

4. Журавлев Б.В., Шумилина А.И., Лосева Т.Н., Рычкова Г.Н. Поведенческие, вегетативные и электрофизиологические корреляты аппарата акцептора результат действия // Вестник АМН СССР. – 1985. – Т. 2. – С. 46-53.
5. Журавлев Б.В., Муртазина Е.П., Ломакина Т.Е. Патент на изобретение №2006103141/14 (003428), «Способ комплексной психофизиологической оценки индивидуально-типологических особенностей целенаправленной деятельности личности». Утвержден патент от 03.02.2006 г.
6. Karanasiou .I, Papageorgiou C., Tsianaka E. et al. Mismatch task conditions and error related ERPs// Behavioral and Brain Functions 2010, 6:14.
7. Baker TE, Holroyd CB. Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200 // Biol Psychol. – 2011. – Apr; 87(1). – С. 25-34.

Статью рекомендовал к опубликованию д.м.н., профессор О.С. Глазачев.

Муртазина Елена Павловна – Федеральное государственное бюджетное учреждение «НИИ нормальной физиологии им. П.К. Анохина» РАМН; e-mail: murtazina@yandex.ru; 125315, г. Москва, ул. Балтийская, 8, к. 228; тел.: 84992310048; лаборатория общей физиологии функциональных систем; к.м.н.; доцент.

Журавлев Борис Васильевич – e-mail: zhuravlev-bv@yandex.ru; 125315, г. Москва, ул. Балтийская, 8, к. 224; тел.: 84992310048; лаборатория общей физиологии функциональных систем, д.м.н.; профессор.

Murtazina Helen Pavlovna – P.K. Anochin Institute of Normal Physiologie Russian Academy of Medical Science; e-mail: murtazina@yandex.ru; r. 228, 8, Baltiisky street, Moscou, 125315, Russia; phone: +74992310048; laboratory of common physiologie of functional systems, assistant professor; cand. of med. sc.; associate professor.

Zhuravlev Boris Vasilevich – e-mail: zhuravlev-bv@yandex.ru; r. 224, 8, Baltiisky street, Moscou, 125315, Russia; phone: 84992310048; laboratory of common physiologie of functional systems; doctor. med. sc.; professor.

УДК 004.934.2

Н.Н. Филатова, К.В. Сидоров

МОДЕЛЬ ИНТЕРПРЕТАЦИИ ЗНАКА ЭМОЦИЙ ПО ЕСТЕСТВЕННОЙ РЕЧИ

Предложена модель интерпретации знака эмоционального состояния человека на основе анализа образцов его речи. Показана возможность применения характеристик нелинейной динамики для оценки информативных показателей эмоционального состояния человека. Предложен новый количественный признак \bar{R}_{\max} – усредненный максимальный вектор реконструкции аттрактора по четырем квадрантам. Показано, что нейтральное состояние характеризуется средними его оценками по сравнению с эмоцией гнева и счастья; закономерность сохраняется независимо от длительности образца. Предложена модель интерпретации знака эмоции по правилу объединения нечетких множеств, характеризующих значения \bar{R}_{\max} для соответствующих эмоциональных состояний.

Эмоция; эмоциональное состояние; речь; речевой сигнал; нелинейная динамика; нечеткое множество.

N.N. Filatova, C.V. Sidorov

THE MODEL FOR THE INTERPRETATION OF SIGN OF EMOTIONS ON NATURAL SPEECH

In this paper we propose a model interpretation of the character's emotional state based on the analysis of samples of his speech. The possibility of using the characteristics of nonlinear dynamics to assess the informative indicators of emotional state. A new quantitative characteristic \bar{R}_{\max} – the average maximum vector of the reconstruction of the attractor for the four quadrants. It is shown that the neutral state is characterized by its average ratings compared with the emotion of anger and happiness; pattern persists regardless of the length of the sample. A model interpretation of the sign of emotion by the rule of union of fuzzy sets that characterize the values \bar{R}_{\max} for the corresponding emotional states.

The emotion; the emotional state; speech; the speech signal; nonlinear dynamics; fuzzy set.

Процесс интерпретации (распознавания) знака эмоций человека по естественной речи является весьма сложной задачей как в плане математической формализации задачи, так и в способах четкой конкретизации эмоционального состояния (ЭС) – однозначного детектирования эмоции по речевому сигналу (РС).

В настоящее время отсутствует универсальная теоретическая модель описания речевых образцов в условиях проявления разных видов эмоций. Модель должна отражать взаимосвязь вида эмоций и объективных характеристик РС. Различные исследования в области акустики, психолингвистики и психофизиологии позволили собрать нам сведения о множестве акустических, просодических и лингвистических характеристик речи, которые можно использовать в качестве информативных признаков при распознавании ЭС, и проявляющихся на уровне речевых сегментов, фонема (отдельных звуков), слогов, целых слов и предложений. Чаще всего используются следующие признаки РС: спектрально-временные, амплитудно-частотные, вейвлет, кепстральные и характеристики нелинейной динамики [1].

В настоящее время в Тверском государственном техническом университете активно ведутся разработки по созданию модели интерпретации знака эмоций по естественной речи с использованием методов нелинейной динамики, позволяющих получить качественную и количественную оценку эмоций. Речевой тракт и процесс речеобразования представляют собой нелинейную диссипативную систему, что позволяет применять методы нелинейной динамики для анализа РС [2].

Методы нелинейной динамики базируются на фундаментальной математической теории, в основе которой лежит теорема Ф. Такенса [3], которая доказывает возможность восстановления (реконструкции) фазового портрета аттрактора по временному ряду или по одной его координате. Для реконструкции аттрактора исследуемый временной ряд x_n, \dots, x_{n-1} подвергается методу задержки координат (рис. 1), т.е. в фазовом пространстве строится последовательность реконструированных векторов, получаемых из элементов временного ряда [4]:

$$y_n = (x_n, x_{n+\tau}, \dots, x_{n+(m-1)\tau}), \quad n = 0, \dots, s-1, \quad s = N - (m-1)\tau, \quad (1)$$

где N – общее число элементов (точек) временного ряда; τ – задержка по времени между элементами временного ряда (временной лаг); m – размерность вложения (размерность лагового пространства).

При построении аттрактора необходимо правильно осуществлять выбор оптимальных значений параметров реконструкции аттрактора – m и τ (1). Величина размерности вложения m может быть определена с помощью корреляционного интеграла, а выбор задержки по времени τ может быть выполнен на основе анализа автокорреляционной функции [5].

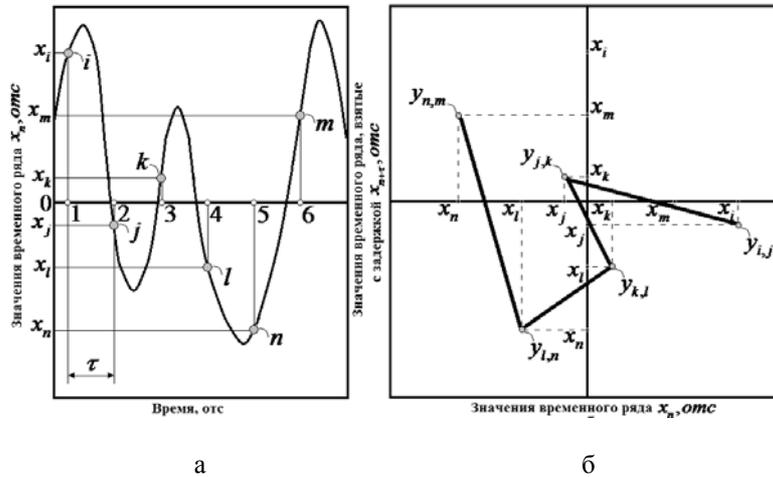


Рис. 1. Временной ряд (а) ($N = 6$); аттрактор (б) ($\tau = 1$ и $m = 2$)

Для проведения исследований использовались записи, взятые из берлинской БД эмоциональной речи EmoDB (Berlin Database of Emotional Speech) [6]. БД состоит из 535 фраз (предложений), произнесенных 10 дикторами (5 мужчин, 5 женщин). Каждый диктор произносил 10 фраз с имитацией одного из 7 эмоциональных состояний: гнев, скука, отвращение, беспокойство/страх, печаль, счастье и нейтральное состояние. Для некоторых эмоций в БД существует несколько вариантов их озвучивания одним и тем же диктором. Авторское исследование данной БД показало, что эмоции распознаются слушателями правильно в 80 % случаев, кроме того, в 60 % случаев признаются естественными. Всего для проведения исследований нами была построена экспериментальная выборка параметрических описаний РС, включающая обучающие выборки ОВ1 (60 предложений) и ОВ2 (150 гласных фонем, полученных из ОВ1). Проведен нелинейный анализ обучающих выборок объектов (предложений) ОВ1 и объектов (фонем) ОВ2 на основе реконструкции аттрактора (рис. 2).

Анализ геометрии полученных визуальных изображений показал (рис. 2), что в большинстве случаев наблюдается взаимосвязь геометрии аттрактора с видом эмоции. Визуально установлено, что эмоция гнева по сравнению с эмоцией счастья имеет большую траекторию разброса реконструкции аттрактора. Следует особо отметить тот факт, что закономерность разброса траектории реконструкции аттрактора одинакова как на уровне предложений речи (ОВ1), так и на уровне отдельных фонем РС (ОВ2).

Разработана подсистема количественной оценки аттракторов РС. Предложен новый признак, определяемый по результатам реконструкции, который существенно снижает размерность описаний речевых образцов и позволяет осуществлять количественно сравнение аттракторов, отображающих различные эмоции – усредненный максимальный вектор реконструкции аттрактора по 4 квадрантам \bar{R}_{\max} (рис. 3).

Вначале находится первый вектор реконструкции аттрактора в I квадранте $R_I^l = \sqrt{x_i^2 + x_{i+\tau}^2}$, где x_i – значение временного ряда в i -й момент времени, τ – временная задержка. Далее вычисляются оставшиеся n -векторов в I квадранте, в результате получается множество значений векторов реконструкции

$R^I = \{R_1^I, R_2^I, \dots, R_n^I\}$. Из множества R^I выбирается максимальный вектор R_{\max}^I . Аналогично находят максимальные вектора реконструкции аттрактора в других квадрантах R_{\max}^{II} , R_{\max}^{III} и R_{\max}^{IV} . Далее рассчитывается усредненный максимальный вектор реконструкции аттрактора по 4 квадрантам \bar{R}_{\max} , который является новым количественным признаком для распознавания i -го речевого образца:

$$\bar{R}_{\max}(i) = 0,25 \sum_{j=1}^4 R_{\max}^j(i), \quad i = 1, \dots, M \quad (2)$$

где j – номер квадранта; i – номер речевого образца (предложение или фонема); $M = 60$ для ОБ1; $M = 150$ для ОБ2.

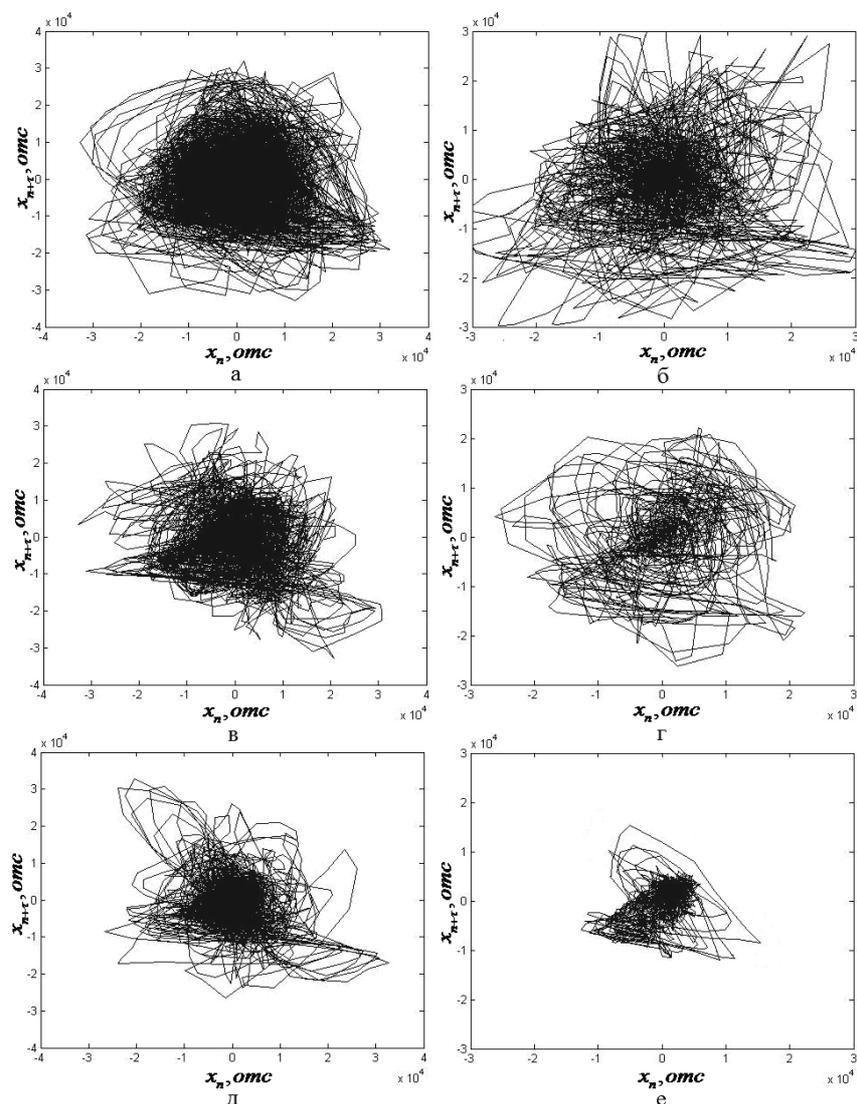


Рис. 2. Примеры реконструкций аттракторов выборок ОБ1 (слева) и ОБ2 (справа): а, б – гнев; в, г – нейтральное состояние; д, е – счастье

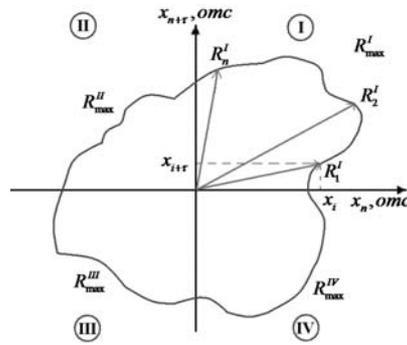


Рис. 3. Усредненный максимальный вектор реконструкции аттрактора \bar{R}_{\max}

Проведена количественная оценка реконструкций аттракторов на выборках речевых образцов разной длительности с помощью \bar{R}_{\max} (табл. 1).

Таблица 1

Усредненный максимальный вектор реконструкции аттрактора \bar{R}_{\max}

Эмоция	Выборка объектов	Признак, отс				
		\bar{R}_{\max}^I	\bar{R}_{\max}^{II}	\bar{R}_{\max}^{III}	\bar{R}_{\max}^{IV}	\bar{R}_{\max}
Гнев	ОВ1	37772	38185	43289	37125	39080
	ОВ2	23562	23440	18988	26407	23099
Нейтральное состояние	ОВ1	31671	33045	34846	40882	35111
	ОВ2	15590	11801	18777	14434	15151
Счастье	ОВ1	28257	34587	29716	39138	32925
	ОВ2	11098	10208	11800	11762	11217

На рис. 4 показаны средние оценки признака \bar{R}_{\max} для выборок ОВ1 и ОВ2. Анализ результатов показывает, что нейтральное состояние характеризуется средними оценками признака \bar{R}_{\max} по сравнению с эмоцией гнева и счастья. Эта закономерность сохраняется как на уровне предложений (ОВ1), так и на уровне отдельных фонем (ОВ2).

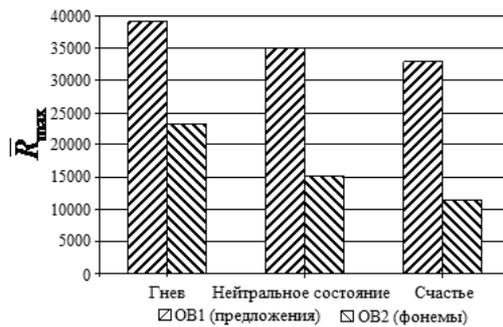


Рис. 4. Средние оценки признака \bar{R}_{\max} для выборок ОВ1 и ОВ2

Учитывая, что количественные оценки речевых образцов могут изменяться у разных дикторов, для интерпретации знака эмоции целесообразно перейти к лингвистическим переменным и использовать аппарат нечетких множеств. Для построения новой модели введем лингвистическую переменную – «Знак эмоции»:

$$P(E) = (\langle \text{Эмоция}(\text{ЗНАК}) \rangle, \{T_i\}, M(T)), \quad (3)$$

где $\{T_i\}$ – терм-множество с вербальными оценками знака эмоции для i -го речевого объекта; $M(T)$ – нечеткие подмножества универсального множества оценок \bar{R}_{\max} [25000;45000] для ОВ1 и [5000;25000] для ОВ2.

Для перехода от значений базовой шкалы (2) к лингвистической шкале (3) построены функции принадлежности (рис. 5). Тогда для каждого речевого объекта оценка знака эмоции будет определяться по правилу объединения нечетких множеств:

$$\mu_{T_j}(\bar{R}_{\max}) = \max_j \{ \mu_{T_1}(\bar{R}_{\max}), \mu_{T_2}(\bar{R}_{\max}), \mu_{T_3}(\bar{R}_{\max}) \} \quad (4)$$

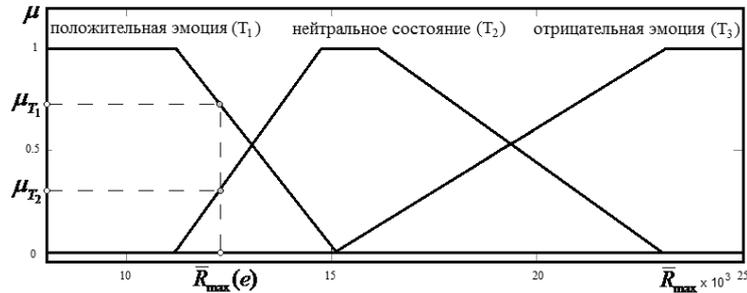


Рис. 5. График функций принадлежности к термам переменной «Знак эмоции» для объектов ОВ2

Например, образец фонемы $\langle e \rangle$ (рис. 2,е) с $\bar{R}_{\max}(e) = 12357$ принадлежит нечеткому множеству «положительная эмоция» (T_1) с функцией принадлежности $\mu_{T_1} = 0,7$ и нечеткому множеству «нейтральное состояние» (T_2) с функцией принадлежности $\mu_{T_2} = 0,3$ (рис. 5). Тогда в соответствии с (3 и 4) интерпретация знака эмоции примет вид

$$P(E) = (\langle \text{Эмоция}(\text{ПОЛОЖИТЕЛЬНАЯ}) \rangle / \mu_{T_1} = 0,7)$$

Таким образом, объединение методов нелинейной динамики и аппарата нечетких множеств позволяет по новому подойти к решению задачи анализа и интерпретации ЭС человека по РС.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Сидоров К.В., Филатова Н.Н. Анализ признаков эмоционально окрашенной речи // Вестник Тверского государственного технического университета. – Тверь, 2012. – Вып. 20. – С. 26-31.
2. Старченко И.Б., Перервенко Ю.С., Борисова О.С., Момот Т.В. Методы нелинейной динамики для биомедицинских приложений // Известия ЮФУ. Технические науки. 2010. – № 9 (110). – С. 42-51.
3. Takens F. Detecting strange attractors in turbulence // Dynamical Systems and Turbulence. – Heidelberg: Springer-Verlag, 1981. – P. 366-381.
4. Гориков В.А., Касаткин С.А. Идентификация временных рядов авиационных событий методами и алгоритмами нелинейной динамики. – М.: Бланк Дизайн, 2008. – 208 с.
5. Малинецкий Г.Г., Потанов А.Б. Современные проблемы нелинейной динамики. – М.: Эдиториал УРСС, 2000. – 336 с.
6. Burkhardt F., Paeschke A., Rolfes M., Sendlmeier W., Weiss B. A Database of German Emotional Speech // Proc. Intern. Conf. Interspeech. – Lissabon, 2005. – URL: <http://pascal.kgw.tu-berlin.de/emodb/index-1280.html>.

Статью рекомендовал к опубликованию к.т.н. Ф.Ф. Филиппов.

Филатова Наталья Николаевна – Тверской государственный технический университет; e-mail: nfilatova99@mail.ru; 170026, г. Тверь, наб. Аф. Никитина, 22, тел.: 84822444488; кафедра автоматизации технологических процессов; д.т.н.; профессор.

Сидоров Константин Владимирович – e-mail: bmisidorov@rambler.ru; тел.: 89607038330; кафедра автоматизации технологических процессов; аспирант.

Filatova Natalia Nicolaevna – Tver State Technical University; e-mail: nfilatova99@mail.ru; 22, Afanasy nikitin street, Tver, 170026, Russia; phone: +74822444488; the department of automation technological processes; dr. of eng. sc.; professor.

Sidorov Constantine Vladimirovich – e-mail: bmisidorov@rambler.ru; phone: +79607038330; the department of automation technological processes; postgraduate student.

УДК 004.93'1

Д.М. Ханеев, Н.Н. Филатова

ПИРАМИДАЛЬНАЯ СЕТЬ ДЛЯ КЛАССИФИКАЦИИ ОБЪЕКТОВ, ПРЕДСТАВЛЕННЫХ НЕЧЕТКИМИ ПРИЗНАКАМИ

Приводится результат исследования пирамидальной сети применительно к задаче классификации объектов, представленных нечеткими признаками в области биомедицинских систем. Рассмотрена методика построения пирамидальной сети и способ выделения из ее структуры нечетких описаний классов объектов, представленных в обучающей выборке. Полученные с помощью сети описания классов могут быть использованы для последующей классификации объектов рабочей выборки. Рассмотрено применение пирамидальной сети в задаче автоматической классификации образцов дыхательных шумов, заданных спектральными признаками.

Классификация; распознавание; дыхательные шумы; нечеткая логика; теория графов.

D.M. Khaneyev, N.N. Filatova

THE PYRAMIDAL NETWORK FOR CLASSIFICATION OF OBJECTS, PRESENTED BY FUZZY FEATURES

This paper presents the research results of the pyramidal network applied to the problem of classification of objects presented fuzzy features in the field of biomedical systems. The technique of building pyramidal network and an extraction method from this structure the fuzzy descriptions of object classes presented in training set is considered. Classes' descriptions which obtained through the network may be used for the working sample objects classifying. The application of a pyramidal network in problem of automatic classification of respiratory noise samples, presented by spectra, is considered.

Classification; pattern recognition; respiratory noises; fuzzy logic; graph theory.

Автоматическая классификация объектов в области биомедицинских систем является сложной задачей в силу множества зависимостей от различных факторов, поэтому ориентироваться на точные значения признаков объектов нельзя. Следовательно, для классификации сложных нестационарных объектов необходимо использовать аппарат нечетких множеств. Такой подход предполагает составление нечетких описаний классов, на основании которых можно судить о степени принадлежности объектов рабочей выборки к тому или иному классу. В качестве структуры, генерирующей нечеткие описания классов, предлагается нейроподобная сеть аналогичная растущей пирамидальной сети (РПС) [1, 2]. Из достоинств пирамидальных сетей следует отметить, что структуру обученной сети, в отличие от большинства искусственных нейронных сетей, можно легко интерпретировать в