исходя из голосовых характеристик абонента. Проведена верификация авторского алгоритма с целью изучения его работоспособности. Приведены выводы, касающиеся эффективности применения алгоритма в измерении эмоций абонента мобильного устройства.

Ключевые слова: эмоция, распознавание, абонент, робот.

SOFTWARE IMPLEMENTATION MEASUREMENT OF EMOTIONS MOBILE PHONE USERS

Mikhailov V. O.

National Research Perm State University, Perm, Russia (614990, Perm, street Bookireva, 15), e-mail: mr.wp@mail.ru

The article describes the relevance of emotion recognition task caller's voice. Given the existing algorithms for recognizing emotions subscribers, and describes the strengths and weaknesses of these algorithms. Authors propose an algorithm to help eliminate the disadvantages of the described algorithms. It is shown that one of the virtues of the author's algorithm based on a general mathematical theory of emotional robots with non-absolute memory, is that you can use it to simulate the emotions of mobile phones and computing their education. This article describes a software application, based on algorithm and determining the numerical value of emotions, based on the characteristics of the voice call. The verification of the author's algorithm in order to study its performance. The conclusions concerning the effectiveness of applying the algorithm to measure emotions subscriber mobile device.

Keywords: emotion, recognition, the user, the robot.

Введение

На современном этапе развития информационных технологий разработка методов автоматического определения эмоционального состояния человека по голосу является актуальной задачей, позволяющей решить ряд экономических, социальных и бытовых проблем и, кроме того, играющей важную роль в вопросах безопасности [3]. Эмоциональный речевой сканер может найти широкое применение в различных транспортных и диспетчерских учреждениях, для ограничения или полного запрета доступа к выполнению служебных обязанностей лиц, находящихся в неустойчивом или неадекватном эмоциональном состоянии. Подобные системы контроля позволят также проводить дополнительную проверку пассажиров авиарейсов в рамках мероприятий по противодействию терроризму [3].

В настоящее время программных продуктов, которые распознают эмоции по голосу, практически не существует, в основном это экспериментальные программы, хотя есть множество алгоритмов, имеющих свои недостатки и достоинства. Рассмотрим некоторые из существующих алгоритмов. Недавно разработан алгоритм под названием «Классификация эмоционально окрашенной речи с использованием метода опорных векторов» (И.Э. Хейдоров, ЯньЦзинбинь, Уши, А. М. Сорока, А. А. Трус).

Авторы, рассматривая опорные вектора для распознавания эмоций, пришли к выводу о том, что применение метода опорных векторов для решения задач классификации эмоционально окрашенной речи позволяет получить высокую точность обученной модели. В сравнении с традиционными статистическими методами классификации [5] влияние методов извлечения векторов признаков на точность классификации обученной модели позволяет предположить, что модернизация этих методов является одним из путей дальнейшего увеличения точности рассмотренного в статье классификатора [5].

В предлагаемой статье рассмотрены проблемы классификации эмоционально окрашенной речи, извлечения векторов признаков, предварительной обработки обучающих выборок, выбора параметров алгоритма и оценки свойств полученного классификатора на основе метода опорных векторов (МОВ).

Согласно исследованиям авторов алгоритма «Классификация эмоционально окрашенной речи с использованием метода опорных векторов» дает точность классификации при правильном выборе оптимальных параметров алгоритма и ядерной функции, составляет 96,2 %. Сложность выбора параметров является одним из минусов алгоритма, так как в алгоритме отсутствует численное значение некоторой характеристики, которое ставится в соответствие значениям эмоции.

Опишем достоинства и недостатки алгоритма «Автоматическое определение изменений эмоционального состояния по речевому сигналу» (Лукьяница А. А., Шишкин А. Г.).

Авторы этого алгоритма описали технику отделения речи от пауз, а затем рассмотрели способы вычисления признаков, основанные на определении частоты основного тона, значениях трёх первых формант, а также на вычислении кепстра [1].

Исследования, проведённые авторами алгоритма, показали большую эффективность метода определения изменений в эмоциональном состоянии человека на основе анализа речевого сигнала. Достоинством этого алгоритма является высокая точность его работы (97.2 %), к недостаткам алгоритма можно отнести сложность настройки алгоритма (так как алгоритм опирается на базы данных с примерами эмоциональных состояний и зависимость от языка абонента, у которого опознается эмоциональное состояние) и невозможность вычислить с помощью алгоритма численное значение амплитуды эмоции.

Алгоритм измерения эмоций абонента мобильного телефона

Для устранения недостатков выше описанных алгоритмов предлагается авторский алгоритм определения эмоционального состояния абонента мобильного устройства, который легко настраивается на абонента и дает возможность возвращать численное значение определяемой эмоции.

Одним из источников эмоций, анализируемых алгоритмом, является речевой сигнал [4]. При изменении эмоционального состояния в человеческом организме происходят сложные процессы, которые в конечном итоге находят отражение в виде мышечных сокращений, в том числе и в голосовом тракте. Это даёт возможность бесконтактного определения эмоционального состояния человека по изменениям в системе речеобразования. Авторский алгоритм основывается на оценке амплитуды звуковой волны и на гипотезе, говорящей о том, что при изменении амплитуды звуковой волны меняется эмоциональное состояние абонента.

Приведем шаги авторского алгоритма, определяющего численные характеристики эмоций мобильного устройства:

- Звуковая волна, поступающая в мобильное устройство через микрофон, разбивается на такты, равные 1 секунде.
- На каждом такте вычисляется максимальное по модулю отклонение от начального значения амплитуды звуковой волны.
- Если отклонение отрицательное, то эмоции приписывают отрицательное значение, иначе – положительное. Формула для определения численного значения амплитуды

C_{i+1} эмоции на такте $[t_i, t_{i+1})$ описывается формулой $C_{i+1} = \frac{P_{i+1} - A_i}{t_{i+1} - t_i}$, где P_{i+1} – максимальное по модулю значение амплитуды звукового сигнала на такте $[t_i, t_{i+1})$, A_i – амплитуда звукового сигнала на такте $[t_{i-1}, t_i)$.

В результате расчетов, основанных на алгоритме, строится эмоциональная кривая, которая характеризует эмоциональное состояние мобильного устройства. Одним из достоинств алгоритма является возможность определения численного значения эмоции устройства в зависимости от звуковых характеристик голоса абонента и независимость значения эмоций мобильного устройства от языка абонента.

Полученное численное значение эмоции можно использовать для определения воспитания мобильного устройства, применяя при этом математические модели теории эмоциональных роботов с неабсолютной памятью, разработанные в Пермском государственном национальном исследовательском университете [2].

Используя эти модели, введем соотношения, определяющие гармоническую эмоцию мобильного устройства [2] в виде

$$M_{i+1}(t) = C_{i+1} * \sin\left(\frac{\pi}{t_{i+1} - t_i}(t - t_i)\right),$$

где $M_{i+1}(t)$ – эмоция мобильного устройства для времени t , отсчитываемого от начала проявления эмоции устройством, и воспитание мобильного устройства R_{i+1} , определяемого формулой

$$R_{i+1} = \frac{2C_{i+1}(t_{i+1} - t_i)}{\pi} \frac{1 - \theta^{i+1}}{1 - \theta},$$

где θ – коэффициент памяти устройства, i – порядковый номер воспитательного такта [2].

Программная реализация алгоритма

Опишем программу, использующую авторский алгоритм и позволяющую определять численные значения эмоций абонента и эмоций, испытываемых мобильным устройством, на основании анализа характеристик звуковой волны, поступающей в устройство через микрофон.

Для работы программы в качестве входных параметров задаются следующие параметры:

1. Файл с расширением *.wav (11025 Hz, 8bit, Mono), содержащий голосовую запись абонента.
2. Коэффициент памяти θ для робота (по умолчанию программа устанавливает коэффициент памяти равный 0,5).

Программа вычисляет и визуализирует:

1. Кривую, описывающую изменение эмоции абонента и построенную для введенного звукового файла – отображается синим цветом (измеряется в бит/с).
2. Кривую, описывающую изменение амплитуды эмоции робота, промоделированную на основе изменения эмоции абонента – отображается красным цветом (измеряется в бит/с).
3. Графическую зависимость воспитания робота от времени, промоделированную на основе гармонической эмоции мобильного устройства – отображается черным цветом (измеряется битами информации).

На основании авторского алгоритма были проведены эксперименты по измерению эмоций абонента по голосу.

В качестве входных данных для работы программы использовались записи голосов различных людей, размещенные в публичной базе данных EmoDB (как мужчин, так и женщин, которые, находясь в состоянии злости или страха, произносили некоторую фразу).

На рис. 1 приведен результат действия программы на примере измерения эмоций абонента мужского пола, произносящего фразу в состоянии злости.

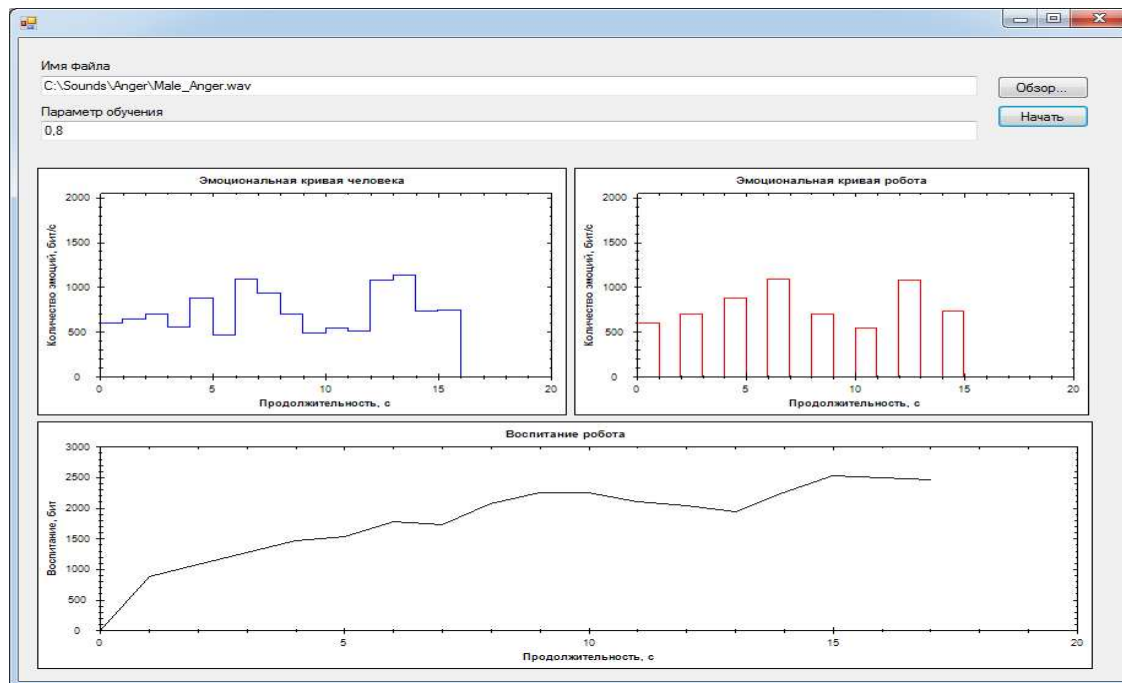


Рисунок 1. Результат действия программы для звукового файла male_Anger.wav с параметром обучения равным 0.8. Результатами программы являются эмоциональная кривая для человека (синий цвет), максимальная амплитуда эмоций робота (красный цвет), кривая воспитания робота (черный цвет)

Для сравнения эмоциональных кривых абонента, произносящего одинаковую фразу в разных состояниях, рассмотрим результат измерения эмоций человека, произносящего фразу в состоянии страха. Соответствующие графические зависимости приведены на рис. 2.

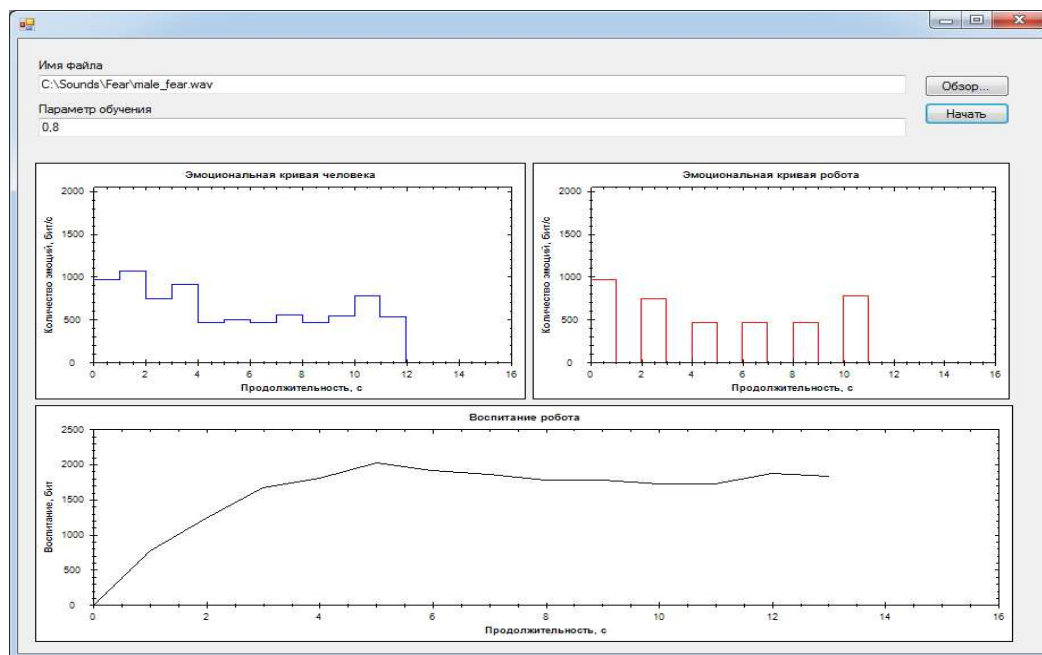


Рисунок 2. Результат действия программы для звукового файла `male_Fear.wav` с параметром обучения равным **0.8**. Результатами программы являются эмоциональная кривая человека (синий цвет), максимальная амплитуда эмоций (красный цвет), кривая воспитания робота (черный цвет)

В результате проведенных экспериментов, удалось установить, что для каждого абонента характерна собственная эмоциональная кривая, определяющая его эмоциональное состояние и позволяющая вычислять амплитуду гармонических эмоций мобильного устройства.

Программа предназначена для использования на персональных компьютерах IBM с операционными системами Windows 7x64, WindowsXPSP3 x86. Для написания программы использовался язык программирования c++. Объем загрузочного модуля 1.67 Мб.

Заключение

На основании предложенного алгоритма, описанного в статье, была реализована программа, измеряющая численное значение эмоции абонента по голосу человека, записанного в звуковой файл. Результатом программы являются эмоциональная кривая состояния человека, произносящего фразу, эмоциональная кривая мобильного устройства, полученная на основе эмоциональной кривой человека, кривая значений воспитания мобильного устройства.

Недостатком описанного в статье алгоритма является то, что алгоритм измеряет численное значение эмоции, но не распознает их. К достоинствам представленного алгоритма можно отнести следующее: алгоритм является достаточно простым (настройка алгоритма происходит автоматически), результаты работы алгоритма не зависят от языка произнесенных человеком фраз, алгоритм возвращает численное значение эмоций мобильного устройства и человека, что является его несомненным преимуществом по отношению к известным алгоритмам.

Список литературы

1. Лукьяница А. А., Шишкин А. Г. Автоматическое определение изменений эмоционального состояния по речевому сигналу // Речевые технологии. – 2009. – №3. – С. 53-61.
2. Пенский О. Г., Черников К. В. Основы математической теории / Перм. гос. ун-т. – Пермь, 2010. – 256 с.
3. Розалиев В. Л. Моделирование эмоциональных реакций пользователя при речевом взаимодействии с автоматизированной системой // Известия ВолгГТУ.–2010.–№6.–С.76-79
4. Розалиев В. Л. Моделирование эмоциональных реакций пользователя при речевом взаимодействии с автоматизированной системой: дис. ... канд. тех. Наук. – Волгоград, 2009.

5. Хейдоров И. Э., Янь Цзинбинь, У Ши, Сорока А. М., Трус А. А. Классификация эмоционально окрашенной речи с использованием метода опорных векторов // Речевые технологии. – 2008. – №3. – С. 53-61.

Рецензенты:

Пенский Олег Геннадьевич, доктор технических наук, профессор кафедры процессов управления и информационной безопасности Пермского государственного национального исследовательского университета, г. Пермь.

Ясницкий Леонид Нахимович, доктор технических наук, профессор кафедры прикладной математики и информатики Пермского государственного национального исследовательского университета, г. Пермь.