

## Фрактальный анализ речевых сигналов

Д.В. Бельков, Е.Н. Едемская

Донецкий национальный технический университет  
belkov65@list.ru

*Бельков Д.В., Едемская Е.Н. Фрактальный анализ речевых сигналов. Анализ речевых сигналов является актуальной задачей поиска объективного показателя эмоционального состояния человека, выполняющего ответственную деятельность. В данной работе для речевых сигналов выполнен фрактальный анализ плотности распределения, автокорреляционных функций и энергетических спектров. Найдены значения показателя Херста. Он может быть показателем эмоционального состояния человека, поскольку степень фрактальности речевых сигналов зависит от эмоционального состояния и гендерных различий.*

**Ключевые слова:** речевой сигнал, плотность распределения, автокорреляционная функция, энергетический спектр, показатель Херста.

### Введение

1. В настоящее время большое внимание уделяется вопросам обработки информации и принятия решений при человеко-компьютерном взаимодействии. Эффективность данного процесса во многом зависит от качества информации, поступающей от пользователя автоматизированной системы и целенаправленности воздействия человека на объекты исследования. Достижение цели диалогового взаимодействия между компьютером и пользователем возможно при учете аспектов, характеризующих речевые потоки, возникающие в процессе общения.

2. Труд человека в системах управления техникой связан с периодическим, возможно длительным и интенсивным воздействием экстремальных значений профессиональных, социальных, экологических факторов, которое сопровождается эмоциями, перенапряжением физических и психических функций, деструкцией деятельности. Одним из источников эмоций является речевой сигнал. Русский язык содержит около 40% эмоционально окрашенных слов. Эмоции кодируются определенными акустическими параметрами в речевом сигнале. Понимание особенностей акустического кодирования эмоций позволит понять механизм восприятия эмоций [1]. Исследования речевых сигналов проводились многими учеными с целью описания как лингвистических, так и технических характеристик речи [2,3]. Однако, проблема автоматического распознавания эмоционально окрашенной речи не является полностью решенной. Эмоционально окрашенная речь находит применение во многих сферах жизнедеятельности человека и является востребованной функцией в современных

автоматизированных системах управления, реабилитации и протезирования, срочного оповещения и т.п. В последние годы усилился интерес к анализу речевых сигналов для поиска объективного показателя эмоционального состояния человека, выполняющего ответственную деятельность. На данный момент отсутствует модель описания речевых образцов в условиях проявления разных видов эмоций. Модель эмоционально окрашенной речи должна отражать взаимосвязь вида эмоций и объективных признаков речевого сигнала. Определение такой взаимосвязи вызывает трудности.

3. Основная задача получения признаков эмоционально окрашенной речи состоит в том, чтобы преобразовать звуковую волну в признаковое пространство, в котором множество объектов одного класса будет сгруппировано вместе, а множество объектов альтернативных классов максимально разнесено. Соотнесение распознаваемого объекта (под объектом понимается фонема речевого сигнала) с базой объектов, которые необходимо идентифицировать, проходит в три этапа: 1) выделение того или иного признака; 2) объединение признаков в группы; 3) выбор предполагаемого значения из ряда альтернатив [4]. Отечественные и зарубежные авторы выделяют четыре группы объективных признаков, позволяющих различать речевые образцы: спектрально-временные, кепстральные, амплитудно-частотные и признаки нелинейной динамики [5-8].

4. В работах [9,10] представлены результаты применения фрактального анализа речевого сигнала при моделировании актерами различных эмоциональных состояний. Выявлены достоверные различия величины фрактальной размерности при моделировании состояния горя,

радости и гнева по сравнению с нейтральным состоянием. Обнаружены гендерные различия динамики показателя фрактальной размерности при моделировании эмоциональных состояний. С фрактальной размерностью  $D$  речевого сигнала связан показатель Херста  $H$ , который характеризует память сигнала и может быть вычислен по формуле  $H=2-D$ .

5. В данной статье предлагается использовать величину  $H$ , как признак эмоционально окрашенной речи. Он относится к группе признаков нелинейной динамики.

6. Целью статьи является исследование структуры речевых сигналов, направленное на выявление их характерных особенностей. В работе решаются следующие задачи: оцениваются плотности распределения, автокорреляционные функции (АКФ), энергетические спектры изучаемых сигналов, определяется степень фрактальности (показатель Херста). Исследования выполнены в среде MatLab.

### Анализ сигналов

7. В работе проведены две серии вычислительных экспериментов. В первой серии выполнен фрактальный анализ двух речевых сигналов. Сигнал  $s1$  - текст читается актрисой без выражения, сигнал  $s2$  тот же текст читается актрисой с выражением. Во второй серии экспериментов анализируются сигнал  $s1$  и сигнал  $s3$  (текст читается актером). Изучаемые сигналы показаны на рисунках 1 – 3.

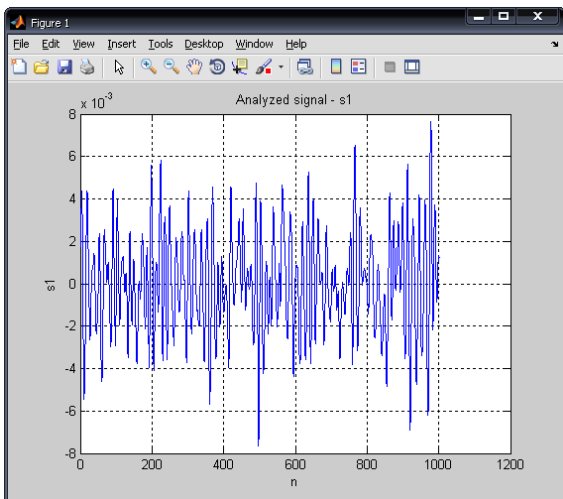


Рисунок 1 – Сигнал  $s1$

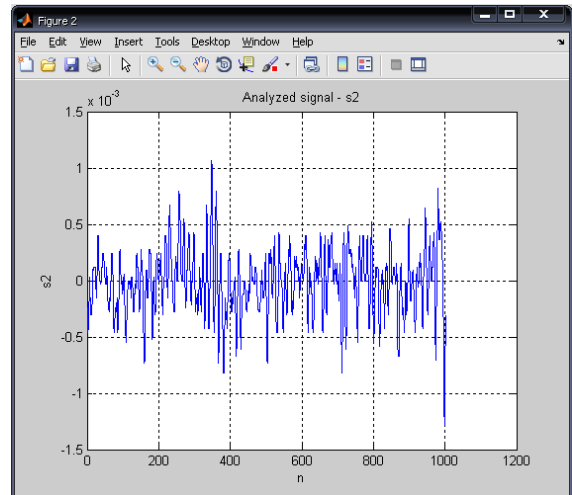


Рисунок 2 - Сигнал  $s2$

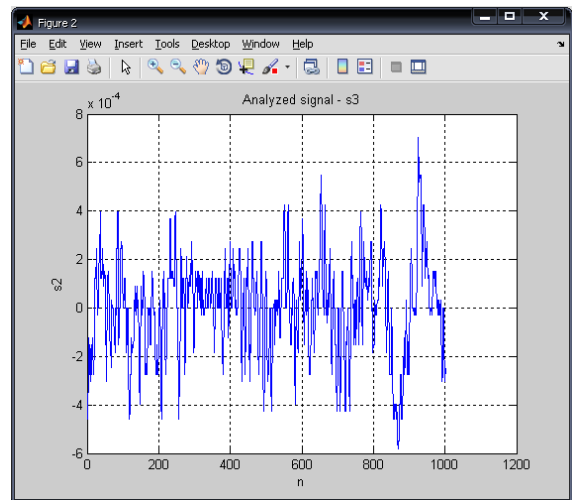


Рисунок 3 – Сигнал  $s3$

8. Оценка плотностей распределения проводится на основании гистограмм относительных частот, показанных на рисунках 4-6. Изучаемые сигналы не подчиняются нормальному распределению.

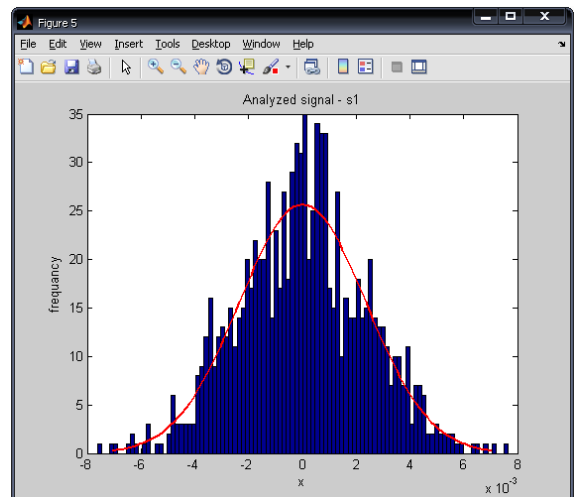


Рисунок 4 – Плотность распределения сигнала  $s1$

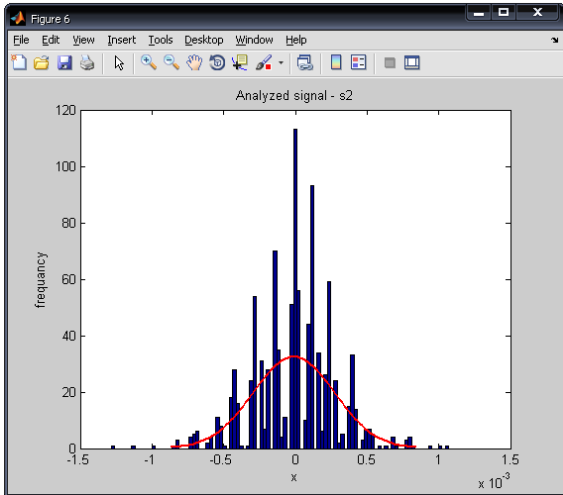


Рисунок 5 – Плотность распределения сигнала s2

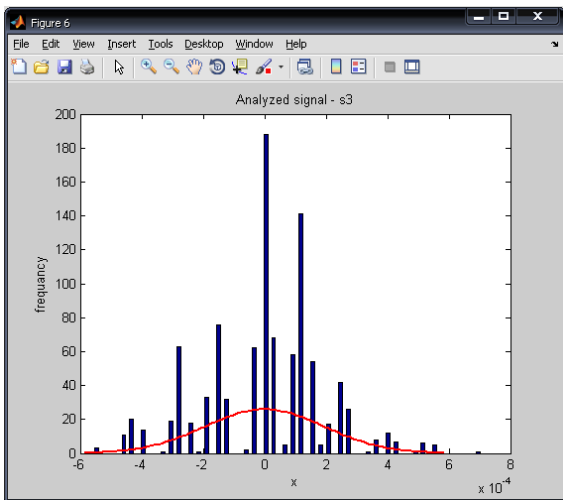


Рисунок 6 – Плотность распределения сигнала s3

9. В работе необходимо определить обладают ли сигналы медленно убывающей зависимостью или быстро убывающей зависимостью.

Процесс  $X$  обладает медленно убывающей зависимостью (МУЗ), если для его АКФ выполняется условие  $r(k) \sim k^{-\beta} \cdot L_1(k) + C_1, k \rightarrow \infty$ , где  $C_1 = const$ ,  $0 < \beta < 1$  и  $L_1$  - медленно меняющаяся на бесконечности функция:  $\lim_{t \rightarrow \infty} \frac{L_1(tx)}{L_1(t)} = 1$  для всех

х. Процессы с МУЗ характеризуются автокорреляционной функцией, которая убывает по степенному закону при увеличении временной задержки (лага). В отличие от процессов с МУЗ, процессы с быстро убывающей зависимостью (БУЗ) обладают экспоненциально спадающей АКФ вида  $r(k) \sim \rho^k, k \rightarrow \infty$ .

Вычисления АКФ по формуле

$$r(k) = \frac{\sum_{i=1}^{N-\tau} (X_i - \bar{X})(X_{i+k} - \bar{X})}{(N - \tau)\sigma^2(X)},$$

где  $\bar{X}$  - выборочное среднее ряда  $X$ ,  $\sigma^2(X)$  - выборочная дисперсия ряда  $X$ ,  $k=0,1,\dots$ . Автокорреляционные функции исследуемых сигналов приведены на рисунках 7-9. Они не обращаются в ноль при больших значениях  $k$ , что говорит о медленном убывании АКФ и присутствии МУЗ.

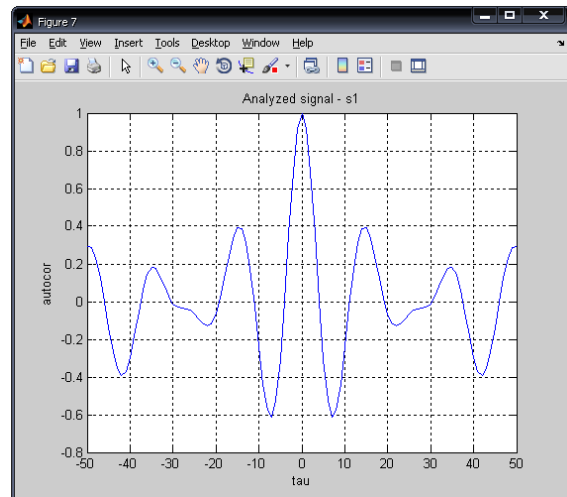


Рисунок 7 – АКФ сигнала s1

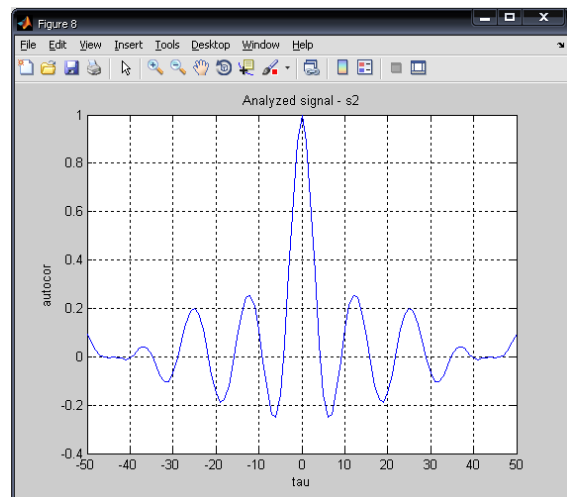


Рисунок 8 – АКФ сигнала s2

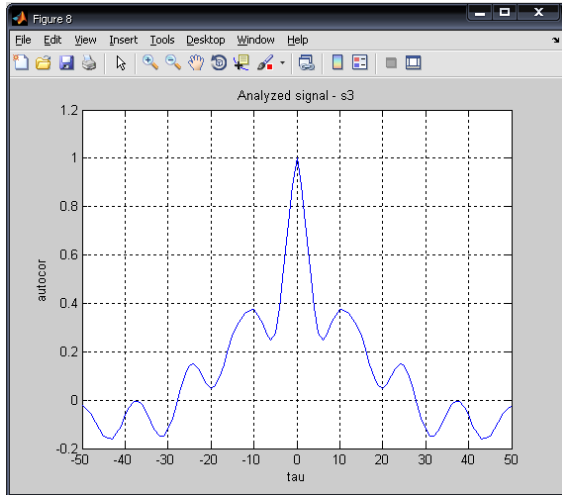


Рисунок 9 – АКФ сигнала s3

10. В частотной области МУЗ приводит к степенному закону поведения спектральной плотности процесса. Энергетические спектры временных рядов представлены на рисунках 10-12.

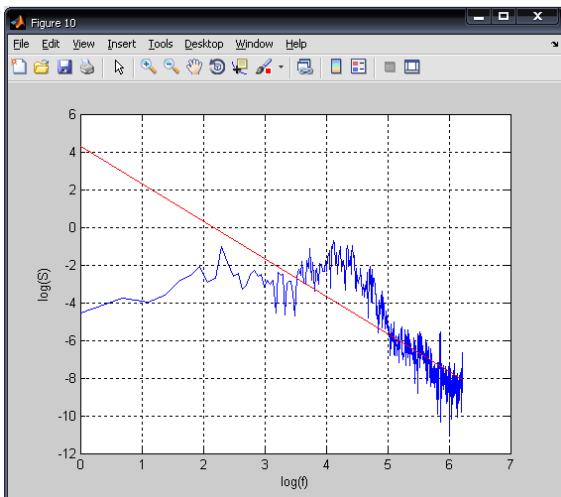


Рисунок 10 – Спектральная плотность сигнала s1

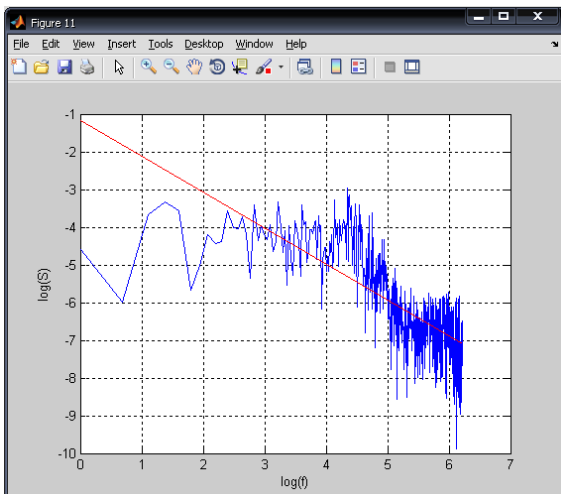


Рисунок 11 – Спектральная плотность сигнала s2

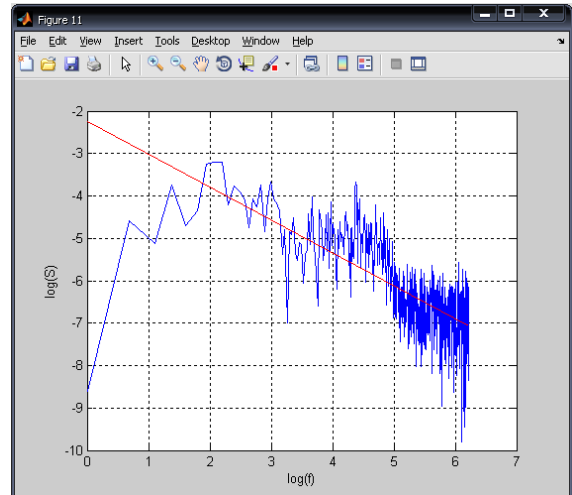


Рисунок 12 – Спектральная плотность сигнала s3

### Показатель Херста

В этом разделе оценивается показатель Херста  $H$ , который является мерой устойчивости статистического явления или мерой длительности долгосрочной зависимости процесса. Значение  $H=0,5$  указывает на отсутствие долгосрочной зависимости. Корреляция между событиями отсутствует. Ряд является случайным, а не фрактальным. Чем ближе значение  $H$  к 1, тем выше степень устойчивости долгосрочной зависимости. При  $0 \leq H < 0,5$  временной ряд является трендоустойчивым (антиперсистентным). Он более изменчив, чем случайный ряд, поскольку состоит из частых реверсов спад-подъем. При  $0,5 < H \leq 1$  ряд трендоустойчив. Тенденция его изменения может быть спрогнозирована. Показатель Херста связан с другими фрактальными параметрами формулами, показанными в таблице 1.

Таблица 1. Фрактальные параметры

Параметр	Расчетная формула
Фрактальная размерность $D$	$D = 2 - H$
Корреляционный параметр $\beta$	$\beta = 2(1 - H)$
Спектральный показатель $b$	$b = 2H + 1$
Фрактальный показатель $\alpha$	$\alpha = 3 - 2H$

Для определения значения  $H$  в работе используется алгоритм R/S-анализа, который состоит в следующем. В исходной реализации  $Z$  последовательно выделяем отрезки  $Z_\tau = z_1, z_2, \dots, z_\tau$ ,  $\tau = 1000, 2000, \dots, 12000$ , для

каждого из которых вычисляем текущее среднее

$$\bar{z}_\tau = \frac{\sum_{i=1}^{\tau} z_i}{\tau}. \text{ Далее для каждого фиксированного } Z_\tau \text{ вычисляем накопленное отклонение для его}$$

$$\text{отрезков длины } t: X_{\tau,t} = \sum_{i=1}^t (z_i - \bar{z}_\tau), t = 1, 2, \dots, \tau,$$

разность (размах) между максимальным и минимальным накопленными отклонениями  $R = R(\tau) = \max_{1 \leq t \leq \tau} (X_{\tau,t}) - \min_{1 \leq t \leq \tau} (X_{\tau,t})$  и стандартное отклонение  $S = S(\tau)$ . Показатель Херста

$$\text{вычисляем по формуле } \frac{R}{S} = \left(\frac{\tau}{2}\right)^H.$$

Логарифмируя обе части этого равенства, получаем последовательность координат точек  $(x_\tau, y_\tau)$ , где  $x_\tau = 1000, 2000, \dots, 12000$ ,

$$y_\tau = H(\tau) = \frac{\lg(R(\tau)/S(\tau))}{\lg(\tau/2)}.$$

Построив график зависимости  $H(\tau)$  и линию тренда, определим аппроксимированное значение показателя Херста  $H$ .

11. На рисунках 13 – 15 показан результат R/S-анализа изучаемых сигналов. Эмпирические значения показателя Херста отклоняются от теоретического значения ( $H=0,5$ ) случайного сигнала. Показатель Херста сигналов показан в таблице 2.

Таблица 2. Показатель Херста

Сигнал	Показатель Херста
s1	0,3745
s2	0,7139
s3	0,6882

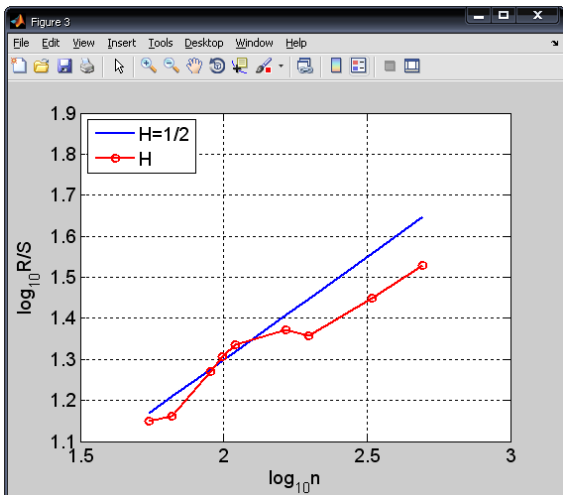


Рисунок 13 – Результат R/S-анализа сигнала s1

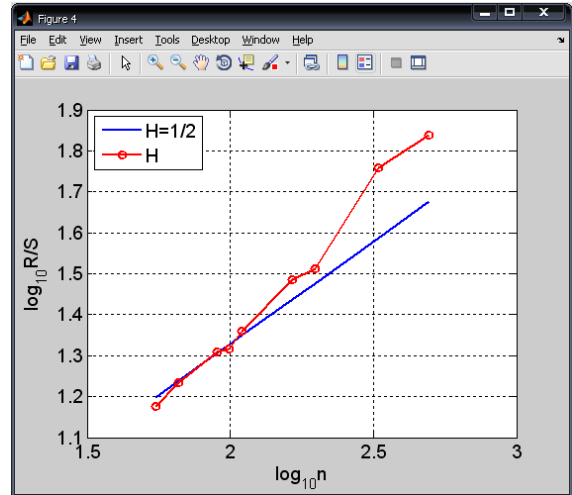


Рисунок 14 – Результат R/S-анализа сигнала s2

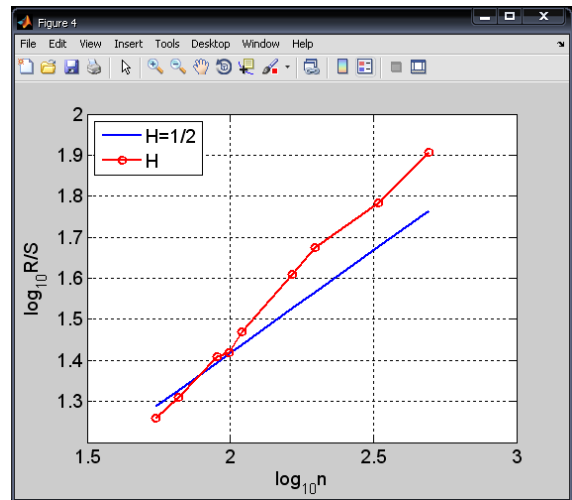


Рисунок 15 – Результат R/S-анализа сигнала s3

## Выводы

12. Анализ речевых сигналов является актуальной задачей поиска объективного показателя эмоционального состояния человека, выполняющего ответственную деятельность.

13. В данной работе для сигналов эмоционально окрашенной речи выполнен фрактальный анализ плотности распределения, автокорреляционных функций и энергетических спектров. Найдены значения показателя Херста. Получены следующие результаты:

14. Изучаемые сигналы не подчиняются нормальному распределению.
15. Изучаемые сигналы не обладают экспоненциально спадающей АКФ, свойственной случайным сигналам.
16. Для сигнала s1 выполняется условие  $0 < H < 0,5$ . Этот ряд является антиперсистентным.
17. Для сигналов s2, s3 выполняется условие  $0,5 < H < 1$ . Эти сигналы являются трендоустойчивыми (персистентными) и обладают долговременной памятью.

## Литература

1. Розалиев, В.Л. Моделирование эмоциональных реакций пользователя при речевом взаимодействии с автоматизированной системой: дис. ...канд. тех. наук / В.Л. Розалиев. Волгоград: ВГТУ, 2009.
2. Хроматиди, А.Ф. Исследование психофизиологического состояния человека на основе эмоциональных признаков речи: дис. ...канд. тех. наук / А.Ф. Хроматиди. Таганрог, 2005.
3. Калюжный, М.В. Система реабилитации слабобудящих на основе настраиваемой сегментарной модели синтезируемой речи: дис. ...канд. тех. наук / М.В. Калюжный. СПб., 2009.
4. Соловьева, Е.С. Методы и алгоритмы обработки, анализа речевого сигнала для решения задач голосовой биометрии: дис. ...канд. тех. наук / Е.С. Соловьева. М., 2008.
5. Chen, Y.T. A study of emotion recognition on mandarin speech and its performance evaluation: Ph. D. dissertation / Y.T. Chen. Tatung, 2008.
6. Morist, M.U. Emotional speech synthesis for a radio dj: corpus design and expression modeling: master thesis MTG-UPF dissertation / M.U. Morist. Barcelona, 2010.
7. Siging, W. Recognition of human emotion in speech using modulation spectral features and support vector machines: master of science dissertation / W. Siging. Kingston, 2009.
8. Сидоров, К.В. К вопросу оценки эмоциональности естественной и синтезированной речи по объективным признакам / К.В. Сидоров, М.В. Калюжный // Вестник Тверского государственного технического университета. Вып. 18. Тверь, 2011. С. 81–85.
9. Н.Н. Лебедева, Р.А. Марагей, А.Н. Котровская. Фрактальный анализ речевого сигнала при различении эмоционального состояния в модельных экспериментах. [//http://www.radiotec.ru/article/7266](http://www.radiotec.ru/article/7266)
10. Кузьмичев С.А. Опыт сравнения речевых и стандартных аудиосигналов в параметрах фрактального анализа // Концепт. – 2015. – Спецвыпуск № 28. [//http://e-koncept.ru/2015/75376.htm](http://e-koncept.ru/2015/75376.htm)

*Бельков Д.В., Едемская Е.Н. Фрактальный анализ речевых сигналов. Анализ речевых сигналов является актуальной задачей поиска объективного показателя эмоционального состояния человека, выполняющего ответственную деятельность. В данной работе для речевых сигналов выполнен фрактальный анализ плотности распределения, автокорреляционных функций и энергетических спектров. Найдены значения показателя Херста. Он может быть показателем эмоционального состояния человека, поскольку степень фрактальности речевых сигналов зависит от эмоционального состояния и гендерных различий.*

**Ключевые слова:** речевой сигнал, плотность распределения, автокорреляционная функция, энергетический спектр, показатель Херста.

*Belkov D.V., Edemskaya E.N. Fractal analysis of speech signals. Analysis of speech signals is actual task of the search of objective index for emocial state of the man. In this work the analysis of closeness of distributing, autocorrelation functions and power spectrums are executed. The values of Hurst index are found. It can be the index for emocial state of the man because the fractal value of speech signals is depending of emocial state and gender difference.*

**Keywords:** speech signal, closeness of distributing, autocorrelation function, power spectrum, the Hurst index.

Статья поступила в редакцию 20.03.2018

Рекомендована к публикации д-ром техн. наук В.Н. Павлышом