

УДК 519.95

ВЕЙВЛЕТ АНАЛИЗ РЕЧИ

© А.П. Зубаков

Ключевые слова: вейвлет-преобразования; речь; распознавание речи.

В настоящей работе предлагается использовать аппарат непрерывных вейвлет-преобразований для анализа речи.

В настоящее время усилия ученых и коллективов исследователей направлены на решение проблемы распознавания речи. Фирмы Microsoft и IBM выделяют значительные средства для реализации командных функций по управлению программным обеспечением на естественном языке. Однако практическая реализация диалога с компьютером оказалась весьма сложной задачей. В нашей стране известны программные продукты, реализующие процедуру распознавания вводимой с микрофона речи, такие как «Горыныч», «Комбат», «Говорящая мышь» и др. Однако реальное качество работы этих систем распознавания не выдерживает критики. При этом в реальной жизни мы постоянно сталкиваемся с практической реализацией процесса распознавания в ходе общения живых индивидуумов.

В настоящей работе предлагается использовать аппарат непрерывных вейвлет-преобразований для анализа речи.

В качестве объекта исследования была использована фраза «ну-ка съешь этих следующих вкусных французских булочек да выпей чайку», произносимая через микрофон и оцифрованная в виде дискретных отсчетов в файлы, соответствующие 19 дикторам. Выбор текста был обусловлен большим фонетическим разнообразием (рис.1).

Для преобразования дискретных отсчетов в непрерывную функцию была произведена их квазилинейная и квадратичная интерполяция, в результате чего были получены две функции, представленные на рис. 2.

Как видно на рис. 2, обе функции достаточно близки друг к другу. Поэтому в дальнейшем использовалась квадратичная интерполирующая функция.

Непрерывный вейвлет-анализ данной функции для вейвлета типа «мексиканская шляпа» с 12×12 коэффициентами позволил получить спектрограмму, представленную на рис. 3.

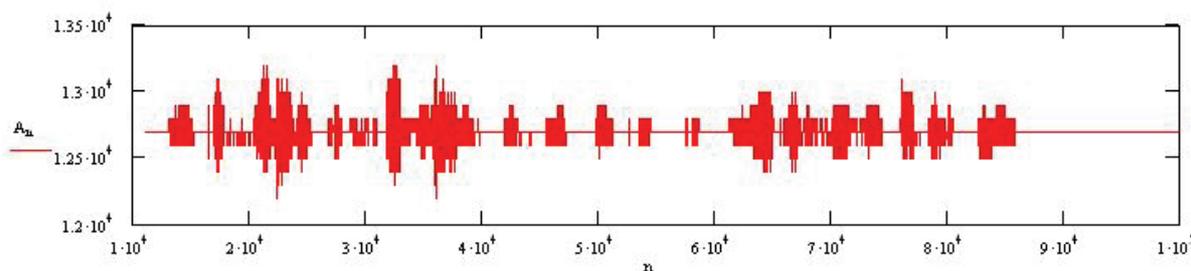


Рис. 1. Оцифрованная реализация тестовой фразы

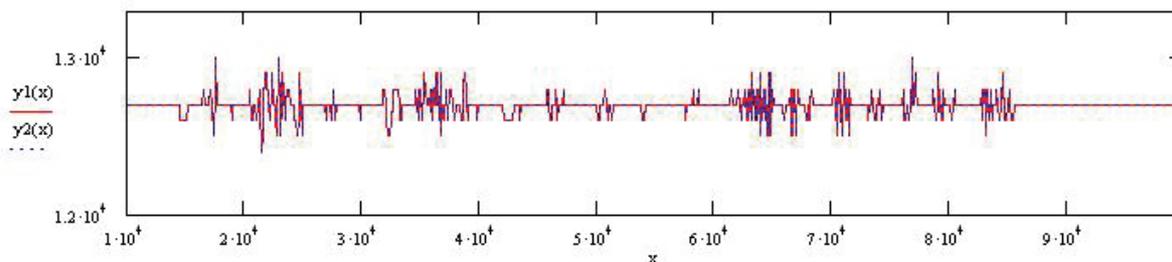


Рис. 2. Квазилинейная - $y_1(x)$ и квадратичная - $y_2(x)$ интерполяция

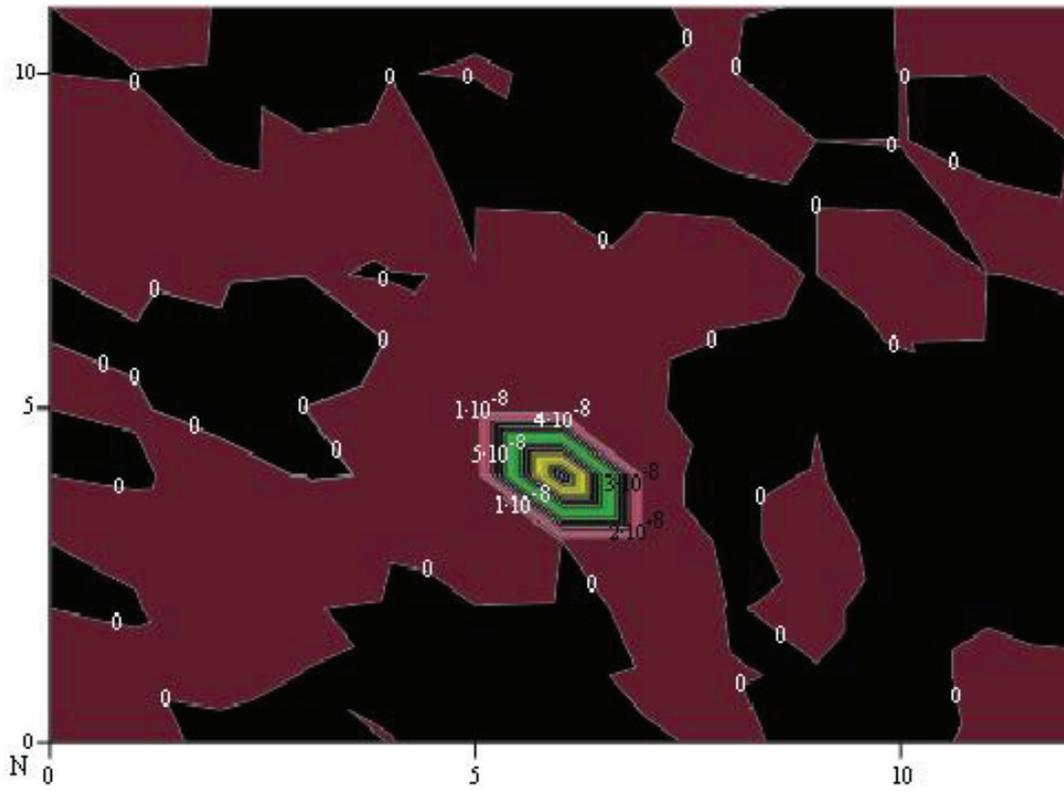


Рис. 3. Спектрограмма с 12×12 вейвлет-коэффициентами

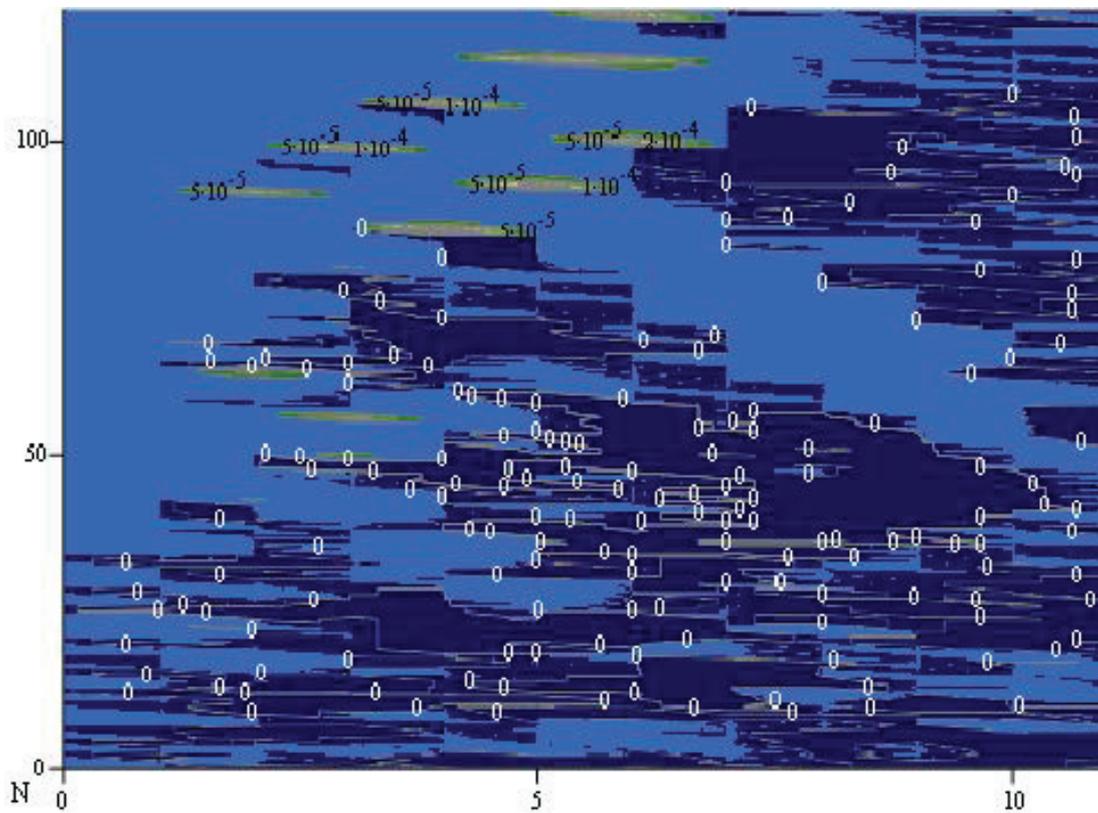


Рис. 4. Спектрограмма с 11×122 вейвлет-коэффициентами

```

N := 10000
inf := 1000
i := sqrt(-1)
Bn-11000 := An
Xn-11000 := n
mhat(t) :=  $\frac{d^2}{dt^2} \exp\left(\frac{-t^2}{2}\right) \rightarrow$ 
 $\rightarrow -\exp\left(\frac{-1}{2} \cdot t^2\right) + t^2 \cdot \exp\left(\frac{-1}{2} \cdot t^2\right)$ 
Nj,b := C(j,b)
b := 0..11
j := 1..12
C(a,b) :=  $\int_{-\infty}^{\infty} \Phi(a,b,t) \cdot y_2(t) dt$ 

```

$$\Phi(a,b,t) := \frac{1}{\sqrt{a}} \cdot \left(\text{mhat}\left(\frac{t-b}{a}\right) \right)$$

Анализ с 11×122 коэффициентами вейвлет-преобразования дал более развернутую картину, представленную на рис. 4.

В дальнейшем представляется необходимым выполнить обратное вейвлет-преобразование и формирование аудио-файлов, чтобы в интерактивном режиме контролировать процесс анализа.

Поступила в редакцию 12 ноября 2010 г.

Zubakov A.P. Wavelet speech analysis

The work offers to use the apparatus of continuous wavelet-transformations for speech analysis.

Key words: wavelet-transformations; speech; speech identification.