

---

Научно-исследовательский институт  
«Центрпрограммсистем»

---

# Программные продукты и системы

МЕЖДУНАРОДНЫЙ НАУЧНО-ПРАКТИЧЕСКИЙ ЖУРНАЛ

№ 3 (111), 2015

Главный редактор  
С.В. ЕМЕЛЬЯНОВ, академик РАН

Тверь

## PROGRAMMNYE PRODUKTY I SISTEMY (SOFTWARE & SYSTEMS)

International research and practice journal

no. 3 (111), 2015

Editor-in-Chief  
S.V. EMEL'YANOV, *Academician of the Russian Academy of Sciences*

Tver  
Russian Federation

---

Research Institute CENTERPROGRAMSYSTEM

---

УДК 004.855.5, 004.853

Дата подачи статьи: 08.06.15

DOI: 10.15827/0236-235X.111.022-027

## **ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ ИНТЕРПРЕТАЦИИ НЕВЕРБАЛЬНОЙ ИНФОРМАЦИИ ПУТЕМ АНАЛИЗА ОБРАЗЦОВ РЕЧИ ИЛИ ЭЛЕКТРОЭНЦЕФАЛОГРАММЫ**

*Н.Н. Филатова, д.т.н., профессор, nfilatova99@mail.ru;*

*К.В. Сидоров, аспирант, bmisidorov@mail.ru;*

*С.А. Терехин, аспирант, rabeenovich69@mail.ru*

*(Тверской государственной технической университет,  
наб. Аф. Никитина, 22, г. Тверь, 170026, Россия)*

Интерпретация невербальной информации, которая содержится в образцах речи и в паттернах электроэнцефалограммы (ЭЭГ), позволяет получить количественные и качественные оценки характеристик эмоциональных реакций человека. В статье рассматривается ПО, ориентированное на решение указанной задачи. Описывается структура программного комплекса, позволяющего выполнять обработку образцов сигналов методами спектрального анализа и методами нелинейной динамики. Можно выполнить расчет спектров мощности или осуществить реконструкцию аттракторов. Для сравнения аттракторов предложены три типа признаков: длина максимальных векторов, описывающих контуры проекций аттракторов, плотность траекторий проекций аттракторов, степень разреженности матрицы плотности траекторий проекций аттракторов. Используя признаки, характеризующие геометрию аттракторов или спектры мощности, пользователь может создавать различные варианты моделей паттернов речи или ЭЭГ-сигналов. В моделях биомедицинских сигналов предусмотрен переход к нечетким признакам. Процедура фаззификации признаков выполняется с использованием треугольной конормы. Модуль интерпретатора эмоций осуществляет оценки трех характеристик: знака эмоции, ее уровня и направления развития (динамики эмоции). Взаимосвязь между этими характеристиками и компонентами моделей сигналов описана в виде набора нечетких правил. Программный комплекс прошел испытания на выборке речевых образцов русской, французской и немецкой речи, а также на выборках ЭЭГ-сигналов. Для русской речи точность интерпретации характеристик эмоций составила 96 %, для немецкой (из базы Emo-DB) – 92 %. Экспериментально доказано совпадение результатов интерпретации эмоций по образцам речи и ЭЭГ, зарегистрированных у одного и того же испытуемого.

**Ключевые слова:** интерпретатор эмоций, анализ речи, анализ ЭЭГ-сигналов, спектр мощности, нечеткое множество, нелинейная динамика, аттрактор.

Междисциплинарные исследования, связанные с формализацией моделей механизма быстрого анализа человеком информации, ведутся уже не один десяток лет.

В последние годы интерес к этому направлению особенно возрос, так как в робототехнике и в области разработки автоматизированных систем управления появилась насущная потребность в создании новых средств быстрой обработки плохо структурированной информации.

Одним из путей решения этой проблемы является имитация биологического механизма эмоциональных реакций, который от природы присущ человеку. Эмоции можно рассматривать как невербальные оценки информации, обрабатываемой мозгом. Эмоциональные реакции обычно превосходят реакции, связанные с вербальными оценками [1].

Одной из первых задач, решаемых в этом направлении, является создание моделей и алгоритмов распознавания некоторых характеристик эмоциональных реакций. На современном этапе развития информационных технологий ее решение связывают с созданием программных комплексов для распознавания эмоций путем интеграции результатов анализа речи, жестов и мимики испытуемого [2, 3]. Успешность этого подхода связана, с одной стороны, с интеграцией большого количе-

ства информации, а с другой – с довольно ограниченным набором жестов и мимики у цивилизованного человека. В целом можно считать довольно успешным решение задачи распознавания валентности (знака) эмоций, особенно при их яркой выраженности. Необходимо отметить, что эти решения тесно привязаны к особенностям конкретной функциональной системы человека (речевой, моторно-двигательный аппарат и др.), что не позволяет создать обобщенную модель механизма эмоций.

Одним из путей решения общей проблемы является создание многоканальной автоматизированной системы для исследования эмоций, которая дает возможность регистрировать реакции человека на внешние эмоционально значимые стимулы по нескольким каналам, включая обязательный мониторинг электрической активности мозга. Рассматриваемая система EEG/S [4] позволяет решить задачи:

– создания БД с образцами *биоэлектрических* (БЭ) сигналов, зарегистрированных при объективном подтверждении знака и уровня эмоциональной реакции испытуемого;

– формирования и испытания моделей интерпретатора характеристик эмоций на основе анализа образцов речи или образцов сигналов *электроэнцефалограммы* (ЭЭГ);

– исследования влияния гендерных, возрастных, а также национальных особенностей языков на параметры модели эмоций.

Архитектура системы EEG/S подробно рассмотрена в работе [4], в данной статье сделан акцент на ее ПО.

### ПО системы EEG/S

При разработке системы максимально использовался принцип открытости, который позволяет объединить программные средства, поставляемые совместно с оборудованием, известные проблемно ориентированные программы, а также оригинальные собственные разработки. В состав ПО системы входят три *программных комплекса* (ПК1–ПК3). ПК1 и ПК2 предназначены для регистрации и предобработки БЭ-сигналов. ПК3 реализует алгоритмы анализа и интерпретации характеристик эмоциональных реакций.

ПК1, функционирующий совместно с компьютерным энцефалографом-анализатором (Энцефалан-131-03), позволяет решать задачи регистрации ЭЭГ по 19 отведениям в формате ASCII, визуализацию, фильтрацию и сегментацию (в интерактивном режиме) сигналов, а также расчет спектров мощности и комплекса из 15 признаков, связанных с ними.

В состав ПК2 включен редактор звука CoolEdit, который обеспечивает регистрацию образцов речи испытуемого в формате ASCII, а также визуализацию и сегментацию (в интерактивном режиме) сигналов.

ПК3 (рис. 1) может работать как совместно с ПК1 и ПК2, так и автономно. В настоящее время он обеспечивает обработку двух типов сигналов (паттернов речи (*речевых сигналов* – РС) и ЭЭГ). Загрузка исходных данных осуществляется из файлов в формате ASCII.

Вторым шагом работы является полностью автоматическая сегментация сигналов, которая решается с целью выделения фрагментов  $x(t)$ :

- не содержащих артефактов (для ЭЭГ-сигналов);
- соответствующих отдельной гласной фонеме (для образцов речи);
- одинаковой длительности ( $L$ ).

Анализ выделенных фрагментов сигналов может осуществляться спектральными методами [5–7] или методами нелинейной динамики. В первом случае для каждого  $x(t)$  рассчитывается спектр мощности  $S(f)$ , дальнейший анализ которого может выполняться с помощью *нейроподобной иерархической структуры* (НИС-классификатора) [6] или методами кластерного анализа. С этой целью формируется описание каждого спектра в виде вектора  $S(f)=\langle sx_1, sx_2, \dots, sx_n \rangle$ , размерность которого зависит от способа задания его составляющих. НИС-классификатор позволяет формировать

правила и распознавать объекты, в описании которых может использоваться от 10 до 1 000 признаков. Если в качестве  $sx_j$  рассматривать относительное значение площади под кривой спектра мощности на ограниченном участке частот ( $\Delta f$ ) и принять  $\Delta f=f_x/Fw$  (где  $f_x$  – частота дискретизации,  $Fw$  – ширина окна FFT), то для описания образца речи понадобится 300 признаков [6]. Но пользователь сам может задать параметр  $\Delta f$  и уменьшить размерность описания. Так, если для описания спектров мощности ЭЭГ-сигналов использовать значения частотных интервалов, связанных с существованием определенных ритмов (альфа, бета и др.), размерность описания сократится до 5–9 признаков.

Применение НИС для анализа и классификации графиков (например спектров мощности) позволяет автоматически выделять интервалы частот, в пределах которых значения первичного признака можно рассматривать как элементы одного нечеткого множества  $\alpha$  уровня (при  $\alpha = 0,5$ ). Это позволяет автоматически определять максимально допустимые значения  $\Delta f$  (модуль выделения информативных интервалов), сокращая число признаков в векторе:  $S(f)=\langle sx_1, sx_2, \dots, sx_n \rangle$  [7].

Для анализа биомедицинских сигналов методами нелинейной динамики используется реконструкция многомерного фазового портрета, которая должна отображать фазовое пространство динамической биосистемы (модуль построения аттракторов). Согласно теореме Ф. Такенса [8], восстановленное пространство может быть сформировано из выборок исходного сигнала, сдвинутых на время задержки  $\tau$ :

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

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

Параметры  $m$  и  $\tau$  оказывают существенное влияние на форму аттрактора и его проекций. Определение временной задержки осуществляется по *автокорреляционной функции* (АКФ):  $\tau$  равна времени первого пересечения нуля АКФ. Величина размерности вложения  $m$  определяется с помощью алгоритма поиска ложных ближайших соседей. При исследовании речи и ЭЭГ-сигналов  $m \geq 2$ . На основе реконструкции аттрактора формируются его проекции (рис. 2).

Для оценки геометрии аттракторов, построенных по экспериментальным данным, обычно используют корреляционную размерность [9]. Однако этот признак позволяет дифференцировать нейтральное состояние человека от эмоционального возбуждения, но проявляет слишком слабые реакции на изменение знака эмоций.

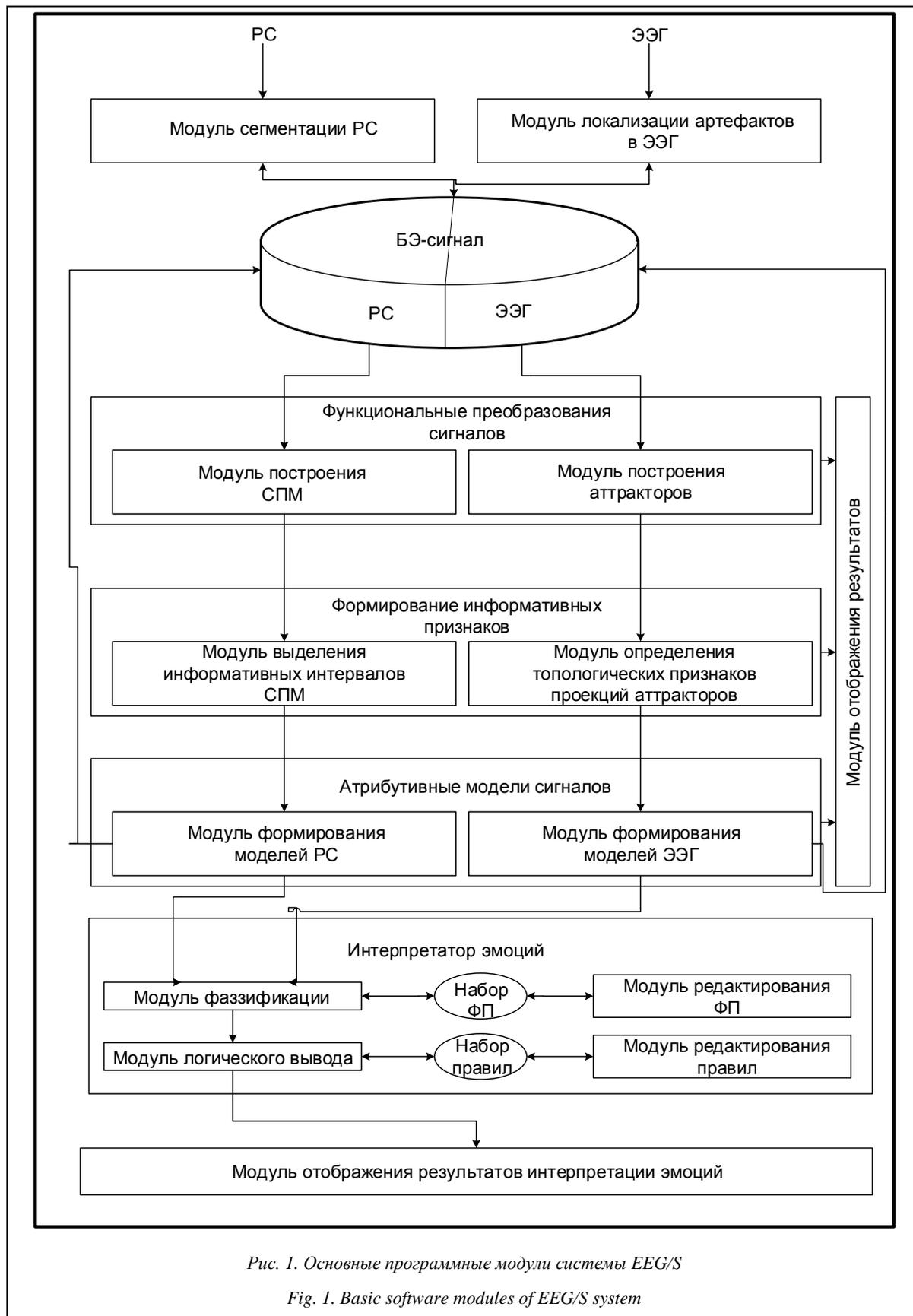
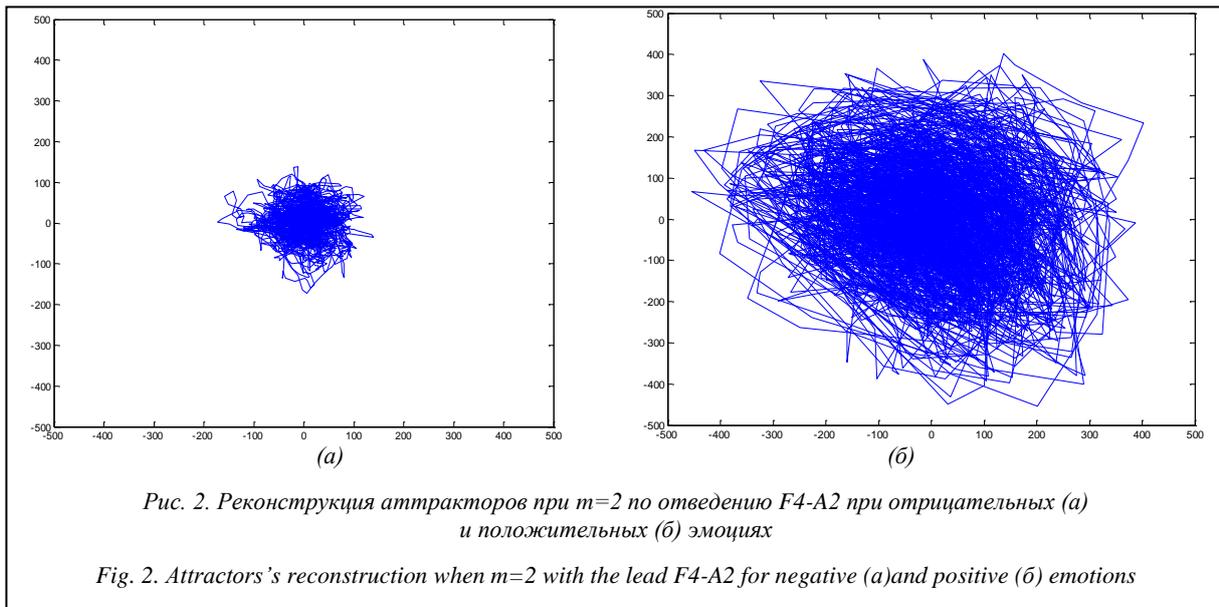


Рис. 1. Основные программные модули системы EEG/S

Fig. 1. Basic software modules of EEG/S system

Включенный в состав ПКЗ модуль определения топологических признаков аттрактора позво-

ляет получить количественные оценки трех новых признаков [10, 11]:



- длины максимальных векторов  $R$ , описывающих контуры проекций аттракторов и характеризующих биомедицинские сигналы по амплитудному составу;
- плотности  $\gamma$  траекторий проекций аттракторов;
- степени разреженности матрицы плотности траекторий проекций аттракторов.

В основу алгоритма их оценки положен последовательный просмотр всех проекций аттрактора, каждая из которых привязывается к собственной системе координат. На каждой проекции выделяются четыре квадранта и определяется вектор  $R$ , например, при  $m=3$ :  $R_i^{k=1} = \sqrt{x(t_i)^2 + x(t_i + \tau)^2}$ ,  $R_i^{k=2} = \sqrt{x(t_i)^2 + x(t_i + 2\tau)^2}$ ,  $R_i^{k=3} = \sqrt{x(t_i + \tau)^2 + x(t_i + 2\tau)^2}$ ,

где  $R_i$  – вектор для  $i$ -й точки аттрактора;  $x(t_i)$  – значение БЭ-сигнала в  $t_i$  момент времени.

Из множества  $R$  для каждого квадранта фазовой плоскости выделяются подмножества  $R^i$ :

$$R^1 = \{R_i \in R \mid (\forall i)x_i, x_{i+\tau} \geq 0 \rightarrow R(x(t_i), x(t_i + \tau)) \cup R^1\}$$

$$R^2 = \{R_i \in R \mid (\forall i)x_i \leq 0, x_{i+\tau} \geq 0 \rightarrow R(x(t_i), x(t_i + \tau)) \cup R^2\}$$

и т.д.

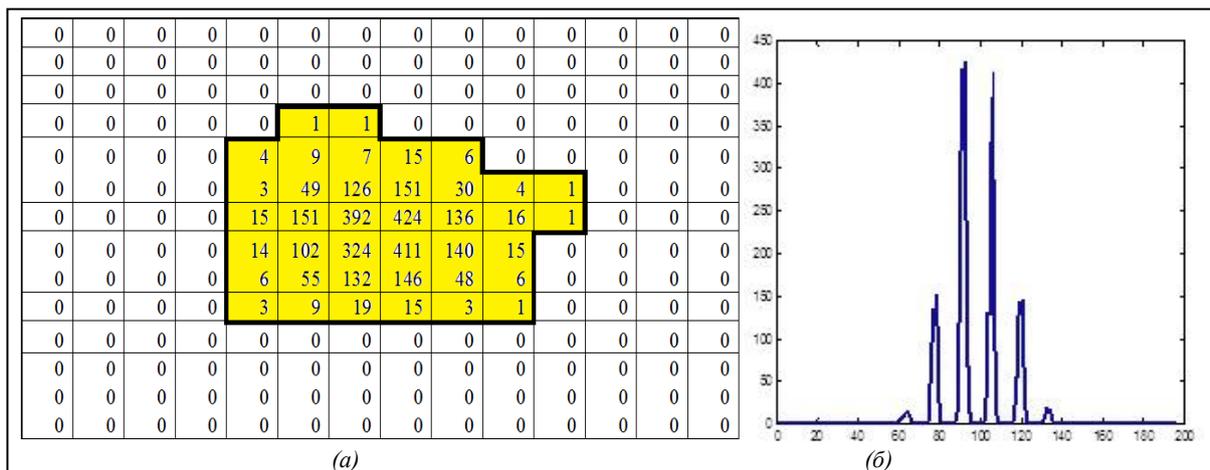
$$R^3 = \{R_i \in R \mid (\forall i)x_i \leq 0, x_{i+\tau} \leq 0 \rightarrow R(x(t_i), x(t_i + \tau)) \cup R^3\}$$

и т.д.

$$R^4 = \{R_i \in R \mid (\forall i)x_i \geq 0, x_{i+\tau} \leq 0 \rightarrow R(x(t_i), x(t_i + \tau)) \cup R^4\}$$

и т.д.

Формируется вектор  $R_{\max}^{kj}$  (максимальный вектор  $j$ -го квадранта  $k$ -й проекции аттрактора), а затем усредненный вектор  $k$ -й проекции аттрактора ( $\bar{R}_{\max}^k$ ). Он имеет четыре составляющие и по-



*Рис. 3. Матрица плотности траекторий аттрактора: (а) – пример матрицы при нейтральном состоянии испытуемого, (б) – график изменения плотности в зависимости от номера ячейки матрицы*

*Fig. 3. A density matrix of attractors' trajectories: (a) – a matrix sample of a testee's neutral state, (б) – a density variation graph depending on matrix element number*

звляет получать грубые оценки площади проекции аттрактора.

Для определения плотности  $\gamma$  проекция аттрактора покрывается регулярной сеткой с постоянным шагом  $d$ . Для каждой ячейки этой сетки можно найти отношение числа точек в ячейке ( $P$ ) к площади ячейки ( $S$ ); точки, оказавшиеся на границах ячейки, распределяются поровну между соседними элементами

$$\gamma_{i,j} = P_{i,j} * S_{i,j}^{-1} \text{ при } P_{i,j} = h_{i,j} + r_{i,j}/k, S_{i,j} = \delta^2,$$

где  $h_{i,j}$  – количество точек, попавших внутрь ячейки с координатами  $(i, j)$ ;  $r_{i,j}$  – количество точек, оказавшихся на границах этой ячейки;  $k$  – число соседних ячеек, имеющих общие точки с ячейкой  $(i, j)$ .

Совокупность оценок  $\{\gamma_{i,j}\}$ , найденная для всех ячеек сетки, представляется матрицей вида

$$M\gamma = [m_{ij}], \forall (i, j) m_{ij} = \gamma_{ij}.$$

Каждый ее элемент характеризует плотность аттрактора на  $(i, j)$ -м участке проекции (рис. 3). Число нулевых ячеек  $k_0$  матрицы  $M\gamma$  используется для оценки разреженности матрицы плотности, а также для оценки равномерности распределения траекторий аттракторов по плоскости проекций.

Используя перечисленные признаки и настройки  $\Delta f$ ,  $Fw$ , пользователь может создавать различные варианты моделей паттернов речи или ЭЭГ-сигналов (модули формирования моделей РС и ЭЭГ).

### Интерпретатор эмоций в составе системы EEG/S

В рамках интерпретатора эмоций используется переход к нечетким признакам в моделях биомедицинских сигналов (модуль фаззификации). Для этого с помощью унифицированного терм-множества вида (Значение признака Малое ( $T_{1\_}$ ), ... Среднее ( $T_{2\_}$ ), ... Большое ( $T_{3\_}$ )) создаются нечеткие множества путем привязки функций принадлежности к границам универсальных множеств значений признаков по шкале соответствующего сигнала (русской речи или ЭЭГ). Процедура фаззификации любого признака осуществляется с использованием треугольной конормы

$$\text{Max}(\mu_A, \mu_B) \Rightarrow A \cup B.$$

Интерпретация эмоций по образцам речи или ЭЭГ-сигналов основана на оценках трех характеристик: знака эмоции ( $Z$ ), ее уровня ( $U$ ) и направления развития (динамики эмоции,  $D$ ). Взаимосвязь между  $Z$ ,  $U$ ,  $D$  и компонентами моделей сигналов описана в виде набора правил.

Модуль логического вывода формирует окончательное решение задачи распознавания класса эмоций на основе результатов применения правил для всех найденных проекций аттрактора. Если по всем проекциям аттрактора определяется один класс, например ( $Lf$ ), его наименование использу-

ется для определения соответствующей характеристики эмоций ( $Z=Lf$ ). Степень соответствия этого заключения принимается равной минимальной функции принадлежности нечетких множеств, построенных для этих проекций. Если совпадение в выводах есть только для  $p$  проекций ( $p < (m-1)$ ), то для анализа выбирается следующий образец сигнала.

Программный комплекс прошел испытания на выборке речевых образцов русской, французской и немецкой речи, а также на выборках ЭЭГ-сигналов. Для русской речи точность интерпретации характеристик эмоций составила 96 %, для немецкой (из базы Emo-DB) – 92 %. Экспериментально доказано совпадение результатов интерпретации эмоций по образцам речи и ЭЭГ, зарегистрированных у одного и того же испытуемого. Реализованные в ПКЗ решения позволяют регистрировать промежуточные и заключительные результаты исследований, обладают чувствительностью к малым изменениям эмоционального состояния испытуемого, не требуют проведения продолжительного тестирования и способствуют уменьшению ошибок при диагностике эмоционального состояния испытуемого.

### Литература

1. Баарс Б., Гейдж Н. Мозг, познание, разум: введение в когнитивные нейронауки: в 2 ч. М.: БИНОМ. Лаборатория знаний, 2014. 1008 с.
2. Picard R.W. Affective Computing. M.I.T Media Laboratory Perceptual Computing Section Technical Report, 1995, no. 321.
3. Заболеева-Зотова А.В., Орлова Ю.А., Розалиев В.Л., Бобков А.С. Развитие системы автоматизированного определения эмоций и возможные сферы применения // Открытое образование. 2011. № 2-2. С. 59–62.
4. Сидоров К.В., Филатова Н.Н. Биотехническая система для анализа эмоций человека // Междунар. конгресс по интеллектуальным системам и информационным технологиям «IS&IT'14»: сб. трудов. В 4-х т. М.: Физматлит, 2014. Т. 2. С. 238–244.
5. Лапшина Т.Н. Психофизиологическая диагностика эмоций человека по показателям ЭЭГ: дис. ... канд. псих. наук. М.: Изд-во МГУ, 2007. 190 с.
6. Филатова Н.Н., Ханеев Д.М., Сидоров К.В. Интерпретатор сигналов на основе нейроподобной иерархической структуры // Программные продукты и системы. 2014. № 1 (105). С. 92–97.
7. Филатова Н.Н., Ханеев Д.М., Сидоров К.В. Алгоритм классификации графиков с последовательным укрупнением признаков // Программные продукты и системы. 2014. № 3 (107). С. 78–86.
8. Takens F. Detecting strange attractors in turbulence. Dynamical Systems and Turbulence, Heidelberg: Springer-Verlag, 1981, pp. 366–381.
9. Анищенко В.С. Лекции по нелинейной динамике. М.–Ижевск: Изд-во НИЦ «Регулярная и хаотическая динамика», 2011. 516 с.
10. Филатова Н.Н., Сидоров К.В. Модель интерпретации знака эмоций по естественной речи // Изв. ЮФУ. Технические науки: тематич. вып.: Медицинские информационные системы. 2012. № 9 (134). С. 39–45.
11. Ребрун И.А., Сидоров К.В., Филатова Н.Н., Ханеев Д.М. Модель проявления эмоций в естественной речи // XIV нац. конф. по искусственному интеллекту с междунар. участием КИИ-2014 (24–27 сентября 2014 г., Казань): тр. конф. Казань: Изд-во РИЦ «Школа», 2014. Т. 2. С. 112–121.

DOI: 10.15827/0236-235X.111.022-027

Received 10.06.15

**A SOFTWARE PACKAGE FOR INTERPRETATION OF NONVERBAL INFORMATION  
BY ANALYZING SPEECH PATTERNS OR ELECTROENCEPHALOGRAM***Filatova N.N., Dr.Sc. (Engineering), Professor, nfilatova99@mail.ru;**Sidorov K.V., Postgraduate Student, bmisidorov@mail.ru;**Terekhin S.A., Postgraduate Student, rabeenovich69@mail.ru**(Tver State Technical University, Nikitin Quay 22, Tver, 170026, Russian Federation)*

**Abstract.** Interpretation of non-verbal information that is contained in speech samples and EEG patterns provides a quantitative and qualitative assessment of human emotional reactions characteristics. The article considers software focused on this problem. The software package allows sample signal processing using both spectral analysis and nonlinear dynamics. It is possible to perform a calculation of a power spectrum or the reconstruction of attractors. The authors propose three types of signs to compare attractors: the maximum length of vectors describing the attractors' projections contours, the density of the attractors' projections trajectories, the degree of sparse matrix density of attractors' projections trajectories. A user can create different versions of speech patterns models or EEG signals using the features that characterize the attractors geometry or power spectra.

The models of biomedical signals provide transition to fuzzy features. Signs fuzzification is carried out using a triangular conorm. The interpreter emotions module evaluates three characteristics: the sign of an emotion, its level and direction of development (dynamics of emotions). The relationship between these characteristics and signal models components is described as a set of fuzzy rules. The software package has been tested on a set of speech samples in Russian, French and German, as well as on the samples of the EEG signals. The accuracy of the Russian speech emotions characteristics interpretation was 96%, for German (from Emo-DB base) it was 92 %. The coincidence of the emotions interpretation results using speech patterns and EEG recorded in the same test is experimentally proved.

**Keywords:** emotions interpreter, speech analysis, EEG signals analysis, power spectrum, fuzzy set, nonlinear dynamics, attractor.

**References**

1. Baars B.J., Gage N.M. *Cognition, Brain, and Consciousness: Introduction to Cognitive Neuroscience*. Academic Press, 672 p. (Russ. ed.: Binom. Laboratoriya znaniy Publ., 2014, 1008 p.).
2. Picard R.W. *Affective Computing*. M.I.T Media Laboratory Perceptual Computing Section Technical Report. 1995, no. 321.
3. Zaboлева-Zotova A.V., Orlova Yu.A., Rosaliev V.L., Bobkov A.S. *Razvitie sistemy avtomatizirovannogo opredeleniya emotsy i vozmozhnye sfery primeneniya* [Development of system of the automated determination of emotions and possible scopes of application]. Open education, 2011, no. 2-2, pp. 59–62 (in Russ.).
4. Sidorov K.V., Filatova N.N. Bioengineering system for human emotions analysis. *Mezhdunar. kongress po intellektualnym sistemam i informatsionnym tekhnologiyam "IS&IT'14": sb. trudov* [Proc. of the Congress on Intelligent Systems and Information Technology "IS&IT'14"]. Moscow, Fizmatlit Publ., 2014, vol. 2, pp. 238–244 (in Russ.).
5. Lapshina T.N. *Psikhofiziologicheskaya diagnostika emotsy cheloveka po pokazatelyam EEG* [Psychophysiological diagnostics of human emotions according to EEG]. PhD thesis. Moscow, MSU Publ., 2007, 190 p.
6. Filatova N.N., Khaneev D.M., Sidorov K.V. Signals interpreter based on neural-like hierarchical structure. *Programmnye produkty i sistemy* [Software & Systems]. 2014, no. 1 (105), pp. 92–97 (in Russ.).
7. Filatova N.N., Khaneev D.M., Sidorov K.V. Diagrams classification algorithm with consequent enlarging of features. *Programmnye produkty i sistemy* [Software & Systems]. 2014, no. 3 (107), pp. 78–86 (in Russ.).
8. Takens F. *Detecting strange attractors in turbulence. Dynamical Systems and Turbulence*. Heidelberg, Springer-Verlag Publ., 1981, pp. 366–381.
9. Anishhenko V.S. *Lektsii po nelineynoy dinamike* [Lectures on Nonlinear Dynamics]. Moscow–Izhevsk, Regulyarnaya i khaoticheskaya dinamika Publ., 2011, 516 p.
10. Filatova N.N., Sidorov K.V. The model for the interpretation of sign of emotions on natural speech. *Izv. YuFU. Tekhnicheskie nauki: tematich. vyp.: Meditsinskie informatsionnye sistemy* [Izvestiya SFedU. Engineering Sciences. A Special Issue "Medical Information Systems"]. 2012, no. 9 (134), pp. 39–45 (in Russ.).
11. Rebrun I.A., Sidorov K.V., Filatova N.N., Khaneev D.M. A model of expressing emotions in natural speech. *14ya nats. konf. po iskusstvennomu intellektu s mezhdunar. uchastiem KII-2014* [Proc. of the 14th National Conf. on Artificial Intelligence with Int. Participation CAI 2014]. Kazan, Russia, RITs Shkola Publ., 2014, vol. 2, pp. 112–121 (in Russ.).