

ФОРМУВАННЯ ОЗНАК МУЗИЧНИХ КОМПОЗИЦІЙ ІЗ ВИКОРИСТАННЯМ МЕТОДУ ЦИФРОВИХ ВІДБИТКІВ

Вінницький національний технічний університет

Анотація

Запропоновано підхід по формуванню цифрових відбитків музичних композицій із використанням мел-частотних кепстральних коефіцієнтів.

Ключові слова: цифрові відбитки, змішані Гаусівські моделі, мел-частотні кепстральні коефіцієнти.

Abstract

The approach on formation of digital prints of musical compositions with use of mel frequency cepstral coefficients is offered.

Keywords: digital prints, mixed Gaussian models, mel frequency cepstral coefficients.

Вступ

На теперішній час разом із швидким поширенням аудіо та мультимедійної інформації зростає і кількість технологій, що використовуються для розпізнавання та відновлення мовлення. Очевидною стає проблема їх індексації для подальшого пошуку та розпізнавання. Системи сегментації звукових файлів та автоматичного розпізнавання мовлення використовуються в різноманітних програмах та додатках для пошуку потрібної композиції або автоматичного набору голосових команд. Одним з ефективних способів ідентифікації мультимедійних матеріалів є технологія, широко відома як «метод цифрових відбитків», названий аналогічно до методу ідентифікації людини по його відбитках пальців [1]. Велике зацікавлення до даної технології проявляють численні компанії, що надають сервіс в області доставки, пошуку і захисту мультимедійних матеріалів та інтелектуальної власності, зокрема із використанням мережі Інтернет. Розгляду одного із підходів по виділенню та формуванню ознак музичних композицій із використанням метода цифрових відбитків для подальшого їх розпізнавання присвячений даний матеріал.

Формування ознак музичних композицій

Існуючі на теперішній час методи отримання аудіовідбитків музичних композицій можна розділити на три великі групи [2]:

- 1) методи, що враховують локальні характеристики звукового сигналу;
- 2) методи, що аналізують довгострокові характеристики звукового сигналу;
- 3) методи на основі прихованих Марківських моделей.

Аналіз цих підходів до формування відбитків музичних композицій показує, що найкраще проблему розпізнавання аудіо матеріалів вирішують такі методи, що засновані на аналізі локальних параметрів звукових сигналів та визначають структуру процесу виділення набору ознак аналізованого аудіо сигналу. У роботах [3, 4] пропонується використати мел-частотні кепстральні коефіцієнти (MFCC) для отримання даних про музичний твір для його подальшого розпізнавання. На основі проведеного огляду різних існуючих методів класифікації для аналізу музичних сигналів було обрано метод класифікації на основі змішаних Гаусівських моделей із використанням мел-частотних кепстральних коефіцієнтів.

Для отримання компактного опису аудіозапису в більшості випадків обчислюють лише молодші кепстральні коефіцієнти, яких зазвичай зберігають тільки 10–15, щоб забезпечити незалежність отриманих параметрів від частоти основного тону, яка присутня в старших коефіцієнтах. Чим більше коефіцієнтів MFCC застосовується, тим точніше описується спектр сигналу. Разом з тим посилюється вплив швидких змін спектра, які корелюються із частотою основного тону та створює небажану мінливість даних. Молодші коефіцієнти описують загальну

форму спектра, тоді як інформація про частоту основного тону та спектральну структуру (spectral fine structure) міститься в старших коефіцієнтах. Коефіцієнти із 6 по 12, які описують вхідний спектр на досить детальному рівні, можуть виявитися надлишковими. Для врахування динамічних властивостей всієї спектральної обвідної використовують похідні кепстральних коефіцієнтів. Так як нас цікавить тільки спектральна обвідна, а не частота основного тону, а також компактність опису аудіозаписів, використання великої кількості коефіцієнтів MFCC не є доцільним.

Обчислення параметрів MFCC відбувається в декілька етапів.

1. Поділ відліків аудіозапису на фрейми.

2. Множення кожного з відліків фрейму на вагову функцію, наприклад, вікно Хеммінга. Застосування вагової функції дозволяє послабити «розтікання» спектру на межі фреймів.

3. Виконання швидкого перетворення Фур'є для обчислення коефіцієнтів амплітуди спектра [6]. Після застосування перетворення Фур'є в частотному поданні основну інформацію несуть низькі частоти. Високі частоти описують шум і несуттєві деталі.

4. Переведення значень частот в мел-шкалу. Шкала мел-частот є лінійною для частот нижче 1 кГц і логарифмічною шкалою для частот вище 1 кГц. Таким чином коефіцієнти MFCC актуалізують низькочастотну інформацію і усереднюють високочастотні складові сигнали.

5. Обчислення потужності спектра шляхом обчислення квадрата абсолютних значень, отриманих після перетворення Фур'є, та накладення на обчислені значення амплітуд спектра набору з трикутних вікон мел-шкали, що перекриваються і рівномірно розміщуються на шкалі мел-частот.

6. Логарифмування зміненого спектру з метою ущільнення його динамічного діапазону.

7. Виконання дискретного косинусного перетворення для зменшення кількості вихідних параметрів і декореляції компонентів.

Розробка ефективного програмного забезпечення є важливою задачею для формування ознак та розпізнавання музичних композицій. Для розробки програмного забезпечення було обрано бібліотеку машинного навчання mlpack, що базується на бібліотеці лінійної алгебри Armadillo та реалізована мовою програмування C++, яку також було використано для створення програми обробки музичних композицій.

Висновок

Запропонований підхід може бути використаний у комп'ютерних системах пошуку, формування ознак та розпізнавання музичних композицій за отриманим цифровим відбитком.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Cano P. Audio Fingerprinting: Concepts and Applications. / Cano P., Gómez E., Batlle E., Gomes L., Bonnet M. // Proceedings of 2002 International Conference on Fuzzy Systems Knowledge Discovery. – Singapore. – 2002.- pp.233–245.
2. Билобров С. В. Обзор методов идентификации аудио материалов с использованием аудио отпечатков / С. В. Билобров, А. А. Каргин // Вісн. Донец. ун-ту. Сер. А. Природн. науки. - 2005. - № 1. - С. 450-454.
3. Chauhan P.M., Desai N.P. Mel Frequency Cepstral Coefficients (MFCC) based speaker identification in noisy environment using wiener filter // Proceedings of International Conference on Green Computing Communication and Electrical Engineering. – Coimbatore, India, 2014. – pp. 1–5.
4. Ткаченко О. М. Ідентифікація музичного твору за його фрагментом / О. М. Ткаченко, О.Ф. Грійо Тукало, Ю. Л. Далекий // Міжнародна науково-практична конференція «Інформаційні технології та комп'ютерна інженерія». Вінниця. 2014. – с. 256-257.

Анастасія Миколаївна Щуровська - студентка групи ІКІ-19м факультету інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: 2ki15b.shchurovska@gmail.com.

Науковий керівник: **Микола Андрійович Очкуров** — старший викладач кафедри обчислювальної техніки, Вінницький національний технічний університет, м. Вінниця.

Anastasiia Shchurovska - students, Department of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: 2ki15b.shchurovska@gmail.com.

Supervisor: **Mykola A. Ochukov** — Senior lecturer of the Computer Techniques Chair, Vinnytsia National Technical University, Vinnytsia.