

# АВТОМАТИЗАЦІЯ ТРАНСКРИБУВАННЯ ПОЛІФОНІЧНИХ МУЗИКАЛЬНИХ КОМПОЗИЦІЙ

Вінницький національний технічний університет

## Анотація

Наведено актуальність задачі автоматизації транскрибування музикальних композицій. Розглянуто існуючі методи розпізнавання акордів та запропоновано основні етапи розв'язання даної задачі. Запропоновано основні можливості розроблюваного веб-додатку з розпізнавання акордів.

**Ключові слова:** музикальний інформаційних пошук, розпізнавання акордів, спектрограма, багатотональний аналіз.

## Abstract

*This paper describes the relevance of an automated musical tracks transcribing problem. The existing chord recognition methods were reviewed. Main stages of solving this task were suggested. Basic features of proposed web-application for chord recognition were listed.*

**Keywords:** music information retrieval, chord detection, spectrogram, multi-pitch analysis.

Аналіз музичних творів, зокрема багатотональний аналіз (Multi-pitch analysis) та розпізнавання акордів є областями, що активно вивчаються та розвиваються. Багатотональний аналіз є фундаментальною проблемою в комп'ютерному аналізі та обробці аудіо сигналів. Він представляє значний інтерес для дослідників, що працюють над проблемою автоматизації транскрибування музики, розділенні джерел (source separation), вилученні мелодії, тощо. В обробці мови, це корисно для розпізнавання розмови багатьох осіб, а також є кроком на шляху до вивчення ефекту «коктейльної вечірки» (cocktail party effect або ефект концентрації уваги на одному об'єкті) [1].

Транскрибування музикальних композицій одна із задач з якою досить часто стикаються музиканти аматори, котрі не маючи професійної підготовки та значного досвіду, часто покидають цю справу. Проте, завдяки розвитку інформаційних технологій, стало можливим автоматизувати, розпізнавання акордів і задачі, які раніше були під силу лише професіоналам, тепер легко можуть виконувати і початківці.

Виявлення (розпізнавання) акордів є автоматизованим процесом присвоєння музичного акорду відповідній частині музичного твору. Воно знаходить застосування в широкому спектрі областей: як інструмент транскрибування музики в нотні записи, як автоматизований спосіб для заповнення бази даних музичними метаданими, які потім можуть бути використані користувачами, для побудови аудіовізуалізацій, що реагують на зміни гармонії, або як частина інших музичних інформаційно-пошукових завдань, таких як ідентифікація пісні, виявлення плагіату, шляхом зіставлення отриманих послідовностей акордів, жанрової або настроєвої класифікації, виявлення поточної позиції в даних нотних записах [2].

Задача розпізнавання акордів в музикальних композиціях складається в отриманні множини послідовності акордів із зазначенням позиції кожного з них. Таке представлення може бути проміжним етапом в роботі інших алгоритмів, а також може представляти цінність само по собі: за його допомогою можна індексувати музикальні твори для пошуку композицій за заданою послідовністю акордів, знаходити різноманітні аранжування однієї і тієї ж композиції. Дану інформацію можна використовувати також для визначення структури композиції, її розділення на більш крупні сегменти. Отже, вищевказана задача має цілком прикладний інтерес, наприклад, для людей, що самостійно навчаються грі на музичних інструментах (зокрема, на гітарі). Так, система розпізнавання акордів у звуці, що працює в реальному часі дозволить контролювати процес навчання та миттєво фіксувати помилки [2].

Об'єктом дослідження є методи та засоби автоматичного транскрибування музикальних композицій. Предметом дослідження є програмне забезпечення для транскрибування музикальних композицій. Метою роботи є створення зручного у використанні веб-додатку транскрибування музикальних композицій, який буде доступний широкому колу користувачів.

Перші спроби розпізнавання акордів були запропоновані ще у 90-х роках. У 1999 р. Т. Фудзісіма запропонував метод розпізнавання акордів без виділення нот, що звучать окремо, який став основою для багатьох існуючих на сьогодні алгоритмів [4]. Основною ідеєю є використання дискретного віконного перетворення Фур'є (ДПФ) для отримання спектрограми звукозапису. Далі, для кожного вектора спектрограми вказується відповідний йому акорд по методу найближчого сусіда. При наявних перевагах, метод має свої недоліки. Наявність шумів також сильно впливає на точність. Подальші роботи базуються на використанні різноманітних методів отримання спектрограми, використовуючи ДПФ, вейвлети тощо. Останнім часом все більше набуває популярності використання нейронних мереж, а в якості моделей класифікації рекурентні нейронні мережі, мережі TDNN [5 – 7].

Протягом останніх десятиліть були розроблені декілька підходів щодо розпізнавання акордів. Вони можуть бути класифіковані як навчання за допомогою зразків, машинне навчання та гібридні підходи. Більшість сучасних підходів щодо рішення таких задач за допомогою машинного навчання засновані на використанні прихованих марковських моделей, динамічних баєсових мереж та методах опорних векторів. Матеріали засновані на вищезгаданих підходах посідали вищі місця за результатами оцінювання у ході проведення щорічної компанії по оцінці алгоритмів музичного інформаційного пошуку MIREX Audio Chord Estimation (Music Information Retrieval Evaluation eXchange, MIREX).

Робота багатьох алгоритмів цифрової обробки звуку складається з нижченаведених етапів [2].

На першому етапі оброблюваний звук представляється у вигляді спектрограми, що показує, як змінюється розподілення звукової енергії по частотам з часом. Крім того, на цьому етапі можна врахувати наявність ритму в аналізованому звукозаписі.

На другому етапі виконуються перетворення, направлені на пригнічення спектру інструментів з невизначеною висотою звучання, а також робиться спроба врахувати наявність повторів у музикальній композиції.

На третьому етапі відбувається перехід від векторів до послідовностей акордів. Найпростішим варіантом тут є співставлення з векторами-шаблонами, а найбільш популярним способом врахувати часто використовувані послідовності акордів є використання прихованих марковських моделей. При цьому слід зазначити, що вибір навчальних композицій впливає на якість роботи і подальше використання таких моделей.

За допомогою розроблюваного веб-додатку автоматичного транскрибування користувач зможе заощадити свій час та витратити його на додаткові заняття, що допоможуть йому краще опанувати інструмент. Основними переваги веб реалізації є те, що: використання веб-платформи, дозволяє виконувати додаток на більшості сучасних пристроїв з мінімальними апаратними вимогами; забезпечується висока мобільність додатку (користувач може користуватися ним всюди, де є доступ до інтернету); не має потреби встановлювати на комп'ютер користувача об'ємне програмного забезпечення.

Для розв'язання задачі автоматизації розпізнавання акордів музикальних композицій ми пропонуємо виконувати такі етапи:

1. Завантаження файлу.
2. Перетворення вхідного файлу в зручну форму для подальшої обробки (декодування вхідного формату; зменшення частоти дискретизації; злиття каналів; конвертація імпульсно-кодової модуляції).
3. Обчислення спектрограми (використання віконної функції; швидке перетворення Фур'є; обчислення магнітуди).
4. Обчислення похибки відстроювання.
5. Обчислення хронограми.
6. Аналіз хромограми за допомогою прихованих моделей маркова.
7. Перетворення результатів у зручний вигляд.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Multi-pitch analysis [Електронний ресурс] – Режим доступу: – <http://www.ece.rochester.edu/~zduan/multipitch/multipitch.html>
2. Глазырин Н. Ю. О задаче распознавания аккордов в цифровых звукозаписях. // Изв. Иркутского гос. Ун-та. Серия «Математика» 2013 том 6, выпуск 2, С. 2 – 17.

3. Khadkevich M. Music signal processing for automatic extraction of harmonic and rhythmic information. / M. Khadkevich – 2011. – 143 с.

4. Fujishima T. Realtime Chord Recognition of Musical Sound: a System Using Common Lisp Music. / T. Fujishima. // In Proceedings of the International Computer Music Conference, Beijing: International Computer Music Association. – 1999.

5. Boulanger-Lewandowski Nicolas. Audio chord recognition with recurrent neural networks. / Nicolas Boulanger-Lewandowski, Bengio Yoshua, Pascal Vincent. // In Proceedings of the International Society for Music Information Retrieval Conference. – 2013.

6. Khadkevich M., Omologo M. Use of hidden hidden Markov models and factored language models for automatic chord recognition. / M. Khadkevich, M. Omologo //10th International Society for Music Information Retrieval Conference (ISMIR 2009). – 2009. – P. 561 – 566.

7. Бардаченко В. Ф. Таймерні нейронні елементи та структури. Монографія / В. Ф. Бардаченко, О. К. Колесницький, С. А. Василецький. – Вінниця : УНІВЕРСУМ-Вінниця, 2005, 126 с.

**Кучеровський Юрій Вадимович** – факультет інформаційних технологій та комп'ютерної інженерії, група 2КН-126, Вінницький національний технічний університет, м. Вінниця, e-mail: yuriy.kucherovskiy@yandex.ua

Науковий керівник: **Арсенюк Ігор Ростиславович** – к. т. н., доцент, доцент кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця.

**Yuriy V. Kucherovskiy** – Department of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: yuriy.kucherovskiy@yandex.ua.

Supervisor: **Igor R. Arsenyuk** – Cand Sc., Assistant Professor of the Chair of Computer Science, Vinnytsia National Technical University, Vinnytsia.