

ЗАСТОСУВАННЯ ДИСКРЕТНОГО ВЕЙВЛЕТ- ПЕРЕТВОРЕННЯ ДЛЯ СТИСНЕННЯ АУДІОДАНИХ У СИСТЕМАХ ОБРОБКИ АУДІОІНФОРМАЦІЇ

¹ Вінницький національний технічний університет

Анотація

Розглянуто застосування дискретного вейвлет-перетворення для стиснення аудіоінформації в системах обробки звуку. Запропоновано метод стиснення аудіоданих, заснований на вейвлет-перетворенні, що дозволяє збільшити коефіцієнт стиснення при мінімальних втратах якості.

Ключові слова: стиснення, вейвлет-перетворення, аудіодані, перетворення Фур'є, кодування.

Abstract

The application of discrete wavelet transform for audio information compression in sound processing systems is considered. An audio data compression method based on wavelet transform is proposed, which allows increasing the compression ratio with minimal quality loss.

Keywords: compression, wavelet transform, audio data, Fourier transform, coding.

Вступ

У наш час із поширенням електронних технологій в побутових приладах та мультимедійній техніці, як цифрові плеєри, фотоапарати, та камери, виникає потреба у зручному зберіганні, передачі та стисненні цифрової інформації [1]. Існує кілька відомих форматів стиснення аудіо, таких як MP3 (MPEG-1 layer 3), Ogg Vorbis, WMA (Windows Media) та RealAudio. На сьогоднішній день MP3 є найбільш поширеним форматом, але це не завжди пов'язано з його кращою якістю звучання чи компресією. Просто він з'явився на ринку раніше інших форматів, а виробники мультимедійних пристроїв не поспішають впроваджувати нові аудіоформати. З тестів стає зрозуміло, що Ogg Vorbis є конкурентним форматом [2,3].

Як MP3, так і Vorbis використовують розбиття сигналу на сегменти та дискретне перетворення Фур'є з використанням психоакустичного фільтру і алгоритмів кодування, але реалізуються вони по-різному.

Метою роботи є підвищення коефіцієнту стиснення аудіоданих шляхом розробки методу стиснення на основі вейвлет-перетворень.

Результати дослідження

Загальна схема алгоритму стиснення сегменту звукового сигналу складається з наступних кроків:

- На вхід процедури стиснення подається два масиви $\{X_k^l\}$ і $\{X_k^r\}$, $k = 0, 1, \dots, N-1$ (розглянемо випадок із стерео сигналом).
- Застосовуємо до кожного з масивів вибране нами перетворення (DWT або DFT) і одержуємо масиви коефіцієнтів $\{C_k^l\}$ і $\{C_k^r\}$
- Застосовуємо стерео з'єднання коефіцієнтів $\{C_k^l\}$ і $\{C_k^r\}$ і одержуємо три масиви $\{e_k\}$, $\{m_k\}$ і $\{\alpha_k\}$.
- Оцінюємо розрядність зберігання величин $\{m_k\}$ і $\{\alpha_k\}$, використовуючи масив експонент.

- Виконуємо RPPM-кодування масиву $\{e_k\}$
 - Виконуємо бітову упаковку масивів $\{m_k\}$ і $\{\alpha_k\}$ з потрібним числом розрядів.
- В результаті виходить блок, що складається з 3-х підблоків (Рис. 1).



Рис. 1 – Три підблоки для стиснення сегменту звукового сигналу

Алгоритм відновлення сигналу симетричний алгоритму стиснення, але з використанням зворотних перетворень.

В ході експерименту було взято два звукових файли формату WAV для оцінки ефективності різних методів стиснення. Обидва файли були оцінені за допомогою однієї і тієї ж функції оцінки розрядності зберігання коефіцієнтів.

У першому експерименті використовувався сильно насичений сигнал, тобто в спектрі сигналу були присутні частоти широкого діапазону. Для цього експерименту була обрана композиція "Inside" від Sting (2003) з альбому "Sacred Love". Результати цього експерименту представлені в табл.1.

Таблиця 1 – Результати першого експерименту

Перетворення	Динамічне програмування	Коефіцієнт стиснення (від початкового об'єму)
DWT	Не використовувалось	22,1%
DWT	Використовувалось	22,0%
DFT	Не використовувалось	23,7%
DFT	Використовувалось	21,5%

У другому експерименті був узятий не сильно насичений сигнал, в тому сенсі, що в спектрі сигналу присутній тільки вузький діапазон частот (частина композиції Falling - Alicia Keys (2001) Songs in A minor). У табл.2 приведені результати проведеного експерименту.

Таблиця 2 – Результати другого експерименту

Перетворення	Динамічне програмування	Коефіцієнт стиснення (від початкового об'єму)
DWT	Не використовувалось	19.9%
DWT	Використовувалось	19.8%
DFT	Не використовувалось	20.7%
DFT	Використовувалось	18.9%

З експериментів видно, що для вейвлет-розкладання динамічне програмування не принесло значного вигаду в об'ємі та якості стиснення. Це через те, що сама структура вейвлет-перетворення локалізує осциляції сигналу незалежно від розбиття на сегменти. Проте, розкладання Фур'є, яке залежить від розбиття сигналу на сегменти, показує певний вигад як в об'ємі, так і в якості звучання за рахунок динамічного програмування.

Щодо реалізації, можна відзначити такі аспекти:

- Масив експонент можна зберігати з точністю 3-4 біта, що дозволяє досягти стиснення майже у два рази без втрат якості. RPPM-кодування масиву експонент в середньому займає лише 8% об'єму початкового сигналу, що значно ефективніше за зберігання масиву без стиснення.

- Допустимі оцінки точності елементів масивів мантис і фаз також можуть бути підібраними для стиснення, зберігаючи сумарний об'єм цих масивів на рівні 16-17% від початкового. При використанні певних алгоритмів кодування, цей об'єм може бути зменшений у два рази, що призводить до загального об'єму стисненого сигналу на рівні близько 15% від початкового, що є конкурентоспроможними результатами на сьогоднішній день порівняно з кращими кодеками, такими як Ogg Vorbis і MPEG Layer-3. Наприклад, при стисненні WAV-файлу в MP3 з якістю 192kbps об'єм MP3-файла становить 14% від початкового об'єму.

Висновки

Запропонований метод стиснення аудіоданих на основі вейвлет-перетворення дозволяє зменшити їх обсяг до 20-25% від початкового, зберігаючи прийнятну якість. Це досягається завдяки стисненню сегментів сигналу, де масив коефіцієнтів розділяється на масив мантис та масив експонент.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Загальні відомості про цифрове аудіо — посібник від А до Я | Sony Ukraine [Електронний ресурс]. – Режим доступу: <https://www.sony.ua/electronics/support/understanding-digital-audio> – Дата звернення: 19.03.2025.
2. Аносов А. О., Проценко М. М., Дубинко О. Л., Павлушко М. Я. Застосування вейвлет-перетворення для аналізу цифрових сигналів / А. О. Аносов, М. М. Проценко, О. Л. Дубинко, М. Я. Павлушко // Сучасний захист інформації. – 2021. – Т. 5, № 2. – С. 12–19. – Режим доступу: <https://journals.dut.edu.ua/index.php/dataprotect/article/view/1792>. – Дата звернення: 19.03.2025.
3. Abduljabbar, Z. S., Ahmed, Z. J., Al-Saleh, A. H. H. (2020). A Survey paper on Lossy Audio Compression Methods. Vol. 15, No. 2, 34-45. Available at: <https://www.semanticscholar.org/reader/bb418de208b7375ca2960aa2a69fa3eaa14793fd>.

Кончинський Андрій Олексійович — студент групи 1АКІТР-24м, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: andrey123210111@gmail.com

Гармаш Володимир Володимирович - канд. техн. наук, доцент кафедри автоматизації та інтелектуальних інформаційних технологій, Вінницький національний технічний університет, м. Вінниця.

Konchynskyi Andrii Oleksiiovych - Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia,, email : andrey123210111@gmail.com

Garmash Volodymyr Volodymyrovych - Ph.D. (Eng), Assistant Professor of Department of Automation and Intelligent Information Technology, Vinnytsia National Technical University, Vinnytsia.