

ШУМООЧИСТКА АУДИОСИГНАЛОВ НА ОСНОВЕ ДИСКРЕТНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ С АДАПТИВНЫМИ ПОРОГАМИ

*Ильяшенко Илья Дмитриевич,
Насретдинов Рауф Салаватович*

Студент, Студент

ИЦТЭФ АлтГУ, Барнаул, Россия

E-mail: ilyashenko.594m@stud.asu.ru, nasretdinov.594m@stud.asu.ru

Научный руководитель — Лепендин Андрей Александрович

Современные технологии улучшения качества записанной речи предполагают разработку новых методов очистки аудиосигналов от шумов различной природы, связанных с наличием фоновых звуков, ревербераций и помех в канале связи. Одним из эффективных методов очистки речевого сигнала от шума является фильтрация на основе вейвлетов. Суть данного метода заключается в декомпозиции сигнала с помощью дискретного вейвлет-преобразования [1], последующем применении пороговой функции к полученным коэффициентам декомпозиции и обратном вейвлет-преобразовании. Однако при использовании такого подхода значения порогов являются фиксированными и не гарантируют наилучшего качества очистки аудиосигнала. В данной работе предложен новый модифицированный метод пороговой фильтрации речевых сигналов на основе дискретного вейвлет-преобразования с обучаемыми значениями порогов.

Суть предлагаемой модификации заключалась в следующем. Вводился вектор пороговых значений T , который подавался на функцию «мягкого» порогового преобразования [2]:

$$y = \text{sign}(x)(|x| - T), \quad (1)$$

где x – значения коэффициентов вейвлет-преобразования сигнала. После обратного восстановления аудиосигнала вычислялась целевая функция S , оценивавшая его среднеквадратичное отклонение от соответствующего чистого сигнала. Полученное значение использовалось для оценки градиента целевой функции по вектору пороговых значений методом обратного распространения ошибки [3] (рисунок 1). Фильтрация зашумленных сигналов осуществлялась двумя способами: с использованием одного и того же значения порога T для всех уровней вейвлет-разложения (глобальный пороговый алгоритм) и с независимыми значениями $T(J)$ для каждого уровня J (поуров-

невый пороговый алгоритм). Предложенный подход сравнивался с методом без адаптации, в котором значения вектора порогов вычислялись по формулам (2) и (3) для глобального и поуровневого пороговых алгоритмов соответственно:

$$T = \sigma \sqrt{2 \log N}, \quad (2)$$

$$T = \sigma \sqrt{2 \log \frac{N}{2^J}}, \quad (3)$$

где σ - интенсивность шума, N - длина сигнала (в отсчетах), J - номер уровня разложения. Для оценки эффективности шумоочистки использовались отношение сигнал-шум и среднеквадратичное отклонение, представленные в формулах (4) и (5) соответственно:

$$\text{SNR} = 10 \log_{10} \frac{\sum_{i=1}^N \hat{Y}_i^2}{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}, \quad (4)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N}}, \quad (5)$$

где Y_i - очищенный от шума сигнал, \hat{Y}_i - чистый сигнал. Апробация проводилась на образцах из голосовой базы данных CSTR VCTK Corpus с аддитивным белым гауссовым шумом, который генерировался из нормального распределения со средним значением 0 и стандартным отклонением 0.03.

Результаты тестирования предложенного метода представлены в таблице 1. Для зашумленных сигналов значения среднеквадратичного отклонения и отношения сигнал-шум были равны $\text{RMSE}=0.03$ и $\text{SNR}=5.8$ соответственно. В качестве базиса применялись вейвлеты Добеши четвертого порядка (db4).

Тип порогового алгоритма	Адаптивный метод	SNR	RMSE
глобальный	Нет	7.21	0.0291
поуровневый	Нет	9.82	0.0216
глобальный	Да	12.93	0.0151
поуровневый	Да	14.53	0.0126

Таблица 1: Сравнение качества работы неадаптивного и адаптивного методов очистки аудиосигналов.

В данной работе было показано, что предложенный новый метод шумоочистки аудиосигналов, продемонстрировал лучшее качество в сравнении с «классическим» подходом с использованием фиксированных значений вектора порогов. Данный метод может найти широкое применение при решении практических задач обработки речевых сигналов в криминалистике, коммуникациях и информационной безопасности.

Иллюстрации

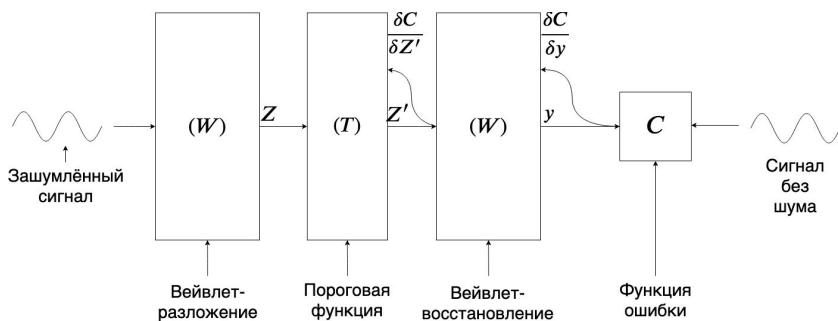


Рисунок 1 – Схема обучения вектора пороговых значений.

W – базис вейвлет-преобразования, T – вектор пороговых значений, Z – вейвлет-разложение, Z' – результат фильтрации, y – восстановленный сигнал, C – функция ошибки

Литература

1. Stephane M. A Wavelet Tour of Signal Processing: The Sparse Way Burlington: Academic Press, 2009.
2. Shemi P. M., Ali M. A. An Improved Method of Audio Denoising Based on Wavelet Transform // In Proceedings of 2015 IEEE International Conference on Power, Instrumentation, Control and Computing, Thrissur, India, 2015, P.1–6.
3. Пыашенко I. D., Nasretdinov R. S., Filin Y. A., Lependin A. A. Trainable wavelet-like transform for feature extraction to audio classification // Journal of Physics: Conference Series. 2019. T. 1333 012136.