

# СИСТЕМА АВТОМАТИЧНОЇ СЕГМЕНТАЦІЇ АУДИОЗАПИСІВ НА ФРАГМЕНТИ МОВЛЕННЯ ТА МУЗИКИ

Вінницький національний технічний університет

## Анотація

Отримав подальшого розвитку метод сегментації звукових файлів, в якому на відміну від інших, запропоновано класифікувати використовуючи дихотомічну класифікацію фрагментів, що дозволило отримати результати наближені до сегментації, що проводяться людиною оператором.

**Ключові слова:** мовлення, музика, сегментація, класифікація, фрагменти.

## Abstract

The method of segmentation of audio files got further development, where unlike the others, it is possible to classify the audio information using the dichotomous classification of fragments, that allows obtaining results close to the segmentation carried out by the human operator.

**Keywords:** speech, music, segmentation, classification, fragments.

## Вступ

Існує багато систем для класифікації вмісту та пошуку зображень; однак, не так багато розробок існує в обробці аудіо та мультимедійного потоку. Хоча в сфері обробки сигналів залучили значну кількість дослідників і досягли успіху за останні роки, і, з легкістю застосовуються новітні підходи на сучасних доступних та швидких комп'ютерах, лише деякі з них торкнулися проблеми сегментації та класифікації музичних та мовленнєвих сигналів[1].

Метою роботи є розширення функціональних можливостей існуючих автоматизованих систем моніторингу радіомовлення за рахунок розробки засобів попередньої сегментації звукових файлів на музичні та мовленнєві фрагменти.

## Результати дослідження

Аудіо-файли, що є записом повноцінної радіо-програми за день, використовувались як вхідні дані[2], а сегментація та класифікація відбувалися одночасно. Отримана точність - це відсоток, що вказує наскільки правильно потрібний клас перебуває в певних межах часу, щ обули визначені людиною-оператором. Довжина правильно класифікованого фрагменту ділиться на загальну довжину програми.

Таблиця 1. Результати роботи алгоритму по кожному з класів.

	Мовлення	Музика
Мовлення	95.4%	4.6%
Музика	1.9%	98.1%

Лівий стовпчик таблиці вказує на клас введеного звуку, а верхній рядок - клас, виведений за алгоритмом. Визначення мовлення дає меншу точність через сегменти спортивних коментарів, які часто класифікуються як музика, тоді як музика в більшості випадків правильно класифікується з точністю в 98,1% часу.

Таблиця 2. Підсумкові дані роботи алгоритму.

	Коректність	Похибка
Загалом	97.3%	2.7%

Загальна точність становить 97,3%, оскільки тестовий матеріал містив більше музики, ніж мовлення. Час обчислень алгоритму залежить від формату та кількості вхідних даних. Час розрахунку не перевищував 1% довжини файлу в будь-якому з тестових файлів.

### Висновки

Встановлено, що запропонований підхід дозволяє підвищити загальну точність визначення мовленнєвих та музичних фрагментів за допомогою автоматичної системи сегментації близьку до людської роботи.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Е. Шейер, С. Малкольм «Побудова та оцінка детектору мовлення та музики на основі стійких ознак». — *IEEE, ICASSP*, 1997. — 6 с.

2. Д. Еліс Мовлення/Музика (Зразки Звуку) — <http://www.ee.columbia.edu/~dpwe/sounds/musp/>, Грудень 2006.

**Баліцький Віталій Васильович** — студент групи ІКІ-17м, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: vitalik.balitskiy@gmail.com

Науковий керівник: **Цирульник Сергій Михайлович** — канд. техн. наук, доцент кафедри обчислювальної техніки, Вінницький національний технічний університет, м. Вінниця, e-mail: sovmsvom@gmail.com

**Balitskiy Vitalii V.** — student of ICE-17m group, Faculty of information technology and computer engineering, Vinnytsia National Technical University, Vinnytsia, email : vitalik.balitskiy@gmail.com

Supervisor: **Tsyurulnyk Serhii M.** — Ph.D., Associate Professor of Computer Engineering, Faculty of information technology and computer engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: sovmsvom@gmail.com