

**ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ МЕТОДІВ СТИСНЕННЯ ЗВУКОВИХ ДАНИХ**

*В статті наведені результати досліджень характеристик основних руйнівних методів і алгоритмів стиснення звукових даних, проведена порівняльна оцінка їх властивостей для визначення найбільш ефективних форматів компресії звукових даних з втратами інформації.*

*Ключові слова: стиснення даних, методи стиснення звукових даних, MP3 (MPEG Layer 3), WMA (Windows Media Audio), OGG Vorbis, AAC (Advanced Audio Coding), WAV (Waveform Audio File Format).*

O.V. NECHYPORENKO

Cherkassy State Technological University, Cherkassy, Ukraine

**RESEARCH EFFICIENCY POPULAR AUDIO COMPRESSION METHOD**

*Abstract - The article presents the research results of basic characteristics of destructive methods and algorithms for compressing audio data. Features compression and lossy audio data marked advantages of compressed digital audio, compactness and durability. The basic properties of the most popular formats of audio compression losses: MP3 (MPEG Layer 3), WMA (Windows Media Audio), OGG Vorbis, AAC (Advanced Audio Coding). Analysis of audio compression was performed using two converters: Freemake Audio Converter and Format Factory. The analysis showed that the compression of the high quality of the first converter was best AAC format, and the second - OGG. Upon compression of low quality to best the first converter proved formats MP3 and AAC. By the time compression was the fastest compression to WMA, but very slow - MP3. However, for high quality WMA got worst compression ratio of both converters.*

*Keywords: data compression audio data compression methods, MP3 (MPEG Layer 3), WMA (Windows Media Audio), OGG Vorbis, AAC (Advanced Audio Coding), WAV (Waveform Audio File Format).*

**Вступ**

Передача та зберігання інформації вимагають досить великих витрат. І чим з більшою кількістю інформації нам доводиться мати справу, тим дорожче це коштує. На жаль, велика частина даних, що потрібно передавати по каналах зв'язку і зберігати, має не саме компактне представлення. Однак таке найбільш просте у використанні подання даних вимагає вдвічі-втричі, а іноді і в сотні разів більше місця для їх збереження та передачі, ніж насправді потрібно [1].

Сучасні канали зв'язку вже давно дозволяють пересилати великі масиви даних за порівняно невеликий час, проте самою повільною і донині залишається передача даних між кінцевим користувачем і постачальником послуг зв'язку. Такі обсяги даних, які займає в нинішній час мультимедійна інформація, передавати за звичними каналами зв'язку доведеться дуже довго.

Збільшенню обсягів інформації також сприяє введення інформаційної надмірності, яка використовується для підвищення надійності функціонування обчислювальних систем. До негативних наслідків введення надмірності відносять підвищення інтенсивності потоку і збільшення ймовірності виникнення помилок, а також зменшення ймовірності їх виправлення [2].

Тому стиснення (компактне представлення) даних – це один з найбільш актуальних напрямків розвитку сучасних комп'ютерних систем і телекомунікацій.

**Аналіз досліджень та публікацій**

Метою стиснення є отримання компактного способу представлення початкових даних, який мінімізує об'єм займаної пам'яті за допомогою деякого їх перетворення. Для кожного з сучасних видів інформації (текст, бази даних, програмні коди, мультимедійна інформація) існують свої найбільш відповідні методи стиснення. Результати досліджень і аналізу методів для стиснення текстових даних наведені в роботах [3, 4], для стиснення мультимедійних даних – в роботах [5, 6].

Всі існуючі методи стиснення інформації умовно поділяють на дві основні групи. Перша з них об'єднує методи стиснення інформації без втрат (lossless compression), а друга група – стиснення з втратами (lossy compression). Методи стиснення з втратою інформації зазвичай забезпечують набагато вищий ступінь стиснення, ніж оборотні методи, але їх не можна застосовувати до текстових документів, баз даних і, тим більше, до програмного коду. Неруйнівні методи стиснення інформації (без втрати) застосовують для стиснення будь-яких типів даних [4].

Загально визнані неруйнівні методи стиснення даних, такі як кодування довжин серій (RLE), статистичні та словникові методи, можуть бути використані для компресії звукових файлів без втрат, але результат істотно залежить від конкретних аудіо даних. Деякі звуки будуть добре стискатися за допомогою RLE, але погано статистичними алгоритмами. Іншим звукам більше підходить статистичне стиснення, а при словниковому підході, навпаки, може статися розширення.

Можна домогтися кращих результатів при стисненні звуку з втратою частини звукової інформації, розвиваючи методи компресії, які враховують особливості сприйняття звуку. Вони видаляють ту частину даних, яка залишається нечутною для органів слуху. Такі підходи (придушення пауз і ущільнення) широко застосовуються на практиці і входять до багатьох міжнародних стандартів стиснення.

Мета роботи – проведення аналізу основних характеристик та оцінки ефективності роботи основних руйнівних методів стиснення звукових даних.

**Матеріали досліджень**

Стиснення (компресія) звукових даних являє собою процес зменшення швидкості цифрового потоку за рахунок скорочення статистичної та психоакустичної надмірності цифрового звукового сигналу.

Основна ідея, на якій засновані всі методики стиснення звукового сигналу з втратами – нехтування тонкими деталями звучання оригіналу, що лежать поза межами сприйняття людського вуха. Тут можна виділити кілька моментів [7]:

- рівень шуму, маскуючий ефект (звукове стиснення базується на факті, що людське вухо фактично втрачає звуки, які лежать в тому ж діапазоні частот, що і більш гучний звук; цей ефект носить назву маскуючого, він змінюється з відмінністю в гучності і частоті звуку);

- розподіл смуги звукових частот на підсмуги, кожна з яких далі обробляється окремо (програма кодування виділяє найгучніші звуки в кожній смузі і використовує цю інформацію для визначення прийнятного рівня шуму для цієї смуги; гучний звук в одній смузі може вплинути на маскуючий ефект і на прилеглі смуги);

- використання психоакустичної моделі (стиснення засновано на видаленні свідомо нечутних частот з більш ретельним збереженням звуків, які добре розрізняються людським вухом);

- використання суміщеного стерео (слуховий апарат людини може визначити напрямок лише середніх частот – високі та низькі звучать як би окремо від джерела. Ці фонові частоти можна кодувати в моно сигнал. Також використовується відмінність в складності потоків в каналах).

На основі цих ідей будуються алгоритми стиснення, що дозволяють досягати ступеня компресії 10:1 або вище практично без втрати в якості звучання. При кодуванні задають необхідний рівень компресії, а алгоритми стиснення отримують необхідне значення рівня стиснення за рахунок втрати якості.

Необхідний рівень стиснення зазвичай вказують у вигляді величини потоку даних (bit rate). Бітрейт – це обсяг інформації в одиницю часу, тобто скільки бітів ми витрачаємо на кодування звуку тривалістю 1 с [8].

Серед існуючих форматів стиснення звукових даних з втратами можна відзначити четвірку найбільш популярних форматів – MP3, AAC, OGG Vorbis і WMA.

MP3 (MPEG Layer 3) – ліцензований формат файлу для зберігання звукової інформації. Це найпопулярніший формат стиснення звукових даних на сьогоднішній день. Формат MP3 заснований на маскуючому ефекті і призначений для значного зменшення обсягу даних, який необхідний для прослуховування запису і збереження якості звуку максимально близького до вихідного. Такий файл із середнім бітрейтом (128 Кбіт/с) призводить до отримання файлу приблизно в 1/11 від початкового файлу, який має середній бітрейт близько 1411 Кбіт/с. Якість отриманого файлу залежить від бітрейту, з яким можуть створюватися mp3-файли. Найбільш прийнятним є бітрейт від 128–320 Кбіт/с, але для нормального якісного звучання достатньо лише 256 Кбіт/с.

Слід враховувати, що у форматі MP3 програми, що стискають звук, не є стандартизованими, тобто кожен грамотний програміст може реалізувати свою схему стиснення. А стандартам підкорюються тільки декодери, це призводить до того, що якість відтворення формату MP3 далеко не завжди залежить від плеєра – програвача файлу.

Особливістю MP3 є Variable Bit Rate – кодер змінює ступінь стиснення залежно від характеру звуку. Такий підхід призводить до зменшення підсумкового розміру файлу або, при збільшенні вимог до якості, при тому ж розмірі файлу дозволяє отримати краще звучання.

AAC (Advanced Audio Coding) – формат звукового файлу з меншою втратою якості при кодуванні, ніж MP3 при однакових розмірах. Формат також дозволяє стискати без втрати якості. Спочатку створювався як наступник MP3 з поліпшеною якістю кодування. Формат AAC, офіційно відомий як ISO / IEC 13818-7, вийшов у світ в 1997 як нова, сьома, частина родини MPEG-2. Існує також формат AAC, відомий як MPEG-4 Частина 3.

OGG Vorbis – вільний формат стиснення звуку, що вийшов влітку 2002 року. OGG Vorbis абсолютно безкоштовний, відкритий і вільний від патентів. OGG Vorbis володіє найбільш сучасною і якісною психоакустичною моделлю, через що співвідношення бітрейт/якість значно нижче, ніж у конкурентів. Як результат – якість звуку краще, а розмір файлу менше. При цьому OGG Vorbis підтримує до 255 окремих каналів з частотою дискретизації до 192 кГц і розрядністю до 32 біт (чого не дозволяє жоден інший формат стиснення звукових даних з втратами).

OGG Vorbis за замовчуванням використовує змінний бітрейт, при цьому його параметри не обмежені якимись жорсткими значеннями, і він може варіюватися навіть на 1 кбіт/с. Найвищий бітрейт чітко не обмежений, тому при максимальній якості кодування він може варіювати від 400 до 700 кбіт/с. Такою ж гнучкістю володіє частота дискретизації – користувачам надається можливість зробити будь-який вибір у межах від 2 до 192 кГц.

WMA (Windows Media Audio) – ліцензований формат файлу, розроблений компанією Microsoft для зберігання і трансляції звукової інформації. Номінально формат WMA характеризується хорошою здатністю стиснення, що дозволяє йому конкурувати за параметрами з форматами MP3, OGG Vorbis і AAC. Якість формату WMA не є однозначно еквівалентною, а перевага навіть перед MP3 однозначною, як це стверджується компанією Microsoft.

Характеристика, основні переваги і недоліки розглянутих популярних форматів стиснення звукових даних наведені в таблиці 1 та таблиці 2 відповідно.

## Характеристика основних форматів стиснення звукових даних

Формат	Квантування, біт	Частота дискретизації, кГц	Число каналів	Бітрейт, Кбіт/с	Ступінь стиснення
MP3	плаваючий	до 48	2	до 320	~11:1 з втратами
AAC	плаваючий	до 96	до 48	до 529	з втратами
OGG Vorbis	до 32	до 192	до 255	до 1000	з втратами
WMA	до 24	до 96	до 8	до 768	2:1, є версія без втрат

Таблиця 2

## Аналіз властивостей основних форматів стиснення звукових даних

Формат	Переваги	Недоліки
MP3	<ol style="list-style-type: none"> <li>1. Заклав основу ідеї спрощення сигналу шляхом використання психоакустики, має кращі позиції за співвідношенням розмір/якість.</li> <li>2. Величезна поширеність і популярність.</li> <li>3. Доступність MP3 практично на всіх існуючих комп'ютерних платформах.</li> <li>4. Використання змінного бітрейта, що дозволяє при меншому розмірі файлу отримати краще звучання.</li> <li>5. Велика кількість кодерів.</li> </ol>	<ol style="list-style-type: none"> <li>1. Обмежений бітрейт (320 Кбіт/с), різке падіння якості при зниженні бітрейта.</li> <li>2. Неможливість професійного застосування (кожне нове збереження розкодованого MP3-запису призводить до погіршення якості).</li> <li>3. Відсутність у стандарті DRM-можливостей (управління авторськими правами).</li> <li>4. Помітне падіння якості для сильно стиснутих музичних файлів.</li> <li>5. Комерціалізованість (кожен виробник, який створює новий MP3-кодер, платить відрахування).</li> </ol>
AAC	<ol style="list-style-type: none"> <li>1. Має безліч доповнень, спрямованих на поліпшення якості вихідного звукового сигналу.</li> <li>2. Дозволяє зберігати в закодованому звуковому сигналі інформацію про авторські права (watermarks).</li> <li>3. Передбачає три профілю кодування (Main, LC і SSR), що впливають на час кодування і якість одержуваного цифрового потоку.</li> <li>4. У порівнянні з MP3 помітно збільшена ефективність компресії, а якість звучання файлу при бітрейті 128 кбіт/с порівняна з якістю 192 кбіт/с MP3.</li> <li>5. Дозволяє створювати багатоканальні файли.</li> <li>6. При кодуванні на низьких бітрейтах є можливість створення файлів AAC HE (високої ефективності).</li> </ol>	<ol style="list-style-type: none"> <li>1. Технологія watermarks є перешкодою на шляху поширення файлів, створених за допомогою AAC.</li> <li>2. Не є зворотно сумісним (NBC – non backwards compatible) з рівнями MPEG-1 не дивлячись на те, що являє собою продовження MPEG-1 Layer I, II, III.</li> <li>3. Велика кількість кодерів AAC несумісні один з одним на тлі тривалого процесу стандартизації.</li> <li>4. Формат AAC підтримує порівняно невелика кількість програми для кодування звукових файлів.</li> </ol>
OGG Vorbis	<ol style="list-style-type: none"> <li>1. Використовує оригінальний математичний алгоритм і власну психоакустичну модель.</li> <li>2. Безкоштовний, відкритий і вільний від патентів.</li> <li>2. Розрахований на стиснення даних на всіх бітрейтах без обмежень в режимі змінного бітрейта. Передбачена можливість зміни бітрейта потоку без його декодування.</li> <li>3. Є можливість кодування декількох каналів аудіо (до 255), можливість редагування вмісту файлів.</li> <li>4. Підтримується потокове відтворення (streaming) в Інтернеті.</li> <li>5. Дає в середньому помітно кращі результати кодування на середніх і високих бітрейтах в порівнянні з MP3 (160 Кбіт/с і вище).</li> </ol>	<ol style="list-style-type: none"> <li>1. Для зберігання даних використовується власний універсальний формат.</li> <li>2. Використання на низьких бітрейтах майже не виправдано.</li> <li>3. Існує всього декілька кодерів OGG.</li> <li>4. Після конвертації MP3 в OGG Vorbis до частин аудіосигналу, відкинутих MP3 кодеком, додаються частини, відкинуті кодеком OGG Vorbis, що призводить до погіршення якості.</li> <li>5. Відмова від патентованих технологій не дозволяє кодеку OGG Vorbis домогтися надвисоких результатів.</li> </ol>
WMA	<ol style="list-style-type: none"> <li>1. Здатний давати більш високу якість звучання, ніж MP3 на порівнянних бітрейтах, особливо, на низьких (нижче 128 Кбіт/с).</li> <li>2. WMA-файли займають менше місця, ніж MP3, при порівнянній якості.</li> <li>3. Наявність вбудованої DRM – цифрової системи управління авторськими правами.</li> <li>4. Передбачено кодування без втрати якості, багатоканальне кодування об'ємного звуку і голосу.</li> <li>5. Висока популярність серед користувачів Windows.</li> </ol>	<ol style="list-style-type: none"> <li>1. Помітно відстає реалізація апаратними програвачами, в порівнянні з MP3.</li> <li>2. Вбудована DRM-система обмежує свободу користувача.</li> <li>3. Кодер WMA всього один.</li> <li>4. Формат WMA має велику кількість помилок різного роду.</li> </ol>

Аналіз стиснення звукових файлів проводився з використанням двох конвертерів: Freemake Audio Converter [9] і Format Factory [10]. Результати експерименту для цих конвертерів наведені в таблиці 3 і таблиці 4 відповідно. Як початкові дані були обрані звукові файли у форматі WAV загальним розміром 71866 Кб.

WAV (Waveform Audio File Format) є основним форматом для багатьох систем відтворення цифрового звуку і використовується як стандартний формат звукових файлів в персональних комп'ютерах. До того ж, він має солідний набір специфікацій. Його повна назва – Microsoft RIFF/WAVE – Resource Interchange File Format/Wave – формат файлів передачі ресурсів / хвильова форма.

Таблиця 3

**Компресія звукових даних Freemake Audio Converter**

Формат стиснення	Бітрейт, Кбіт/с	Об'єм початкового файла, Кб	Об'єм стисненого файла, Кб	Коефіцієнт стиснення, %	Час стиснення, с
MP3	128	71866	2995	4,16	25
MP3	320	71866	7487	10,41	25
WMA	128	71866	3023	4,20	5
WMA	320	71866	7539	10,49	5
OGG	128	71866	3343	4,65	16
OGG	320	71866	7139	9,93	17
AAC	128	71866	3027	4,21	22
AAC	320	71866	3589	4,99	23

Таблиця 4

**Компресія звукових даних Format Factory**

Формат стиснення	Бітрейт, Кбіт/с	Об'єм початкового файла, Кб	Об'єм стисненого файла, Кб	Коефіцієнт стиснення, %	Час стиснення, с
MP3	128	71866	2996	4,16	11
MP3	320	71866	7489	10,42	9
WMA	128	71866	3227	4,49	5
WMA	320	71866	12898	17,94	5
OGG	128	71866	4703	6,54	19
OGG	320	71866	4703	6,54	18
AAC	128	71866	2996	4,16	7
AAC	320	71866	7482	10,41	6

Результати аналізу показали, що при стисненні з високою якістю (320 Кбіт/с) для першого конвертера найкращим виявився формат AAC, а для другого – OGG з показниками коефіцієнта стиснення 4,99% і 6,54% відповідно. При стисненні з низькою якістю (128 Кбіт/с) для першого конвертера найкращими виявилися формати MP3 і AAC. За часом компресії найшвидшою виявилася компресія в WMA, а самою повільною – MP3. Однак для 320 Кбіт/с WMA отримав найгірший показник коефіцієнта стиснення і в першому конвертері (10,49%) і в другому (17,94%).

**Висновки**

1. Розглянуто особливості компресії звукових даних з втратами. Очевидні переваги стисненого цифрового звуку: компактність (малі габарити сучасних носіїв цифрових даних) і довговічність (при зберіганні на цифровому носії дані на ньому не спотворюються з плином часу).

2. Проаналізовано основні властивості найбільш популярних форматів стиснення звуку, таких як MP3, WMA, OGG Vorbis і AAC. Поширення і подальший розвиток технологій lossy-кодування відкриває широкі можливості поширення та зберігання звукових даних.

3. Результати аналізу показали, що MP3 вимагає максимального значення бітрейта для максимальної якості треку, в такому випадку MP3 не зрізає частотний діапазон і вносить найменшу кількість спотворень у вихідний сигнал. WMA менше за інших схильний до кліпінга – хрипів на максимальній гучності, але й має найгірший коефіцієнт стиснення. OGG має хороші показники і навіть при низькому значенні бітрейту намагається зберегти весь спектр частот вихідної композиції. На першому місці за поширеністю є MP3 формат, друге місце займає WMA і OGG Vorbis, а третє – формат AAC. Розглянуті технології стиснення звукової інформації MP3, OGG, AAC і WMA – зручні та ефективні алгоритми, широко поширені і заслужено впроваджені в багатьох електронних звукових пристроях.

**Література**

1. Шелухин О.И. Сжатие аудио и видео информации [Электронный ресурс] : учебное пособие / О.И. Шелухин, А.В. Гузеев. – М., 2012. – Режим доступа : <http://kibia.ru/teachers/sheluhin/4.pdf>
2. Нечипоренко О. В. Синтез специализированных цифровых устройств по условию надежности на

основе систем счисления с постоянным числом единиц : монография / Нечипоренко О. В. – Черкассы : ВУЭМ, 2009. – 109 с.

3. Тарасов О.В. Дослідження ефективності блочно-статистичного методу стиснення інформації / О.В. Тарасов, Є.В. Онопко // Системи обробки інформації. – 2012. – Випуск 4 (102), том 1. – С. 72–75.

4. Нечипоренко О. В. Системный анализ и оценка методов сжатия данных для баз данных лазерных технологических комплексов / О. В. Нечипоренко, С. А. Миценко // Вісник Хмельницького національного університету. – 2014. – № 1. – С. 94–100.

5. Корпань Я. В. Методи та алгоритми компактного представлення графічної інформації в комп'ютерних системах / Я.В. Корпань // Технологічний аудит та резерви виробництва. – 2015. – № 3/2 (23). – С. 32–36.

6. Корпань Я. В. Аналіз методів та алгоритмів компресії-декомпресії цифрових відеоданих / Я. В. Корпань // Вісник Хмельницького національного університету. – 2015. – № 3. – С. 175–179.

7. Алгоритмы сжатия [Електронний ресурс]. – Режим доступу : [http://mf.grsu.by/UchProc/livak/en/po/comprsite/theory\\_souvid.html](http://mf.grsu.by/UchProc/livak/en/po/comprsite/theory_souvid.html)

8. Методы сжатия и компрессии аудио-сигналов (виды, отличия, использование) [Електронний ресурс]. – Режим доступу : <http://studopedia.org/3-100199.html>

9. Freemake Audio Converter [Електронний ресурс]. – Режим доступу : <http://www.freemake.com/ru/downloads/>

10. Format Factory [Електронний ресурс]. – Режим доступу : <http://format-factory.ru.softonic.com/>

#### References

1. O.I. Sheluhin, A.V. Guzeev, Szhatie audio i video informatsii. Uchebnoe posobie. M. 2012. URL: <http://kibia.ru/teachers/sheluhin/4.pdf>

2. Nechiporenko O. V. Sintez spetsializirovannyih tsifrovyyih ustroystv po usloviyu nadezhnosti na osnove sistem schisleniya s postoyannym chisлом edinit: monografiya, Cherkassy: VUEM, 2009. 109 s.

3. O.V. Tarasov, E.V. Onopko, Doslidzhennya efektyvnosti blochno-statystychnoho metodu stysnennya informatsiyi, Systemy obrobky informatsiyi, 2012, vypusk 4 (102), tom 1. S. 72-75

4. O. V. Nechyporenko, S. A. Mitsenko, Sistemnyiy analiz i otsenka metodov szhatiya dannyih dlya baz dannyih lazernyih tehnologicheskikh kompleksov, Herald of Khmelnytskyi National University. 2014. № 1. S. 94–100.

5. Korpan' Ya. V. Metody ta alhorytmy kompaktnoho predstavlennya hrafichnoyi informatsiyi v komp'yuternykh systemakh, Tekhnolohichnyy audyt ta rezervy vyrobnytstva, 2015. № 3/2 (23). S. 32-36.

6. Korpan' Ya. V. Analiz metodiv ta alhorytmiv kompresiyi-dekompresiyi tsyfrovyykh videodanykh. Herald of Khmelnytskyi National University. 2015. № 3. S. 175–179.

7. Algoritmyi szhatiya. URL: [http://mf.grsu.by/UchProc/livak/en/po/comprsite/theory\\_souvid.html](http://mf.grsu.by/UchProc/livak/en/po/comprsite/theory_souvid.html)

8. Metodyi szhatiya i kompressii audio-signalov (vidyi, otlichya, ispolzovanie). URL: <http://studopedia.org/3-100199.html>

9. Freemake Audio Converter. URL: <http://www.freemake.com/ru/downloads/>

10. Format Factory. URL: <http://format-factory.ru.softonic.com/>

Рецензія/Peer review : 10.7.2015 р.

Надрукована/Printed : 30.8.2015 р.

Рецензент: д.т.н., проф., Рудницький В.М.