

Современные технологии сжатия аудиосигналов

Адиканова С., Бейбит Ж.

*Восточно-Казахстанский университет им. С. Аманжолова, г. Усть-Каменогорск
(Казахстан)*

Ersal_7882@mail.ru, Jarhin_mn@mail.ru

Аннотация

В настоящее время можно наблюдать взрывной рост в использовании сети Интернет и мобильных телефонов, а сближение двух этих технологий открывает широкий диапазон новых возможностей на уже процветающем рынке мультимедиа. Эти возможности побуждают к проведению исследований, которые могут и должны выявить недостатки существующих методов обработки цифровых данных и показать пути их (методов) оптимизации для удовлетворения современных нужд рынка.

Ключевые слова: сжатия информации, вейвлет-преобразования, обработка нестационарных сигналов, кодирование.

1. Введение

Несомненно, бурное развитие средств вычислительной техники привело к появлению различных мультимедийных приложений и программ, в которых используются тексты, изображения, анимированные фрагменты и звук. Для хранения звука требуется места меньше, чем для видео, но больше, чем для текста и изображений. В данном контексте очень большой интерес представляют вопросы сжатия звуковой информации, которые в настоящее время являются весьма актуальными, о чем говорят большое количество практических исследований и постоянная разработка и опубликование стандартов в этой области [1]. Этим и обуславливается актуальность данной проблемы в медиапространстве. Особенно интересны в данной области исследования и работы А.Ю. Тропченко, А.А. Тропченко, Д. Сэломона, Ю.А. Ковалгина, Э.И. Вологодина, Е.В. Пиуновского. Однако, несмотря на подобное внимание к данной проблеме, распространенные алгоритмы сжатия аудиоданных которые предлагают IT специалисты, все еще содержат ряд существенных недостатков которые мы попытаемся описать в данной работе. Очевидно, что для их устранения необходимо внедрение новых технологий в процесс компрессии звука. Одной из таких перспективных технологий являются вейвлет-преобразования (ВП). По мнению исследователей, можно без преувеличения сказать, что вейвлеты (wavelets) произвели революцию в области теории и практики обработки нестационарных сигналов [2]. Вейвлет-функции уже достаточно долго используются в различных областях не только науки, но и техники, что связано с большими преимуществами, которые обещает их применение. Но некоторые возможности их приложения все еще вызывают вопросы о целесообразности и споры в научных сообществах. Одним из таких приложений является сжатие аудиоданных с использованием ВП. Таким образом, целью данной работы является: рассмотреть и показать преимущества современных технологий сжатия аудиосигналов, в частности вейвлет- преобразования.

Задачи:

1. Раскрыть понятие “сжатие данных”, охарактеризовать принципы сжатия данных;
2. Проанализировать предлагаемые подходы к использованию ВП в процессе кодирования;

Итак, сжатие данных - (англ. *data compression*) – алгоритмическое преобразование данных, производимое с целью уменьшения занимаемого ими объёма. Применяется для более рационального использования устройств хранения и передачи данных. Обратная процедура называется восстановлением данных (распаковкой, декомпрессией). Сжатие основано на устранении избыточности, содержащейся в исходных данных. Простейшим примером избыточности является повторение в тексте фрагментов (например, слов естественного или машинного языка). Подобная избыточность обычно устраняется заменой повторяющейся последовательности ссылкой на уже закодированный фрагмент с указанием его длины. Другой вид избыточности связан с тем, что некоторые значения в сжимаемых данных встречаются чаще других. Сокращение объёма данных достигается за счёт замены часто встречающихся данных короткими кодовыми словами, а редких – длинными (энтропийное кодирование). Сжатие данных, не обладающих свойством избыточности (например, случайный сигнал или белый шум, зашифрованные сообщения), принципиально невозможно без потерь.

Сжатие без потерь позволяет полностью восстановить исходное сообщение, так как не уменьшает в нем количество информации, несмотря на уменьшение длины. Такая возможность возникает, только если распределение вероятностей на множестве сообщений не равномерное, например часть теоретически возможных в прежней кодировке сообщений на практике не встречается.

2. Принципы сжатия данных.

В основе любого способа сжатия лежит модель источника данных, или, точнее, модель избыточности. Иными словами, для сжатия данных используются некоторые априорные сведения о том, какого рода данные сжимаются. Не обладая такими сведениями об источнике, невозможно сделать никаких предположений о преобразовании, которое позволило бы уменьшить объём сообщения. Модель избыточности может быть статической, неизменной для всего сжимаемого сообщения, либо строиться или параметризоваться на этапе сжатия (и восстановления). Методы, позволяющие на основе входных данных изменять модель избыточности информации, называются адаптивными. Неадаптивными являются обычно узкоспециализированные алгоритмы, применяемые для работы с данными, обладающими хорошо определёнными и неизменными характеристиками. Подавляющая часть достаточно универсальных алгоритмов являются в той или иной мере адаптивными.

Все методы сжатия данных делятся на два основных класса:

- Сжатие без потерь;
- Сжатие с потерями.

При использовании сжатия без потерь возможно полное восстановление исходных данных, сжатие с потерями позволяет восстановить данные с искажениями, обычно несущественными с точки зрения дальнейшего использования восстановленных данных. Сжатие без потерь обычно используется для передачи и хранения текстовых данных, компьютерных программ, реже – для сокращения объёма аудио- и видеоданных, цифровых фотографий и т.п., в случаях, когда искажения недопустимы или нежелательны. Сжатие с потерями, обладающее значительно большей, чем сжатие без потерь, эффективностью, обычно применяется для сокращения объёма аудио- и видеоданных и цифровых фотографий в тех случаях, когда такое сокращение является приоритетным, а полное соответствие исходных и восстановленных данных не требуется.

3. Вейвлеты в кодировании звука.

Одной из перспективных технологий являются вейвлеты. В пользу возможности применения ВП в области кодирования звука говорят следующие факты:

– многократные заявления разработчиков кодека Vorbis, известного как Ogg Vorbis (Xiph.Org Foundation), о скором включении дискретного вейвлет-преобразования (ДВП) в процесс сжатия в данном формате;

– попытки применения вейвлетов для кодирования аудиосигналов (для разнообразных применений) фирмой Adapted Wave Technologies;

– поддержка Институтом инженеров по электротехнике и радиоэлектронике (IEEE, Institute of Electrical and Electronics Engineers) исследований по разработке адаптивных вейвлет-кодеров для звуковых данных [3];

– наличие нескольких успешных разработок в области сжатия аудиоданных с помощью ВП на низких скоростях передачи битов (bit rates, битрейтах) [3, 4].

Два последних факта представляют особый интерес, так как в них представлены реальные результаты попыток внедрения вейвлетов в процесс кодирования аудиоданных, вейвлет-технологии, используемые в каждой из этих работ, различны, и оба этих исследования впоследствии послужили основанием для ряда научных статей в области нововведений в цифровой обработке звука. Результаты анализа исследований представлены в таблице 1. Представить сравнительные характеристики технологий сжатия звука на базе ВП.

Таблица 1

Сравнительные характеристики технологий сжатия звука на базе ВП

Используемые преобразования и технологии	Прозрачное сжатие аудио на низких битрейтах с использованием адаптированных вейвлетов [4]	Сжатие аудио высокого качества с использованием адаптивной декомпозиции вейвлет-пакетов и психоакустической модели
Вейвлет-технология (замена традиционному МДКП)	ДВП	Пакетное ДВП (ПДВП)
Базисная вейвлет-функция	Не указана (предлагается адаптивный подбор функции)	Биортогональный вейвлет (5-го порядка)
Психоакустическая модель	Упрощенная	Модифицированная психоакустическая модель 2 стандарта MPEG ISO/IEC 11172-3
Дополнительные технологии	Устранение статистической избыточности с использованием динамического словаря	Алгоритм распределения бит на основе нуль-дерева
Способ проверки качества кодирования	Субъективное тестирование	Субъективное тестирование
Битрейт, на котором достигается прозрачное кодирование	48–66 кбит/с (с применением дополнительных технологий); 64–70 кбит/с	45 кбит/с

Комментируя результаты сравнения, представленные в таблице, необходимо уточнить, что под “прозрачным” кодированием понимается такое сжатие, при котором слушатель (участник “субъективного тестирования”) не может отличить сжатый звуковой сигнал от оригинала (их качество кажется ему идентичным). Здесь сразу становится очевидной сложность процесса оценки качества сжатия аудиоинформации, ее субъективность. Но процесс тестирования представлен в рассмотренных работах весьма основательно, и

для сравнения следует заметить, что разработчики формата МР3 заявляли о достижимости прозрачного кодирования при использовании соответствующего алгоритма лишь на битрейте, равном 128 кбит/с (и даже это заявление подвергается частой критике).

Как видно из таблицы, рассмотренные работы схожи по общему подходу к модификации процесса сжатия звука, но различаются реализацией отдельных его шагов. Отталкиваясь от того, что понятие “вейвлет”, более всего нас интересующее, выделено даже в названиях обеих статей, можно отметить следующие общие недостатки, присущие данным исследованиям (именно при акцентировании внимания на ВП):

- использование дополнительных технологий (что мешает точному определению вклада вейвлетов в улучшение результатов сжатия);

- недостаточно глубокий подход к выбору ВП, используемого при кодировании.

В то же время можно заметить, что при подобной расстановке приоритетов и выделении представленных выше недостатков работа [3] является все же более предпочтительной: дополнительные технологии, применяемые в ней, заимствованы (и проверены другими исследователями), в ней осуществлен переход от ДВП к более сложным и эффективным ПДВП. Скорее всего, именно с этим и связан достигнутый в ней более высокий результат (низкий битрейт).

Таким образом, современные методы сжатия аудиоданных не лишены недостатков и поэтому могут быть улучшены. В данной работе были показаны основные проблемы пространственных алгоритмов кодирования звука. Также был проведен анализ исследований, в которых была доказана возможность улучшения существующих технологий звуковой компрессии с помощью вейвлетов, являющихся одной из самых перспективных технологий в области цифровой обработки сигналов.

Список литературы

1. Ковалгин Ю.А., Вологдин Э.И. Цифровое кодирование звуковых сигналов. — СПб. : КОРОНА-принт, 2004.
2. Воробьев В.И., Грибунин В.Г. Теория и практика вейвлет-преобразования. — СПб. : Изд-во ВУС, 1999.
3. Srinivasan P., Jamieson L. High Quality Audio Compression Using an Adaptive Wavelet Packet Decomposition and Psychoacoustic Modeling // IEEE Transactions on Signal Processing. — 1998. — Vol. 46, no. 4.
4. Sinha D., Tewfik A. Low Bit Rate Transparent Audio Compression Using Adapted Wavelets // IEEE Trans. ASSP. — 1993. — Vol. 41, no. 12.
5. Сэлмон Д. Сжатие данных, изображений и звука. — М. : Техносфера, 2004.
6. Тропченко А.Ю., Тропченко А.А. Методы сжатия изображений, аудиосигналов и видео. — СПб. : СПбГУ ИТМО, 2006.
7. Тропченко А.Ю., Курносенков И.Н. Анализ современных стандартов сжатия видеоданных // Научно-технический вестник СПбГУ ИТМО. — 2006. — № 32. — С. 17–21.
8. Пиуновский Е.В., Тропченко А.А. Современные технологии сжатия аудиосигналов // Научно-технический вестник СПбГУ ИТМО. — 2010. — № 1(65).