

УДК 004.8

АНАЛІЗ АУДІОСИГНАЛІВ ЗА ДОПОМОГОЮ МЕЛ-ЧАСТОТНИХ КЕПСТРАЛЬНИХ КОЕФІЦІЄНТІВ

Ракілов К.Г., студент, k17_rakilov.kh@server.odessa.ua, МАУП «ДП»
Шибасва Н.О., к.т.н., доцент, shibaeva@server.odessa.ua, Коледж «Сервер»

Звуковий тон та гармонія є основними елементами звуку, відображаючи його природу та сприйняття людьми. Представлення в часово-частотному просторі, такі як спектрограми та короткочасне перетворення Фур'є, допомагають аналізувати характеристики звукових сигналів у різних контекстах. Звукові шкали, такі як рівнотемперовані, впливають на спосіб, як звук сприймається та використовується в різних музичних та звукових контекстах. Логарифмічні спектрограми та інші інструменти аналізу розширюють нашу здатність розуміти та використовувати звук, надаючи більш точні та повні зображення його характеристик у різних областях. [1]

Розробка програмного забезпечення для аналізу аудіосигналів з використанням мел-частотних кепстральних коефіцієнтів на сьогоднішній день стоїть на передньому краї індустрії звукової обробки. Цей сегмент досліджень є одним з найгарячіших у галузі. Історія його розвитку має свої видатні моменти, включаючи винайдення фонографа в XIX столітті та наступний бум цифрової телефонії у XX столітті. Варто відмітити, що в сучасному світі широкий спектр застосування аудіотехнологій стає ще актуальнішим завдяки прогресу у розробці програмного забезпечення, спрямованого на аналіз аудіосигналів за допомогою мел-частотних кепстральних коефіцієнтів. Від використання у кінематографії з тривимірним звуком до особистих пристроїв відтворення, ця технологія стає не просто важливою, але й визначальною в повсякденному житті [2].

Найпоширеніша задача кепстрального аналізу ставить метою порівняння подібності тембру двох аудіо файлів. Для отримання необхідних характеристик спочатку треба використати формулу перетворення Фур'є:

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-\frac{2\pi i}{N}kn} \quad (1)$$

Наступним кроком є обчислення паспортизації аудіосигналу:

$$M(f) = 2595 \cdot \log_{10} \left(1 + \frac{f}{700} \right) \quad (2)$$

Після цих розрахунків застосовується дискретне косинусне перетворення (DCT):

$$C_q = \sum_{n=0}^{N-1} E'(n) \cdot \cos\left(\frac{\pi q}{N} \cdot \left(n + \frac{1}{2}\right)\right) \quad (3)$$

Результат рекурсивного пошуку подібних за тембром аудіофайлів за допомогою кепстрального аналізу зображено на рисунку 1.

File Path	Timbre Similarity
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\drum_break.wav	
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\drum_break_slower.flac	98.81%
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\cut.mp3	90.41%
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\cookies.wav	90.07%
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\funky.wav	88.80%
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\fx1\FX_12.flac	87.75%
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\fx\FX_4.wav	87.53%
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\fx\FX_10.flac	87.01%
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\worm.wav	86.41%
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\think.wav	84.33%
C:\Users\nakil\Desktop\МАУП\7 Курс\Диплом\src\audios\fx\FX_9.wav	83.68%

Рисунок 1 – Результат порівняння файлів за допомогою кепстрального аналізу

Висновок. У результаті проведеного дослідження було надано алгоритм пошуку подібних за тембром аудіофайлів за допомогою мел-частотних кепстральних коефіцієнтів. Алгоритм працює з найпоширенішими форматами аудіофайлів та має розширення для порівняння за темпом та тональністю.

Список використаних джерел

1. Müller, Meinard & Ellis, Daniel & Klapuri, Anssi & Richard, Gaël. Signal Processing for Music Analysis. Selected Topics in Signal Processing, IEEE Journal of. 5, №10; 2011: p.2 — 5.
2. Gaël Richard, Paris Smaragdis, Sharon Gannot, Patrick A Naylor, Shoji Makino, et al.. Audio Signal Processing in the 21st Century. IEEE Signal Processing Magazine №X, 2023: p.1 — 10.