

Е. В. Пиуновский, А. А. Тропченко

АНАЛИЗ АУДИОДАНЫХ С ПОМОЩЬЮ ВЕЙВЛЕТ-ФУНКЦИЙ

Рассмотрены области применения вейвлет-функций при цифровой обработке звука. Показаны преимущества вейвлет-преобразований по сравнению с традиционными методами анализа. Приведены результаты исследования возможностей вейвлет-анализа при фильтрации и сжатии аудиоданных.

Ключевые слова: анализ, звук, вейвлет, сжатие, аудиоданные.

Введение. Вейвлет-преобразование (ВП) — относительно новая эффективная технология, позволяющая проводить обработку сигналов различного типа. Обладая рядом преимуществ по сравнению с традиционными видами преобразований, ВП нашли применение в области кодирования видеоданных и изображений. Кроме того, существует множество исследований по использованию ВП для сжатия звука, в ходе которых было показано, что данный вид функций позволяет выделять различные характеристики аудиосигналов. Это свойство обуславливает возможность применения ВП для анализа звуковых данных с последующим использованием полученных сведений не только для их компрессии, но и для решения ряда других задач. Одной из таких задач является так называемый поиск музыкальной информации (англ. MIR — Music Information Retrieval), иными словами — выделение основных сигналов в музыкальном фрагменте. Это направление получило в последнее время широкое распространение [1].

Выделение информации в звуковых сигналах (особенно, в музыке) с помощью преобразований позволяет выполнять ряд сложных действий без вмешательства человека. В первую очередь, это касается поиска музыкальной информации по ее содержанию, классификации и жанровой принадлежности. В настоящее время существует достаточное количество публикаций, в которых описываются разные подходы к решению данной задачи на основе ВП [1—3]. Тем не менее исследования по применению ВП при обработке звука с последующим его сжатием или шумоподавлением не теряют своей актуальности и перспективности, учитывая разнообразие доступных функций.

Вейвлет-преобразование и преобразование Фурье. Вейвлет-функции являются набором функций, получаемых с помощью сжатий (растяжений) и сдвигов по временной оси исходной функции (материнского вейвлета). Математически это можно выразить следующим образом:

$$\Psi_{ab}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right),$$

где $\Psi(t)$ — материнский (исходный) вейвлет, t — время, b — параметр, характеризующий сдвиг по времени, a — параметр масштаба, $1/\sqrt{a}$ — множитель, обеспечивающий нормализацию.

Видно, что уже в самом определении вейвлетов заложена идея масштабируемости. Это позволяет подбирать функции, способные обеспечить наиболее точный анализ, для каждого вида сигнала. Кроме того, вейвлеты имеют ряд преимуществ по сравнению с традиционными методами работы со звуком. Так, преобразование Фурье не обладает временным разрешением, а использование специально разработанного оконного преобразования Фурье (ОПФ) позволяет фиксировать время с определенной точностью, однако сигнал при этом делится на окна одинакового размера, что лишает процедуру анализа гибкости. ВП, напротив, имеет частотно-временное окно изменяемого размера, площадь которого (как и в случае с ОПФ) определяется согласно принципу неопределенности Гейзенберга [4]. Применительно к практике можно выделить два основных преимущества вейвлетов для анализа сигналов, а именно: компактный носитель (хорошее временное разрешение) и полосчатый спектр (хорошее частотное разрешение).

Виды анализа звука с помощью вейвлетов. Один из видов анализа аудиосигналов — их графическое отображение с помощью различных способов. Как правило, выделяют четыре основных вида представления звуковой волны:

- сигналограмму (амплитудно-временная зависимость);
- спектр сигнала (амплитудно-частотная зависимость);
- спектрограмму (часто называемую „сонограммой“ — псевдотрехмерный график с отображением амплитуд частотных составляющих);
- трехмерную спектрограмму (с отображением амплитуд частотных составляющих по третьей координатной оси).

По мнению специалистов, спектрограмма обеспечивает наилучшее визуальное представление спектра сигнала и позволяет в подробностях анализировать динамику его развития [5]. Именно использование ВП позволяет построить наиболее наглядные графики подобного типа, на которых, в соответствии со свойствами вейвлетов, определена четкая привязка сигнала ко времени.

При числовом анализе звука можно выделить два основных подхода на базе вейвлетов: дискретное ВП (ДВП) и пакетное дискретное ВП (ПДВП). Если рассматривать такие преобразования как наборы фильтров, то ДВП можно характеризовать как пропускание сигнала через набор низкочастотных фильтров, что приводит к „однобокому“ дереву преобразования. Особенность такого подхода заключается в неявном предположении о большем количестве информации об исходном сигнале, содержащемся в области низких частот, что оправдано для многих реальных сигналов и соответствует особенностям человеческого слуха. При ПДВП производится полное каскадирование сигнала как в низкочастотной, так и в высокочастотной области, в результате чего формируется „сбалансированное“ дерево, путь по которому может быть выбран в зависимости от предъявляемых требований к степени сжатия сигнала. Таким образом, пакетное ДВП является более гибким средством декомпозиции сигнала и подходит для решения особенно сложных задач [4, 6]. При этом следует отметить дополнительное особое преимущество ПДВП при обработке музыкальных данных, которое позволяет определить набор вейвлет-коэффициентов, совпадающих с октавами и нотами [2].

Обработка аудиоданных с помощью вейвлет-анализа. Экспериментальная часть проведенных авторами исследований заключалась в анализе набора музыкальных файлов в системе компьютерного моделирования MatLab, которая, имея в своем составе дополнение Wavelet Toolbox, является одним из лучших средств исследования вейвлетов [6]. В ходе экспериментов применялось моделирование работы блока анализа аудиоданных на основе ДВП (в терминах MatLab: One-Dimensional Wavelet Analysis — одномерный анализ с помощью дискретного преобразования) и ПДВП (в терминах MatLab: One-Dimensional Wavelet Packet Analysis — одномерный анализ с помощью вейвлет-пакетов). В первом случае осуществля-

лась фильтрация звука в целях шумоподавления, во втором — его сжатие. В качестве анализируемых данных использовались три фрагмента со следующим звуковым наполнением:

- 1) звучание (сигнал) отдельного музыкального инструмента (пианино);
- 2) звучание (сигнал) симфонического оркестра;
- 3) зашумленный музыкальный сигнал (с уровнем шума около 35 дБ).

Были установлены следующие постоянные параметры работы системы:

- 5-й уровень декомпозиции для вейвлетного дерева (подобранный для получения лучшего соотношения времени обработки и качества анализа);
- 4-й порядок функций в случае семейств вейвлетов (на основе статистических сведений об эффективности базовых функций [4, 6], а также исследования, описанного в работе [7]);
- степень сжатия звука, равная 10 (в системе MatLab задается параметром „number of zeros“ [6]); столь высокая степень сжатия задана для удобства сравнения результатов, которые при более низкой степени сжатия менее резко изменяются при переходе от одного базового вейвлета к другому.

В качестве базовых функций преобразований использовались следующие семь наборов вейвлетов — Хаара (Haar), Добеши (Db4), симлет (sym4), койфлет (coif4), биортогональный (bior4.4), обратный биортогональный (rbior4.4) и Мейера (Meur).

На первом этапе эксперимента осуществлялось сжатие аудиосигналов с помощью ПДВП. Результаты данной операции для всех функций представлены на рис. 1. Как показывает анализ рисунка, огибающие результатов сжатия сигналов для первого и второго музыкальных фрагментов (кривые 1 и 2) схожи между собой, но различаются лишь качеством сжатия (это проявляется в смещении соответствующих графиков по вертикали). Наиболее сложным для компрессии оказался третий фрагмент (кривая 3). Это связано с неоднородностью и нестационарностью данного сигнала, что вызвано высоким уровнем зашумленности и большой амплитудой высокочастотной составляющей по сравнению с другими сигналами. Лучшие результаты сжатия аудиосигнала наблюдались при использовании в качестве базовых функций биортогонального вейвлета и вейвлета Мейера.

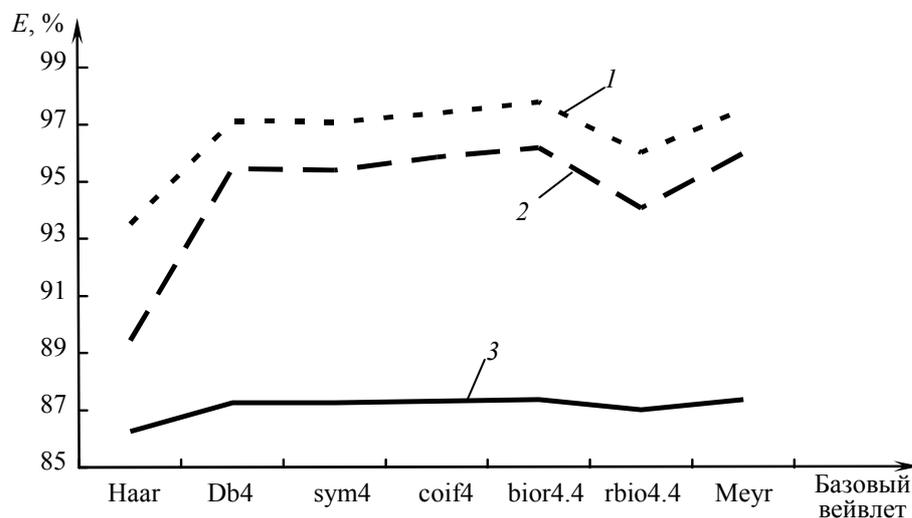


Рис. 1

На следующем этапе проводились фильтрация данных с помощью ДВП и последующее сжатие полученных сигналов посредством ПДВП (как и на первом этапе). Результаты показаны на рис. 2 (обозначения кривых соответствуют приведенным на рис. 1). Форма огибающей изменилась только для зашумленного сигнала (кривая 3), для него же заметно значительное улучшение качества сжатия при анализе с помощью определенных вейвлетов (в частности, биортогонального вейвлета с увеличением сохраненной энергии E на 2 %). Целесообразно проиллюстрировать полученный график, приведя средние показатели улучшения качества

для каждого звукового фрагмента: 0,26 % — для отдельного инструмента, 0,81 % — для оркестра и 0,87 % — для зашумленного сигнала.

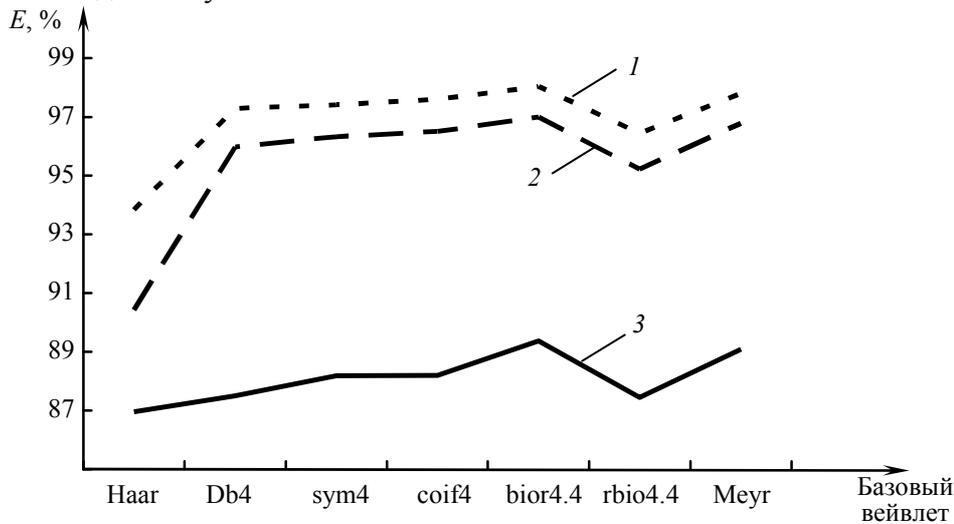


Рис. 2

Изменение качества сжатия сигналов в зависимости от применяемого базового вейвлета можно показать на другом графике (рис. 3): результаты анализа фрагментов 1 и 2 похожи — наибольшего повышения качества сжатия достигли симлет и обратный биортогональный вейвлет; совершенно другая ситуация с зашумленным сигналом — при его анализе уже отмеченные биортогональный вейвлет и вейвлет Мейера показали наилучший результат.

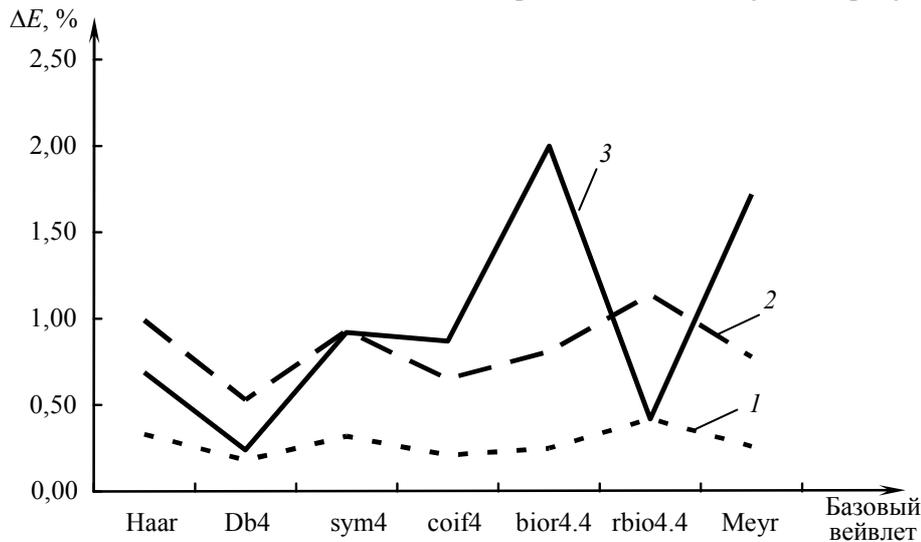


Рис. 3

Заключение. Рассмотренные в настоящей статье технологии анализа звуковых данных с помощью вейвлет-функций являются перспективными методами, способными превзойти и дополнить традиционные подходы в области цифровой обработки аудиосигналов. Обладая рядом преимуществ по сравнению с преобразованием Фурье и ОПФ, вейвлеты позволяют осуществлять более гибкий и масштабируемый анализ, что приводит к очень точному выделению свойств сигналов как при графическом, так и при числовом представлении получаемых коэффициентов.

Представленные результаты позволяют выделить ряд направлений для дальнейших исследований — например, подбор базовых функций и типов преобразований (и их комбинаций) для решения традиционных задач, а также применение вейвлетов в актуальной и перспективной области поиска музыкальной информации.

СПИСОК ЛИТЕРАТУРЫ

1. *Tzanetakis G., Essl G., Cook P.* Audio analysis using the discrete wavelet transform // Proc. Conf. in Acoustics and Music Theory Applications, 2001 [Электронный ресурс]: <http://soundlab.cs.princeton.edu/publications/2001_ama_aadwt.pdf>.
2. *Grimaldi M., Cunningham P., Kokaram A.* A wavelet packet representation of audio signals for music genre classification using different ensemble and feature selection techniques // Proc. of Multimedia Information Retrieval. 2003. P. 102—108 [Электронный ресурс]: <<http://dl.acm.org/citation.cfm?id=973281>>.
3. *La Cour-Harbo A., Endelt L.* Time-frequency distributions of music based on sparse wavelet packet representations // Control Eng. Working Paper, Ser. 4850. Aalborg: Dep. of Control Engineering, Aalborg Univ., 2005 [Электронный ресурс]: <http://vbn.aau.dk/files/6353816/4850_Endelt.pdf>.
4. *Воробьев В. И., Грибунин В. Г.* Теория и практика вейвлет-преобразования. СПб: Изд-во Военного ун-та связи, 1999. 204 с.
5. *Радзишевский А. Ю.* Основы аналогового и цифрового звука. М.: Изд. дом „Вильямс“, 2006. 288 с.
6. *Misiti M. et al.* Wavelet Toolbox User's Guide. USA, MA: The Mathworks, Inc., 2010. 454 p. [Электронный ресурс]: <http://www.mathworks.com/help/pdf_doc/wavelet/wavelet_ug.pdf>.
7. *Пиуновский Е. В., Тропченко А. А.* Возможности сжатия звука с использованием вейвлет-преобразований // Интеллектуальные и информационные системы: Материалы Всерос. науч.-техн. конф. Тула, 2009.

Сведения об авторах

- Евгений Витальевич Пиуновский** — аспирант; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники; E-mail: evgriun@gmail.com
- Андрей Александрович Тропченко** — канд. техн. наук, доцент; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники; E-mail: zayka_98rus@mail.ru

Рекомендована кафедрой
вычислительной техники

Поступила в редакцию
03.03.11 г.