

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ВІННИЦЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ

Грійо Тукало Оксана Франсисківна

УДК 621.39

**ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПОШУКУ ЗАДАНИХ
ФРАГМЕНТІВ В АРХІВІ АУДІОЗАПИСІВ ІЗ
ЗАСТОСУВАННЯМ KD-ДЕРЕВ**

05.13.06 – інформаційні технології

Автореферат

дисертації на здобуття наукового ступеня
кандидата технічних наук

Вінниця – 2016

Дисертацією є рукопис.

Роботу виконано у Вінницькому національному технічному університеті Міністерства освіти і науки України.

Науковий керівник

кандидат технічних наук, доцент
Ткаченко Олександр Миколайович,
Вінницький національний технічний університет,
доцент кафедри обчислювальної техніки.

Офіційні опоненти:

доктор технічних наук, професор
Казимир Володимир Вікторович,
Чернігівський національний технологічний університет,
проректор з наукової роботи

доктор технічних наук, професор
Поворознюк Анатолій Іванович,
Національний технічний університет
"Харківський політехнічний інститут",
професор кафедри обчислювальної техніки та програмування.

Захист відбудеться « 04 » березня 20 16 р. о 10.00 годині на засіданні спеціалізованої вченої ради Д 05.052.01 у Вінницькому національному технічному університеті за адресою: 21021, м. Вінниця, вул. Хмельницьке шосе, 95, ГНК, ауд. 210.

З дисертацією можна ознайомитись у бібліотеці Вінницького національного технічного університету за адресою: 21021, м. Вінниця, вул. Хмельницьке шосе, 95, ГНК.

Автореферат розісланий « 02 » лютого 20 16 р.

Учений секретар

спеціалізованої вченої ради

С. М. Захарченко

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Розвиток Інтернету, цифрових ЗМІ, мультимедійних технологій та нових способів обробки звукової інформації викликав великий інтерес і увагу до шляхів, якими інформаційні технології можуть бути застосовані до такого контенту. З можливістю доступу до великих архівів мультимедіа, практично в будь-якому місці і в будь-який час, необхідно було запропонувати нові способи навігації та взаємодії з цими архівами аудіозаписів.

Суть пошуку на основі аудіоконтенту полягає в тому, щоб автоматично визначати аудіозаписи, що містять частини, подібні до заданого аудіофрагменту. Більшість існуючих систем поєднує в собі такі етапи обробки аудіозаписів: аналіз вхідного аудіосигналу, виділення певних параметрів (ознак), вибір міри їх порівняння та визначення аудіозапису або деякої категорії аудіозаписів, найближчих до аудіофрагменту за обраною мірою.

Важливим завданням для систем пошуку на основі аудіоконтенту є масштабованість, оскільки розміри мультимедійних архівів (музики, фільмів, теле- та радіопередач, стенограм судових засідань тощо) сягають сотень тисяч і навіть мільйонів аудіозаписів та зростають постійно. Це, в свою чергу, також вимагає використання компактних параметрів з низькою складністю обчислень та ефективних методів пошуку в архівах аудіозаписів. Відповідно в процесі розробки методів пошуку в великих архівах аудіозаписів потрібно досягти компромісу між обчислювальною складністю та повнотою пошуку.

Незважаючи на великі успіхи, досягнуті у цій галузі в останні роки, сучасним інформаційно-пошуковим системам притаманні певні недоліки, зокрема

- недостатня продуктивність під час роботи з великими архівами;
- неможливість їхнього безпосереднього застосування для розв'язання таких задач, як моніторинг трансляції на радіостанціях (телебаченні) для визначення кількості повторень в ефірі рекламного продукту, музичного твору тощо.

Отже, створення інформаційної технології пошуку заданих фрагментів в архівах аудіозаписів є актуальною науково-практичною задачею.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконувалася відповідно до планів науково-дослідних робіт кафедри обчислювальної техніки Вінницького національного технічного університету, зокрема дисертаційне дослідження проводилось в рамках держбюджетної теми "Технології побудови інтелектуальних аналого-цифрових систем моніторингу та аналізу мультимедійної інформації" (№ держреєстрації 0115U001119).

Мета і задачі дослідження. Метою дисертаційної роботи є підвищення повноти та релевантності результатів і швидкості пошуку заданих аудіофрагментів в архіві аудіозаписів за рахунок розроблення нової інформаційної технології.

Для досягнення поставленої мети в дисертації розв'язуються такі наукові та практичні задачі:

- аналіз існуючих систем і методів автоматичного оцінювання подібності аудіозаписів для розв'язання задач пошуку на основі аудіоконтенту; визначення основних вимог до параметрів опису аудіозаписів, на основі яких здійснюється вибір параметрів математичної моделі та дослідження їх основних характеристик для розв'язання задачі пошуку аудіофрагментів;

- удосконалення методу кластеризації на основі k-середніх з метою отримання розв'язку, наближеного до глобального мінімуму похибки кластеризації;

- розробка моделі корпусу аудіозаписів, яка містить масив файлів, базу параметрів та метаданих і динамічне kd-дерево для виконання в ньому пошуку аудіозаписів за аудіофрагментами з різною тривалістю звучання;

- удосконалення методу пошуку на основі kd-дерева з удосконаленою оцінкою міри близькості, яка дозволяє досягнути підвищення повноти та релевантності результатів пошуку;

- розробка методу комбінованого пошуку аудіозапису за заданим фрагментом на основі аналізу основних факторів впливу на продуктивність пошуку та масштабованість системи для підвищення швидкості пошуку;

- розробка алгоритмічного та програмного забезпечення й практична апробація інформаційної технології пошуку заданих фрагментів в архівах аудіозаписів.

Об'єктом дослідження є процеси оброблення та пошуку аудіозаписів в електронних архівах.

Предметом дослідження є модель корпусу аудіозаписів та методи пошуку в корпусі аудіозаписів заданого аудіофрагмента.

Методи досліджень: цифрова обробка сигналів для визначення параметрів математичної моделі опису аудіосигналу; кластерний аналіз при формуванні корпусу даних; математична статистика для обґрунтування можливості пошуку аудіозапису за фрагментом; теорія алгоритмів і методів обчислень для розробки методів пошуку; теорія дослідження операцій для оцінювання обчислювальної складності запропонованих методів; комп'ютерне моделювання для дослідження розроблених методів пошуку заданого аудіофрагмента в корпусі.

Наукова новизна одержаних результатів

1. **Вперше запропоновано** модель корпусу аудіозаписів, яка містить масив файлів, базу параметрів та метаданих і динамічне kd-дерево, що дозволило зменшити тривалість обраного для пошуку аудіофрагмента та реалізувати інформаційну технологію пошуку аудіозаписів з різною тривалістю звучання.

2. **Вперше запропоновано** метод комбінованого пошуку аудіозапису за заданим аудіофрагментом, який базується на виконанні наближеного пошуку у базі параметрів зменшеної розмірності із застосуванням kd-дерева кількох найближчих аудіозаписів на першому етапі пошуку, серед яких на другому етапі відбувається вибір релевантного аудіозапису, що дозволило досягти зменшення часу пошуку порівняно з точним пошуком на основі kd-дерева.

3. **Удосконалено** метод пошуку на основі kd-дерева, який, на відміну від існуючих, використовує оцінку міри близькості за зваженою кількістю

потрапляють в список найближчих центроїдів, що дозволяє підвищити повноту та релевантність результатів пошуку.

4. **Отримав** подальшого розвитку метод кластеризації k -середніх, який відрізняється від існуючих удосконаленою процедурою відбору вектора на позицію вставки нового центроїда шляхом послідовного запуску методу k -середніх, що дозволяє отримати розв'язок, наближений до глобального мінімуму похибки кластеризації.

Практичне значення одержаних результатів. Розроблено інформаційну технологію пошуку заданих аудіофрагментів в архіві аудіозаписів, яка містить:

- алгоритм та програму для реалізації методу кластеризації на основі послідовного запуску k -середніх з удосконаленим вибором вектора на позицію вставки нового центроїда;
- алгоритм та програму пошуку заданого аудіофрагмента на основі математичного очікування похибки кластеризації;
- алгоритм та програму швидкого пошуку на основі kd -дерев;
- методику використання системи автоматичного пошуку аудіозаписів.

Розроблено алгоритми та програмні засоби для дослідження методу швидкого пошуку заданих аудіофрагментів в архіві аудіозаписів із застосуванням kd -дерев. Запропоновані методи та розроблені на їх основі програми пошуку дозволяють зменшити складність обчислень в 28 разів порівняно з точним пошуком на основі kd -дерева.

Результати досліджень впроваджено, що підтверджено відповідними актами, в Інституті електроніки та зв'язку Української академії наук (м. Київ, акт впровадження від 05.02.2015), а також у навчальний процес у Вінницькому національному технічному університеті з дисципліни "Цифрова обробка сигналів" (м. Вінниця, акт впровадження від 25.05.2015).

Особистий внесок здобувача. Всі результати, що складають основний зміст дисертації, отримані здобувачем самостійно. В опублікованих у співавторстві наукових працях автору належать: [1] – експериментальне дослідження та аналіз методів швидкого пошуку векторів, [2] – розробка методу кластеризації на основі послідовного запуску k -середніх з удосконаленим вибором вектора на позицію вставки нового центроїда, [3] – розробка та аналіз різних методів побудови kd -дерева, [4] – застосування kd -дерев для визначення кандидатів на вибір нового центроїда, [5, 6] – програмна реалізація та оцінювання ефективності методів пошуку на основі kd -дерева за кількістю операцій, [7] – обґрунтування можливості пошуку аудіозапису за його фрагментом на основі математичного очікування похибки кластеризації та визначення мінімальної тривалості аудіофрагмента, [8] – розробка методу пошуку на основі kd -дерева та аналітичних співвідношень для оцінювання міри близькості невідомого аудіофрагмента з аудіозаписами в корпусі, [9] – експериментальне дослідження методу пошуку на основі kd -дерева, [10] – оцінювання необхідної кількості кандидатів на першому етапі пошуку, [11, 12] – аналіз основних факторів впливу на ефективність пошуку, [13, 14] – експериментальна перевірка методу пошуку заданого аудіофрагмента на основі математичного очікування похибки

кластеризації (приведеної власної відстані), [15] – обґрунтування виконання швидкого пошуку найближчого аудіозапису в корпусі на основі kd-дерева та аналіз чинників впливу на продуктивність методу пошуку на основі kd-дерева, [16] – розробка та програмна реалізація методу кластеризації з обчисленням відстаней лише до активних центрів, [17, 18] – розробка способів швидкого пошуку в багатовимірному просторі, [19] – аналіз параметрів математичної моделі опису аудіосигналів та методів швидкого пошуку, розробка та експериментальне дослідження методів побудови kd-дерева, аналіз чинників впливу на ефективність пошуку, розробка методу пошуку на основі kd-дерева, розробка двоетапного методу пошуку, розробка алгоритмів та програмного забезпечення для реалізації запропонованих методів швидкого пошуку та їх експериментальне дослідження.

Апробація результатів дисертації. Основні результати наукових досліджень за темою роботи доповідалися та обговорювалися на 12 науково-технічних конференціях: III Міжнародній науково-практичній конференції "Інформаційні технології та комп'ютерна інженерія" (Вінниця, 2012); II Міжнародній науково-практичній конференції "Актуальные проблемы современной науки" (Ставрополь, Росія, 2013); IV Міжнародній науково-практичній конференції "Інформаційні технології та комп'ютерна інженерія" (Вінниця, 2014); XII Міжнародній конференції "Контроль і управління в складних системах" (Вінниця, 2014); XII Всеукраїнській міжнародній конференції "Оброблення сигналів і зображень та розпізнавання образів" (Київ, 2014); XLII, XLIII та XLIV щорічних конференціях професорсько-викладацького складу, співробітників та студентів Вінницького національного технічного університету (Вінниця, 2013–2015).

Публікації. За темою дисертації опубліковано 19 наукових праць, серед яких 8 статей надруковано у наукових фахових виданнях України [1 – 8], 7 матеріалів доповідей опубліковано у збірниках матеріалів міжнародних конференцій [9 – 15], отримано свідоцтво про державну реєстрацію авторського права на твір (комп'ютерну програму) [16], 2 патенти України на корисну модель [17, 18] та монографія [19]. До міжнародних наукометричних баз входять 7 публікацій, зокрема [1] – входить до CSA, CISA, INIS, Inspec, ВІНІТІ РАН, роботи [2, 3, 12] та [6 – 8] – індексуються в РІНЦ.

Структура роботи та її обсяг. Дисертаційна робота загальним обсягом 194 сторінки складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Основний зміст дисертації викладено на 153 сторінках, робота містить 65 рисунків, 16 таблиць, 136 найменувань бібліографічних джерел та акти впровадження.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі обґрунтовано актуальність теми дисертаційної роботи, показано її зв'язок з науковими програмами, визначено об'єкт, предмет і методи дисертаційного дослідження, сформульовано мету та задачі роботи, наведено

наукову новизну, практичне значення, апробацію та структуру роботи, зазначено особистий внесок здобувача.

У першому розділі дисертації розглянуто сфери застосування пошуку на основі аудіоконтенту, описано задачу оцінювання подібності аудіозаписів, сформульовано основні етапи розв'язання задачі автоматичного пошуку аудіозаписів, проведено аналіз параметрів математичної моделі аудіосигналу, методів класифікації та швидкого пошуку, найбільш придатних для розв'язання даної задачі, наведено основні підходи та методи для оцінювання подібності мультимедійного контенту та проведено огляд сучасних систем пошуку аудіозаписів за аудіоконтентом.

На підставі проведеного аналізу зроблено висновок, що, незважаючи на великі успіхи, сучасним інформаційно-пошуковим системам притаманні певні недоліки, зокрема:

- недостатня продуктивність під час роботи з великими архівами;
- неможливість їхнього безпосереднього застосування для розв'язання таких задач, як моніторинг трансляції на радіостанціях (телебаченні) для визначення кількості повторень в ефірі рекламного продукту, музичного твору тощо.

Це обумовлює актуальність дисертаційної роботи.

Обґрунтовано, що, у зв'язку з великими розмірами мультимедійних архівів (музики, фільмів, теле- та радіопередач, стенограм судових засідань тощо), в процесі розробки інформаційної технології пошуку аудіофрагментів важливо забезпечити її масштабованість, що вимагає забезпечення низької складності обчислень під час пошуку. Також в ході виконаного огляду існуючих методів і основних підходів до автоматичного оцінювання подібності аудіозаписів визначено, що для побудови реальної масштабованої системи пошуку аудіозаписів необхідним є застосування кластерного аналізу для побудови швидкого пошукового індексу в архіві та методів швидкого пошуку.

У другому розділі визначено вимоги до параметрів математичної моделі опису аудіосигналу, на основі яких обґрунтовано вибір як параметрів мел-частотних кепстральних коефіцієнтів (MFCC – Mel Frequency Cepstral Coefficients), а також проведено дослідження їхніх основних характеристик для розв'язання задачі пошуку аудіофрагмента в архіві аудіозаписів. Обґрунтовано доцільність застосування кластерного аналізу для формування бази параметрів корпусу аудіозаписів. Розроблено метод кластеризації на основі послідовного запуску k-середніх з удосконаленим вибором кандидата на нову позицію вставки центроїда. Проведено дослідження kd-дерев як засобу структуризації бази параметрів для пошуку аудіофрагмента в корпусі аудіозаписів та аналіз основних чинників впливу на продуктивність пошуку.

В основу запропонованої інформаційної технології пошуку заданих фрагментів в архівах аудіозаписів покладено пошук на основі аудіоконтенту, тобто оцінювання подібності аудіозаписів виконується виключно на основі інформації, яку можна отримати з аудіофайлу.

На основі аналізу літературних джерел визначено такі основні етапи в процесі виконання автоматичного пошуку аудіозаписів: 1) обчислення параметрів опису аудіосигналу; 2) зменшення розмірності параметрів; 3) впорядкування параметрів; 4) безпосередньо виконання пошуку за обраною мірою оцінювання подібності аудіозаписів.

Визначено вимоги до параметрів математичної моделі, що використовується для інформаційного пошуку аудіозаписів.

1. Компактність та некорельованість, що дозволить зменшити необхідні обсяги пам'яті для їх зберігання та в подальшому впливатиме на швидкість пошуку в архівах аудіозаписів.

2. Підсилення корисного сигналу і, відповідно, стійкість до шумів за рахунок застосування нерівномірної шкали частот та нелінійного сприйняття гучності (децибели), що актуалізують низькочастотну інформацію (корисний сигнал) і усереднюють високочастотні складові сигналу (де, в основному, зосереджується шумова складова сигналу).

3. Можливість застосування для аналізу різних звуків (мовлення, музика, оточуючі звуки), що дозволить забезпечити універсальність підходу для розв'язання різних задач на основі визначення подібності аудіозаписів.

4. Невелика обчислювальна складність отримання параметрів.

5. Можливість застосування простих метрик для визначення близькості між наборами параметрів.

6. Можливість застосування кластерного аналізу.

За результатами аналізу параметрів в першому розділі, а також відповідно до наведених вище вимог вирішено здійснювати параметризацію аудіозаписів на основі MFCC:

$$MFCC_i = \sum_{j=0}^{N_{FB}-1} \ln(s_j) \cos\left(\frac{\pi i(j+0,5)}{N_{FB}}\right), i = 1, \dots, N_{MFCC}.$$

де N_{FB} – кількість трикутних фільтрів, що перекриваються і рівномірно розміщуються на шкалі Мел-частот; s_j – спектральна енергія; N_{MFCC} – кількість коефіцієнтів (розмірність) вектора параметрів MFCC.

Визначено величини основних характеристик, необхідних для обчислення MFCC-параметрів в задачі пошуку аудіофрагментів у корпусі аудіозаписів, а саме: частота дискретизації аудіофайлу – 44,1 кГц, довжина фрейму, для якого обчислюється вектор MFCC, – 20 мс, розмір перекриття фреймів – 50% (10 мс), кількість трикутних фільтрів – 40, кількість коефіцієнтів (розмірність вектора) MFCC – 13.

Пошук невідомого аудіозапису виконується шляхом порівняння його параметричного подання з базою параметрів аудіозаписів, що містяться в корпусі, критерієм порівняння є відстань. Загальну схему пошуку аудіозапису наведено на рисунку 1.

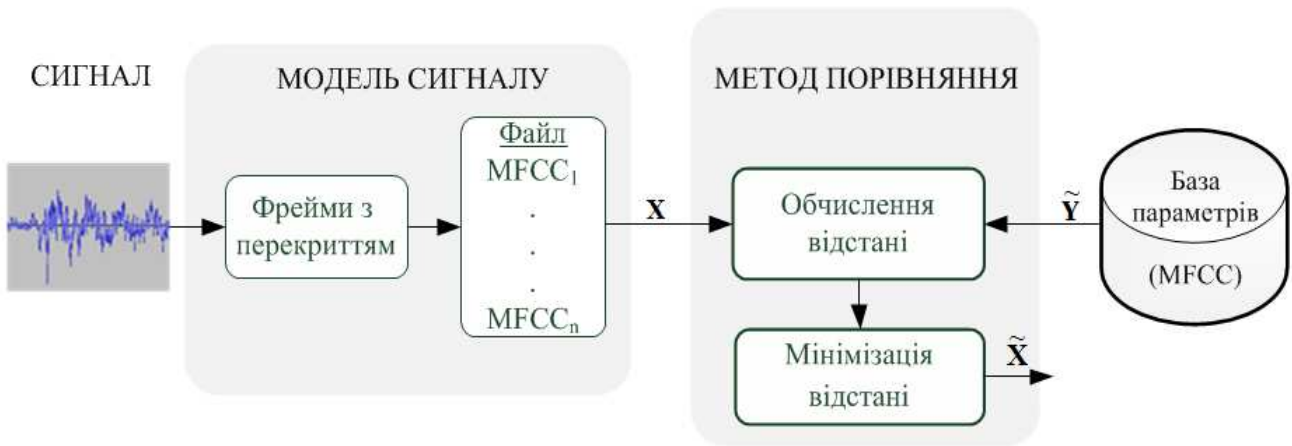


Рисунок 1 – Загальна схема пошуку аудіозапису

Подання аудіозапису за допомогою параметрів MFCC (векторів) дозволило компактно описати аудіозапис і тим самим прискорити процес пошуку в десятки разів. Подальше зменшення обсягів пам'яті, необхідної для зберігання бази параметрів, було досягнуто завдяки застосуванню методів кластерного аналізу для формування корпусу аудіозаписів.

Аудіозаписи в корпусі подано центроїдами кластерів, обчислених для параметрів MFCC. Кластеризація виконувалася за методом k-середніх окремо для кожного файлу MFCC-параметрів, що описує аудіозапис корпусу. Відповідно відстань між файлами параметрів, що описують один і той же аудіозапис в корпусі до кластеризації і після, буде додатною та дорівнюватиме величині сумарної похибки кластеризації E^2 для k центроїдів:

$$E^2 = \sum_{i=1}^k e_i^2 = \sum_{i=1}^k \sum_{j=1}^{N_i} D_{Eu}^2(\mathbf{x}_j, \mathbf{c}_i) = \sum_{i=1}^k \sum_{j=1}^{N_i} \sum_{m=1}^d (x_{jm} - c_{im})^2, \quad (1)$$

$$e_i^2 \rightarrow \min \Rightarrow E^2 \rightarrow \min,$$

Для подолання високої складності обчислень, притаманної глобальному методу k-середніх, коли послідовно виконується запуск k-середніх для кожного з $1, 2, \dots, k$ центроїдів, формування бази параметрів (центроїдів) для аудіозаписів корпусу здійснювалось за допомогою вдосконаленого методу кластеризації k-середніх, запропонованого в дисертаційній роботі.

Основна ідея запропонованого методу кластеризації:

- здійснювати запуск k-середніх не для кожного вхідного вектора $\mathbf{x}_i \in \mathbf{S}$, а для певної множини векторів $\mathbf{X} = \{\mathbf{x}_0\} \subset \mathbf{S}$, що будуть використовуватися як кандидати на позицію вставки нового центроїда, що дозволяє зменшити складність обчислень;

- процедура вибору вектора $\mathbf{x} \in \{\mathbf{x}_0\}$ з числа кандидатів на позицію вставки нового центроїда полягає у виборі того вектора, який забезпечує мінімальне спотворення $E^2 = \sum_{\mathbf{x} \in \mathbf{S}} D^2(\mathbf{x}, \mathbf{c})$ при додаванні його як початкової позиції нового

центроїда замість запуску k -середніх для кожного вектора, що дозволяє досягти значного зменшення часу роботи;

- операція перерахунку нового положення центроїда виконується лише для активних центроїдів $SC^{(a)}$, тобто тих, для яких змінилась кількість векторів, що до них належать, після додавання нового центроїда; що реалізується шляхом зберігання множини з m найближчих центроїдів $W = \{c_1, c_2, \dots, c_m\}$ та відстаней до них $r_i = D(x, c_i)$ для кожного з векторів $x \in S$, цим також досягається зменшення складності обчислень методу.

Обчислення сумарного спотворення згідно з формулою (1) вимагає виконання такої кількості операцій: $N_{op} = 2 \cdot d \cdot N_i \cdot k$, де N_i – кількість точок, що належать центроїду c_i . При застосуванні запропонованого методу кластеризації було досягнуто значного зменшення кількості операцій $N_{op} = (M_i + 5) \cdot d \cdot k$, оскільки $M_i \ll N_i$; де M_i – кількість точок c_i , що відійдуть центроїду c_{k+1} .

Таким чином, запропонований метод відрізняється від відомих удосконаленням вибором вектора з числа кандидатів для вставки чергового центроїда.

Результати експериментальних досліджень показали, що більшої ефективності можна досягти, використовуючи $|C| = 4 \cdot k$ кандидатів для отримання k центроїдів (тобто 4000 кандидатів для генерації 1000 центроїдів), оскільки це дозволяє досягнути мінімуму спотворення без значного уповільнення роботи методу.

Проведено порівняння запропонованого методу кластеризації на основі вибору вектора-кандидата, який забезпечує мінімальне спотворення, з перерахунком нового положення центроїдів (RC, Recalculate Centroids) при $|C| = 16000$ з двома модифікаціями k -середніх, що показано на рисунку 2.

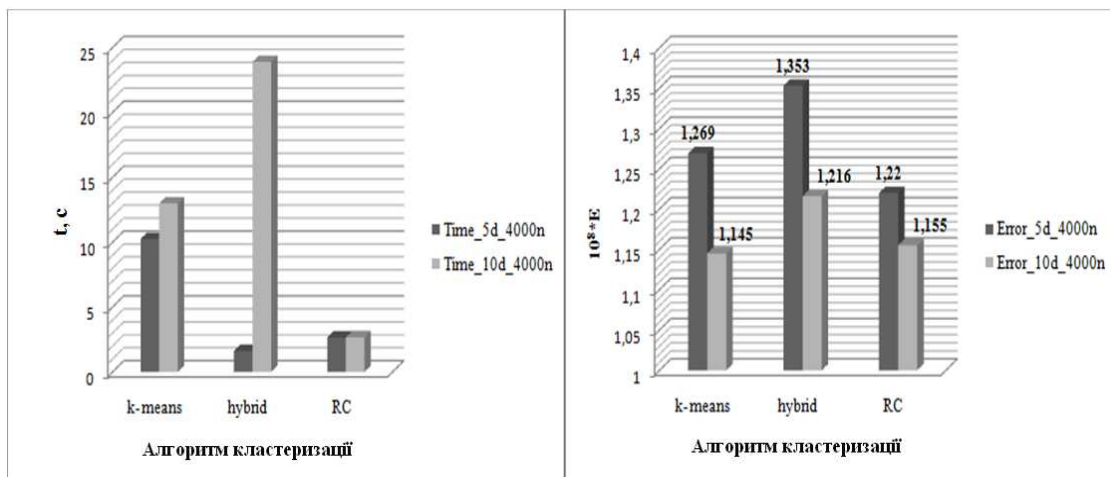


Рисунок 2 – Порівняння методів за часом роботи t (зліва) та сумарним спотворенням E (справа): $k = 4000$

Результати показали, що для обох розмірностей найбільше значення сумарного спотворення спостерігається при застосуванні методу кластеризації на

основі kd-дерев (hybrid), найменше ж спотворення отримано для RC для $d = 5$ та класичного k-means (MATLAB) для $d = 10$. Найбільша швидкодія для $d = 5$ досягається методом з використанням kd-дерев, проте зі збільшенням розмірності для hybrid спостерігається стрімке зростання часу роботи, що зумовлено експоненціальним характером залежності часу від розмірності. Ефективність запропонованого методу кластеризації практично не залежить від розмірності.

У великих архівах аудіозаписів виконання пошуку шляхом повного перебору, тобто порівняння обраного аудіофрагмента з усіма центроїдами аудіозаписів стає неможливим, оскільки час пошуку (кількість обчислень відстаней) пропорційний $O(dN)$, де N – кількість центроїдів-векторів в корпусі, d – розмірність векторів. З огляду на це та на основі проведеного аналізу методів швидкого пошуку в розділі 1 запропоновано виконувати пошук аудіофрагмента в корпусі аудіозаписів з застосуванням kd-дерев, тобто впорядкувати центроїди (d-вимірні вектори) бази параметрів корпусу у вигляді ієрархічної структури – k-вимірного дерева пошуку.

Метод пошуку найближчого центроїда у корпусі аудіозаписів, впорядкованого на основі kd-дерева, ілюстрацію якого для двовимірного випадку показано на рисунку 3, має дві стадії пошуку:

- пряма – спуск по дереву до термінальної вершини, яка вимагає $O(\log_2 N)$ операцій;
- обернена – пошук серед елементів дерева в заданому діапазоні (радіусі).

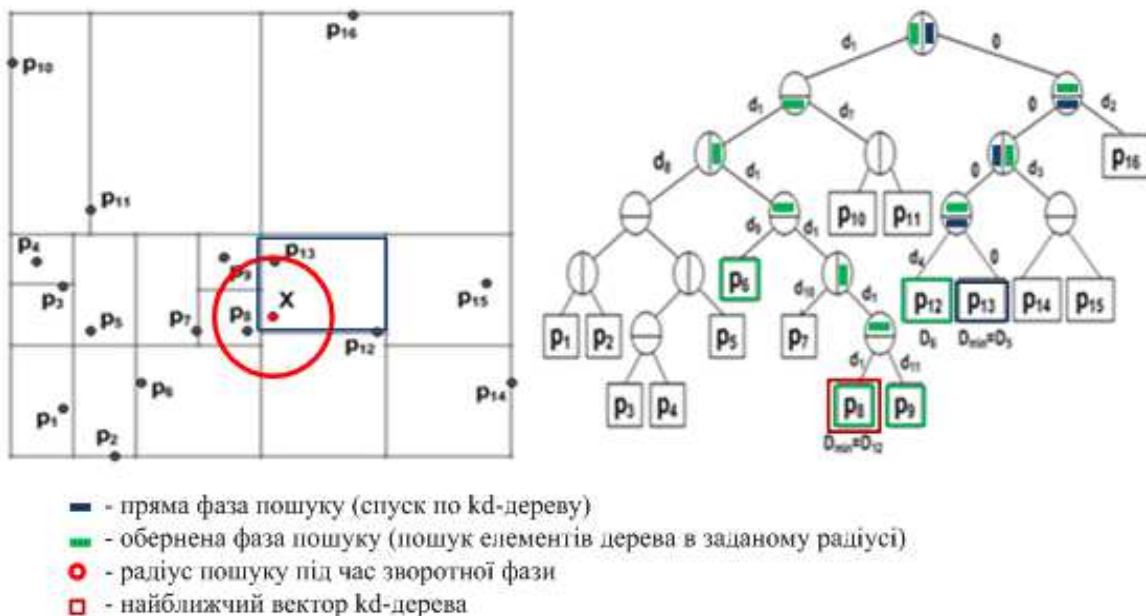


Рисунок 3 – Ілюстрація пошуку найближчого центроїда по kd-дереву для двовимірного випадку

Під час прямого пошуку фіксуються всі відстані до вузлів d_i та запам'ятовуються у порядку їх зростання, за рахунок чого під час оберненої фази здійснюється їхній обхід у порядку зростання відстані до x . Пряма фаза

завершується обчисленням відстані $D_{\min} = D_k$ до відповідної термінальної вершини (на рисунку 3 – прямокутника), в якій лежить невідомий вектор, а також потенційно найближчий елемент дерева (центроїд в корпусі), що задає радіус пошуку під час оберненої фази. Після цього починається обернена фаза пошуку, під час якої виконується рекурсивний обхід вузлів, що лежать у межах радіуса пошуку $d_i < D_{\min}$. Якщо на оберненій фазі пошуку досягнуто листа, для якого виконується умова $D_k < D_{\min}$, радіус пошуку коригують $D_{\min} = D_k$. Таким чином буде здійснено обхід кожного листа дерева, відстань до якого менша, ніж до найближчого поточного вектора.

Проаналізувавши основні чинники впливу на продуктивність пошуку, було визначено, що на час пошуку суттєво впливають:

- розмірність векторів параметрів d ;
- коефіцієнт розповсюдження оберненої фази пошуку r .

Зроблено висновок, що значно зменшити складність обчислень, яка є критичним фактором для великих архівів, можна шляхом застосування наближеного пошуку, коли знайдений вектор не обов'язково буде найближчим, але досить близьким, що визначається коефіцієнтом розповсюдження оберненої фази пошуку:

$$r = \begin{cases} (0;1), e > 0 & \text{– наближений пошук;} \\ 1, e = 0 & \text{– точний пошук.} \end{cases}$$

Центроїд корпусу \mathbf{p} можна вважати r -найближчим вектором до \mathbf{x} , якщо відстань до комірки дерева, якій \mathbf{p} належить, дорівнює:

$$d(\mathbf{x}, \mathbf{p}) \leq r \cdot D(\mathbf{x}, \mathbf{p}^*), \mathbf{p}, \mathbf{p}^* \in \mathbf{Q}, r = \frac{1}{(1+e)}, e > 0, 0 < r < 1,$$

де \mathbf{p}^* – найближчий центроїд, знайдений до цього моменту; r – коефіцієнт розповсюдження оберненої фази пошуку.

Експериментальні дослідження підтвердили, що наближений пошук r -близьких векторів ($0 < r < 1$) дозволяє досягти значного зменшення часу пошуку:

при $r = \frac{1}{3}$ пошук r -найближчого вектора дозволив досягти скорочення часу пошуку в середньому в 50 разів порівняно з пошуком дійсно найближчого вектора.

У третьому розділі обґрунтовано можливість визначення аудіозапису за його аудіофрагментом, запропоновано метод пошуку найближчого аудіозапису в корпусі за приведеною відстанню, розглянуто загальний підхід до реалізації пошуку в корпусі аудіозаписів, впорядкованому на основі kd-дерева, виконано порівняння запропонованих оцінок міри близькості невідомого аудіофрагмента з аудіозаписами в корпусі, удосконалено метод пошуку аудіофрагмента на основі kd-дерева з оцінкою міри близькості за зваженою кількістю потраплянь в список найближчих центроїдів, досліджено чинники впливу на продуктивність пошуку із застосуванням kd-дерев, запропоновано метод комбінованого пошуку аудіозапису

за заданим аудіофрагментом, наведено результати експериментальних досліджень запропонованих методів та їх порівняння з іншими методами.

Запропоновано оцінювати близькість до аудіозапису корпусу $\tilde{\mathbf{Y}} = \{\tilde{\mathbf{y}}\}, |\tilde{\mathbf{Y}}| = m$, $\tilde{\mathbf{y}} = (\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_d)$ за допомогою приведеної відстані (ПВ – $D_{ПВ}$), яка дорівнює частці від ділення відстані до аудіозапису в корпусі на кількість фреймів n аудіозапису, заданого для пошуку $\mathbf{X} = \{\mathbf{x}\}, |\mathbf{X}| = n$, $\mathbf{x} = (x_1, x_2, \dots, x_d)$:

$$D_{ПВ} = \frac{D(\mathbf{X}, \tilde{\mathbf{Y}})}{n} = \frac{\sum_{l=1}^n D \min^2}{n} = \frac{\sum_{l=1}^n \min(\sum_{i=1}^d (x_i - \tilde{y}_{ji})^2)}{n}, j = \overline{1, m}. \quad (2)$$

Відстань до власного аудіозапису (приведена власна відстань, ПВВ – $D_{ПВ(B)}$) є математичним очікуванням (МО) похибки кластеризації:

$$D_{ПВ(B)} = \frac{D(\mathbf{X}, \tilde{\mathbf{X}})}{n} = M_E = \frac{E^2}{n}. \quad (3)$$

Таким чином, в дисертаційній роботі обґрунтовано можливість визначення аудіозапису за його аудіофрагментом на основі приведеної власної відстані, оскільки $D_{ПВ}$ не залежить від кількості фреймів (тривалості аудіозапису) та практично не змінюється в часі для різних фрагментів аудіозапису, що підтвердила статистична перевірка (рисунок 4).

Відзначено важливість мінімізації тривалості аудіофрагмента, який необхідно визначити, при збереженні достатньої тривалості аудіофрагмента, щоб, з одного боку, зменшити складність обчислень, а з іншого – забезпечити прийняття правильного рішення під час пошуку.

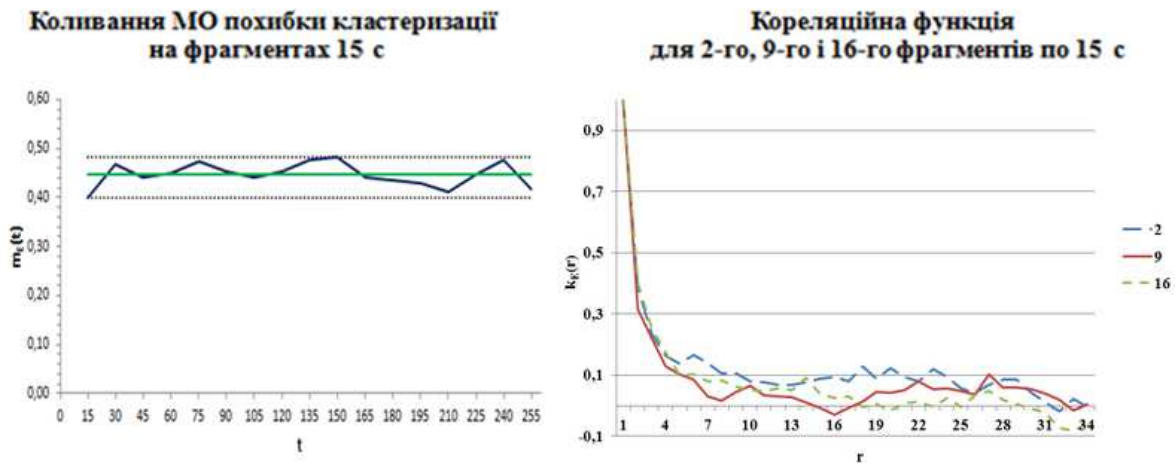


Рисунок 4 – Результати статистичної перевірки похибки кластеризації на стаціонарність

Для проведення експериментальних досліджень було сформовано корпус з 1000 аудіозаписів, усі аудіозаписи в форматі wav (mono) з частотою дискретизації 44,1 кГц, попередньо з аудіофайлів було видалено тишу з початку та кінця записів. Кількість центроїдів для опису кожного аудіозапису в корпусі складала $m=1000$ (в середньому на кластер припадає близько 20 векторів). З 1000

аудіозаписів корпусу випадковим чином було обрано аудіофрагменти (тривалістю 1, 5 та 15 с) 100 аудіозаписів для їхнього пошуку.

Результати досліджень оцінювання близькості аудіозаписів за методом пошуку власного аудіозапису на основі середнього значення похибки кластеризації D_{IV} підтвердили коректність такого підходу. Запропонований підхід дозволив скоротити тривалість аудіофрагмента та реалізувати інформаційну технологію оцінювання подібності музичних творів з різною тривалістю звучання.

В дисертаційній роботі для підвищення швидкості пошуку запропоновано об'єднати центроїди аудіозаписів корпусу в межах одного файлу, щоб впорядкувати їх на основі kd-дерева, що зображено на рисунку 5, оскільки, коли кожен аудіозапис в корпусі $\tilde{Y}_q, q=1, Q$ є окремим файлом, зі зростанням Q кількість операцій зростатиме лінійно $N_{op} = O(c \cdot Q \cdot \log m)$, $c = const$, що є неприйнятним для великих архівів аудіозаписів, тому що час пошуку стане занадто великим.

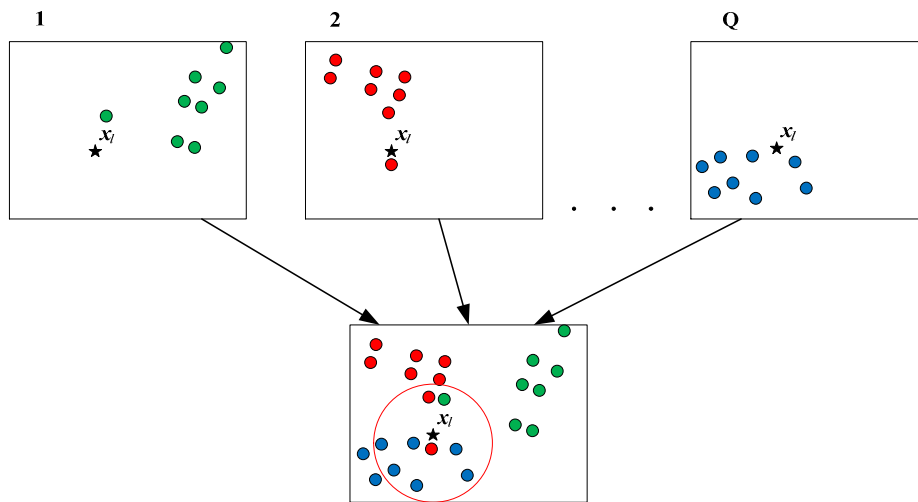


Рисунок 5 – Об'єднання файлів з центроїдами MFCC-параметрів

Проведені дослідження показали, що в результаті виконання пошуку в базі параметрів, впорядкованій на основі kd-дерева, в списку знайдених для невідомого аудіофрагмента X найближчих центроїдів відсоток центроїдів власного аудіозапису є більшим відносно частки центроїдів решти аудіозаписів корпусу, однак центроїд власного аудіозапису в корпусі не завжди знаходиться на першій позиції в списку найближчих. У зв'язку з цим зроблено висновок, що доцільно виконувати пошук не одного найближчого центроїда, оскільки це призводитиме до помилкових результатів, а деякої множини $C, |C|=k$ центроїдів. Відповідно, загальна кількість операцій складає:

$$N'_{op} + N''_{op} = O(c \cdot \log(Q \cdot m)) + O(k \cdot \log(Q \cdot m)) = O(c + k) \cdot \log N, N'_{op} \ll N_{op}.$$

Отже, послідовність кроків при виконанні пошуку відповідного для аудіофрагмента аудіозапису в корпусі на основі kd-дерева така.

1. Виконання швидкого пошуку в упорядкованій на основі kd-дерева базі параметрів корпусу аудіозаписів, в процесі якого за Евклідовою метрикою відбирається множина \mathbf{C} векторів-центроїдів аудіозаписів (кандидатів на найближчий вектор), упорядкованих за зростанням відстані:

$$\mathbf{C} \subset \mathbf{T}, \mathbf{C} = \{\tilde{\mathbf{Y}}_1, \tilde{\mathbf{Y}}_2, \dots, \tilde{\mathbf{Y}}_k\}, |\tilde{\mathbf{Y}}| = m, |\mathbf{C}| = k, |\mathbf{T}| = N = m \cdot Q, k \leq N,$$

де k – кількість векторів аудіозаписів у списку найближчих до фрейму невідомого аудіофрагмента; m – кількість центроїдів (векторів) аудіозапису,

Q – загальна кількість аудіозаписів у корпусі; N – загальна кількість центроїдів-векторів корпусу аудіозаписів.

Для кожного з k найближчих центроїдів множини \mathbf{C} зберігаються такі дані: відстань до фрейму невідомого аудіофрагмента $D_p^2(\mathbf{x}, \tilde{\mathbf{y}}_{qj})$ та індекс відповідного аудіозапису i_{pq} в корпусі:

$$\mathbf{I} = \{i_{pq}\}, \mathbf{D} = \{D_p^2(\mathbf{x}_l, \tilde{\mathbf{y}}_{qj})\}, D_p^2(\mathbf{x}_l, \tilde{\mathbf{y}}_{qj}) = \sum_{i=1}^d (x_i - \tilde{y}_i)^2, p = \overline{1, k}, l = \overline{1, n}, j = \overline{1, m}, q = \overline{1, Q}.$$

2. Визначення для аудіофрагмента власного аудіозапису в корпусі за обраною мірою оцінювання близькості.

Запропоновано та досліджено декілька варіантів оцінювання міри близькості невідомого аудіофрагмента з аудіозаписами, які базуються на обчисленні відстані від фрагмента до заданої кількості k найближчих центроїдів та кількості потраплянь в список найближчих. На рисунку 6 показано результати порівняння заданого для пошуку аудіофрагмента з аудіозаписами в корпусі (власний аудіозапис під № 3) за запропонованими оцінками міри близькості, та зведена гістограма, на якій наведено всі варіанти оцінювання міри близькості у відносних одиницях для досягнення єдиного масштабу.

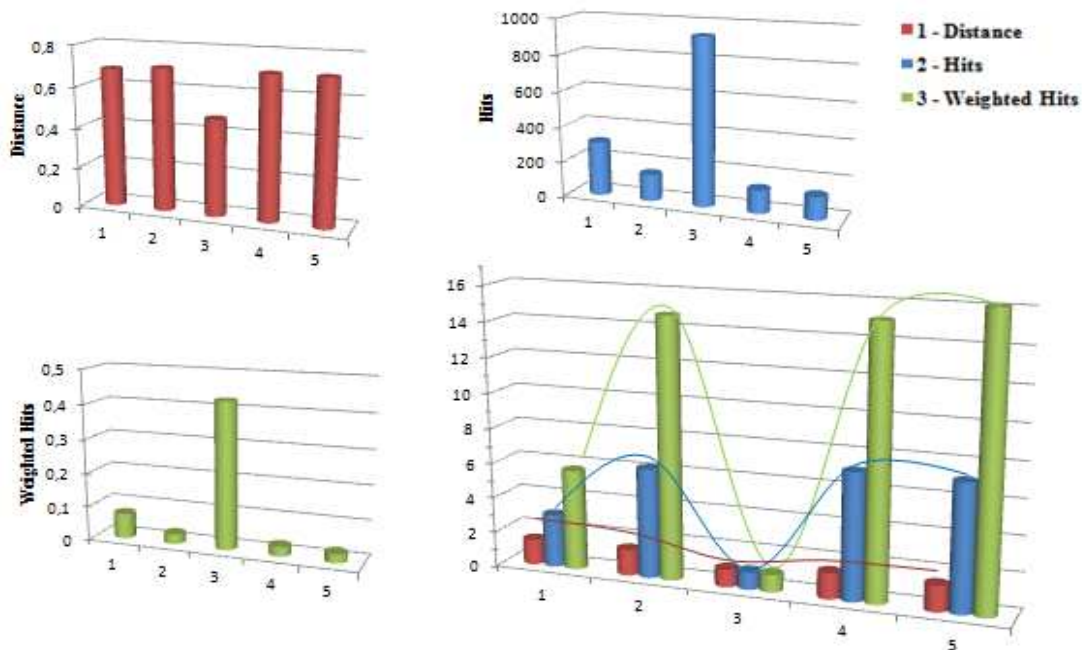


Рисунок 6 – Порівняння оцінок міри близькості невідомого аудіофрагмента з аудіозаписами корпусу

- Оцінка міри близькості за приведеною відстанню (Distance):

$$WH(\mathbf{X}, \tilde{\mathbf{Y}}) = \min_q \sum_{l=1}^n \frac{\mathbf{Weighted Hits}_q}{n}, \mathbf{Distance}_q = \begin{cases} \min(D_p^2), \forall q \in \mathbf{I} \\ D_p^2, p = k, \forall q \notin \mathbf{I} \end{cases} \quad (4)$$

- Оцінка міри близькості за кількістю потраплянь h в список k найближчих центроїдів (Hits):

$$H(\mathbf{X}, \tilde{\mathbf{Y}}) = \max_q \sum_{l=1}^n \frac{\mathbf{Hits}_q}{n}, \mathbf{Hits}_q = \sum_{p=1}^k h_p, h_p = \begin{cases} 1, \forall q \in \mathbf{I} \\ 0, \forall q \notin \mathbf{I} \end{cases} \quad (5)$$

- Оцінка міри близькості за зваженою кількістю потраплянь в список k найближчих центроїдів (Weighted Hits):

$$WH(\mathbf{X}, \tilde{\mathbf{Y}}) = \max_q \sum_{l=1}^n \frac{\mathbf{Weighted Hits}_q}{n}, \mathbf{Weighted Hits}_q = \sum_{p=1}^k (D_k^2 - D_p^2) \cdot h_p. \quad (6)$$

Відповідно до отриманих результатів найвищий рівень розрізнення власного та чужих аудіозаписів в корпусі забезпечує оцінювання міри близькості за зваженою кількістю потраплянь аудіозапису у список k найближчих (Weighted Hits), для якої оцінка відстані до власного аудіозапису при $k=20$ в середньому в 8 разів перевищує аналогічну оцінку відстані до інших аудіозаписів.

Таким чином, запропоновано метод пошуку власного аудіозапису з удосконаленою оцінкою міри близькості, що дозволило збільшити рівень розрізнення власного аудіозапису та решти аудіозаписів у корпусі в процесі пошуку. Для аудіофрагмента тривалістю 5 с повнота пошуку становить 99,8%, релевантність – 97,4%, для аудіофрагмента 1 с – 99,6% та 94,3% відповідно, що перевищує аналогічні показники існуючих методів. Разом з тим час пошуку аудіофрагмента (3 с і 0,75 с) є меншим за його тривалість (відповідно 5 с і 1 с), що дозволяє виконувати пошук у реальному масштабі часу.

Проведено дослідження головних чинників, від яких залежить обчислювальна складність пошуку, оцінювання якої виконувалось за часом пошуку та кількістю операцій (враховувались арифметичні операції, де використовувались координати векторів параметрів та виконувалось обчислення відстаней).

Отримані експериментальні результати показали, що кількість операцій при зростанні кількості аудіозаписів корпусу в 100 разів зростає лише в 10 разів. Час пошуку найближчого аудіозапису в корпусі з 1000 аудіозаписів для 20 кандидатів (найближчих центроїдів) становить 3 с, що менше тривалості аудіофрагмента (5 с) та значно менше порівняно з повним пошуком.

Показано, що зменшення розмірності простору параметрів (рисунок 7) та виконання наближеного пошуку на основі kd-дерева з заданим коефіцієнтом розповсюдження оберненої фази пошуку ($0 < r < 1$) (рисунок 8) дозволяє значно зменшити необхідну кількість операцій для визначення власного аудіозапису.

На рисунку 7 праворуч показано кількість випадків, у яких власний аудіозапис зайняв першу позицію у списку найближчих (Right, зелений колір), кількість випадків потрапляння власного аудіозапису в список п'яти найближчих

(Clarify, помаранчевий колір) або не потрапляння до числа п'яти найближчих аудіозаписів (Wrong, червоний колір).

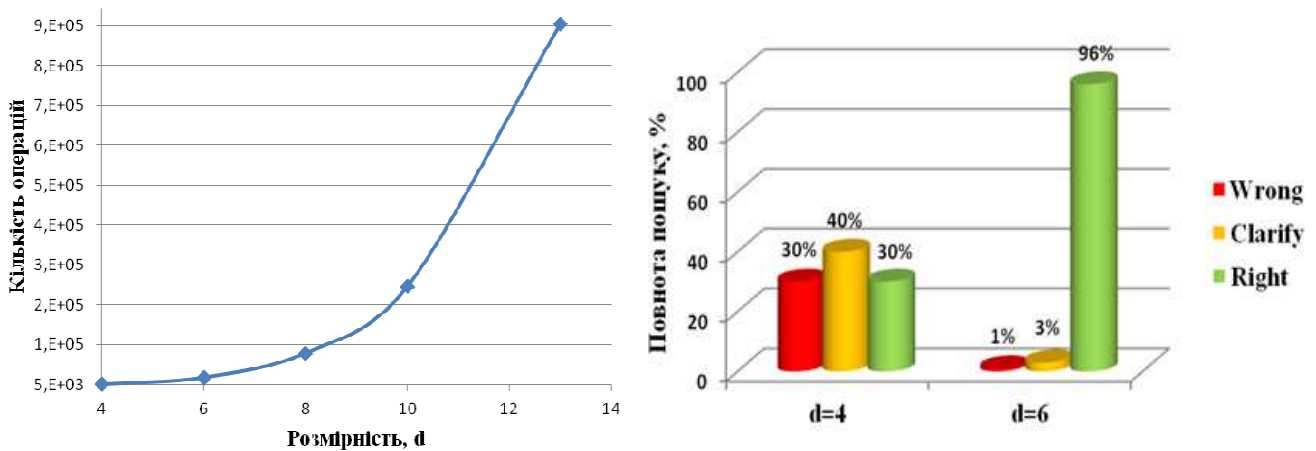


Рисунок 7 – Залежність кількості операцій пошуку найближчого аудіозапису (зліва) та повноти пошуку (справа), % від розмірності параметрів d

За отриманими результатами досліджень при зменшенні розмірності до $d=6$ середній час пошуку власного аудіозапису в корпусі склав менше 0,1 с, причому аудіофрагмент не було визначено правильно у 4% випадків.

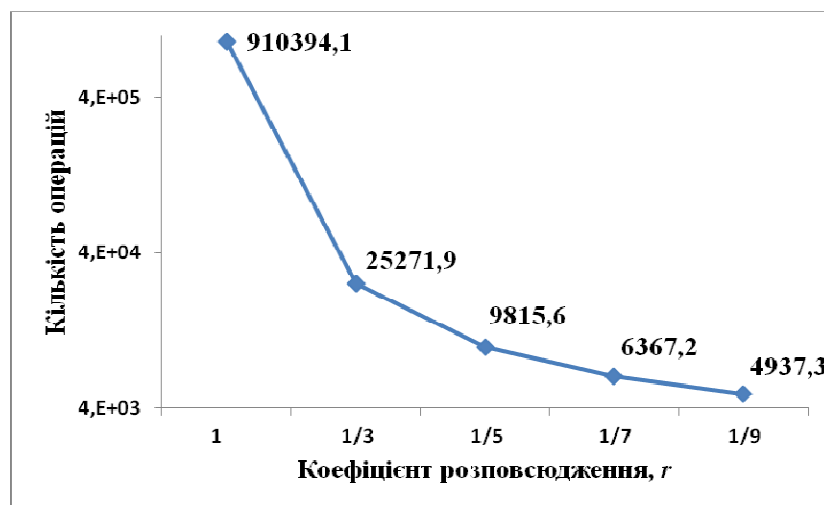


Рисунок 8 – Залежність кількості операцій пошуку найближчого аудіозапису від коефіцієнта r для $d=13$

Визначено, що компроміс між обчислювальною складністю методу та повнотою і релевантністю пошуку досягається при $d=6$ та $r = \frac{1}{5}$: сукупне зниження розмірності простору параметрів та виконання наближеного пошуку ($d=6$, $r = \frac{1}{5}$) дозволяє досягти зменшення складності обчислень в 388 разів при повноті пошуку 90%.

В дисертаційній роботі для досягнення подальшого зменшення складності обчислень при збереженні повноти та релевантності результатів на основі аналізу основних чинників впливу на продуктивність пошуку запропоновано метод комбінованого пошуку аудіозапису за заданим аудіофрагментом.

Суть методу формулюється таким чином: спочатку виконується швидкий пошук (з коефіцієнтом розповсюдження оберненої фази пошуку $0 < r < 1$) в корпусі зниженої розмірності, упорядкованому на основі kd-дерева, після чого серед кількох найбільш подібних до заданого аудіофрагмента аудіозаписів здійснюється повний пошук, під час якого виконується порівняння параметрів MFCC (а не центроїдів) кожного з цих аудіозаписів з параметрами невідомого аудіофрагмента, причому для виконання результативного пошуку достатньо лише декількох фреймів.

Оскільки кількість музичних треків ($Q' = 5$, $Q' \ll Q$) та тривалість аудіофрагмента ($F'_{fragment} = 10$ фреймів), для яких виконується уточнення результату, є невеликими та сталими, етап уточнення не потребує великої кількості операцій. Порівнюючи складність обчислень першого та другого етапів комбінованого пошуку можна сказати, що це сумірні значення.

Згідно з отриманими результатами при зниженні розмірності до $d=6$ повнота пошуку, виключно на основі kd-дерева, склала 96%, 88% і 75% для різних оцінок близькості, застосування методу комбінованого пошуку вже дозволило правильно визначити 99%, 97% та 96% аудіозаписів корпусу.

Застосування комбінованого методу пошуку порівняно з точним пошуком на основі kd-дерева ($d=13$, $r=1$) дозволяє знизити складність обчислень приблизно в 12 разів при $d=13$ ($r = \frac{1}{5}$), приблизно в 17 разів при $d=6$ ($r=1$), та приблизно в 28 разів при $d=6$ та $r = \frac{1}{5}$.

Отже, запропонований метод комбінованого пошуку придатний для виконання пошуку в великих архівах аудіозаписів та реальних додатках, де критичною є швидкість пошуку.

У **четвертому розділі** розроблено структуру, алгоритми та програмне забезпечення для реалізації інформаційної технології пошуку заданих фрагментів в архіві аудіозаписів, наведено відомості щодо застосування інформаційної технології для розв'язання задачі пошуку заданих аудіофрагментів і моніторингу радіомовлення та наведено оцінювання ефективності розробленої в роботі інформаційної технології пошуку заданих фрагментів в архіві аудіозаписів, сформульовано методичні рекомендації щодо використання розробленої системи пошуку, а також вимоги до апаратної та програмної частини комплексу.

На рисунку 9 показано загальну структуру розробленої системи пошуку.

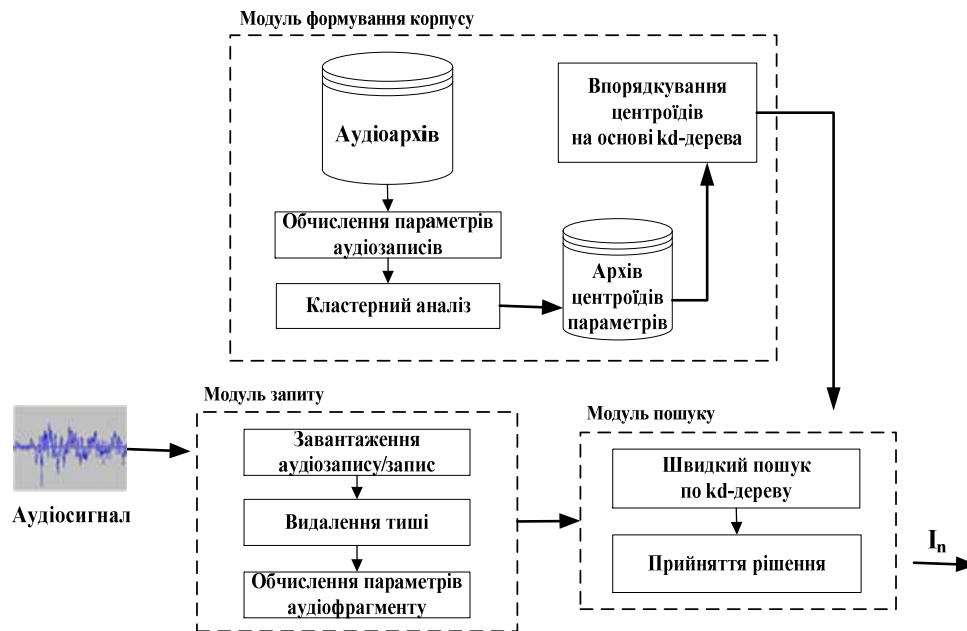


Рисунок 9 – Структура системи пошуку аудіофрагмента в корпусі аудіозаписів

Основні компоненти інформаційної системи: взаємодія з користувачем (графічний інтерфейс); завантаження/запис аудіофрагмента; попередня обробка аудіофрагмента з метою його підготовки до пошуку; зберігання архіву аудіозаписів, бази параметрів аудіозаписів та їх метаданих (корпус); порівняння параметричного подання аудіофрагмента з базою параметрів аудіозаписів корпусу та повернення результатів пошуку у вигляді метаданих про знайдений аудіозапис в корпусі та безпосередньо самого аудіозапису з можливістю його відтворення (підсистема пошуку аудіофрагмента). Таким чином, структуру інформаційної технології як сукупності компонентів, що утворюють єдине ціле, показано на діаграмі, зображеній на рисунку 10.

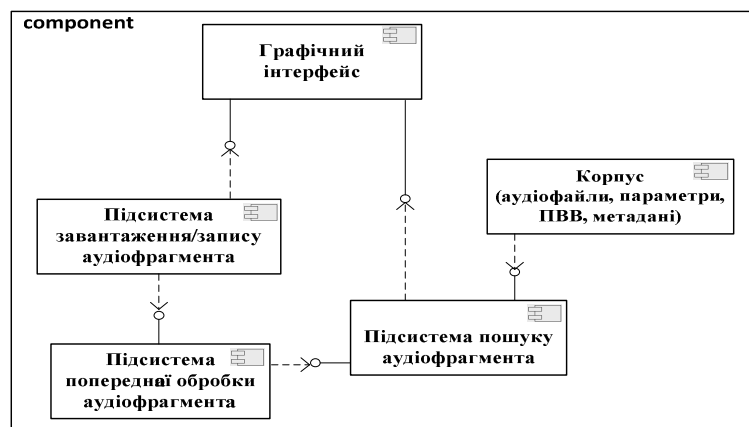


Рисунок 10 – UML-діаграма компонентів інформаційної технології пошуку заданих аудіофрагментів в архіві аудіозаписів

Розроблена інформаційна технологія пошуку заданих аудіофрагментів в архіві аудіозаписів містить алгоритми та програми для реалізації методів,

запропонованих в розділах 2 та 3, та складається з двох великих етапів: етап формування корпусу аудіозаписів та етап безпосереднього пошуку аудіофрагмента в корпусі аудіозаписів.

Корпус аудіозаписів займає важливе місце у створенні інформаційної технології, оскільки від його структури істотно залежить швидкодія та стійкість інформаційної технології пошуку аудіофрагмента. Запропоновано модель корпусу аудіозаписів, яка містить: масив аудіофайлів; базу параметрів, тобто текстові файли з параметрами MFCC (вектори розмірністю $d=13$) для кожного аудіозапису корпусу, єдиний текстовий файл з центроїдами векторів параметрів MFCC усіх аудіозаписів (1000 – для кожного аудіозапису в корпусі), а також значення ПВВ для кожного аудіозапису; динамічне kd-дерево – файл з центроїдами векторів параметрів MFCC для усіх аудіозаписів корпусу, впорядкованих (центроїдів) на основі структури даних kd-дерева; метадані (наприклад, назву, тривалість аудіозапису), що дозволила реалізувати інформаційну технологію пошуку аудіозаписів з різною тривалістю звучання.

Розроблене програмне забезпечення можна використовувати для отримання відомостей про певний аудіофрагмент (наприклад, записаний з радіо чи з іншого джерела), або визначення чи зустрічається аудіофрагмент в певному архіві аудіозаписів. Середній час пошуку (Windows7, Intel Core i5-4200M CPU @ 2.5GHz, 6GB RAM) становить менше 1 с. Тестова версія корпусу містить 1000 аудіозаписів, однак методи, що реалізовані в даному програмному продукті, забезпечують можливості розширення архіву аудіозаписів для пошуку в корпусі.

В дисертаційній роботі розглянуто застосування запропонованих методів пошуку для моніторингу радіомовлення, що полягає в необхідності визначення кількості повторень аудіофрагмента (реклами, музичного треку тощо) в радіоефірі за певний період часу та фіксації інформації про трансляцію (назва аудіофайлу, класифікаційна категорія, час трансляції, тривалість відтворення тощо). На даний час для офіційного телерадіомоніторингу аналогового та цифрового мовлення є актуальною задача автоматичного пошуку заданих аудіофрагментів (заставок, реклам, музичних відрізків і т. д.), що прозвучали в теле- та радіопередачах, їх контрольного прослуховування оператором та внесення відповідних записів в базу даних офіційного моніторингу телерадіопрограм. Об'єктом аналізу та пошуку заданих аудіофрагментів є файли телерадіопрограм, що фактично є архівом аудіозаписів телерадіомоніторингу. Аналіз здійснюється з метою планового і оперативного контролю відповідності змісту програм Законам України про телерадіомовлення, рекламу та іншим чинним документам, які стосуються діяльності Національної ради України з питань телебачення та радіомовлення.

З метою оцінювання ефективності розробленої технології для задачі моніторингу радіомовлення було здійснено запис 25 год. ефірного часу місцевої радіостанції. Було отримано 100 аудіозаписів, кожен тривалістю 15 хв, з яких було сформовано корпус шляхом обчислення MFCC-параметрів та їх кластеризації. З кожного аудіозапису випадковим чином було обрано фрагмент

тривалістю 5 с. Порівняння проводилося з системою з відкритим кодом EchoPrint. Результати порівняння наведено на рисунку 11.



Рисунок 11 – Порівняння результатів пошуку аудіозаписів в архіві з системою EchoPrint для задачі моніторингу радіомовлення

Отримані значення повноти (92%) та релевантності (87%) результатів пошуку аудіофрагментів для задачі моніторингу радіомовлення продемонстрували перевагу запропонованої технології порівняно з системою EchoPrint (86% – повнота, 78% – релевантність). Час пошуку дорівнює 0,1 с для обох систем. Таким чином, розроблена інформаційна технологія дозволяє здійснювати автоматичне відстеження факту відтворення аудіофрагмента в ефірі радіостанції.

В **додатках** наведено основні лістинги програмного забезпечення, розробленого для реалізації інформаційної технології та методик щодо її використання для пошуку заданих аудіофрагментів в архіві аудіозаписів, документи та відомості про впровадження результатів, що отримані в дисертаційному дослідженні.

ВИСНОВКИ

В дисертаційній роботі розв'язано актуальну науково-технічну задачу підвищення повноти, релевантності та швидкості пошуку за рахунок розроблення інформаційної технології пошуку заданих фрагментів в архіві аудіозаписів із застосуванням kd-дерев.

1. Проведено огляд існуючих методів і основних підходів до автоматичного оцінювання подібності аудіозаписів. Розглянуто сфери застосування пошуку на основі аудіоконтенту. Проведено аналіз параметрів математичної моделі опису аудіосигналів, методів класифікації та методів швидкого пошуку. Обґрунтовано вибір MFCC як параметрів математичної моделі опису аудіозаписів, а також проведено дослідження їхніх основних характеристик для розв'язання задачі пошуку аудіозаписів.

2. Запропоновано модель корпусу аудіозаписів, яка містить масив файлів, базу параметрів та метаданих і динамічне kd-дерево, що дозволило зменшити

тривалість обраного для пошуку аудіофрагмента та реалізувати інформаційну технологію пошуку аудіозаписів з різною тривалістю звучання. Показано, що мінімальна тривалість аудіофрагмента, яка забезпечує повноту результатів пошуку аудіозапису 99,4%, має складати 1 с, що у 5–10 разів менше порівняно з іншими методами.

3. Розроблено метод комбінованого пошуку аудіозапису за заданим аудіофрагментом, який базується на виконанні швидкого наближеного пошуку у просторі параметрів зменшеної розмірності із застосуванням kd-дерева кількох найближчих аудіозаписів на першому етапі пошуку, серед яких на другому етапі відбувається вибір власного аудіозапису. Застосування методу дозволило досягти зменшення складності обчислень у 28 разів порівняно з точним пошуком на основі kd-дерева. Метод комбінованого пошуку дозволяє скоротити час пошуку порівняно з методом LMSST в 1,4 раза та, порівняно з методом MIVTMP, у 2,2 раза без погіршення повноти та релевантності пошуку.

4. Отримав подальшого розвитку метод пошуку на основі kd-дерева з удосконаленою оцінкою міри близькості за зваженою кількістю потраплянь в список найближчих центроїдів. Запропонований метод дозволив підвищити повноту та релевантність результатів пошуку, а саме: для аудіофрагмента тривалістю 5с повнота пошуку становить 99,8%, релевантність – 97,4%, для аудіофрагмента 1с – 99,6% та 94,3% відповідно, що перевищує аналогічні показники існуючих методів.

5. Обґрунтовано можливість та доцільність формування корпусу аудіозаписів на основі кластерного аналізу, що дозволяє значно скоротити обсяги пом'яті, необхідні для зберігання аудіозаписів, і, відповідно, забезпечує можливість подальшого розширення корпусу. Удосконалено метод кластеризації k-середніх з удосконаленою процедурою вибору вектора на позицію вставки нового центроїда шляхом послідовного запуску методу k-середніх, що дозволило зменшити похибку кластеризації на 11%.

Розроблено алгоритми та програмне забезпечення для реалізації інформаційної технології пошуку заданих аудіофрагментів в архіві аудіозаписів. Розроблено методика та програмні засоби для дослідження методів швидкого пошуку. Запропоновані методи та розроблені на їх основі алгоритми й програмні засоби дозволяють проводити пошук заданого аудіофрагмента у реальному масштабі часу. Результати досліджень впроваджено в Інституті електроніки та зв'язку Української академії наук, а також у навчальний процес у Вінницькому національному технічному університеті з дисципліни "Цифрова обробка сигналів".

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Ткаченко А. Н. Метод направленного поиска векторов в кодовых книгах / А. Н. Ткаченко, О. Ф. Грийо Тукало // Электронное моделирование. – 2010. – № 2, Т. 32. – С. 77–85. – ISSN 0204-3572.
2. Ткаченко О. М. Метод кластеризації на основі послідовного запуску

к-середніх з удосконаленим вибором кандидата на нову позицію вставки [Електронний ресурс] / О. М.Ткаченко, О. Ф. Грійо Тукало, О. В. Дзись, С. М. Лаховець // Наукові праці ВНТУ. – 2012. – № 2, В. 2. – . Режим доступу до журн.: <http://praci.vntu.edu.ua/index.php/praci/article/view/323/321>. – ISSN 2307-5376.

3. Ткаченко О. М. Пошук векторів у кодових книгах при ущільненні мовлення на основі бінарного дерева / О. М. Ткаченко, О. Ф. Грійо Тукало // Інформаційні технології та комп'ютерна інженерія. – 2011. – № 1. – С. 38–44. – ISSN 1999-9941.

4. Ткаченко О. М. Метод кластеризації на основі послідовного запуску к-середніх з обчисленням відстаней до активних центроїдів / О. М. Ткаченко, О. Ф. Грійо Тукало, Н. О. Біліченко, О. В. Дзись // Реєстрація, зберігання і обробка даних. – 2012. – № 1, Т. 14, В. 1. – С. 25–34. – ISSN 1560-9189.

5. Ткаченко А. Н. Метод дельта-квантования параметров речевого сигнала с быстрым поиском ближайшего вектора в кодовой книге / А. Н. Ткаченко, О. Ф. Грійо Тукало, А. В. Дзись // Международный научный журнал "Управляющие системы и машины". – 2013. – № 4. – С. 3–11. – ISSN 0130–5395.

6. Ткаченко О. М. Метод швидкого пошуку найближчого сусіда з обчисленням відстані за зваженою евклідовою метрикою / О. М. Ткаченко, О. Ф. Грійо Тукало // Вісник Вінницького політехнічного інституту. – 2013. – № 1. – С. 116–122. – ISSN 1997-9266.

7. Ткаченко О. М. Підхід до оцінювання тривалості фрагмента для пошуку музичного твору за заданим шаблоном / О. М. Ткаченко, О. Ф. Грійо Тукало // Інформаційні технології та комп'ютерна інженерія. – 2014. – № 1.– С. 31–40. – ISSN 1999-9941.

8. Ткаченко О. М. Метод підвищення швидкості пошуку фрагменту аудіозапису із застосуванням kd-дерев / О. М. Ткаченко, О. Ф. Грійо Тукало // Інформаційні технології та комп'ютерна інженерія. – 2014. – № 3.– С. 57–66. – ISSN 1999-9941.

9. Ткаченко О. М. Пошук найближчого вектора у кодових книгах на основі бінарного дерева / О. М. Ткаченко, О. Ф. Грійо Тукало // Інформаційні технології та комп'ютерна інженерія: I Міжнар. наук.-практ. конф., 19–21 травня 2010 р. : тези доповідей. – Вінниця, 2010. – С. 160–161. – 2 с. – ISBN 978-966-641-356-0.

10. Ткаченко О. М. Двоетапна стратегія пошуку найближчого сусіда у векторних кодових книгах / О. М. Ткаченко, О. Ф. Грійо Тукало // Методи та засоби кодування, захисту й ущільнення інформації: III Міжнар. наук.-практ. конф., 20–22 квітня 2011 р. : тези доповідей – Вінниця, 2011. – С. 188–189. – 2 с. – ISBN 978-966-641-406-2.

11. Ткаченко О. М. Метод швидкого пошуку векторів у кодових книгах на основі бінарного дерева / О. М. Ткаченко, О. Ф. Грійо Тукало // Інформаційні технології та комп'ютерна інженерія: III Міжнар. наук.-практ. конф., 29–31 травня 2012 р. : тези доповідей. – Вінниця, 2012.– С. 224–225. – 2 с. – ISBN 978-966-641-465-9.

12. Ткаченко А. Н. Метод поиска ближайшего вектора с заданной

вероятностью / А. Н. Ткаченко, О. Ф. Грийо Тукало // Актуальные проблемы современной науки: II Междунар. науч.-практ. конф., 13–16 березня 2013 р. – №2, В.2. – Россия, Ставрополь, 2013. – С. 195–198. – 4 с. – ISSN 2227-1392.

13. Ткаченко О. М. Ідентифікація музичного твору за фрагментом / О. М. Ткаченко, О. Ф. Грийо Тукало, Ю. Л. Далекий // Інформаційні технології та комп'ютерна інженерія: IV Міжнар. наук.-практ. конф., 28–30 травня 2014 р. : тези доповідей. – Вінниця, 2014. – С. 256–257. – 2 с. – ISBN 978-966-641-465-9.

14. Ткаченко О. М. Ідентифікація аудіозапису за коротким фрагментом на основі приведеної власної відстані / О. М. Ткаченко, О. Ф. Грийо Тукало // Контроль і управління в складних системах: XII Міжнар. конф., 14–16 жовтня 2014 р. : тези доповідей. – Вінниця, 2014. – С. 94. – ISBN 978-966-2462-66-1.

15. Ткаченко О. М. Ідентифікація фрагмента музичного твору на основі приведеної власної відстані / О. М. Ткаченко, О. Ф. Грийо Тукало // Оброблення сигналів і зображень та розпізнавання образів: XII Всеукр. міжнар. конф., 3–7 листопада 2014 р. – Київ, 2014. – С. 23–26. – 4 с. – ISBN 978-966-479-069-4.

16. Ткаченко О. М. Комп'ютерна програма "Програма кластеризації на основі послідовного запуску k-середніх з обчисленням відстаней до активних центрів" / О. М. Ткаченко, О. Ф. Грийо Тукало, О. В. Дзісь, С. М. Лаховець // Свідectво про реєстрацію авторського права на твір № 45264. – К.: Державний департамент інтелектуальної власності України. – Дата реєстрації: 21.08.2012 р.

17. Патент України на корисну модель № 49814, МПК (2009) G10L 19/00, G10L 21/00. Спосіб спрямованого пошуку векторів при ущільненні мовних сигналів./ Ткаченко О. М., Грийо Тукало О. Ф., заявник і патентовласник Вінницький національний технічний університет. – № 200909012 ; заявл. 31.08.2009; опубл. 10.03.2010, Бюл № 5.

18. Патент України на корисну модель № 70762, МПК (2012) G10L 21/00. Спосіб двоетапного пошуку векторів під час ущільнення мовних векторів. / Ткаченко О. М., Грийо Тукало О. Ф., заявник і патентовласник Вінницький національний технічний університет. – № 201114324; заявл. 05.12.2011; опубл. 25.06.2012, Бюл № 12.

19. Грийо Тукало О. Ф. Быстрый поиск векторов при сжатии речевых сигналов : монография / О. Ф. Грийо Тукало, А. Н. Ткаченко // LAP LAMBERT Academic Publishing. – 2014. – 108 с. – ISBN 978-3-659-50369-6.

АНОТАЦІЯ

Грийо Тукало Оксана Франсисківна. Інформаційна технологія пошуку заданих фрагментів в архіві аудіозаписів із застосуванням kd-дерев. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології. – Вінницький національний технічний університет, Вінниця, 2016.

Дисертаційну роботу присвячено підвищенню повноти і релевантності результатів та швидкості пошуку за рахунок розроблення інформаційної технології пошуку заданих фрагментів в архіві аудіозаписів із застосуванням kd-дерев.

З розвитком Інтернету, цифрових ЗМІ, мультимедійних технологій та можливістю доступу до великих архівів мультимедіа виникла необхідність створення нових способів навігації та взаємодії з цими архівами аудіозаписів. Незважаючи на великі успіхи, досягнуті у цій галузі в останні роки, сучасним інформаційно-пошуковим системам притаманні певні недоліки, зокрема:

- недостатня продуктивність під час роботи з великими архівами;
- неможливість їхнього безпосереднього застосування для розв'язання таких задач, як моніторинг трансляції на радіостанціях (телебаченні) для визначення кількості повторень в ефірі рекламного продукту, музичного твору тощо.

Запропоновано модель корпусу аудіозаписів, яка містить масив файлів, базу параметрів та метаданих і динамічне kd-дерево, що дозволило зменшити тривалість обраного для пошуку аудіофрагмента та реалізувати інформаційну технологію пошуку аудіозаписів з різною тривалістю звучання. Показано, що мінімальна тривалість аудіофрагмента, яка забезпечує повноту результатів пошуку аудіозапису 99,4%, має складати 1 с.

Отримав подальшого розвитку метод кластеризації k-середніх, який відрізняється від існуючих удосконаленою процедурою відбору вектора на позицію вставки нового центроїда шляхом послідовного запуску методу k-середніх, що дозволяє отримати розв'язок, наближений до глобального мінімуму похибки кластеризації.

Удосконалено метод пошуку на основі kd-дерева, який, на відміну від існуючих, використовує оцінку міри близькості за зваженою кількістю потраплянь в список найближчих центроїдів, що дозволяє підвищити повноту та релевантність результатів пошуку.

Запропоновано метод комбінованого пошуку аудіозапису за заданим аудіофрагментом, який базується на виконанні швидкого наближеного пошуку у базі параметрів зменшеної розмірності із застосуванням kd-дерева кількох найближчих аудіозаписів на першому етапі пошуку, серед яких на другому етапі відбувається вибір релевантного аудіозапису, що дозволило досягти зменшення часу пошуку порівняно з точним пошуком на основі kd-дерева.

На основі результатів теоретичних досліджень розроблено інформаційну технологію пошуку заданих аудіофрагментів в архіві аудіозаписів, яка містить: алгоритм та програму для реалізації методу кластеризації на основі послідовного запуску k-середніх з удосконаленим вибором вектора на позицію вставки нового центроїда; алгоритм та програму пошуку заданого аудіофрагмента на основі математичного очікування похибки кластеризації; алгоритм та програму швидкого пошуку на основі kd-дерев; методика використання системи автоматичного пошуку аудіозаписів. Запропоновані методи та розроблені на їхній

основі алгоритми та програмні засоби дозволяють проводити пошук заданого аудіофрагмента у реальному масштабі часу.

Результати досліджень впроваджено в Інституті електроніки та зв'язку Української академії наук, а також у навчальний процес у Вінницькому національному технічному університеті з дисципліни "Цифрова обробка сигналів".

Ключові слова: інформаційні технології, автоматичний пошук аудіозапису за аудіофрагментом, наближений пошук найближчого вектора, kd-дерево, архів аудіозаписів, кластерний аналіз, спектральне спотворення, похибка кластеризації, параметризація, векторне квантування, мел-частотні кепстральні коефіцієнти, зменшення розмірності параметрів.

АННОТАЦІЯ

Грийо Тукало Оксана Франсыковна. Информационная технология поиска заданных фрагментов в архиве аудиозаписей с применением kd-деревьев. – На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.06 – информационные технологии. – Винницкий национальный технический университет, Винница, 2016.

Диссертация посвящена повышению полноты и релевантности результатов и скорости поиска за счет разработки информационной технологии поиска заданных фрагментов в архиве аудиозаписей с применением kd-деревьев.

С развитием Интернета, цифровых СМИ, мультимедийных технологий и возможностью доступа к большим архивам мультимедиа возникла необходимость предложить новые способы навигации и взаимодействия с этими архивами аудиозаписей. Несмотря на большие успехи, достигнутые в этой области в последние годы, современным информационно-поисковым системам присущи определенные недостатки, в частности:

- недостаточная производительность при работе с большими архивами;
- невозможность их непосредственного применения для решения таких задач, как мониторинг трансляции на радиостанциях (телевидении) для определения количества повторений в эфире рекламного продукта, музыкального произведения и т. д.

Предложена модель корпуса аудиозаписей, которая содержит массив файлов, базу параметров и метаданных, динамическое kd-дерево, что позволило уменьшить длительность выбранного для поиска аудиофрагмента и реализовать информационную технологию поиска аудиозаписей с разной продолжительностью звучания. Показано, что минимальная продолжительность аудиофрагментов, которая обеспечивает полноту результатов поиска аудиозаписи 99,4%, должна составлять 1 с.

Получил дальнейшее развитие метод кластеризации k-средних, который отличается от существующих улучшенной процедурой выбора вектора на

позицию вставки нового центроида путем последовательного запуска метода k-средних, что позволяет получить решение, приближенное к глобальному минимуму ошибки кластеризации, и лежит в основе формирования базы параметров корпуса аудиозаписей.

Усовершенствован метод поиска на основе kd-дерева, который, в отличие от существующих, использует оценку степени близости по взвешенному количеству попаданий в список ближайших центроидов, что позволяет повысить полноту и релевантность результатов поиска.

Предложен метод комбинированного поиска аудиозаписи по заданному аудиофрагменту, основанный на выполнении быстрого приближенного поиска в базе параметров уменьшенной размерности с применением kd-дерева нескольких ближайших аудиозаписей на первом этапе поиска, среди которых на втором этапе происходит выбор релевантной аудиозаписи, что позволило добиться уменьшения времени поиска по сравнению с точным поиском на основе kd-дерева.

На основе результатов теоретических исследований разработана информационная технология поиска заданных аудиофрагментов в архиве аудиозаписей, которая включает: алгоритм и программу для реализации метода кластеризации на основе последовательного запуска k-средних с усовершенствованным выбором вектора на позицию вставки нового центроида; алгоритм и программу поиска заданного аудиофрагмента на основе математического ожидания погрешности кластеризации; алгоритм и программу быстрого поиска на основе kd-деревьев, методику использования системы автоматического поиска аудиозаписей.

Предложенные методы и разработанные на их основе алгоритмы и программные средства позволяют проводить поиск заданного аудиофрагмента в реальном масштабе времени.

Результаты исследований внедрены в Институте электроники и связи Украинской академии наук, а также в учебный процесс в Винницком национальном техническом университете по дисциплине "Цифровая обработка сигналов".

Ключевые слова: информационные технологии, автоматический поиск аудиозаписи по аудиофрагменту, приближенный поиск ближайшего вектора, kd-дерево, архив аудиозаписей, кластерный анализ, спектральное искажение, погрешность кластеризации, параметризация, векторное квантование, мел-частотные кепстральные коэффициенты, уменьшение размерности параметров.

SUMMARY

Oksana Griyo Tukalo. Information technology for search of the specified audio fragments in the archive of audio recordings using kd-tree. – Manuscript.

Thesis for the scientific degree of the Candidate of Technical Sciences, speciality 05.13.06 – Information Technology. – Vinnytsia National Technical University, Vinnytsia, 2016.

The Dissertation is devoted to improving recall and precision of the search results and increasing the speed on the basis of information technology for search of the specified audio fragments in the archive of audio recordings using kd-tree.

With the development of the Internet, digital media, multimedia technology and access to the large multimedia archives became necessary to create new ways of navigation and interaction with the audio archives. Despite the great progress achieved in this area in the recent years, modern information retrieval systems have some shortcomings, in particular:

- the insufficient performance with regard to the large archives;
- the impossibility of the direct application to solve such tasks as broadcast monitoring on the radio (television) to determine the repetition number of some audio recording (e.g. an advertising product, a musical piece, etc.) during translation.

The offered model of the audio body, which contains an array of audiofiles, a database of parameters and metadata, and dynamic kd-tree, allowed to reduce the audio fragment duration, selected for the search, and implement information technology for search of audio recordings with the different duration. It is shown that the lowest duration of the audio fragment that ensures the recall of the search results at 99.4%, should be 1s.

The method of k-means clustering got the further development, differs from the existing ones by the improved procedure of selecting the vector as a new centroid by successive run of k-means that provides a solution, close to the global minimum of the clustering error.

The search method based on kd-tree is improved, which unlike the existing, uses the improved estimation of the measure proximity, enhancing the rollback and precision of search results.

The method of the combined search of audio recording for a specified audio fragment is proposed, based on the implementation of the fast approximate kd-tree search of the several nearest recordings in the parameter database of the reduced dimension on the first stage of the search, among them the relevant audio recording is defined on the second stage, which resulted in reducing the time search compared to an accurate kd-tree search.

Based on the results of the theoretical research the information technology for search of the specified audio fragments in the archive of audio recordings is developed that includes an algorithm and a program for implementing the clustering method based on the successive run of k-means with the improved procedure of selecting the vector as a new centroid; an algorithm and a program for search of the specified audio fragment on the basis of the mean value of the clustering error; an algorithm and a program for the fast search on the basis of kd-tree, a technique of using the automatic search of audio recordings.

The proposed methods, as well as algorithms and software, developed on their basis, allowed to search the specified audio fragment in real time.

The research results implemented at the Institute of Electronics and Communications of the Ukrainian Academy of Sciences and in the educational process of VNTU at the course "Digital Signal Processing".

Keywords: information technology, automatic search of an audio recording by the specified audio fragment, approximate search for the nearest vector, kd-tree, archive of audio recordings, cluster analysis, spectral distortion, clustering error, parameterization, vector quantization, mel-frequency cepstral coefficients, reducing the dimension of parameters.

Підписано до друку 12.11.2015 р. Формат 29,7x42¼
Наклад 100 пр. Зам. № 2016-024

Віддруковано в комп'ютерному інформаційно-видавничому центрі
Вінницького національного технічного університету
Вінниця, вул. Хмельницьке шосе, 95. Тел. 59-81-59.
publish.vntu.edu.ua; email: kivc.vntu@gmail.com
Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.