

Рисунок 3 – Спектрограммы оригинального и обработанного сигналов

Список использованных источников:

1. Simpson A. Frequency-Lowering Devices for Managing High-Frequency Hearing. *Ear Hearing*, 2009, vol. 13, no. 2, pp. 87-106.
2. Paarmann L. Nonlinear Spectrum Compression for the Hearing Impaired via a Frequency-Domain Processing Algorithm. *28th IEEE EMBS Annual International Conference*, 2006, vol.10, pp. 239-243.
3. Wang Q., Zhao L., Zou C. Piecewise-Linear Frequency Shifting Algorithm for Frequency Resolution Enhancement in Digital Hearing Aids. *MDPI Conference*, 2017, vol.1, pp. 1-13.

ВЫЯВЛЕНИЕ ПАТОЛОГИЙ В ГОЛОСЕ НА ОСНОВЕ РАЗЛОЖЕНИЯ СИГНАЛА НА ЭМПИРИЧЕСКИЕ МОДЫ

Семенюк Н.А.

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Вашкевич М.И. – к.т.н., доцент

В работе рассмотрена задача анализа голосового сигнала на основе декомпозиции на эмпирические моды с целью выявления в нем признаков патологии. В качестве характерных признаков, которые могут быть использованы для детектирования патологии, были выявлены – число мод, на которые раскладывается сигнал и – отношение сигнал/шум, рассчитанное на основе того, что часть полученных мод относились в шумовой составляющей голоса.

Декомпозиция сигнала на эмпирические моды (ДЭМ) – это способ анализа сигналов различной природы, который становится все более распространенным в экспериментальных исследованиях [1-2]. Уже первые попытки его применения позволили установить, что с помощью декомпозиции на эмпирические моды можно обнаружить модуляцию с близкими значениями частот моделирующего и моделируемого процессов, уменьшать ширину спектральных линий и т.п. ДЭМ позволяет учитывать локальные особенности сигнала (экстремумы и нули сигнала), а также шумовые, периодические и трендовые его компоненты. Было показано, что ДЭМ подходит для кодирования речевого сигнала, определения мгновенных параметров сигнала, редактирования акустического шума [3] и проч.

В процессе анализа данных полученных экспериментальным путем одним из этапов предобработки часто является приведение сигнала к нулевому среднему уровню, что позволяет избавиться от проблем, связанных с некорректным определением мгновенной частоты. Однако данная процедура не всегда может быть корректно выполнена для нестационарных процессов, поскольку их среднее значение может меняться и принимать отличные от нуля величины на локальных участках. Для корректного определения мгновенной частоты компонент, на которые раскладывается сигнал, в рамках метода ДЭМ предлагается обеспечить выполнение следующих условий:

- 1) локальное среднее значение огибающих, определяемых локальными максимумами функции c_j и локальными минимумами этой функции должно быть равно нулю;
- 2) число пересечений нулевого уровня функцией и число локальных максимумов (или минимумов) должно совпадать или отличаться не более чем на единицу.

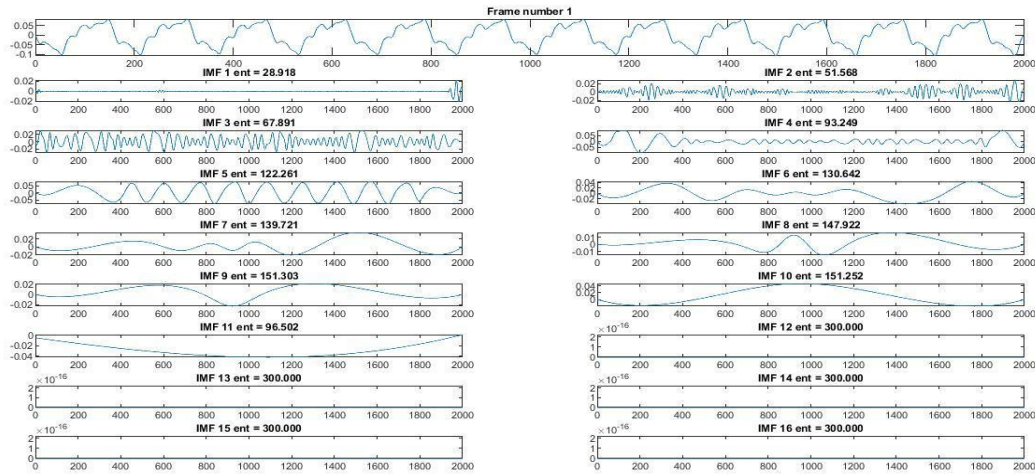


Рисунок 1 – Пример декомпозиции сгенерированного сигнала на эмпирические моды

Сам процесс отсеивания состоит из нескольких этапов (Рисунок 1):

- 1) Определение локальных экстремумов (максимумов и минимумов) сигнала $h_{j,i-1}(t)$;
- 2) Определение верхней $U_{j,i-1}(t)$ и нижней $L_{j,i-1}(t)$ огибающих сигнала с помощью кубического сплайна-интерполяции по найденным локальным экстремумам $h_{j,i-1}(t)$;
- 3) Вычисление среднего значения огибающих сигнала:

$$\mu_{j,i-1}(t) = \frac{U_{j,i-1}(t) + L_{j,i-1}(t)}{2}$$

- 4) Вычисление остатка от сигнала:
- 5) Вычисление критерия остановки:

$$h_{j,i}(t) = h_{j,i-1}(t) - \mu_{j,i-1}(t), \quad i = i + 1$$

$$SD(i) = \sum \frac{[h_{j,i-1}(t) - h_{j,i}(t)]^2}{h_{j,i-1}^2(t)}$$

- 6) Повторение алгоритма до тех пор, пока $SD \geq 0,25$. Когда SD станет меньше, чем 0,25, $h_{j,i-1}(t)$ станет функцией внутренней моды.

Недостатки метода ДЭМ:

- 1) часто происходит смешивание нескольких мод, которые интерпретируются как одна;
- 2) получение ЭМ, состоящих из участков сигнала несоизмеримых масштабов или участков соизмеримых масштабов, но находящихся в различных областях ЭМ.

Существует такие заболевания гортани, первое проявление которых отражается на изменениях в голосе. Раскладывая голос на компоненты, было обнаружено, что показатель сигнал/шум (SNR показатель) и число мод, получаемых в процессе разложения, для здоровых голосов и голосов с патологиями отличаются.

Применив ДЭМ для анализа речевых сигналов, соответствующих устойчивой гласной «а» реальных данных (20 здоровых голосов и 20 голосов с патологиями), получили следующие диаграммы размаха (рис. 2 и рис. 3). Запись голосов производилась в РНПЦ Отоларингологии (г. Минск) под контролем врача фониагра.

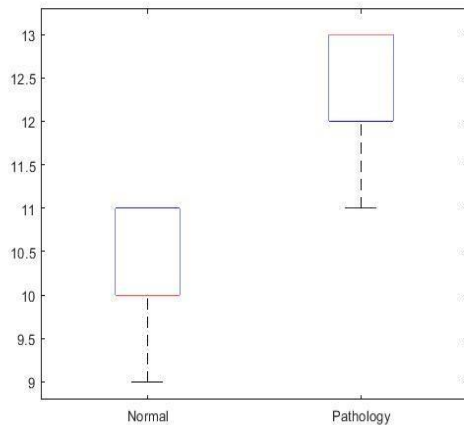


Рисунок 2 – Число мод для здорового голоса и голоса с патологией

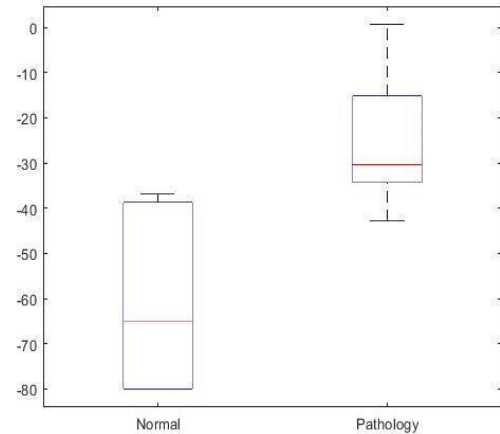


Рисунок 3 – Отношение сигнал/шум (SNR) для здорового голоса и голоса с патологией

По графикам видно, что значения этих двух показателей явно отличаются, что говорит о возможности использования этого метода в области патологий ЛОР-органов, в частности заболеваний гортани, и необходимости дальнейшего развития этого метода.

Список использованных источников:

1. A.-O. Boudraa, J.-Ch. Cexus, "EMD-Based Signal Filtering", IEEE Transactions on instrumentation and measurement, vol. 56, no. 6, pp. 2196-2202, 2007.
2. J. V. Sloten, P. Verdonck, M. Nyssen, J. Haueisen (Eds.): ECIFMBE 2008, IFMBE Proceedings 22, pp. 252–255, 2008.
3. И.А. Воронецкий, Анализ сигналов методом декомпозиции на эмпирические моды и его применение в обработке речевых сигналов / Доклады БГУИР. – 2012. – №7(69). – С. 18 – 24.
4. Juan Rafael Orozco-Arroyave, Elkyn Alexander Belalcazar-Bolaños, Julián David Arias-Londoño, Jesús Francisco Vargas-Bonilla, Sabine Skodda, Jan Ruzs, Khaled Daqrouq, Florian Hönig, and Elmar N'oth, Characterization Methods for the Detection of Multiple Voice Disorders: Neurological, Functional, and Laryngeal Diseases, IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 19, NO. 6, 2015
5. Применение комплементарной множественной декомпозиции на эмпирические моды для анализа речевых сигналов / А. К. Алимуратов, Ю. С. Квитка // Измерение. Мониторинг. Управление. Контроль. – 2014. – № 4 (10). – С. 69–75.

ГЛУБОКАЯ НЕЙРОННАЯ СЕТЬ ДЛЯ ДЕТЕКТИРОВАНИЯ ЛИЦ В УСЛОВИЯХ ОГРАНИЧЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ

Суша А.В.

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Вашкевич М.И. – к.т.н., доцент

В работе приводятся приёмы проектирования глубоких нейронных сетей, показывающих высокую производительность на встраиваемых и мобильных платформах, на примере решения задачи детектирования лиц. Детектор лиц основан на концепции You Only Look Once, которая заключается в том, что изображение анализируется только один раз при прямом проходе нейронной сети. Приведенные методики проектирования позволили существенно повысить скорость детектирования без серьезной потери в точности и полноте.

Применение глубоких нейронных сетей для детектирования лиц на изображении позволяет достичь высокого уровня точности и инвариантности к множеству факторов, затрудняющих детектирование: разный масштаб лиц, частичное перекрытие лиц другими объектами, поза людей в момент захвата изображения, выражения лиц, засветка и др. – всё это существенно понижает точность детектирования классическими методами [1]. Несмотря на то, что в настоящее время детекторы на глубоких нейронных сетях показывают высокую скорость детектирования на персональных компьютерах, оснащёнными мощными ЦПУ или ГПУ, всё ещё существуют определенные сложности в достижении высокой производительности в масштабе реального времени