

РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ АВТОМАТИЗОВАНОЇ РОЗМІТКИ ЗВУКОВИХ ФАЙЛІВ

Вінницький національний технічний університет

Анотація

Описано особливості розробки програмного забезпечення для автоматизованої розмітки звукових файлів. Основну увагу приділено визначенню типу звуків (мовлення, спів, музика, шум) та їх часових проміжків.

Ключові слова: аудіоаналіз, звукові файли, автоматизація, розпізнавання мовлення, програмне забезпечення.

Abstract

Features of software development for automated audio file annotation are described. The main focus is on identifying the type of sounds (speech, singing, music, noise) and their time intervals.

Keywords: audio analysis, sound files, automation, speech recognition, software.

У сучасному світі обсяг аудіоінформації, що потребує обробки, зростає експоненційно. Це зумовлено розвитком медіатехнологій, стрімінгових сервісів, подкастів, а також необхідністю автоматизації транскрипції та архівування звукових даних. Для ефективного використання аудіоінформації потрібно її структурувати, що вимагає якісної розмітки звукових файлів (ЗФ). Вручну цей процес займає багато часу та ресурсів, що робить його малоефективним. Саме тому автоматизовані системи аналізу звуку стають актуальним рішенням у цій сфері [1].

Сучасні технології аудіоаналізу дають змогу визначати типи звукових сигналів за допомогою методів спектрального аналізу та алгоритмів машинного навчання. Автоматизована розмітка ЗФ дозволяє сегментувати звук на мовлення, музику, шум та спів, що значно спрощує подальше використання даних. Це важливо не лише для медіаіндустрії, а й для систем автоматичного субтитрування, розпізнавання голосу та аудіоархівування. Крім того, алгоритми зниження шуму можуть значно покращити якість мовлення, що особливо корисно для записів із низькою якістю аудіо.

Автоматизована система розмітки ЗФ може базуватися на технологіях глибокого навчання, спектрального аналізу та методах виявлення голосової активності. Такі підходи вже знаходять застосування у різних сферах, зокрема в голосових асистентах, телефонних центрах, системах безпеки та медіаархівуванні. Розроблений програмний застосунок вирішує проблему розмітки аудіофайлів шляхом автоматичного аналізу їхнього вмісту, виділення часових проміжків для різних типів звуків та можливого покращення якості мовлення.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Gadandeep Singh. Best Practices for labeling Speech Recognition Custom Data. Електронний ресурс. Режим доступу: <https://gaganmanku96.medium.com/best-practices-for-labeling-speech-recognition-custom-data-25865294e5d2>.

Ткаченко Олександр Миколайович — к.т.н., доцент кафедри програмного забезпечення, Вінницький національний технічний університет.

Дажура Олег Романович – студент групи ЗПІ-21б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет.

Oleg Dazhura – student of group ЗПІ-21b, Faculty of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, oleg.dazhura@gmail.com.

Oleksandr Tkachenko — Cand. Sc. (Eng.), assistant professor of the Software Chair, Vinnytsia National Technical University, Vinnytsia, alextk1960@gmail.com.