

Анализ звукового сигнала с помощью вейвлета Морле

Т.Д. Шлякова¹

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

Аннотация. В данной работе выполнено сравнительное исследование и анализ особенностей звукового сигнала с помощью вейвлет-преобразования и Фурье преобразования, сравнение информативности признаков. Проведено сравнение спектральных последовательностей, полученных в результате обоих преобразований, по информативности признаков с точки зрения кластеризации двух типов голосов: женского и мужского.

1. Введение

Спектральный анализ звуковых сигналов является довольно популярным направлением в век развития информационных технологий, и используется в нейронных сетях для задачи распознавания речи в акустике, электрических сигналов в радиотехнике, кардиограмм в медицине и т.д.

Хотя спектральный анализ включает в себя достаточное количество различных методов, наиболее популярным из них является преобразование Фурье. Но в последнее время также активно применяется вейвлет-преобразование как мощный инструмент частотно-временного анализа и кодирования сигналов, предназначенный для анализа сложных нестационарных сигналов [1]. Данное преобразование не только нашло применение во многих областях, но служит в некоторой степени альтернативой преобразованию Фурье в обработке изображений, анализе статистических данных, обработке физических данных и в других приложениях. В качестве примера, вейвлет-разложение может использоваться для фильтрации сигнала, чтобы устранить нежелательные частотные составляющие в звуковом сигнале [2]. Одно из основных преимуществ вейвлет-анализа заключается в том, что он позволяет заметить хорошо локализованные изменения сигнала, тогда как анализ Фурье этого не дает – в коэффициентах Фурье отражается поведение сигнала за все время его существования [3].

В данной работе рассматривается преобразование Фурье, непрерывное вейвлет-преобразование и базисные вейвлет-функции, в частности вейвлет Морле и мексиканская шляпа. Эти преобразования применяются к аудиосигналам для формирования спектральных последовательностей. Основное внимание уделяется сравнительному анализу результатов преобразования Фурье и вейвлет-преобразования с точки зрения возможности кластеризации двух типов голосов (женского и мужского) по сформированным признакам.

2. Теоретические основы

Пусть есть пространство функций, заданных на конечном интервале $(0, N)$. Возьмем произвольную функцию $f(t)$. Тогда норма в пространстве определяется как:

$$\|f(t)\|^2 = \langle f(t), f(t) \rangle = \int_0^N f(t)\bar{f}(t)dt, \quad (1)$$

где $\bar{f}(t)$ – комплексно-сопряженная функция с $f(t)$.

Если данная норма конечна, то значит функция принадлежит пространству квадратично интегрируемых функций $L^2(0, N)$, и её энергия конечна.

Базис пространства может быть образован любой ортогональной системой функций [4,5]. Прямое преобразование Фурье является проекцией сигнала на систему комплексной экспоненциальной функцией:

$$F(u) = \frac{1}{T} \int_0^T f(t)e^{-2\pi i u t} dt = FT[f(t)]; \quad u \in (-\infty, +\infty). \quad (2)$$

Известны также, так называемые, оптимальные базисы. Среди них базис Карунена-Лоэва [6], которое обеспечивает некоррелированные коэффициенты [7] разложения для случайных сигналов [8, 9]. Кроме того, при передаче информации в ограниченных оптических системах рассматривают коммуникационные моды [10-12]. Заметим, однако, что оптимальные базисы не имеют аналитического представления [13], поэтому для их расчета требуется решать задачи на собственные значения [14-17]. Причем при изменении характеристик системы требуется новый пересчет оптимального базиса. Некоторым компромиссом между детерминированным базисом и оптимальным является вейвлет-разложение.

“Wavelet” переводится с английского как “всплеск”. Вейвлет – преобразование сигнала это представление сигнала в виде интеграла Фурье или обобщенной суммы с использованием базисной функции. Базисная функция как раз называется вейвлетом и представлена в виде:

$$\psi_{a,b}(t) = \psi\left(\frac{t-b}{a}\right), \quad (3)$$

где t – вектор времени, $\psi(t)$ – функция вейвлета (ядро преобразования), – переменная, характеризующая изменения масштаба по времени, – переменная, характеризующая сдвиг по времени. Основное отличие вейвлет-разложения от разложения по детерминированному базису в том, что имеется возможность подбирать функцию вейвлета (2), т.е. оптимизировать разложение в соответствии с задачей.

Как видно из формулы, переменная обеспечивает локализацию по частоте (с увеличением значения происходит переход от высоких частот к низким), в то время как обеспечивает локализацию по времени. Эти функции являются компактными как во временной, так и в частотной области. Вейвлеты непосредственно связаны с кратнo-масштабным анализом сигналов [18].

В качестве исходных вейвлет - функций $\psi(t)$ могут использоваться любые функции, удовлетворяющие свойствам:

1. Ограниченность:

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty.$$

2. Локализация: процесс сужения и расширения вейвлета по временной оси должен линейно изменять его частоту и ширину спектра.

3. Нулевое среднее значение:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0.$$

4. Самоподобие: вейвлет $\psi_{a,b}(t)$ должен иметь то же число осцилляций, что и его исходный $\psi(t)$.

Наиболее популярными семействами вейвлетов являются те, которые конструируются на основе производных функций Гаусса:

$$g(t) = e^{-\frac{t^2}{2\sigma^2}}, \quad (4)$$

где t – вектор времени, σ – переменная, характеризующая ширину спектра. Это обусловлено тем обстоятельством, что функция Гаусса имеет наилучшие показатели локализации как во временной, так и в частотной областях [7].

Из них чаще всего на практике используется вейвлет Морле. Запишем его в виде:

$$\psi(t) = \cos(\mu t) e^{-\frac{t^2}{2\sigma^2}}, \quad (5)$$

где t – вектор времени, σ – переменная, характеризующая ширину спектра, μ – переменная, характеризующая частоту.

По формуле (1.2) с помощью непрерывных a и b из исходного вейвлета $\psi(t)$ построим базис $\psi_{a,b}(t)$. Тогда по определению запишем прямое непрерывное вейвлет-преобразование от звукового сигнала $y(t)$:

$$W(a, b) = \int_{-\infty}^{\infty} y(t) \psi_{a,b}(t) dt. \quad (6)$$

Из формулы (4) видно одно из главных отличий вейвлет-преобразования от преобразования Фурье, которое обеспечивает лучшую локализацию – вейвлет-спектр $W(a, b)$ зависит от двух аргументов, один из которых показывает зависимость от частоты, а второй от времени.

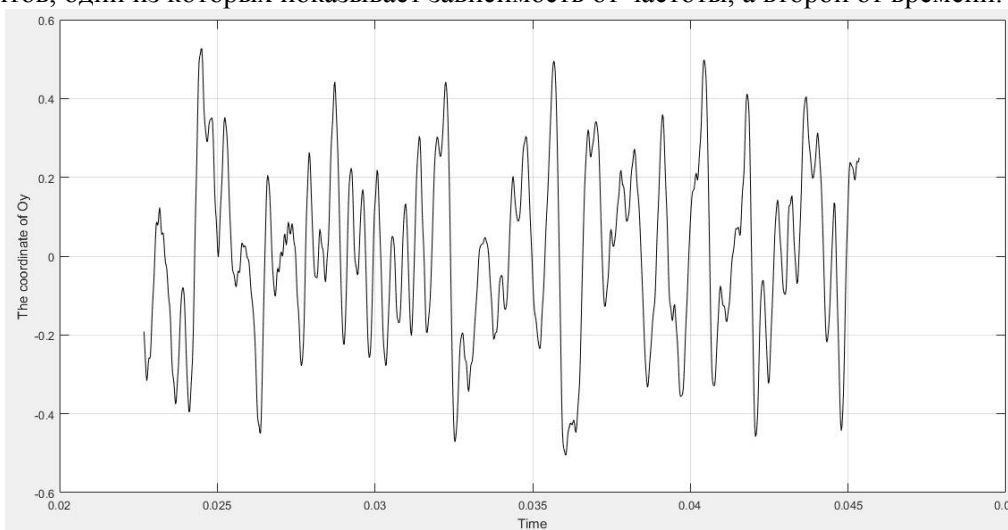


Рисунок 1. График боя курантов.

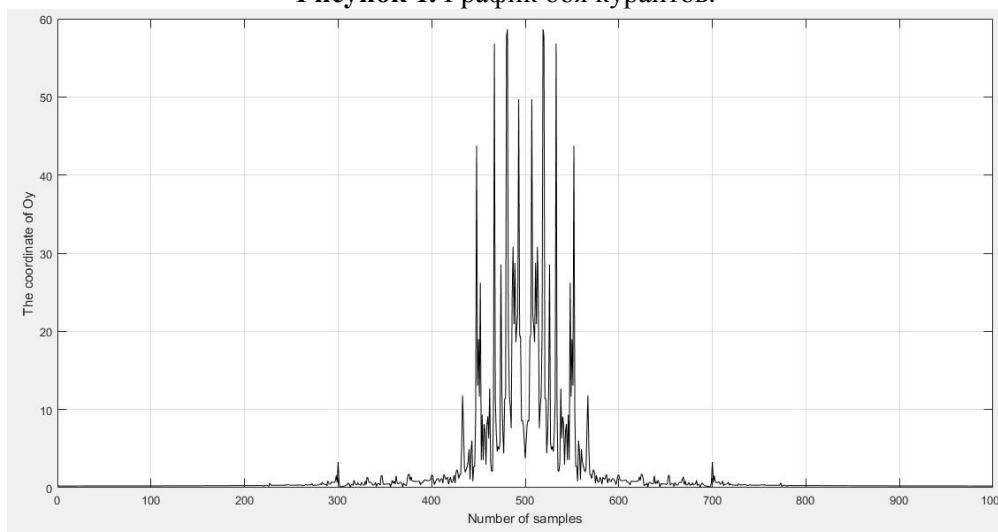


Рисунок 2. График Фурье-спектра боя курантов.

Построим графики боя курантов и всех его преобразований. Графики показаны на рисунках 1, 2, 4 и 5. Так в качестве входного сигнала подавалась действительная функция, график Фурье преобразование можно обрезать. Приближенный график Фурье-спектра показан на рисунке 3.

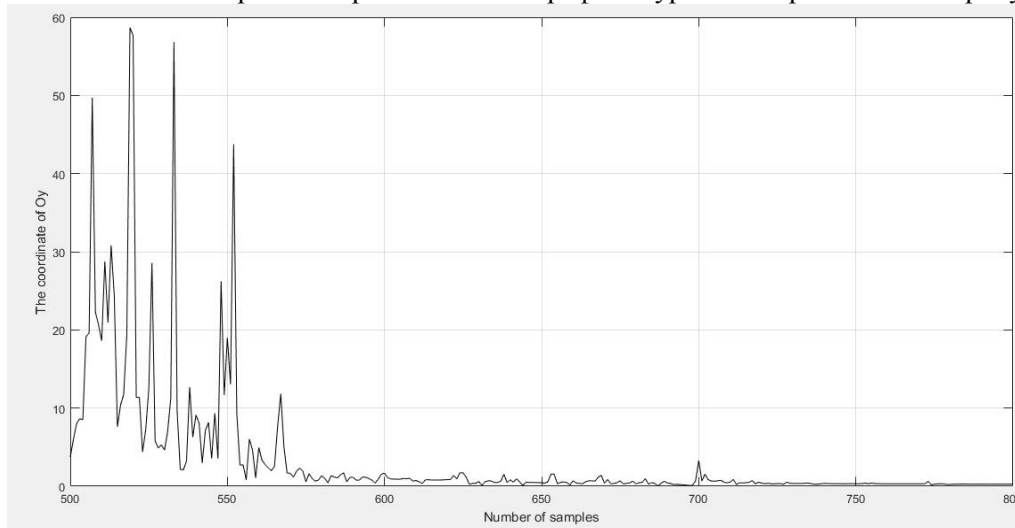


Рисунок 3. Увеличенный график Фурье-спектра боя курантов.

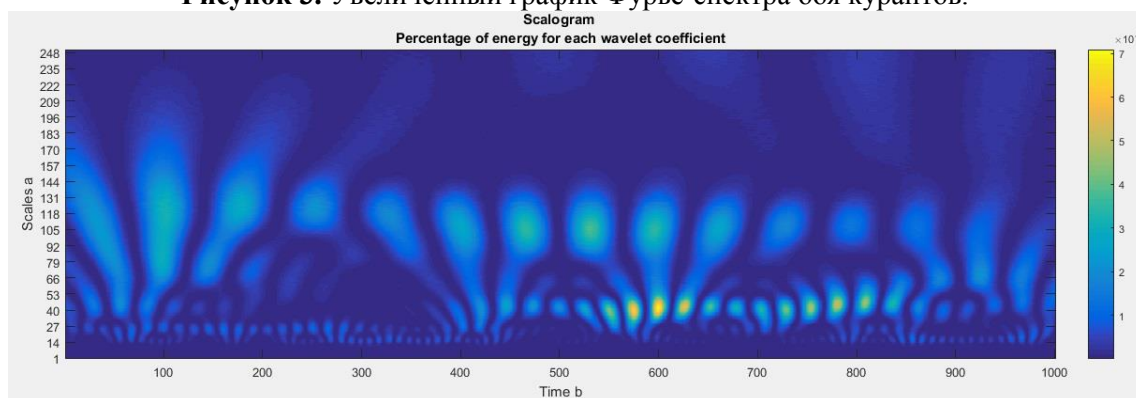


Рисунок 4. График вейвлет-спектра боя курантов, вычисленный с помощью вейвлета Морле.

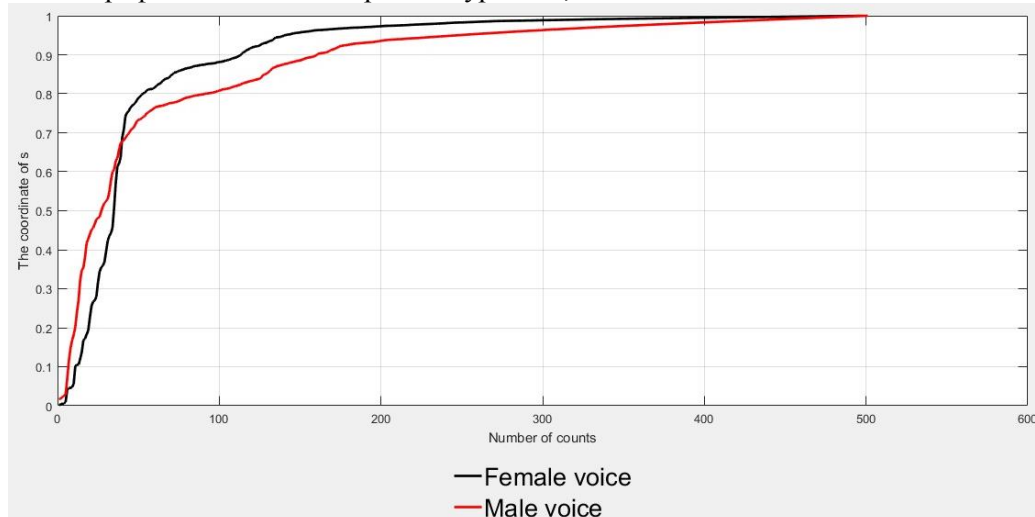


Рисунок 5. Графики зависимости параметра s от количества отсчетов для женского и мужского голосов.

Вейвлет-спектр показывает как частоты изменяются со временем в отличие от Фурье спектра, который просто показывает какие частоты присутствуют в сигнале.

Также было произведено усреднение Фурье спектров пятидесяти женских и пятидесяти мужских голосов. Для дополнительной обработки вводился дополнительный параметр s , вычисляемый как:

$$s = \frac{\sum_{p=0}^n F_p}{\sum_{p=0}^N F_p}, \quad (7)$$

где n – это количество отсчетов, N – это размер спектра, F_p – это значение Фурье спектра в заданной точке. График зависимости данного параметра от количества отсчетов представлен на рисунке 5.

Также на основе усредненных спектров был построен простой классификатор голосов, определяющий кому принадлежит голос: женщине или мужчине. Он тестировался на пятнадцати наборах голосов не входящих в усреднение. и показал правильное детектирование в 80% случаев.

3. Заключение

В данной работе был выполнен сравнительный спектральный анализ для обработки звуковых сигналов с использованием преобразования Фурье и непрерывного вейвлет-преобразования. Сравнительное исследование показало, что преобразования Фурье не дает излишней информации о звуковых сигналах в отличие от вейвлет-преобразования. Это важно при выполнении кластеризации с целью дальнейшего распознавания. Спектральные Фурье-последовательности, рассчитанные по выборке из пятидесяти представителей, использовались для распознавания двух типов голосов: женского и мужского. С помощью усредненных спектров построен простой классификатор тембра человеческого голоса, показавший успешное распознавание на тестовой выборке.

4. Литература

- [1] Смоленцев, Н.К. Основы теории вейвлетов. Вейвлеты в MATLAB / Н.К. Смоленцев – М.: ДМК Пресс, 2009. – 448 с.
- [2] Яковлев, А.Н. Введение в вейвлет-преобразования: Учеб. Пособие / А.Н. Яковлев – Новосибирск: НГТУ, 2003. – 104 с.
- [3] Астафьева, Н.М. Вейвлет-анализ: основы теории и примеры применения // Успехи физических наук. – 1996. – № 11. – С. 1145-1170.
- [4] Давыдов, А.В. Основы вейвлет-преобразования сигналов / А.В. Давыдов. – 2010. – 18 с.
- [5] Агафонова, Н.А. Математические основы частотных методов теории и практики автоматического управления: Учеб. Пособие / Н.А. Агафонова, Ю.С. Тверской – Иваново: ГОУВПО «Ивановский государственный энергетический университет имени В.И. Ленина», 2006. – 280 с.
- [6] Wang, L. Karhunen-Loeve Expansions and their Application – Ann Arbor: ProQuest, 2008. – 292 p.
- [7] Soifer, V.A. Decorrelated features of images extracted with the aid of optical Karhunen-Loeve expansion / V.A. Soifer, M.A. Golub, S.N. Khonina // Pattern Recognition and Image Analysis. – 1993. – Vol. 3(3). – P. 289-295.
- [8] Fukunaga, K. Representation of Random Processes Using the Finite Karhunen-Loeve Expansion / K. Furunaga, W.L.G. Koontz // Information and Control. – 1970. – Vol. 16(1). – P. 85-101. DOI: 10.1016/S0019-9958(70)80043-2.
- [9] Huang, S.P. Convergence study of the truncated Karhunen-Loeve expansion for simulation of stochastic processes / S.P. Huang, S.T. Quek, K.K. Phoon // International Journal for Numerical Methods in Engineering. – 2001. – Vol. 52. – P. 1029-1043. DOI: 10.1002/nme.255.
- [10] Miller, D.A.B. Communicating with waves between volumes: evaluating orthogonal spatial channels and limits on coupling strengths // Appl. Opt. – 2000. – Vol. 39. – P. 1681-1699.
- [11] Броварова, М.А. Повышение разрешающей способности с помощью вытянутых сфероидальных волновых функций / М.А. Броварова, С.Н. Хонина // Компьютерная оптика. – 2001. – Т. 21. – С. 53-57.

- [12] Kirilenko, M.S. Coding of an optical signal by a superposition of spheroidal functions for undistorted transmission of information in the lens system / M.S. Kirilenko, S.N. Khonina // Proc. SPIE. – 2014. – Vol. 9156. – P. 91560J-9. DOI: 10.1117/12.2054214.
- [13] Хонина, С.Н. Приближение сфероидальных волновых функций конечными рядами // Компьютерная оптика. – 1999. – Т. 19. – С. 65-70.
- [14] Хонина, С.Н. Метод вычисления собственных значений вытянутых сфероидальных функций нулевого порядка / С.Н. Хонина, С.Г. Волотовский, В.А. Сойфер // Доклады Академии наук. – 2001. – Т. 376, № 1. – С. 30-33.
- [15] Кириленко, М.С. Расчёт собственных функций изображающей двухлинзовой системы в условиях осевой симметрии / М.С. Кириленко, С.Н. Хонина // Компьютерная оптика. – 2014. – Т. 38, № 3. – С. 412-417.
- [16] Кириленко, М.С. Вычисление собственных функций ограниченного дробного преобразования Фурье / М.С. Кириленко, Р.О. Зубцов, С.Н. Хонина // Компьютерная оптика. – 2015. – Т. 39, № 3. – С. 332-338. DOI: 10.18287/0134-2452-2015-39-3-332-338.
- [17] Хонина, С.Н. Формирование заданных распределений на основе разложения по вихревым собственным функциям ограниченного непараксиального оператора распространения / С.Н. Хонина, С.Г. Волотовский, М.С. Кириленко // Компьютерная оптика. – 2019. – Т. 43, № 2. – С. 184-192. DOI: 10.18287/2412-6179-2019-43-2-184-192.
- [18] Воробьев, В.И. Теория и практика вейвлет-преобразование / В.И. Воробьев, В.Г. Грибунин – С.-Петербург: ВУС, 1999. – 204 с.
- [19] Sun, S. Automatic moment segmentation and peak detection analysis of heart sound pattern via short-time modified Hilbert transform // Computer methods and programs in biomedicine. – 2014. – Vol. 114. – P. 219-230.

Sound signal analysis using Morlet wavelet

T.D. Shlyakova¹

¹Samara National Research University, Moskovskoye shosse 34A, Samara, Russia, 443086

Abstract. In this work, a comparative study and analysis of the features of an audio signal using a wavelet transform and Fourier transform, a comparison of the information content of the features. The spectral sequences obtained as a result of both transformations are compared according to the information content of the signs from the point of view of clustering of two types of voices: female and male.