

УДК 004.032.3

В. М. ШУТКО, доктор техн. наук, професор;

Г. І. ГАЙДУР, канд. техн. наук, доцент;

М. І. ПОПОВ, аспірант;

Є. В. ПРИЛЄПОВ, аспірант,

Державний університет телекомунікацій, Київ

## ДОСЛІДЖЕННЯ НОВІТНІХ АЛГОРИТМІВ СТИСНЕННЯ АУДІОДАНИХ НА БАЗІ ЛІНІЙНИХ ПЕРЕТВОРЕНЬ

**Розглянуто роль і значення методів стиснення, або компресії, аудіоданих у сучасних телекомунікаціях. Проаналізовано покладені в основу зазначених методів алгоритми компресії, що використовують лінійні перетворення, і подано порівняльну характеристику найважливіших методів стиснення щодо їх застосовності на практиці. За результатами дослідження наведено обґрунтування перспективності подальшого розвитку цього напрямку наукових пошуків.**

**Ключові слова:** MP3; AAC; Vorbis Ogg; Muse Pack; VBR; стиснення аудіоданих; компресія; кодек; лінійні перетворення; банк фільтрів; психоакустична модель; бітрейт; частота дискретизації; кодування.

### Вступ

Сьогодні використовуються численні аудіокодеки, зручні для виконання найрізноманітніших завдань. Проте здебільшого в основу їх покладено традиційні методи стиснення, або компресії. Тому для дослідження новітніх алгоритмів у цій сфері відкриваються широкі можливості.

### Аналіз останніх досліджень і публікацій

Аналіз джерел за темою даного наукового пошуку показує, що проблема дослідження та створення алгоритмів аудіостиснення була і лишається предметом пильної уваги вітчизняних науковців, таких як В. М. Шутко, М. О. Шутко, О. О. Колганова, О. Д. Пономарчук, А. Ю. Тропченко, О. М. Амброс, В. М. Бондаренко. З огляду на дедалі вищу швидкість розвитку техніки та появу нових галузей застосування стиснення аудіоданих постає потреба проаналізувати наявні вирішення та дослідити їх спроможність задовольнити усі потреби сучасних технологій.

### Постановка проблеми

Незважаючи на стрімкий прогрес технологій передавання даних і пристроїв для їх зберігання, пошук нових методів стиснення не втрачає актуальності. Із розширенням можливостей компресії підвищуються очікування щодо якості аудіосигналів та швидкості їх передавання — характеристик, особливо важливих, коли йдеться про використання радіосистем у галузі авіації. Адже під час польоту обмін супровідною інформацією має бути гранично чіткий, із мінімально припустимою затримкою.

**Мета статті** — подати огляд і аналіз сучасних методів аудіокомпресії на базі лінійних перетворень, дослідити їх універсальність і обґрунтувати необхідність розробки нових алгоритмів компресії.

### Виклад основного матеріалу

Стиснення звукових даних — це тип стиснення (кодування) даних, що застосовується з метою скорочення обсягу аудіофайлів або зменшення смуги пропускання для потокового аудіо [1]. Алгоритми стиснення звукових файлів реалізуються в комп'ютерних програмах, які називають аудіокодеками. Загальні алгоритми стиснення в разі звукових даних неефективні, вони унеможливають роботу в реальному часі. Тому знадобились спеціальні методи аудіокомпресії [2].

Нині маємо кілька лідерів серед аудіоформатів. Найбільшого поширення набув формат MP3. Проте з моменту його появи минуло багато часу, з'явилися нові, значно досконаліші методи стиснення.

Технічні характеристики популярних форматів подано в таблиці.

Що ж до вибору кодека, то він істотно залежить від конкретного завдання. Визначальними в цьому питанні стають система та програмне забезпечення, обсяг пам'яті та характеристики смуги пропускання даних. Узявши до уваги весь комплекс наявних характеристик та кожен з них зокрема, можна зробити вдалий вибір.

Розвиток методів стиснення даних та психоакустики поступово призвів до того, що стандарт MP3 став обмежувати можливості реалізації нових ідей у відеокодуванні. З огляду на це Інститут Фраунгофера (Fraunhofer IIS), а також компанії Dolby, AT&T, Sony та Nokia 1997 року розробили новий формат стиснення аудіоінформації — **Advanced Audio Coding (AAC)**, який увійшов до складу стандартів MPEG-2 і MPEG-4 [3].

Основні особливості, що вигідно відрізняють AAC від MP3, такі:

1) підтримка повнішого набору форматів (до 48 каналів) і частот дискретизації звуку (від 8 до 96 кГц);

Технічні характеристики аудіоформатів

Формат стиснення аудіоданих	Алгоритм стиснення	Частота дискретизації, кГц	Бітрейт, кбіт/с	Затримка, мс	Режим кодування			Мультиканальність
					CBR	VBR	Stereo	
MP3 (MPE G-1; 2; 2.5; Audio Layer III)	MDCT, Hybrid Subband	8; 11.025; 12; 16; 22.05; 24; 32; 44.1; 48	8; 16; 24; 32; 40; 48; 56; 64; 80; 96; 112; 128; 144; 160; 192; 224; 256; 288; 320	>100	Так	Так	Так: Dual, Mid/Side, Intensity	Ні
AAC	MDCT, Hybrid Subband (AAC-HE)	8 ... 192	8 ... 529 (stereo)	20 ... 405	Так	Так	Так: Dual, Mid/Side, Intensity, Parametric	До 48 каналів
Vorbis Ogg	MDCT	Від 1 Гц до 200 кГц	Змінний	>100	Так/ABR	Так	Так: Dual, Lossless, Phase, Point (Intensity)	До 225 каналів
Muse Pack	Subband	32; 37.8; 44.1; 48	3 ... 1300	?	Ні	Так	Так	До 8 каналів

2) значно ефективніший банк фільтрів (гібридний банк фільтрів MP3 було замінено простішим MDCT — модифікованим дискретним косинусним перетворенням);

3) вихід за колишні межі варіації частотно-часового розширення банку фільтрів, що дало змогу поліпшити кодування транзєнтів (перехідних процесів сигналу) та стаціонарних ділянок;

4) вища якість кодування частот понад 16 кГц;

5) додаткові опції стандарту, які збагачують його можливості: технологія формування шуму в часовій області (TNS), прогнозування MDCT-коефіцієнтів у часі (long term prediction), режим параметричного кодування стереосигналу (parametric stereo), синтез шумів (perceptual noise substitution), технологія відновлення високих частот (SBR).

Завдяки переліченим властивостям формат AAC дозволяє досягати достатньо гнучкого та ефективного кодування звуку, але через високу популярність MP3 досі не набув значного поширення, хоча активно використовується у продуктах багатьох великих компаній.

**Vorbis Ogg** — порівняно новий (2002 р.) універсальний формат аудіокомпресії. Психоакустична модель, яка використовується в Ogg, схожа на MP3, але математична обробка та практична реалізація цієї моделі істотно відмінні.

Формат має чимало переваг, однією з яких є велика кількість аудіоканалів. Ogg підтримує до 225 окремих каналів із частотою дискретизації до 192 кГц та розрядністю до 32 біт. Завдяки цьому кодек ідеально забезпечує кодування 6-канального звуку DVD-Audio. Vorbis Ogg як sample accurate формат гарантує відсутність зсуву чи втрат окремих семплів після кодування та декодування. Це

дуже важливо при застосуванні в системах із потоковими аудіоданими.

Vorbis Ogg використовує змінний бітрейт, нижнє значення якого може варіювати навіть до 1 кбіт/с, а максимальне — досягати 400...700 кбіт/с. Таку саму гнучкість має частота дискретизації — користувачеві дозволяється зробити вибір від 2000 до 192 000 Гц [4].

**Muse Pack** (mpp, mp+, mpc, MPEG+) — формат для стиснення аудіоданих, який підпадає під GNU General Public License. Забезпечує якісне стиснення (не менш 160 кбіт/с) на високих бітрейтах, значно перевищуючи MP3 [5]. Головні переваги формату такі:

1) кодек не виконує другого dct-перетворення і фактично не обтяжений артефактами pre-echo, на відміну від форматів, згаданих раніше;

2) високоефективні алгоритми змінного бітрейту [6] (для простіших ділянок кодер відводить невеликий бітрейт, а для складніших — значно більший); кодування тиші з мінімальними втратами — менш ніж 1 кбіт/с (для MP3 у режимі VBR це значення досягає 32 кбіт/с, а для Vorbis Ogg — 2 кбіт/с);

3) використання гнучкої психоакустичної моделі, перевагою якої є динамічний фільтр для низьких частот на базі фреймів;

4) найкраща модель кодування оптимізованих таблиць Хаффмана.

Проте цей формат не позбавлений і недоліків. До них можна віднести відсутність підтримки аудіоданих розрядністю понад 24 біт та невисоку якість на низьких та середніх (до 160 кбіт/с) бітрейтах.

Вибір методу стиснення аудіоданих залежить від багатьох факторів, таких як характеристики

системи, смуга пропускання та вимоги щодо якості сигналу.

Виходячи з очікувань стосовно якості та розбірливості інформації, можна здійснити порівняльний аналіз методів стиснення на певних бітрейтах.

Коли йдеться про режим кодування True VBR, тобто режим із цільовою якістю, в ідеалі на виході отримуються аудіодані зі змінним бітрейтом (для складних частин сигналу потрібно більше бітів, а для простих навпаки) і постійною якістю [7].

Для низького (25...128 кбіт/с) рівня бітрейту найкращим вибором може стати AAC. Завдяки своїй гнучкості та прогнозованості MDCT-кофіцієнтів цей формат забезпечує добрі показники стиснення, досягаючи значно вищої, ніж у конкурентів, якості мовного сигналу. Для середнього (128...192 кбіт/с) рівня бітрейту підходить як AAC, так і Ogg, причому різниця між ними стрімко зменшується зі зростанням бітрейту. У разі високого (192...320 кбіт/с) бітрейту відмінність між AAC та Vorbis майже непомітна, але тут незаперечним лідером стає Muse Pack, саме за високого бітрейту виявляються його властивості як принципово відмінного від усіх інших алгоритмів алгоритму стиснення та психоакустичної моделі.

#### Висновки

Вибір кодека для стиснення аудіоданих залежить від багатьох параметрів системи та програмного забезпечення. Важливу роль відіграють також характеристики смуги пропускання даних (для стиснення потокового аудіо) та вимоги щодо якості й розбірливості мовного сигналу чи музичної композиції.

Стиснення аудіоданих не втрачає актуальності, незважаючи на вагомий результати теоретичних

досліджень і практичних розробок у цій сфері. Кожний метод стиснення поряд із перевагами, має й недоліки, але всі вони знаходять застосування за певних умов. Звідси випливає нове актуальне питання щодо створення *універсального методу стиснення*, який однаково задовольнятиме всі типи систем із різними характеристиками.

#### Список використаної літератури

1. Тропченко, А. Ю. Методы сжатия изображений, аудиосигналов и видео: учеб. пособие / А. Ю. Тропченко, А. А. Тропченко.— СПб.: СПбГУ ИТМО, 2009.— 108 с.

2. Шутко, В. М. Методи та засоби стиснення інформації: навч. посібник / [В. М. Шутко, М. О. Шутко, О. О. Колганова, О. Д. Пономарчук].— К.: НАУ, 2012.— 168 с.

3. Advanced Audio Coding [Електронний ресурс].— Режим доступу:

[www.audiocogs.org/codecs/aac](http://www.audiocogs.org/codecs/aac)  
<http://www.audiocogs.org/codecs/aac/>

4. Vorbis audio compression [Електронний ресурс].— Режим доступу:

[www.xiph.org/vorbis](http://www.xiph.org/vorbis)  
<http://www.xiph.org/vorbis>

5. What is Musepack? [Електронний ресурс].— Режим доступу:

[www.musepack.net](http://www.musepack.net)  
<http://www.musepack.net/>

6. Брузицкий, Т. О современных форматах кодирования аудио [Електронний ресурс].— Режим доступу:

[www.websound.ru](http://www.websound.ru)

7. Сэлмон, Д. Сжатие данных, изображений и звука / Д. Сэлмон.— М.: Техносфера, 2004.— 368 с.

**Рецензент:** доктор техн. наук, професор В. Г. Сайко, Державний університет телекомунікацій, Київ.

*В. М. Шутко, Г. И. Гайдур, М. И. Попов, Е. В. Прилепов*

#### ИССЛЕДОВАНИЕ НОВЕЙШИХ АЛГОРИТМОВ СЖАТИЯ АУДИОДАНЫХ НА БАЗЕ ЛИНЕЙНЫХ ПРЕОБРАЗОВАНИЙ

Рассмотрена роль и значение сжатия аудиоданных, приведены анализ существующих алгоритмов компрессии с использованием линейных преобразований. На основе полученной информации проведено сравнение и исследованы применимость каждого из существующих методов. По результатам исследования приведены обоснования перспективности дальнейшего развития этого направления.

**Ключевые слова:** MP3; AAC; Vorbis Ogg; Muse Pack; VBR; сжатие аудиоданных; компрессия; кодек; линейные преобразования; банк фильтров; психоакустическая модель; битрейт; частота дискретизации; кодирование.

*V. M. Shutko, G. I. Gaydur, M. I. Popov, E. V. Pryleпов*

#### THE INVESTIGATION OF NEW ALGORITHMS COMPRESSION OF AUDIO DATA ON THE BASE LINEAR TRANSFORMATIONS

This article discusses role and significance of audio compression. It shows analysis of the existing compression algorithms with linear transformations usage. Done comparison and researched applicability of each of the existing methods based on the analysis. Research substantiates promising of the further development of this area.

**Keywords:** MP3; AAC; Vorbis Ogg; Muse Pack; VBR; audio compression; compression; codec; linear transformations; filter bank; psychoacoustic model; bitrate; sample rate; coding.