

В.Л. РОЗАЛИЕВ, А.В. ЗАБОЛЕЕВА-ЗОТОВА
Волгоградский государственный технический университет

РАСПОЗНАВАНИЕ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА ПО ЕГО РЕЧИ

Проблема автоматического распознавания эмоционального состояния человека на основе акустических характеристик речи представляет интерес, как в теоретическом плане, так и для решения различных прикладных задач. Эта проблем на данный момент не является решенной. В данной работе предложена модель, описывающая эмоциональное состояние пользователя.

Интерес к распознаванию эмоций обусловлен тем, что при создании любого коммерческого продукта важной его характеристикой является конкурентоспособность. Знания об эмоциональном состоянии человека могут значительно повысить эту характеристику.

Эмоции есть отражение мозгом человека и животных какой-либо актуальной потребности и вероятности ее удовлетворения [2]. Каждому уровню потребностей, соответствует определенная преобладающая эмоция [1]. Следовательно, опираясь на потребности, эмоция является своеобразной «психической силой», направляющей поведение человека [3]. У человека, переживающего эмоцию, можно зафиксировать изменение электрической активности мышц лица [2], изменения в речи, некоторые изменения наблюдаются и в электрической активности мозга, в функционировании кровеносной и дыхательной систем [4]. Зная, какую эмоцию испытывает человек в тот или иной момент, можно предугадать его поведение.

Модель, описывающая эмоциональное состояние пользователя, представима в виде: $M = \langle X, Y, Z, f \rangle$, где X – вектор параметров речевого сигнала; Y – множество эмоциональных состояний; Z – множество доопределяющее X , строго говоря, являющееся необязательным; f – функция переходов, задается уравнением вида $Y = f(X, Z)$.

Под множеством X понимаются следующие акустические параметры:

$X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12})$, где x_1 – математическое ожидание частоты основного тона (ЧОТ, F_0), x_2 – дисперсия ЧОТ, x_3 – максимум ЧОТ, x_4 – минимум ЧОТ, x_5 – математическое ожидание интенсивности речи, x_6 – максимум интенсивности, x_7 – минимум интенсивности, x_8 – дисперсия интенсивности, $x_9, x_{10}, x_{11}, x_{12}$ – математическое ожидание формант F_1, F_2, F_3, F_4 . Выходными параметрами модели являются эмоциональные состояния пользователя.

$Y=(y_1, y_2, y_3, y_4, y_5, y_6)$, где y_1 –страх; y_2 –грусть; y_3 –гнев; y_4 –стыд; y_5 –радость; y_6 – бесстрашие. Множество Z – это множество, состоящее из $Z = (R, K)$, множества $R = (r_1, r_2, r_3)$ – видео параметров, изменения губ говорящего: r_1 – углы между нормалью к оси лица и самыми удаленными точками контуров губ (СУТКГ), r_2 – расстояния от оси лица до СУТКГ, r_3 – внутренние радиусы контуров, очерчивающих половинки губ человека, от оси лица до СУТКГ, и множества $K = (k_1, k_2)$ – множество параметров клавиатурного подчёрка: k_1 – параметры мыши $k_1=\{\text{угол перемещения, расстояние перемещения}\}$; k_2 – параметры клавиатуры $k_2=\{\text{время ввода символа, среднее время ввода (математическое ожидание), отклонение от среднего (дисперсия), изменение скорости набора (разность между соседними временными интервалами, делящимися между отдельными нажатиями)}\}$.

Таким образом, по множеству речевых параметров, становится возможным определение эмоционального состояния говорящего. Первоначальное предположение выносится на основании акустических параметров, в дальнейшем после распознавания фонем, построения из них фраз и определения ключевых (с наибольшей интенсивностью) слов, делается окончательный вывод об эмоциональном состоянии говорящего. Если говорящий является пользователем ЭВМ, то можно воспользоваться доопределяющими параметрами. При этом распознается, из поступающей с камеры картинки, лицо человека; проводится ось лица, разделяющая его на две равные половины; распознаются губы; контур губ разделяется на два контура, проходящих от оси лица до самых удаленных точек губ. Так же производится запись клавиатурного подчёрка пользователя. Получаемые данные с клавиатуры и мыши, соотносятся со сделанными выводами об эмоциональном состоянии. В дальнейшем при совпадении данных, делается вывод о том в каком эмоциональном состоянии находится пользователь.

Список литературы

1. К классификации эмоций / <http://www.voppsy.ru/issues/1991/914/914096.htm>
2. Симонов, П.В. Эмоциональный мозг: Физиология: Нейроанатомия: Психология эмоций. М., 1981.
3. Фоминых И.Б. Эмоции как аппарат оценок поведения интеллектуальных систем // Десятая национальная конференция по искусственному интеллекту с международным участием: Тр. конф. М.: Физматлит, 2006.
4. Эмоции / <http://www.inftech.webservis.ru/it/ii/books/book001/07g.htm>