Носулько Ігор Віталійович,
судовий експерт сектору звуко- та відеозапису відділу досліджень у сфері інформаційних технологій Харківського науково-дослідного експертно-криміналістичного центру МВС України

ВПЛИВ СТИСНЕННЯ ЗВУКОВОЇ ІНФОРМАЦІЇ НА ІДЕНТИФІКАЦІЙНЕ ДОСЛІДЖЕННЯ ДИКТОРА

Вид стиснення звукових даних одна з найголовніших речей при ідентифікаційному дослідженні звукової інформації (голос та мовлення диктора, акустичних сигналів та середовищ). Стиснення звукових даних використовується задля зменшення обсягу аудіофайлів або для можливості зменшення смуги пропускання потокового аудіо (наприклад сингали GSM). Алгоритми стиснення звукових файлів реалізуються в комп'ютерних програмах, які називаються аудіокодеками. Як і загальному випадку стиснення даних, розрізняють стискування звуку без втрат, що робить можливим відновлення вихідних даних без спотворень, і стиск із втратами, у якому таке відновлення неможливе. Алгоритми стиснення з втратами дають більший ступінь стиснення, наприклад audio CD може вмістити не більше години «нестиснутої» музики, при стисненні без втрат CD вмістить майже 2 години музики, а при стисненні з втратами при середньому бітрейті – від 7 до 10 годин.

Стиснення без втрат. Складність стиснення звуку без втрат у тому, що звук є надзвичайно складним у своїй структурі. Одним із методів стиснення є пошук зразків та їх повторень, проте цей метод не ефективний для більш хаотичних даних, якими є, наприклад, оцифрований звук або фотографії. Найбільш поширеними форматами стиснення без втрат є: Free Lossless Audio Codec (FLAC), Apple Lossless, MPEG-4 ALS, Monkey's Audio, і TTA.

Стиснення із втратами. Стиснення із втратами має надзвичайно широке застосування. Крім комп'ютерних програм, стиск із втратами використовується в потоковому аудіо в DVD, цифровому телебаченні та радіо, потоковому медіа в інтернеті, та при записі на відеокамери.

Новацією цього методу стиснення стало використання психоакустики виявлення компонентів звучання, які сприймаються слухом людини. Прикладом можуть бути або високі частоти, які сприймаються лише за достатньої їх потужності, або тихі звуки, що виникають одночасно або відразу після гучних звуків і тому

піддаються ефекту маскування – тихі звуки маскуються гучними – такі компоненти звучання можуть бути передані менш точно, або взагалі не передані.

Принцип стиснення полягає у зниженні точності частин звукової інформації, що практично невиразне для слуху більшості людей. Звуковий сигнал розбивається на рівні за тривалістю відрізки, кожен із яких після обробки упаковується у свій кадр (кадр). Розкладання в спектр вимагає безперервності вхідного сигналу, у зв'язку з цим для розрахунків використовується попередній і наступний кадр. У звуковому сигналі є гармоніки з меншою амплітудою і гармоніки, що лежать поблизу більш інтенсивних – такі гармоніки відсікаються, тому що середньостатистичне вухо людини не завжди зможе визначити присутність або відсутність таких гармонік. Така особливість слуху називається ефектом маскування. Також можлива заміна двох і найближчих піків одним усередненим (що, як правило, і призводить до спотворення звуку). Критерій відсікання визначається вимогою до вихідного потоку. Оскільки весь спектр актуальний, високочастотні гармоніки не відсікаються, лише вибірково видаляються, щоб зменшити потік інформації рахунок розрідження спектра. Після спектральної «зачистки» застосовуються математичні методи стиснення та упаковка у кадри.

Стиснення в певній частотній області може полягати в тому, що замасковані або нульові компоненти взагалі не запам'ятовуються, або кодуються з меншим дозволом. Наприклад, частотні компоненти до 200 Гц і більше 14 кГц можуть бути закодовані з 4-бітною розрядністю, тоді як компоненти в середньому діапазоні - з 16 бітною. Результатом такої операції стане кодування із середньою розрядністю 8-біт, проте результат буде значно кращим, ніж при кодуванні всього діапазону частот з 8-бітною розрядністю. Однак очевидно, що перекодовані з низькою роздільною здатністю фрагменти спектру вже не можуть бути відновлені в точності, і, таким чином, втрачаються безповоротно.

Головним параметром стиснення з втратами є бітрейт, що визначає ступінь стиснення файлу та, відповідно, якість. Розрізняють стиснення із постійним бітрейтом (англ. *Constant BitRate*- CBR), змінним бітрейтом (англ. *Variable BitRate*- VBR) і усередині бітрейтом (англ. *Average BitRate*- ABR).

Найбільш поширеними форматами стиснення із втратами є: AAC, ADPCM, ATRAC, Dolby AC-3, MP2, MP3, Musepack Ogg Vorbis, WMA та інші.

Приклади впливу стиснення звукової інформації:

- **втрата деталей голосу**: при стисненні голосової інформації відбувається втрата важливих деталей голосу, таких як тембр, інтонація та динаміка;
- **зміна частотних характеристик**: процес стиснення може викликати зміни в частотних характеристиках голосу. Деякі діапазони частот можуть бути відфільтровані, що може призвести до спотворень відтворення голосу. Така «модифікація» звукової інформації диктора, впливає на результат ідентифікаційного дослідження;
- **вплив на фонематичну розбірливість**: стиснення може впливати на чіткість артикуляції та вимови звуків, особливо на високих рівнях стиснення. Це може призвести до меншої розбірливості та менш точного відтворення мовлення;
- **поява артефактів та шумів**: втрати, які виникають під час стиснення, можуть призводити до появи артефактів і шумів у голосовому сигналі. Це може бути специфічним для різних алгоритмів стиснення і погіршує вихідне значення сигнал/шум звукової інформації диктора для ідентифікаційного дослідження.

Спираючись на наведене, можемо зробити висновок, що для найбільш коректної ідентифікації диктора, необхідна наявність якомога більшої точності спектральної складової звукового сигналу, в ідеалі використання форматів без стиснення чи стиснення без втрат. При використанні звукової інформації при стисненні з втратами, доречно використовувати фонограми тільки з високим бітрейтом.

Список використаних джерел

1. Бабак В.П., Хандецький В.С., Шрюфер Е. Обробка сигналів, К. Либідь, 1996, 392 с.

2. Müller C. Speaker Classification I: Fundamentals, Features, and Methods. Springer London, Limited, 2007.

3. Khalifeh A. F., Al-Tamimi A.-K., Darabkh K. A. Perceptual evaluation of audio quality under lossy networks. *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, м. Chennai, 22–24 берез. 2017 р. 2017. URL: https://doi.org/10.1109/wispnet.2017.8299900 (дата звернення: 09.04.2024).

4. Записування і відтворення інформації. Терміни та визначення: ДСТУ 2737–94. [Чинний від 31.08.94]. К.: Держстандарт України, 1994.