

УДК 57.08

Разработка БТС распознавания изменённого психоэмоционального состояния диктора по речевому сигналу

07, июль 2012

Чанина М.А.

*Студент,
кафедра «Биомедицинские технические системы»*

*Научный руководитель: Спиридонов И.Н.,
д.т.н., профессор кафедры «Биомедицинские технические системы»*

МГТУ им. Н.Э. Баумана
zujikka@mail.ru

Эмоции и речь тесно взаимосвязаны и играют огромную роль в общении. В связи с этим, автоматическая и объективная диагностика эмоционального состояния человека по его речи представляет большой практический интерес. Исследованием данного феномена занимаются различные научные и коммерческие организации.

Психоэмоциональное состояние — особая форма психических состояний человека с преобладанием эмоционального реагирования по типу доминанты. Эмоциональные проявления в реагировании на действительность необходимы человеку, так как они регулируют его самочувствие и функциональное состояние. Дефицит эмоций снижает активность центральной нервной системы и может явиться причиной снижения работоспособности. Чрезмерное влияние эмоциогенных факторов может вызвать состояние нервно-психического напряжения и срыв высшей нервной деятельности. Так как психоэмоциональное состояние отличается от психического состояния именно выраженным преобладанием эмоций, необходимо более подробно остановиться на их определении и функциях.

Эмоции – это особый класс психических процессов и состояний, связанных с инстинктами, потребностями и мотивами, отражающих в форме непосредственного переживания (удовлетворённости, радости, страха и т.д.) значимость действующих на индивида явлений и ситуаций для осуществления его жизнедеятельности. Эмоциональные состояния или эмоции представляют собой сложные и богатые реакции человека от взрывов страстей до тонких оттенков настроений.

В данной работе исследуется изменение психоэмоционального состояния человека, переживаемого человеком в процессе лжи и характер изменения невербальных характеристик его речи в этом состоянии.

Отмечено, что в процессе лжи человек испытывает ряд эмоций стенического характера (таких как: напряжённое внимание, тревога, страх и даже гнев). Эти эмоциональные переживания определяются как эмоциональное напряжение или более распространённым термином – эмоциональная напряжённость.

Эмоциональная напряженность, по мнению отечественных исследователей, имеет различную степень выраженности, которая зависит от характера конкретной эмоциогенной ситуации, в которой находится человек и индивидуальных особенностей самого человека.

В состоянии эмоциональной напряженности в речи говорящего появляется ряд характерных и достаточно устойчивых изменений, которые можно рассматривать как признаки этого состояния у человека.

По литературным данным и данным собственных исследований эти признаки следующие:

1. Акустические характеристики речи:

- возрастание среднего уровня громкости речи(рис.1.);
- возрастание динамического диапазона уровня речи;
- возрастание средней частоты основного тона голоса(рис.1.);
- возрастание динамического диапазона основного тона;

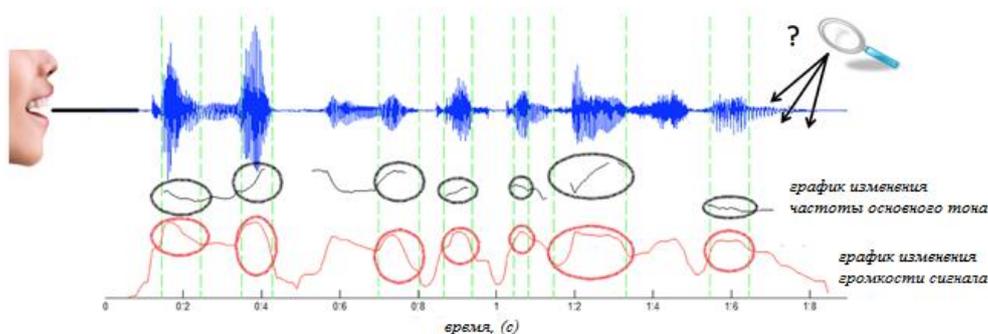


Рис.1. Изменение частоты основного тона и громкости в состоянии эмоциональной напряженности

- изменение огибающей интонационного контура.

2. Временные характеристики речи:

- изменение латентного времени речевой реакции;
- возрастание общего времени (количества и длительности) пауз;
- уменьшение длительности звучания гласных звуков.

В соответствии с этим, происходит и изменение ряда производных показателей речи, включающих временные характеристики:

- возрастает коэффициент паузации,
- возрастает темп артикуляции,
- уменьшается соотношение вокализованной и невокализованной речи.(рис.2.)

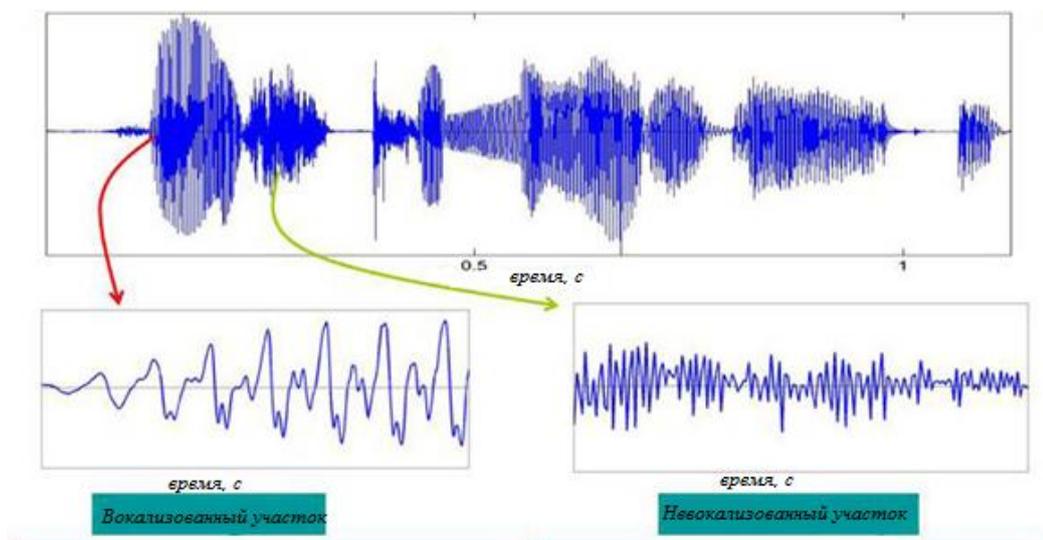


Рис. 2. Разделение речи на вокализованные и невокализованные участки

В ходе работы над данной тематикой планируется проведение следующих экспериментов.

1. Из стопки перевернутых карточек с цифрами от 0 до 9 человек вытаскивает одну карточку. Запоминает нарисованную на ней цифру. Кладет карточку обратно. После этого оператор последовательно спрашивает у испытуемого, какую карточку он достал. На что испытуемый всегда отвечает: «Нет». Таким образом, один из ответов – ложный, остальные – истинные. Цель программы: распознать ложь, узнать, какую карточку вынул испытуемый.

В результате этого эксперимента была собрана база данных из 10 спикеров.

7 были включены в группу обучения, 3 – в группу контроля.

2. В базу данных заносятся названия улиц, на которых проживают испытуемые. После этого оператор последовательно перечисляет улицы и спрашивает, на этой ли улице он проживает. На что испытуемый всегда отвечает: «Нет». Таким образом, один из ответов – ложный, остальные – истинные. Цель программы: распознать ложь, узнать, на какой улице проживает испытуемый.

В результате этого эксперимента была собрана база данных из 10 спикеров.

7 были включены в группу обучения, 3 – в группу контроля.

3. В базу данных заносятся возраста испытуемых. После этого оператор последовательно спрашивает у испытуемого, его ли это возраст. На что испытуемый всегда отвечает: «Нет». Таким образом, один из ответов – ложный, остальные – истинные. Цель программы: распознать ложь, узнать, каков возраст испытуемого.

В результате этого эксперимента была собрана база данных из 10 спикеров.

7 были включены в группу обучения, 3 – в группу контроля.

Планируется следующий способ обработки записей из базы данных.

- Построение столбца признаков
- Построение таблицы группы обучения. Результат – массив, включающий в себя 2 класса, соответствующих «лжи» и «истине». Каждый класс «лжи» представлен $7 \times 3 = 21$ записью, класс «истин» - $9 \times 7 \times 3 = 189$ записей.

Формирование матрицы перехода в LDA-пространство

- Нахождение среднего каждого класса
- Нахождение общего среднего
- Вычисления матрицы внутриклассовых различий

- Вычисление матрицы межклассовых различий
- Выделение собственных векторов, соответствующих наибольшему собственному числу, диагонализировать матрицу, матрица Z
- Отображение матрицы внутриклассовых различий в пространство Z , матрица mU_Sw_U
- Диагонализировать матрицу mU_Sw_U , найти векторы, соответствующие наибольшему значению собственных чисел, матрица U_vec
- Диагонализированное перемножение матриц Z и U_vec является матрицей перехода

Формирование группы тестирования. В группу контроля вошли 3 спикера. Вычисление меры различий тестируемого изображения с каждым из группы обучения. Тестируемый объект, представляющий собой вектор, отображается в LDA-пространство путем умножения на матрицу перехода. Результат – вектор-столбец. Каждый элемент группы обучения также отображается в LDA-пространство. Результат – вектор-столбец. Вычисляется Евклидово расстояние между тестируемым объектом и каждым из элементов группы обучения.

Отнесение тестируемого объекта к какому-либо классу. Тестируемому объекту присваивается класс, равный классу элемента группы обучения, расстояние до которого оказалось наименьшим.

Список литературы:

1. Дж.Фланаган. Анализ, синтез и восприятие речи. М.: Связь, 1968г.- 396с.
2. Чистович Л.А., Венцов А.В., Гранстрем М.П. Физиология речи. Восприятие речи человеком.-Л., «Наука», 1976.-388 с.
3. Алдошина И. / Основы психоакустики. Слух и речь. Части 1, 2 // Звукорежиссер // <http://rus.625-net.ru>
4. Фант Г. Акустическая теория речеобразования. М.: Наука, 1964.
5. Создание комплекса: Отчет о НИР / ООО «Нейроботикс»; Соловьева Е.С., - Шифр темы «Эскиз». -М. 2005.-65с.
6. Создание аппаратно-программных средств и методик оценки измененного психофизиологического состояния человека: Отчет об опытно-конструкторской работе / ООО «ЦРТ»; Руководитель А.Н.Араев, -Шифр темы «Цензура-06». - СПб., 2006.-125с.
7. Хроматиди А.И. Исследование психофизиологического состояния человека на основе эмоциональных признаков речи: Дис. канд. техн. наук.-Таганрог, 2005.-158с.
8. Носенко Э.Л. Особенности речи в состоянии эмоциональной напряженности. Днепропетровск. Издательство Днепропетровского Университета. 1975. С.132
9. Галунов В.И., Манеров В.Х., Тарасов В.И. Слуховой анализ речи, полученной при использовании различных методов моделирования эмоциональных состояний // Речь и эмоции. Л., 1979
10. Беловол Е.В. Проявление свойств темперамента в акустических характеристиках речи. Дис. канд. психол. наук: 19.00.01: Москва, 1999. РГБ ОД, 61:00-19/48-5
11. El Ayadi M., Kamel M.S. and Karray F. Survey on speech emotion recognition: Features, classification schemes, and databases // Pattern Recognition, 44(3), 2011. pp. 572-587.
12. Hozjan V. and Kacic Z. Context-independent multilingual emotion recognition from speech signal // Int. J. Speech Technol, 6, 2003. pp. 311–320.
13. Пахшина А.Х. К проблеме распознавания эмоционального контекста звуковой речи

14. Экман П. Психология лжи // Пер. с англ. под ред. В.В. Знакова. СПб: Питер. 2000. С. 272.
15. Ekman, P. (1981). Mistakes when deceiving. *Annals of the New York Academy of Sciences*, 364, 269-278.
16. Noguerras A., Moreno A., Bonafonte A., Marino J. B., "Speech emotion Recognition Using Hidden Markov Models", *Eurospeech 2001*, 2001.
17. Hawkins, S., House, J., Huckvale M., Local J., Ogden R. "ProSynth: An Integrated Prosodic Approach to Device-Independent, Natural sounding Speech Synthesis", *International Conference Speech and Language Processing*, 1998.
18. Горшков Ю.Г. «Новые решения речевых технологий безопасности» *Журнал «СПЕЦИАЛЬНАЯ ТЕХНИКА»* № 4, 2006. <http://www.ess.ru>.