

УДК 004.934.2

## ГЕНЕРАТИВНЫЙ МЕТОД ПОЛУЧЕНИЯ СПЕКТРАЛЬНЫХ ОГИБАЮЩИХ В КОНТЕКСТЕ ЗАДАЧИ АНАЛИЗА СОСТОЯНИЯ ГОЛОСОВОЙ ФУНКЦИИ ЧЕЛОВЕКА

Д.С. ЛИХАЧЁВ, М.И. ВАШКЕВИЧ, Н.А. ПЕТРОВСКИЙ, И.С. АЗАРОВ

*Белорусский государственный университет информатики и радиоэлектроники  
(Минск, Беларусь)*

**Аннотация.** В работе предлагается генеративный метод получения спектральных огибающих, предназначенный для использования в задачах анализа и оценки состояния голосовой функции человека. Оценка состояния голосовой функции является определяющей для диагностики различных заболеваний, связанных с изменениями в звучании голоса. В предлагаемом методе получения спектральной оценки используется дискретное преобразование Фурье (ДПФ). В качестве входных данных для ДПФ используются специальным образом сгенерированные последовательности, а не входной сигнал. Входные данные для ДПФ формируются с использованием параметров авторегрессионной модели таким образом, чтобы обеспечить строгую наилучшую периодичность анализируемой последовательности. Как показали результаты экспериментов, спектральные огибающие, полученные с помощью предложенного метода, имеют меньшую среднеквадратичную ошибку представления истинного спектра модельного сигнала по сравнению с остальными. Таким образом, использование предложенного метода для вычисления спектральных огибающих имеет хорошую перспективу для соответствующих задач цифровой обработки речевых сигналов. Получаемые огибающие могут использоваться в качестве признаков для систем машинного обучения.

**Ключевые слова:** генеративный метод, спектральная огибающая, авторегрессионная модель, преобразование Фурье, диагностика заболеваний по голосу.

## GENERATIVE METHOD OF OBTAINING SPECTRAL ENVELOPES FOR SPEECH SIGNAL ANALYSIS AND PROCESSING TASKS

DENIS.S. LIKHACHOV, MAXIM.I. VASHKEVICH, NICK.A. PETROVSKY, ELIAS.S. AZAROV

*Belarusian State University of Informatics and Radioelectronics  
(Minsk, Belarus)*

**Abstract.** The paper proposes a generative method for obtaining spectral envelopes intended for use in the tasks of analyzing and evaluating the state of human voice function. Assessment of the state of voice function is crucial for the diagnosis of various diseases associated with changes in the sound of the voice. The proposed method for obtaining a spectral estimate uses a discrete Fourier transform (DFT). Specially generated sequences are used as input data for the DFT, and not the input signal. The input data for the DFT are generated using the parameters of the autoregressive model in such a way as to ensure the strict best periodicity of the analyzed sequence. As the results of the experiments have shown, the spectral envelopes obtained using the proposed method have a smaller root-mean-square error in representing the true spectrum of the model signal compared to the others. Thus, the use of the proposed method for calculating spectral envelopes has a good prospect for the corresponding tasks of digital processing of speech signals. The resulting envelopes can be used as features for machine learning systems.

**Keywords:** generative method, spectral envelope, autoregressive model, Fourier transform, diagnosis of diseases by voice.

### Введение

Оценка состояния голосовой функции человека является определяющей для диагностики различных заболеваний, связанных с изменениями в звучании голоса. Перспективным подходом для решения этой задачи является использование методов

машинного обучения на основе спектральных признаков, которые представлены, в том числе, и в виде спектральных огибающих [1]. Подавляющее большинство применяемых в настоящее время методов получения спектральных признаков используют преобразование Фурье. Это обусловлено прежде всего тем, что данное преобразование хорошо подходит для частотно-временного анализа периодических или частично периодических сигналов, таких как, речевые и аудио сигналы. При использовании ДПФ для получения точной спектральной оценки с точки зрения соответствия истинному спектру необходимо выполнение двух основных требований. Во-первых, анализируемый сигнал должен быть строго периодическим и стационарным на протяжении всего интервала анализа. Во-вторых, период анализируемого сигнала должен точно совпадать с размером ДПФ. Для речевых и аудиосигналов два вышеприведенных требования в строгом виде не выполняются. Это обуславливает необходимость применение временного окна для обработки входных данных [2]. Обработка окном хотя и позволяет ослабить влияние боковых спектральных лепестков, но в тоже время приводит к значительным искажениям в полученной спектральной оценке по отношению к истинной. Также необходимо отметить, что при получении огибающей на основе уже имеющегося спектра возникает проблема его корректной дискретизации и сглаживания.

### **Генеративный метод получения спектральных огибающих**

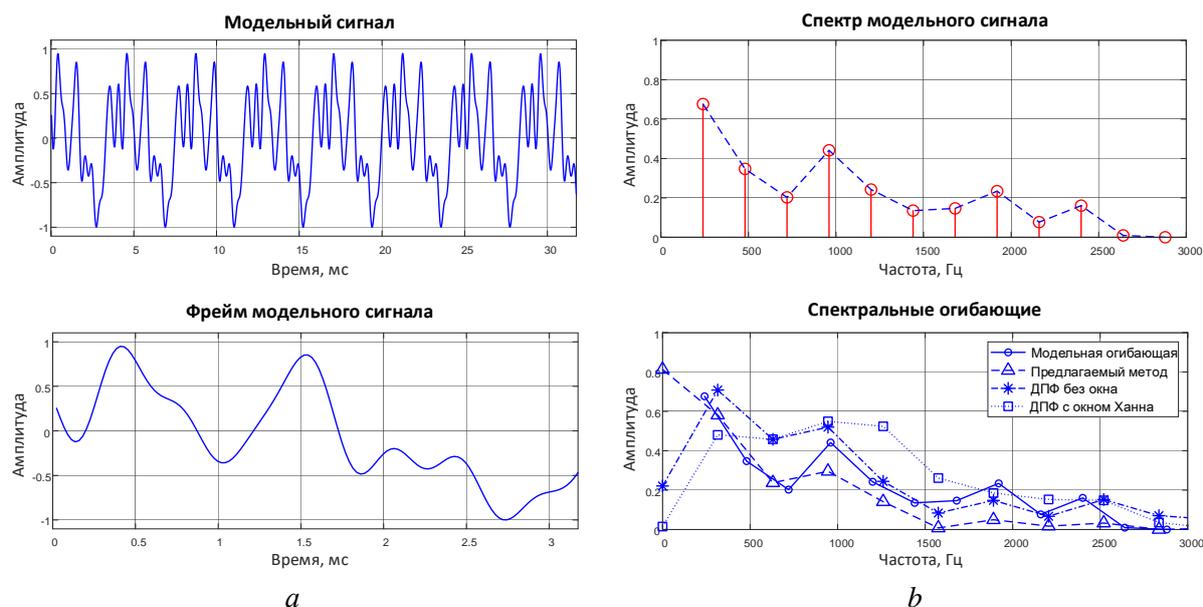
Основная идея предлагаемого метода получения спектральной оценки в форме огибающей состоит в том, чтобы использовать ДПФ без применения оконной функции для обработки входных данных, обеспечив при этом согласованность периода анализируемого сигнала и размера ДПФ. С этой целью предлагается в качестве входных данных для ДПФ использовать специальным образом сгенерированную последовательность, которая имела бы равный размеру ДПФ период и соответствующий анализируемому сигналу спектральный состав.

Основные этапы предлагаемого метода вычисления спектральной огибающей следующие: 1) нахождение параметров авторегрессионной модели по входному фрейму анализируемого сигнала; 2) генерация с помощью авторегрессионной модели анализируемого процесса периодической последовательности с периодом, равным целевому размеру ДПФ; 3) вычисление дискретного спектра одного периода последовательности с помощью ДПФ.

Пример спектральных огибающих, вычисленных для одного фрейма модельного сигнала приведен на рис. 1.

### **Результаты и их обсуждение**

Был проведен сравнительный анализ эффективности представления истинного спектра модельного сигнала с помощью спектральных огибающих, полученными тремя способами: 1) предложенным генеративным методом; 2) ДПФ без применения окна; 3) ДПФ с применением окна Ханна. В качестве входного анализируемого сигнала использовался синтетический речеподобный сигнал с 11 гармониками с периодом основного тона от 85 до 240 Гц. Анализ проводился для окон анализа длительностью от 1.5 до 20 мс. В качестве критерия эффективности представления спектра модельного сигнала с помощью вычисленных огибающих использовалась среднеквадратичная ошибка (MSE). Для корректного сопоставления частотной сетки спектра модельного сигнала и полученных огибающих использовалась линейная интерполяция. Все огибающие при сравнении нормировались по энергии. Полученные результаты отображены в табл. 1.



**Рис. 1.** Пример спектральных огибающих, определенных для одного фрейма модельного сигнала: *a* – модельный сигнал и фрейм модельного сигнала для вычисления спектральных огибающих; *b* – истинный спектр модельного сигнала и полученные спектральные огибающие

**Таблица 1.** Среднеквадратичная ошибка (MSE) представления спектра модельного сигнала с помощью спектральных огибающих

Метод получения спектральной огибающей	Длительность фрейма анализа, мс				
	1,5	3,2	5,0	10	20
предложенный метод	<b>0,041503</b>	<b>0,038539</b>	<b>0,038408</b>	<b>0,040169</b>	<b>0,044601</b>
ДПФ без применения окна	0,069208	0,055167	0,055008	0,055956	0,067225
ДПФ с применением окна Ханна	0,061174	0,049126	0,041851	0,045179	0,062380

Как показали результаты экспериментов, спектральные огибающие, полученные с помощью предложенного генеративного метода имеют меньшую среднеквадратичную ошибку представления истинного спектра модельного сигнала по сравнению с остальными.

### Заключение

Опираясь на полученные экспериментальные результаты можно сделать вывод о том, что использование предложенного метода для вычисления спектральных огибающих имеет хорошую перспективу для задач обработки речевых сигналов, в которых имеет важное значение точность представления спектральных характеристик анализируемого сигнала и возможность отслеживания динамики их изменения во времени при использовании относительно малого количества параметров. Примером такой задачи может служить классификация голосового сигнала на основе оценки состояния голосовой функции человека с помощью методов машинного обучения [1].

### Список литературы

1. Вашкевич М.И., Лихачёв Д.С., Азаров И.С. Система анализа и классификации голосового сигнала на основе пертурбационных параметров и кепстрального представления в психоакустических шкалах. Доклады БГУИР. 2022;20(1):73-82. <https://doi.org/10.35596/1729-7648-2022-20-1-73-82>
2. F. J. Harris. On the use of windows for harmonic analysis with the discrete Fourier transform / Proceedings of the IEEE, vol. 66, no. 1, pp. 51-83, Jan. 1978, doi: 10.1109/PROC.1978.10837.