

**Хоткін І.О**

## **ОТРИМАННЯ АУДІО-ВІДБИТКІВ**

*Аудіо-відбиток - це компактний підпис на основі вмісту, який підсумовує аудіозапис. Технології аудіо відбитків пальців привернули увагу, оскільки вони дозволяють ідентифікувати звук незалежно від його формату. Ця стаття має на меті дати бачення звукового відбитку. Наведено обґрунтування, а також відмінності водяного маркування. Описано основні вимоги систем аудіо-відбитків. Прокоментовано основні режими використання аудіо-відбитків, а саме ідентифікація, автентифікація, створення секретного ключа на основі контенту для водяного маркування та аудіо на основі вилучення вмісту.*

**Ключові слова:** аудіо дактилоскопія, контент-ідентифікація на основі вмісту, водна маркування, перевірка цілісності, пошук аудіоінформації, надійне хешування

**Вступ.** Аудіо-відбиток відомий завдяки його здатності пов'язувати неназначений звук до відповідних метаданих (наприклад, ім'я виконавця та назва пісні), незалежно від формату звуку. Аудіо-відбитки витягують перцептивний довідник аудіовмісту, тобто відбитка, і зберігають його в базі даних. Коли він подається з маркованим аудіо, його відбиток обчислюється і порівнюється з тими, що зберігаються в базі даних. Використовуючи відбитки та алгоритми відповідності, спотворені версії запису можуть бути ідентифіковані як однаковий звуковий вміст.

**Постановка проблеми.** Джерело труднощів при автоматичному визначенні аудіоконтенту впливає з його високої розмірності, значної розбіжності аудіоданих для сприйняття подібного вмісту та необхідності ефективного порівняння відбитків із величезною колекцією зареєстрованих відбитків у базі даних. Найлегшим підходом буде пряме порівняння оцифрованої форми хвилі, але такий спосіб не є ефективним. Більш ефективна реалізація цього підходу може використовувати хеш-метод, такий як MD5 (Message Digest 5) або CRC (Cyclic Redundancy Checking), щоб отримати компактне представлення двійкового файлу. У цій структурі можна порівнювати хеш-значення замість цілих файлів. Однак значення хешу є крихкими, одного перевертання біта достатньо, щоб хеш повністю змінився. Звичайно, ця установка не є надійною для стиснення або мінімальних спотворень будь-якого виду, і насправді вона не може розглядатися як ідентифікація на основі контенту, оскільки вона не розглядає вміст, зрозумілий як інформацію, лише біти.

**Мета статті.** Дати бачення аудіо-відбитку і розглянути пов'язані з ним складнощі.

**Основний текст.** Аудіо-відбиток - це стислий підпис на основі вмісту, який підсумовує аудіозапис. Аудіо-відбитки привернули багато уваги своїми можливостями аудіо-ідентифікації. Технології аудіо-відбитків витягують акустичні характеристики звукового вмісту та зберігають їх у базі даних. Коли вони представлені з невстановленим фрагментом звукового вмісту, характеристики цього фрагмента обчислюються та співпадають із тими, що зберігаються в базі даних. Використовуючи відбитки та відповідні алгоритми, деформовані версії одного запису можуть бути ідентифіковані як однаковий аудіозапис.

Такий спосіб відрізняється від альтернативного існуючого рішення для моніторингу аудіовмісту: водяного маркування аудіо. У водяному маркуванні [1] дослідження психоакустики проводяться так, що довільне повідомлення може бути вбудоване в запис, не змінюючи сприйняття звуку. Відповідні пристрої можуть перевірити наявність водяного знаку, перш ніж приступати до операцій, які можуть призвести до порушення авторських прав. При влученні аудіо-відбитків, повідомлення автоматично виводиться з найбільш релевантних компонентів звуку. У порівнянні з водяним маркуванням, такий спосіб ідеально менш вразливий до атак та деформацій, оскільки спроба змінити відбиток, означає зміну якості звуку.

Недоліком буде складність аудіо-відбитку: вона вища, ніж водяне маркування, і є необхідність підключення до сховища аудіо-відбитків. Крім того, на відміну від водяного маркування, повідомлення не є незалежним від вмісту. Тому, наприклад, неможливо розрізнити перцептивно однакові копії запису. Як і у технології водяних знаків, у знятті аудіо-відбитків є більш корисне використання, ніж просто ідентифікація. Зокрема, вони також можуть бути використовуватись для перевірки цілісності вмісту; аналогічно крихким водяним знакам.

Нижче наведено бажані властивості хорошої, надійної та ефективною аудіо-системи відбитків пальців [2] [3]:

**Точність.** Вимірює ступінь правильності результатів ідентифікації. Загальна кількість правильних, пропущених та неправильних ідентифікацій (помилкових позитивних результатів).

**Складність:** це стосується обчислювальних накладних витрат і витрат, пов'язаних з вилученням різних відбитків пальців, які включають розмір відбитка пальця, складність використовуваного алгоритму пошуку, складність порівняння відбитків пальців та їх відповідність, витрати на додавання нових елементів до бази даних тощо. Це надзвичайно важливий і релевантний параметр для систем, які потребують роботи і

розгортання в режимі реального часу, а також для систем з обмеженими обчислювальними ресурсами або пропускнуою здатністю.

**Розмір відбитків:** це кількість елементів або бітів, вилучених за секунду чи пісню. Для забезпечення масштабованості системи та бази даних розмір відбитків пальців повинен бути якомога меншим. Розмір відбитків пальців безпосередньо пов'язаний з кількістю відбитків, які можна представити, і з деталізацією. Чим більша швидкість відбитків, тим дрібніше зернистість.

**Зернистість:** мінімальна довжина аудіо, необхідна для надійної ідентифікації. На основі невеликого фрагмента звукову доріжку можна точно визначити. Коли система є дрібнозернистою, це означає, що система здатна надійно ідентифікувати аудіо, використовуючи невеликі уривки.

**Надійність:** різні методи оцінки наявності запиту в колекції предметів для ідентифікації мають головне значення для створення списку відтворення для правозахисних організацій. У таких ситуаціях, якщо пісня ще не транслювалася, її не слід ідентифікувати як збіг, навіть ціною відсутності справжніх збігів.

**Міцність:** здатність точно ідентифікувати предмет, незалежно від рівня стиснення та деформації, або перешкод у каналі передачі та протистояти ефекту операцій обробки сигналу.

**Масштабованість:** аудіо-система повинна бути масштабованою, щоб зберігати велику кількість відбитків. На це впливають як параметри бази даних (швидкість пошуку, ефективність пошуку, структури індексації), так і самі параметри відбитків (скільки відбитків можна розрізнити надійно).

Більш детальне перерахування вимог може допомогти розрізнити різні підходи [4].

Удосконалення певної вимоги часто означає втрату працездатності в деяких інших. Як правило, відбиток повинен бути:

- Перцептивний довідник запису. Відбиток повинен зберігати максимум акустично-відповідної інформації. Цей довідник повинен дозволяти дискримінувати велику кількість відбитків. Це може суперечити іншим вимогам, таким як складність та надійність.
- Незмінний до викривлень. Це впливає з вимоги надійності. Програми цілісності вмісту, однак, знімають це обмеження для збереження вмісту деформацій, щоб виявити навмисні маніпуляції.
- Компактний. Представлення невеликого розміру цікаве за складністю, оскільки потрібно зберігати та порівнювати велику кількість (можливо мільйони) відбитків. Однак занадто коротке представлення може бути недостатньо для дискримінації між записами, що впливає, таким чином, на точність та надійність.
- Легко обчислюється. Вилучення відбитків не повинно займати багато часу.

Незалежно від конкретного підходу до отримання компактного підпису на основі вмісту, може бути розроблена загальна архітектура для опису функціональності відбитків при використанні для ідентифікації.

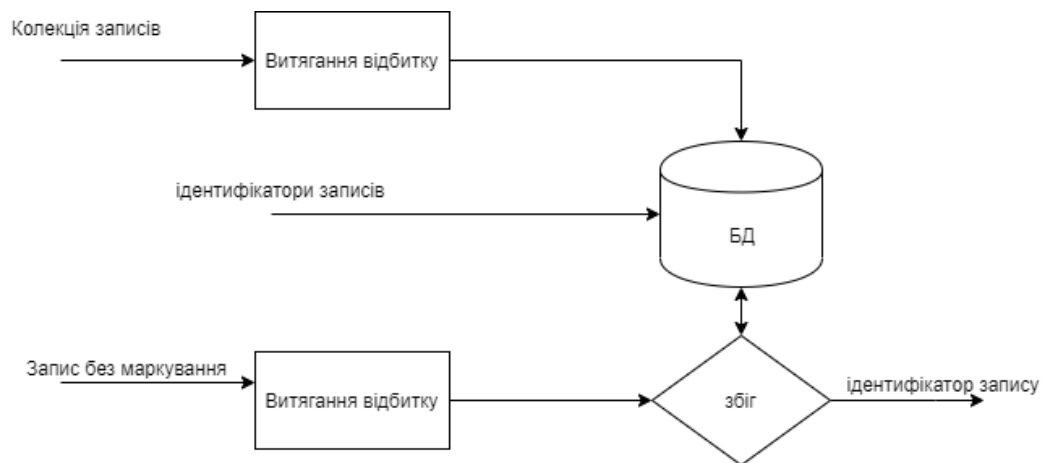


Рисунок 1.1 Контент-система аудіоідентифікації на основі вмісту

Перевірка цілісності спрямована на виявлення зміни даних. Загальна функціональність (рис. 1.1) аналогічна ідентифікації. Спочатку відбиток витягується з оригінального аудіо. На етапі перевірки аудіо-відбиток, витягнутий з тестового сигналу, порівнюється з відбитком оригіналу. В результаті виводиться звіт, який вказує, чи був оброблений сигнал. Дані перевірки, які повинні бути значно меншими, ніж аудіодані, можуть надсилатися разом із вихідними аудіоданими або зберігатися в базі даних. Метод, відомий як самовстановлення, дозволяє уникнути потреби в базі даних або спеціально виділеному заголовку шляхом вбудовування контенту на основі вмісту в аудіодані за допомогою водяного маркування.

Фактичні реалізації звукового відбитку зазвичай відповідають наведеній схемі (рис. 1.2), з різницями щодо акустичних особливостей, що спостерігаються, та моделювання звуку, а також алгоритмів відповідності та пошуку. Найпростішим підходом було б пряме порівняння файлів. Спосіб ефективною реалізації цієї ідеї полягає у використанні хеш-методу, такого як MD5 (Message Digest 5) або CRC (Cyclic Redundancy Checking) для отримання компактного представлення двійкового файлу. У цьому налаштуванні можна порівняти

компактні підписи замість цілих файлів. Такий підхід не є надійним для стиснення або деформації будь-якого виду, і навіть не може розглядатися як ідентифікація аудіо за контентом, оскільки не заснована на аналізі вмісту, а лише на маніпуляції, що виконуються на двійкових даних. Цей підхід був би невідповідним для моніторингу потокового або аналогового аудіо; однак, він встановлює основу для класу методів відбитків: надійне або перцептивне хешування [5] [6]. Ідея надійного хешування - це включення акустичних функцій у хеш-функцію так, що остаточний хеш-код є надійним для аудіо-маніпуляцій, доки вміст зберігається. У літературі можна знайти багато особливостей звукової характеристики, такі як енергія, гучність, нульова швидкість перетину, крок, гармонійність, спектральна площинність [7] мел-частотні кепстральні коефіцієнти (MFCC).

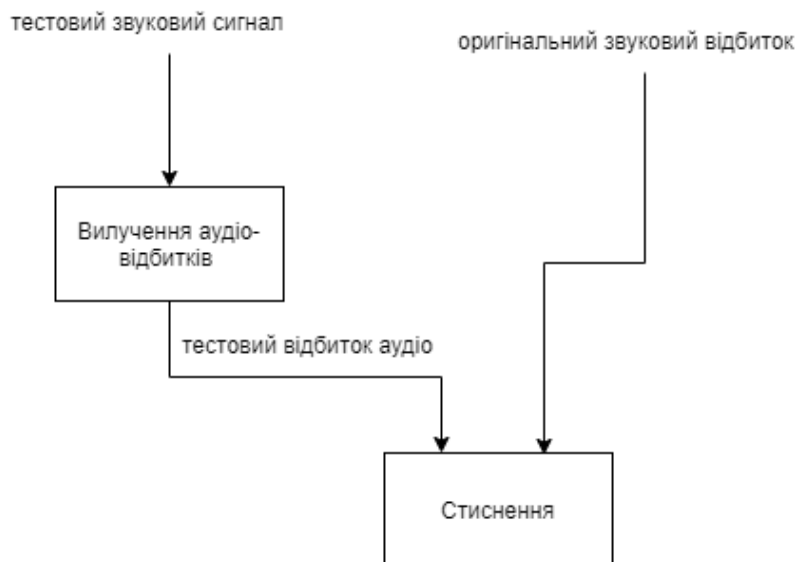


Рисунок 1.2 Структура перевірки цілісності

Притаманна складність завдання ідентифікації звуку, головним чином, пов'язана з різницею між оригінальними заголовками та широкомовними: пісня може бути частково передана, динамік може говорити поверх різних сегментів пісні, твір може відтворюватися швидше і кілька ефектів маніпуляції можуть бути застосовані для посилення психоакустичного впливу слухача (компресори, вирівнювання, підсилювач басів тощо).

Проте система також повинна бути швидкою, оскільки вона повинна проводити порівняння з кількома тисячами (порядком 100 000) пісень в режимі он-лайн. Це впливає на пам'ять та обчислювальні реквізити, оскільки система повинна спостерігати за кількома радіостанціями, давати результати в Інтернеті і не повинна бути дуже дорогою з точки зору обладнання.

Ідеальна система аудіо-відбитків повинна відповідати декільком вимогам. Вона повинна мати можливість точно ідентифікувати предмет, незалежно від рівня стиснення та спотворення або перешкод у каналі передачі. Залежно від програми, вона повинна мати можливість ідентифікувати заголовки з уривків лише за кілька секунд. Система відбитків також повинна бути ефективною в обчисленні. Ефективність має вирішальне значення в реальному застосуванні як при обчисленні відбитків невідомого аудіо, так і в пошуку найкращого збігу у величезному сховищі аудіо-відбитків. Ця обчислювальна вартість пов'язана з розміром відбитків, складністю алгоритму пошуку та складністю вилучення відбитків.

**Висновки.** проаналізовано способи вилучення звукових відбитків, а також деякі можливі сценарії використання та контексти додатків. Переглянуті бажані вимоги в схемі відбитків, які підтверджують наявність компромісу між ними. Більшість додатків отримують можливість пов'язувати вміст із маркованим аудіо, але у цієї технології є більше варіантів для використання.

## Література

1. L. Boney, A. Tewfik, K. Hamdy, Цифрові водяні знаки для аудіосигналів / L. Boney, A. Tewfik, K. Hamdy 473-480с., 1996.
2. Моделювання аудіо відбитків: Структура, деформація, ємність / P. J. O Doets
3. Огляд аудіо-відбитків / Pedro Cano, Eloi Battle
4. T. Kalker, Програми та виклики для аудіо-відбитків / 111th AES Конвенція, Нью Йорк, 2001
5. M.K. Mihchak, R. Venkatesan, Перцептивний алгоритм звучання аудіо. Праці 4-го семінару з питань приховування інформації, Пітсбург. / PA, Ар. 2001

6. J. Haitsma, T. Kalker, and J. Oostveen, "Надійне хешування звуку для ідентифікації контенту / Праці Міжнародного семінару з питань контентної мультимедійної індексації, Brescia, Italy, Sept. 2001.
7. E. Allamanche, J. Herre, O. Helmuth, B. Froba, " T. Kasten, and M. Cremer, Зміст ідентифікації аудіоматеріалів за допомогою опису низького рівня MPEG-7 / Праці Міжнародного симпозиуму пошуку музичної інформації, Bloomington, IN, Oct. 2001.

### References

1. L. Boney, A. Tewfik, and K. Hamdy, "Digital Watermarks for Audio Signals," IEEE Proceedings Multimedia, 473-480, 1996.
2. Modeling Audio Fingerprints : Structure, Distortion, Capacity by P. J. O Doets
3. A Review of Audio Fingerprinting by Pedro Cano AndEloi Battle.
4. T. Kalker, "Applications and Challenges for Audio Fingerprinting", presentation at the 111th AES Convention, New York, 2001
5. M.K. Mihchak and R. Venkatesan, "A Perceptual Audio Hashing Algorithm: a Tool for Robust Audio Identification and Information Hiding," Proceedings of the 4th Workshop on Information Hiding, Pittsburg, PA, Apr. 2001.
6. J. Haitsma, T. Kalker, and J. Oostveen, "Robust Audio Hashing for Content Identification," Proceedings of the International Workshop on Content-Based Multimedia Indexing, Brescia, Italy, Sept. 2001.
7. E. Allamanche, J. Herre, O. Helmuth, B. Froba, " T. Kasten, and M. Cremer, "Content-Based Identification of Audio Material Using MPEG-7 Low Level Description," Proceedings of the International Symposium of Music Information Retrieval, Bloomington, IN, Oct. 2001.

### Обзор отпечатков аудио

Аудио-отпечаток пальца - это компактный справочник, полученный из воспринимаемых аспектов записи. Технологии аудио-отпечатков привлекли внимание, поскольку они позволяют идентифицировать звук независимо от его формата. Цель данной статьи - дать видение звукового отпечатка. Приведены обоснования, а также различия по водяной маркировке. Описаны основные требования систем аудио-отпечатков. Прокомментированы основные режимы использования аудио-отпечатков, а именно идентификация, аутентификация, создание секретного ключа на основе контента для водяной маркировки и аудио на основе извлечения содержания.

**Ключевые слова:** аудио дактилоскопия, контент-идентификация на основе содержания, водная маркировка, проверка целостности, поиск аудиоинформации, надежное хеширование.

### A Review of Audio Fingerprinting

An audio fingerprint is a compact digest derived from perceptually relevant aspects of a recording. Fingerprinting technologies have attracted attention since they allow the identification of audio independently of its format. This paper aims to give a vision on Audio Fingerprinting. The rationale is presented along with the differences with respect to watermarking. The main requirements of fingerprinting systems are described. The main modes of using audio fingerprints, namely identification, authentication, the creation of a secret key based on content for watermarking and audio based on the extraction of content, have been commented on.

**Keywords:** audio fingerprinting, content-based audio identification, watermarking, integrity verification, audio information retrieval, robust hashing.

**Хоткін І.О.** – магістр групи КН-18Дм кафедри комп'ютерних наук та інженерії Східноукраїнського національного університету імені Володимира Даля, e-mail: khotk.ilya@gmail.com