

## Розробка модуля розпізнавання музики для мобільного додатку

Вінницький національний технічний університет

### Анотація

*Розглядаються алгоритми аналізу та розпізнавання музичних композицій, а також способи їх застосування в рамках мобільного додатку.*

**Ключові слова:** мобільний додаток, Android, аналіз музичних звуків, спектрограма, розпізнавання музики.

### Abstract

*Algorithms of music analysis are considered. Also described the way how these technologies can be applied to developed mobile application.*

**Keywords:** mobile application, Android, musical sounds analysis, spectrogram, music recognition.

### Вступ

Сьогодні існує велика кількість додатків для створення музики, проте вони потребують досить глибоких знань теорії. Якщо користувач має бажання “відцифрувати” мелодію, яку щойно вигдав, буде складно знайти простий у використанні інструмент, який дозволить швидко перетворити награну голосом мелодію у музичну послідовність у цифровому форматі, що буде відіграна за допомогою вказаного музичного інструменту. Розроблюваний додаток вирішує дану проблему, а також дозволяє редагувати та доповнювати щойно створену мелодію для досягнення бажаного результату.

Метою є підвищення рівня автоматизації процесу створення музичних композицій шляхом поєднання алгоритмів розпізнавання музичних звуків та технологій їх синтезу, що забезпечує можливість програвання мелодій, які сприймає мікрофон пристрою за допомогою обраних музичних інструментів.

Об’єктом дослідження постають сучасні технології аналізу розпізнавання музичних звуків.

Предметом дослідження є алгоритмічне та програмне забезпечення для автоматизації процесів розпізнавання музичних композицій.

Головною задачею є розробка автоматизованої мобільної системи з використанням технологій аналізу та розпізнавання музичних звуків.

### Результати дослідження

Модуль розроблюваного додатку, що розглядається, призначений для розпізнавання існуючих музичних композицій для отримання їх метаданих, а саме: назва композиції, інформація про виконавця.

Під час запису сигналу за допомогою мікрофону отримується масив, який є представленням звукового сигналу в часовій області. Для отримання частотних характеристик сигналів, розвернутих у часі, необхідно використовувати дискретне перетворення Фур’є (ДПФ). ДПФ – це математичний метод аналізу Фур’є для дискретних сигналів. З його допомогою можна перетворити скінченний набір зразків сигналу, взятих з рівними проміжками часу, в список коефіцієнтів скінченної комбінації комплексних синусоїд, впорядкованих за частотою, беручи до уваги те, що ці синусоїди були дискретизовані з однією і тією ж самою частотою.

Дискретне перетворення Фур’є визначається за формулою 1

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} nk} \quad (1)$$

Позначення:

- $N$  – кількість значень сигналу, виміряних за період, а також кількість компонент розкладення;
- $x_n, n=0, \dots, N-1$ , - виміряні значення сигналу (в дискретних часових точках з номерами  $n=0, \dots, N-1$ ), які являються вхідними даними для прямого перетворення і вихідними для оберненого;
- $X_k, k=0, \dots, N-1$ , -  $N$  комплексних амплітуд синусоїдальних сигналів, з яких складається вхідний сигнал; являються вихідними даними для прямого перетворення і вхідними для оберненого; оскільки амплітуди комплексні, то за ними можна обчислити одночасно і амплітуду, і фазу;
- $\frac{|X_k|}{N}$  - звичайна амплітуда  $k$ -го синусоїдального сигналу;
- $\arg(X_k)$  - фаза  $k$ -го синусоїдального сигналу (аргумент комплексного числа);
- $k$  – індекс частоти. Частота  $k$ -го сигналу рівна  $\frac{k}{T}$ , де  $T$  – період часу, впродовж якого брались вхідні дані.

Одним з найбільш популярних чисельних алгоритмів для обрахування ДПФ є швидке перетворення Фур'є (ШПФ). Насправді, ШПФ представлено цілим набором алгоритмів. Серед них частіше за все використовуються варіанти алгоритму Кулі-Тьюкі. В основі цього алгоритму лежить принцип “Розділяй і володарюй”. В ході обчислень використовується рекурсивне розкладання вихідного ДПФ на дрібні частини. Пряме обчислення ДПФ для деякого набору даних  $n$  потребує  $O(n^2)$  операцій, а використання алгоритму Кулі-Тьюкі дозволяє вирішити ту саму задачу за  $O(n \log n)$  операцій (2).

$$X_k = \begin{cases} E_k + e^{-\frac{2\pi i}{N}k} O_k & \text{for } 0 \leq k < N/2 \\ E_{k-N/2} + e^{-\frac{2\pi i}{N}k} O_{k-N/2} & \text{for } N/2 \leq k < N. \end{cases} \quad (2)$$

$E_k$  – ДПФ вхідних даних парних позицій;

$O_k$  – ДПФ вхідних даних непарних позицій;

У результаті процесу ідентифікації процедура розпізнавання музики здійснюється за методом аналізу та синтезу музичних звуків, модель якого наведена на рис. 1:

1. Запис початкового звукового сигналу.
2. Перетворення сигналу в частотне подання.
3. Створення відбитку для порівняння.
4. Збереження у базу даних у необхідному форматі.



Рисунок 1 – Модель розпізнавання музичних композицій

Порівняння повних спектрограм погіршує продуктивність роботи програми, адже необхідно оброблювати декілька мегабайт нестиснених даних. Наявність сторонніх шумів також негативно

відображаються на ефективності роботи алгоритму. Проте, при відтворенні музичних доріжок з різного роду шумами, найкраще зберігаються піки спектрограм, наприклад, точки локального максимуму амплітуди. В результаті зі спектрограм отримуються “сузір’я” з такими піками. Тому при пошуку результату для запитуваного відрізка мелодії, необхідно знайти в базі даних відповідний трек, в якому співпадіння піків спектрограм виявилось найбільшим.

### Висновки

Таким чином, розглянуті алгоритми розпізнавання та аналізу музичних звуків забезпечують оптимізацію процесів автоматизації синтезу та аналізу музичних звуків. Розроблений додаток дозволяє спростити процес створення музичних композицій за допомогою використання розглянутого модуля розпізнавання музичних звуків. Такі послідовності можуть бути відіграні за допомогою обраних музичних інструментів, тому процес створення музичних композицій може стати досить зручним та інтуїтивно зрозумілим для користувача.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. High-Perfomance Audio [Електронний ресурс]. Режим доступу: URL: <https://developer.android.com/ndk/guides/audio/index.html>
3. Спектрограма [Електронний ресурс]. Режим доступу: URL: <https://ru.wikipedia.org/wiki/%D0%A1%D0%BF%D0%B5%D0%BA%D1%82%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B0>
4. An Industrial-Strength Audio Search Algorithm [Електронний ресурс]. Режим доступу: URL: <http://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf>
5. Как Яндекс распознаёт музыку с микрофона [Електронний ресурс]. Режим доступу: URL: <https://habrahabr.ru/company/yandex/blog/181219/>

**Ставицький Павло Валерійович**, студент групи ІПІ-17м, факультет інформаційних технологій та комп’ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: [morfly3000@gmail.com](mailto:morfly3000@gmail.com)  
Науковий керівник: **Войтко Вікторія Володимирівна**, доцент кафедри програмного забезпечення, Вінницький національний технічний університет, Вінниця, e-mail: [dekanfki@i.ua](mailto:dekanfki@i.ua)

**Pavlo Stavitskiy**, student of group IPI-17m, Faculty for Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: [morfly3000@gmail.com](mailto:morfly3000@gmail.com)  
Supervizor: **Viktoriia Voitko**, Associate Professor of Software Chair, Vinnytsia National Technical University, Vinnytsia, e-mail: [dekanfki@i.ua](mailto:dekanfki@i.ua)