

УДК 621.391

О.О. Колганова, к.т.н.
В.М. Шутко, д.т.н.
Я.А. Рибачук**ГІБРИДНІ МЕТОДИ СТИСНЕННЯ ДАНИХ**

Національний авіаційний університет

*Стаття присвячена аналізу гібридних методів стиснення звукових сигналів, вивчення їх особливостей та оцінці можливостей такого підходу у стиснення зображень***Ключові слова:** *гібридні методи, стиснення звукових сигналів.*

Вступ. Стрімкий розвиток систем зв'язку міжнародних та національних телекомунікаційних систем характеризується зростанням кількості високошвидкісних цифрових каналів передачі інформації та широкомасштабним підвищенням ефективності роботи пристроїв цифрового стиснення каналів, що зумовлене стрімким зростанням об'єму передач інформації в мережах телекомунікацій. Сьогодні все частіше для підвищення ефективності використання цифрових каналів застосовуються методи стиснення інформації. Необхідно відзначити, що найбільший обсяг інформації, що передається цифровими каналами зв'язку, припадає на мультимедійний трафік [1]. Тому актуальним є стиснення саме мультимедійних потоків даних, зокрема мовних і відео сигналів, а також графічних даних.

Сьогодні розроблено декілька десятків алгоритмів стиснення даних, які структурно та семантично відрізняється один від одного [2]. Наприклад, для стиснення відео існують відеокодеки H.264, VP-8, MPEG-4; фотокодеки JPEG, JPEG2000 – для стиснення графічних даних. Сучасні практичні потреби висувають ряд вимог до алгоритмів стиснення сигналів, зокрема, якість мовлення, швидкість передавання сигналів, складність виконання, алгоритмічні затримки, завадостійкість. У зв'язку з цим, останнім часом з'явилося ряд нових алгоритмів стиснення таких мовних сигналів, наприклад CS-ACELP [3], MELP [4], NetCoder [5].

Проведений аналіз алгоритмів стиснення показує, що вибір алгоритму для стиснення даних є достатньо складним завданням, вирішення якого залежить від ряду характеристик алгоритму стиснення та вимог системи, в якій алгоритм буде функціонувати. Основною та найбільш істотною характеристикою є якість. Крім того, до важливих характеристик, на які потрібно звертати увагу при виборі алгоритму, належать швидкість передачі, алгоритмічні затримки, що вносяться алгоритмом при стисненні сигналу, та складність виконання алгоритму.

Задачею будь-якого алгоритму стиснення сигналів є одержання цифрової послідовності, яка потребує мінімальної швидкості передачі і з якої декомпресор зможе відновити вхідний сигнал з мінімальними втратами. Насамперед, слід зазначити, що представлення сигналів цифрової формі, яке містить дискретизацію та квантування [3], передбачає втрату деякої частини інформації вхідного сигналу (стиснення з втратами).

Аналіз літературних джерел. У літературі поняття «гібридні» методи стиснення, насамперед, стосується мовних кодеків. Під кодуванням мається на увазі перетворення мовного сигналу в деякий «інший» сигнал, який можна представити з меншою кількістю розрядів, що в підсумку підвищить швидкість передачі даних.

Кодування звичайно підрозділяють на три види:

1. кодування безпосередньо реалізації мовного (Wave Form Codec);
2. вимір, кодування і передача на прийомну сторону параметрів, за якими вже на приймальній стороні проводиться синтез цього (штучного) сигналу. Такі системи називають вокодерними (Source Codec);
3. гібридні способи кодування, тобто поєднання першого і другого способів кодування.

Перераховані вище методи використовуються залежно від особливостей створення та відтворення мовного сигналу. При створенні мовного сигналу до уваги беруться такі властивості сигналу:

1. під час тривалих змін:
 - зміна амплітуди;
 - поділ мови на дзвінки та глухі звуки, а також паузи мовчання;

- особливості мовної та фонетичної структури.

2. під час короткочасних змін:

- короткочасна кореляція;
- особливості структури тону звуку (для дзвінких звуків);
- особливості структури шуму (для глухих звуків).

При відтворенні мовного сигналу до уваги беруть локальний спектральний динамічний діапазон та слухове маскування. Якщо не використовується жодна з перелічених властивостей, то до сигналу можна застосовувати прості методи стиснення форми сигналу для одержання високоякісного мовлення.

Алгоритми стиснення, побудовані згідно з методом гібридного стиснення, є найновішими та найефективнішими серед сучасних алгоритмів стиснення. Вони використовують найдоступнішу інформацію про мовний сигнал для покращання якості мовлення та зниження швидкості результуючого коду. Зокрема вони використовують замаскований слуховий шум, слухову частоту розширення, слухову фазу нечутливості, складові зміни енергії, довгострокові звукові властивості тракту та інформацію про крок під час квантування [6].

При гібридному методі стиснення, мовний сигнал дискретизується, з метою одержання необхідних параметрів мови. Однак замість безпосереднього стиснення висоти, модуляції та інших характеристик мовного сигналу, ці параметри використовуються для синтезу фрагменту мови, з якого вони були одержані. Синтезований фрагмент мови, звичайно тривалістю 20 мс, порівнюється з початковим зразком. Якщо вони збігаються з прийнятним допуском, то параметри мови залишаються без змін. Якщо ж реальний та синтезований фрагменти відрізняються більш ніж на задану величину, то параметри мови коректуються для досягнення необхідного збігу. Кінцевим етапом цього механізму зворотного зв'язку є виконання процедури аналізу шляхом синтезу [7]: спроба відредувати параметри мови для забезпечення синтезу форми сигналу, якомога більш наближеної до початкової форми. Після остаточного визначення значень параметрів мови, компресор зіставляє одержані характеристики з характеристиками кодової книги. При виявленні збігу замість значення параметра використовується його положення в кодовій книзі, що суттєво зменшує обсяг інформації, яку потрібно передати [8]. Описаний вище алгоритм лежить в основі побудови та функціонування сімейства алгоритмів CELP (Code Excited Linear Predictive Coder) [9, 10]. Компресори CELP працюють на швидкостях нижче 4800 Кб/с та забезпечують задовільну якість мовлення. Головним недоліком компресорів типу CELP є їх висока складність та значні обчислювальні потреби.

Гібридний метод стиснення. Сьогодні найперспективнішим методом є гібридне стиснення мовних сигналів. Тому актуальним є створення нових алгоритмів стиснення мовних сигналів, які дозволять забезпечити прийнятну та високу якість мовлення при низьких швидкостях. Перспективним також є реалізація алгоритмів стиснення мовних сигналів на спеціалізованих процесорах. Особливу увагу необхідно приділити дослідженню та розробці універсальних алгоритмів стиснення мовних сигналів для кожного методу стиснення.

Розглянемо схеми гібридного звичайного аудіо-кодерів (рис. 1, 2).

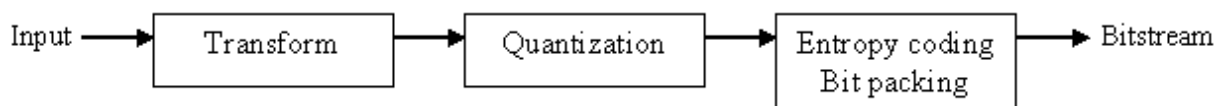


Рис. 1. Стандартний аудіо-кодер

Аналізуючи схему легко побачити суттєві недоліки традиційної схеми кодування. По-перше кодування форми хвилі. Це означає, що сигнали з різною формою хвилі можуть звучати однаково: шумові сигнали, інвертований сигнал, зміщений сигнал. Другий недолік – незалежне кодування фреймів, адже музика є сукупністю повторюваних видозмінюваних звуків.

Основна ідея реалізована у гібридному кодері (рис. 2) – поділ сигналу на три компоненти і використання різних типів кодування у залежності від специфіки компоненти:

1. Гармонічну, використовуючи перетворення, що добре локалізують частоти, оснований на перетворенні Фур'є;
2. Ударні, використовуючи вейвлет-перетворення, що має кращу часову локалізацію;
3. Шумову, застосовуючи кодування енергетичних огинаючих спектру.

До гармонічної компоненти застосовуються такі два підходи:

1. Представлення, основане на MDCT- масці:

- гармоніка – локально стаціонарна за часом виділена частина MDCT-квасіспектра;

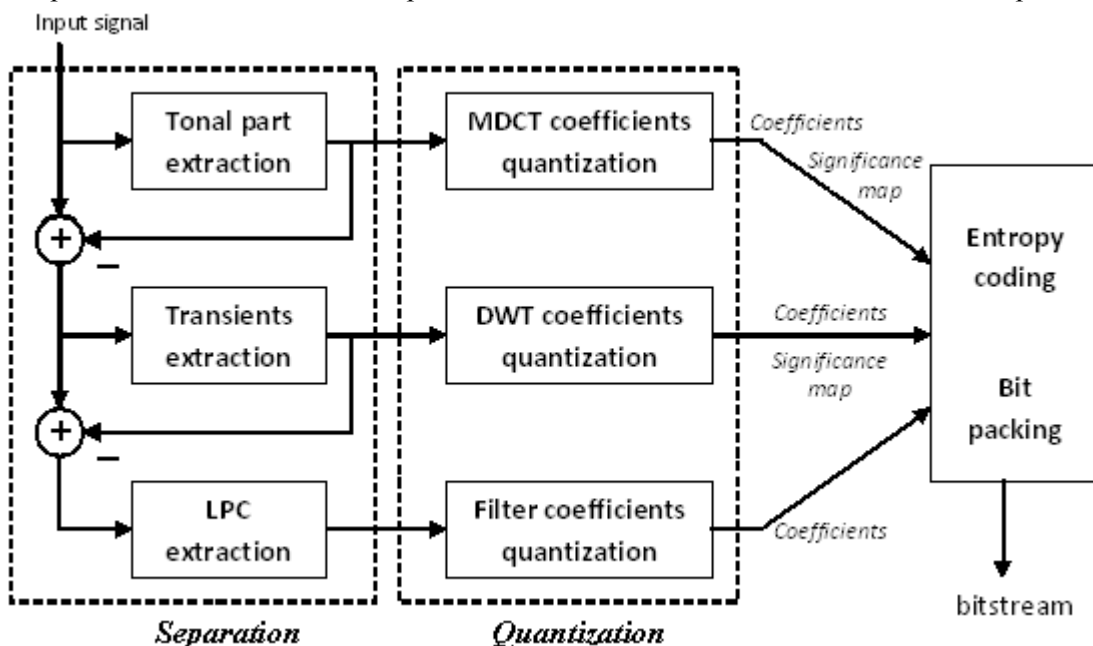


Рис.2. Гібридний кодек

- обнулення не ‘гармонічних’ коефіцієнтів;
- традиційне стиснення гармонічної частини.

2. Векторне представлення стиснення гармонік - при відновленні використовується інтерполяція (рис.3).

Перехідні сигнали (удари) проходять наступну обробку (рис.4). Проводиться виділення високоамплітудних вибросів у сигналі віддаленими гармоніками. Далі відбувається розкладі стиснення даних на базі одновірних ортогональних вейвлетів.

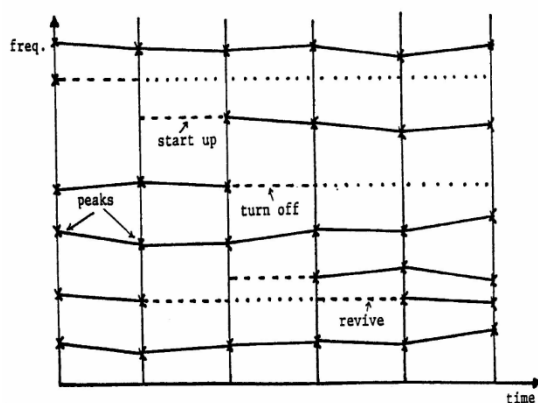


Рис. 3. Інтерполяція даних для відновлення гармонік

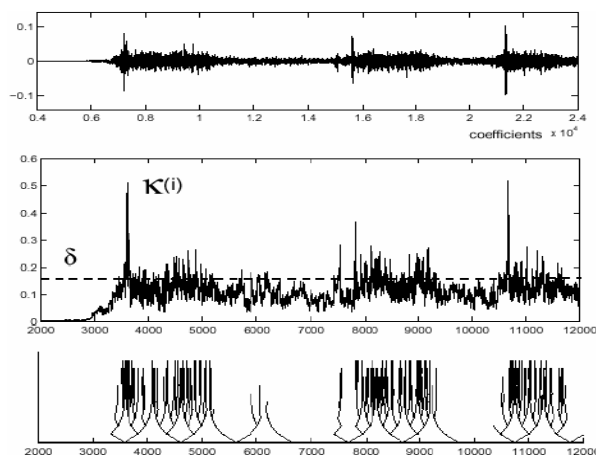


Рис.4. Виділення ударних гармонік

Залишок сигналу знаходиться за наступною формулою:

$$\text{Залишок} = \text{Сигнал} - \text{Гармоніки} - \text{Перехідні сигнали.}$$

Висувається гіпотеза, що залишок – це шумовий сигнал. Для кодування шуму застосовується LPC кодування спектральної огинаючої сигналу. Для реконструкції використовується фільтрація білого шуму.

Після розкладу сигналу одержують компоненти такої форми, як представлено на рис. 5.

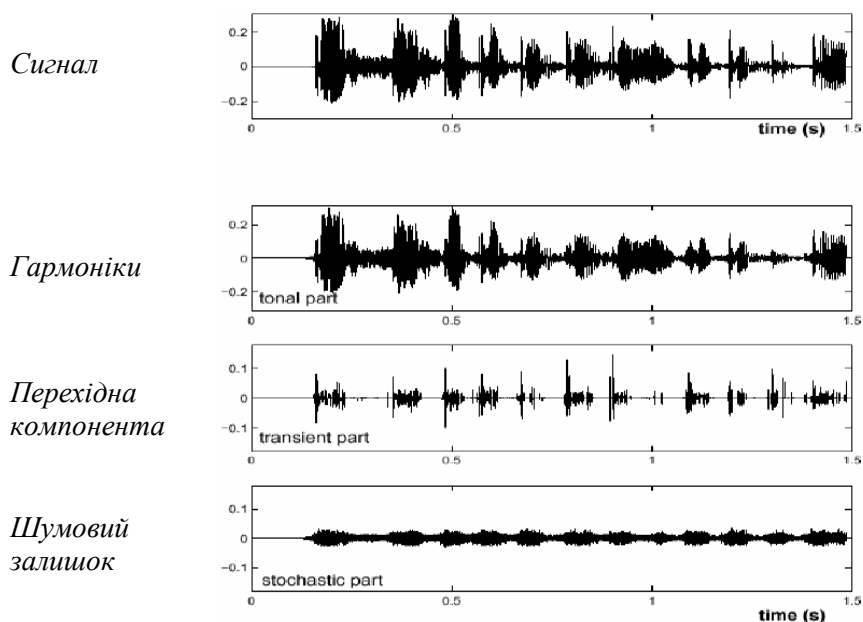


Рис. 5. Компоненти звукового сигналу

Висновки. З проведеного аналізу можемо зробити висновок, що основною перевагою гібридного кодування є адаптивне кодування, яке залежить від конкретних властивостей компонент сигналу. Але є й недоліки такого підходу:

- надлишкове представлення;

- адитивний синтез шуму передбачає не стійкість при ітераційному застосуванні.

Хоча гібридне кодування орієнтоване на велике стиснення, але дотеперішнього часу, абсолютно прозоре кодування з CD- якістю не досягнуто. Це спонукає вчених до пошуку можливостей вдосконалення описаного методу.

Гібридне стиснення безперечно має багато перспектив, хоча і потребує ще певних доробок. Цікавим є застосування такого методу розкладання сигналу й у стисненні графічних зображень. Адже будь-яке зображення теж має області плавних переходів кольорів та різкі стрибки амплітуд. Такий підхід може дати кращу якість стиснення. Звісно це призведе до ускладнення алгоритму і підвищення часу обробки даних, але такий метод все одно буде мати певну область застосування.

Літературні джерела:

1. Гольдштейн В.С., Пинчук А.В., Суховицкий А.Л. IP-Телефония. – М.: Радио и Связь, 2001. – 336 с.
2. Marwan Jabri. Design hits media standards wall // Electronic Engineering Times. – 2002.
3. ITU-T Recommendation G.729, Coding of speech at 8 kbit/s using conjugate-structure algebraic code-excited linear prediction (CS-ACELP), March 1996.
4. Lynn M. Supplee Alan. V. McCree. MELP: the new federal standard at 2400 bps. ICASSP 97.
5. www.audiocodes.com.
6. Li W., Sridhar A., Teng T. Comparison of Speech Coding Algorithms: ADPCM, CELP and VSELP. – Fall, 1999.
7. Gersho A. and Wang S. Recent Trends and Techniques in Speech Coding // Proc. 24th. Asilomar Conf., Pacific Grove, Nov. 1990.
8. Хелд Г. Сокращение задержки голоса по IP. – LAN, 2000. – № 07.
9. Schroeder M.R. and Atal B. Code-Excited Linear Prediction (CELP): High Quality Speech at Very Low Bit Rates // Proc. ICASSP-85. – Tampa, Apr. 1985. – P. 937.
10. Un C.K. and Magill D.T. The Residual-Excited Linear Prediction Vocoder with Transmission Rate below 9.6 kbits/s // IEEE Trans. COM-23(12). – Dec. 1975. – P. 1466.
11. Ватолин Д. Сжатие аудиоданных. Общие принципы и устройство MP3. - <http://compression.ru/video>