

**Василь Васильович Веремєнко<sup>1</sup>, Андрій Іванович Роговенко<sup>2</sup>**

<sup>1</sup>аспірант кафедри інформаційних та комп'ютерних систем  
Національний університет «Чернігівська політехніка» (Чернігів, Україна)  
E-mail: [vasiliy.veremeenko@gmail.com](mailto:vasiliy.veremeenko@gmail.com). ORCID: <https://orcid.org/0009-0004-2742-6367>

<sup>2</sup>кандидат технічних наук, доцент кафедри інформаційних та комп'ютерних систем  
Національний університет «Чернігівська політехніка» (Чернігів, Україна)  
E-mail: [arogovenko@stu.cn.ua](mailto:arogovenko@stu.cn.ua). ORCID: <https://orcid.org/0000-0003-4594-5692>  
ResearcherID: [G-3926-2014](https://orcid.org/0000-0003-4594-5692). Scopus Author ID: [57484900000](https://orcid.org/0000-0003-4594-5692)

## ЗАСТОСУВАННЯ АУДІОАНАЛІЗУ ДЛЯ ПОЄДНАННЯ РОЗРІЗНЕНИХ ПОДІЙ У СИСТЕМАХ СПОСТЕРЕЖЕННЯ БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ

У статті розглянуто проблему поєднання розрізнених подій у гетерогенних системах мультимедійного спостереження БПЛА. Запропоновано використання аудіоаналізу як додаткового каналу для кореляції подій у випадках обмежень відеоспостереження. Досліджено методи спектрального аналізу та формування аудіовідбитків, зокрема хеш-орієнтований підхід і Mel-спектральний профіль. Експерименти показали, що хеш-підхід забезпечує кращу стійкість до шуму та чіткіше розділення схожих і різних сигналів. Результати підтверджують доцільність інтеграції аудіоаналізу в системи спостереження для підвищення ефективності виявлення загроз.

**Ключові слова:** аудіоаналіз; аудіовідбитки; БПЛА; спектральний аналіз; STFT; Mel-спектрограма; кореляція подій; мультимедійне спостереження.

Рис.: 4. Табл.: 1. Бібл.: 8.

**Актуальність теми дослідження.** Сучасні системи відеоспостереження та охорони об'єктів дедалі частіше експлуатуються в умовах динамічних і технологічно ускладнених загроз, зокрема через застосування безпілотних літальних апаратів (БПЛА) для несанкціонованого спостереження, проникнення або підготовки диверсійних дій. Для об'єктів критичної інфраструктури такі інциденти мають підвищений ризик, а вимоги до раннього виявлення та коректної інтерпретації подій зростають.

Практика впровадження відеоспостереження показує, що в реальних інсталяціях камери часто не забезпечують перекриття полів зору, а окремі події фіксуються фрагментарно та можуть помилково сприйматися як не пов'язані між собою. За таких умов використання лише відеоданих є недостатнім, особливо в ситуаціях зашумленого або низькоякісного відеопотоку (погодні умови, низька освітленість, дальні дистанції, невеликий розмір цілі). Тому перспективним є побудова гетерогенних мультимедійних систем спостереження, що поєднують різні типи сенсорів та канали отримання інформації.

Залучення аудіоданих як додаткового інформаційного каналу створює можливість отримання додаткових ознак для кореляції подій між різними сенсорами, зокрема у випадках, коли візуальна інформація неповна або неоднозначна. Звуки роботи двигунів БПЛА мають характерні спектральні компоненти та можуть бути використані для автоматизованого встановлення зв'язку між аудіофрагментами, записаними в різні моменти часу і з різних пристроїв. У цьому контексті актуальним є дослідження застосовності алгоритмів аудіовідбитків та спектральних характеристик для зіставлення аудіозаписів і поєднання розрізнених подій у єдиний інцидентний ланцюг у системах спостереження.

Отже, тема дослідження є актуальною з огляду на необхідність підвищення ефективності виявлення загроз з боку БПЛА та розвитку практичних методів інтеграції аудіоаналізу у гетерогенні системи мультимедійного спостереження для покращення кореляції подій і підтримки прийняття рішень у реальному часі.

**Аналіз останніх досліджень і публікацій.** У сучасних системах відеоспостереження проблема виявлення та відстеження безпілотних літальних апаратів (БПЛА) розглядається як одна з ключових задач безпеки. Це пов'язано з тим, що дрони є малорозмірними, малококонтрастними та швидко рухомими об'єктами, що значно ускладнює їх виявлення у відеопотоці.

У роботі [1] розглядається саме задача виявлення БПЛА у відео з камер спостереження. Автори показують, що через малий розмір та слабку візуальну виразність дронів класичні методи неефективні, тому пропонується використовувати просторово-часову інформацію та оптичний потік. Це дозволяє аналізувати рух об'єкта між кадрами, фактично переходячи від статичного детектування до відстеження траєкторії руху. Таким чином, у роботі прямо обґрунтовується необхідність використання часової інформації (траєкторії), оскільки лише за рахунок одного кадру виявлення є ненадійним.

У статті [2] описано систему відеоспостереження, яка виконує послідовно детекцію та відстеження БПЛА. Після виявлення дрона формується область інтересу (ROI), яка далі відстежується за допомогою оптичного потоку. Такий підхід дозволяє значно зменшити обчислювальні витрати та підвищити точність, оскільки обробляється лише релевантна частина зображення. У цій роботі відстеження траєкторії використовується як механізм стабілізації детекції та забезпечення безперервного спостереження, що прямо відповідає задачі побудови шляху об'єкта у відеопотоці.

Важливим напрямком є використання кількох сенсорів або камер, що фактично відповідає багатокамерним системам відеоспостереження. У роботі [3] досліджується підхід, при якому поєднуються дані з камер і радарів. Автори показують, що використання одного сенсора має суттєві обмеження (шум, пропуски, помилки), тоді як їх об'єднання дозволяє підвищити точність локалізації та відстеження. Для моделювання руху використовується фільтр Калмана, який безпосередньо працює з траєкторією об'єкта як часовим процесом. У роботі прямо зазначається, що поєднання сенсорів дозволяє компенсувати недоліки окремих джерел і забезпечити більш надійне відстеження дронів у реальних умовах.

Сучасні умови експлуатації об'єктів спостереження характеризуються появою нових типів загроз, пов'язаних із несанкціонованим проникненням, у тому числі із застосуванням безпілотних літальних апаратів, що ускладнює завдання своєчасного виявлення та інтерпретації подій зовнішнього втручання. Таким чином при побудові систем відеоспостереження побудова гетерогенних мультимедійних систем спостереження, що поєднують різні типи сенсорів та канали отримання інформації, покращує можливості виявлення, класифікації та збільшення точності розрахунку ризиків при втручанні. Такий підхід був запропонований авторами, зокрема, на конференції [4].

Аналіз публікацій пов'язаних з відеоспостереженням та особливо зі зростанням кількості загроз, пов'язаних із використанням безпілотних літальних апаратів для несанкціонованого спостереження або проникнення на об'єкти критичної інфраструктури показує важливість та актуальність виявлення та класифікації загроз з боку БПЛА. Практика використання систем відеоспостереження показує, що в багатьох випадках камери не мають взаємного перекриття полів зору, унаслідок чого зафіксовані події можуть сприйматися як незалежні та не пов'язані між собою. За таких умов використання лише відеоінформації є недостатнім. Залучення аудіоданих як додаткового джерела інформації дозволяє отримати нові ознаки для встановлення зв'язку між подіями, зафіксованими різними сенсорами, та сформуванню більш цілісного уявлення про перебіг зовнішнього втручання. Особливо актуальним є аналіз звукових сигналів, що супроводжують роботу двигунів БПЛА, оскільки такі сигнали мають характерні спектральні особливості та можуть бути використані для автоматизованої кореляції подій у мультимедійних системах спостереження.

У статті розглядаються методи порівняння аудіофрагментів, що дає можливість побудови спільного подієвого ланцюга в гетерогенних системах мультимедійного спостереження. Предметом дослідження є методи аналізу та порівняння звукових сигналів, зокрема алгоритми формування аудіовідбитків і спектральних характеристик, що застосовуються для встановлення зв'язку між аудіозаписами роботи двигунів безпілотних літальних апаратів, отриманими з різних джерел.

**Мета дослідження.** Метою дослідження є аналіз можливостей застосування алгоритмів аудіовідбитків для встановлення зв'язку між розрізненими подіями в гетерогенних системах мультимедійного спостереження на основі звукових сигналів роботи двигунів БПЛА.

Для досягнення поставленої мети у роботі розв'язано такі завдання: проаналізовано спектральні характеристики звукових сигналів двигунів БПЛА; реалізовано алгоритм формування аудіовідбитків на основі спектрального аналізу; виконано порівняння хеш-орієнтованого підходу та методу Mel-спектрального профілю; оцінено можливість використання аудіоаналізу для кореляції подій, зафіксованих різними сенсорами.

**Виклад основного матеріалу.** Методи дослідження включають методи спектрального аналізу звукових сигналів, алгоритми формування аудіовідбитків на основі виділення локальних максимумів амплітуд у частотній області, метод Mel-спектрального профілю з використанням косинусної міри подібності, а також експериментальні методи комп'ютерного моделювання та порівняльного аналізу результатів. Реалізацію алгоритмів та обробку експериментальних даних виконано із застосуванням мови програмування Python та відповідних програмних бібліотек.

Одним із підходів до поєднання таких подій є використання аудіовідбитків, сформованих на основі спектрального аналізу звукових сигналів. Принцип побудови аудіовідбитків ґрунтується на виділенні локальних максимумів амплітуд у частотній області та формуванні пар частотних компонент, що описано у праці Avery Li-Chun Wang [5]. Загальну схему та основні етапи цього алгоритму наведено на рис. 1.

Зазначений підхід дозволяє виконувати зіставлення аудіо-фрагментів без прямого зіставлення відліків аудіо-сигналів, а також забезпечує можливість аналізу записів різної тривалості та підвищує стійкість до шумових завад. У межах даної роботи аналіз зосереджений на звукових сигналах, що супроводжують роботу двигунів БПЛА, без розгляду задач локалізації джерела звуку або ідентифікації конкретної моделі апарата.

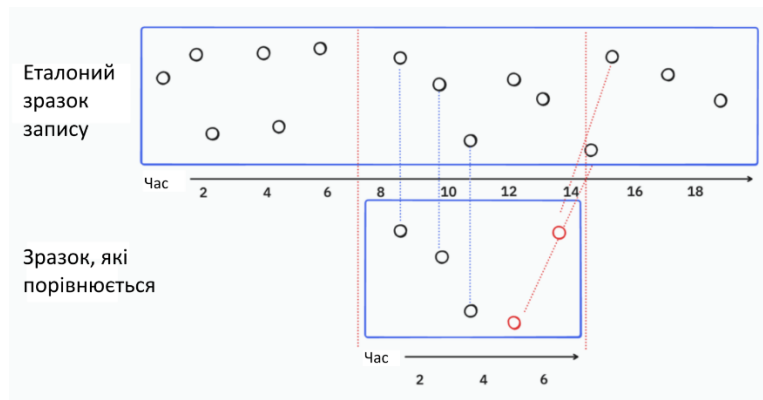


Рис. 1. Загальна схема та основні етапи алгоритму формування аудіовідбитків  
Джерело: [5].

Подальший аналіз включає формування аудіовідбитків для вибраних звукових фрагментів, їх зіставлення між собою за визначеними критеріями подібності та оцінювання можливості встановлення зв'язку між записами, отриманими з різних джерел. Такий підхід дозволяє перевірити придатність аудіоаналізу для поєднання подій у гетерогенних системах мультимедійного спостереження.

Для перевірки гіпотези можливості застосування алгоритму побудови відбитку була здійснена програмна реалізація алгоритму на мові Python. Згідно з алгоритмом знаходились максимальні значення амплітуд і формувала графіки за цими даними. Розподілення

точок, які позначають максимальне значення амплітуди можна побачити на рис. 2. У ролі досліджуваних даних власні файли спостереження БПЛА.

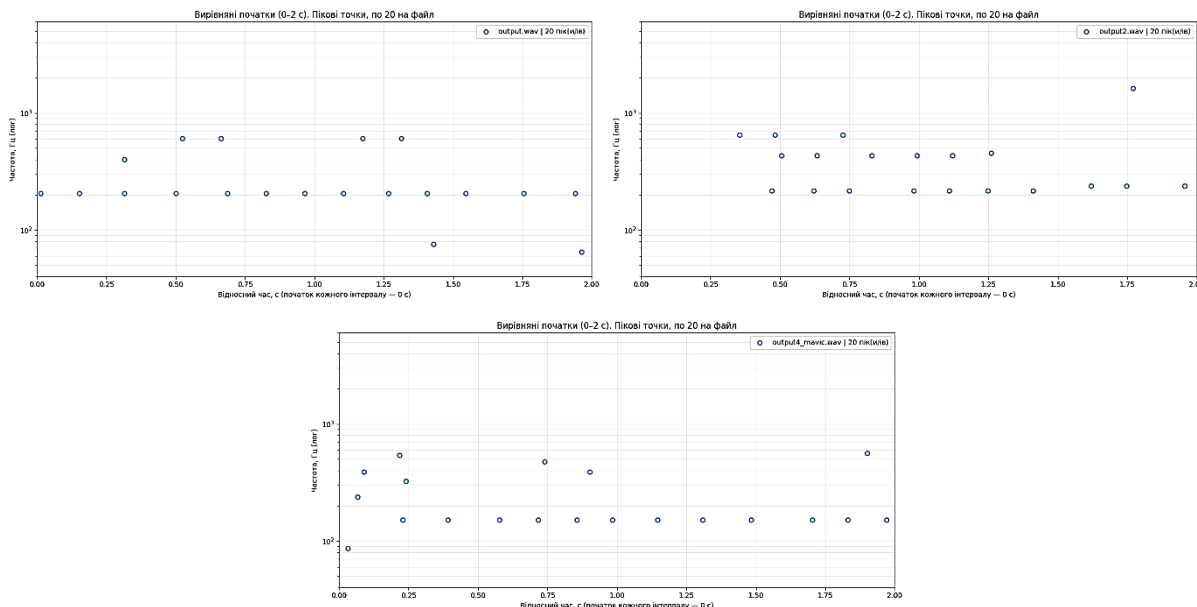


Рис. 2. Приклад амплітудно-частотних характеристик аудіофрагментів двигуна БПЛА

Джерело: розроблено авторами.

Порівнюючи отримані амплітудно-частотні графіки, видно, що на відміни від музичних композицій, складно знайти характерні максимуми амплітуди, які б давали чітку картинку спорідненості джерел звуку. Але також присутній на всіх трьох варіантах на фоні шумового засмічення, присутня стала рівномірна пряма, яка відповідає гулу мотору, оскільки мотор видає звукову хвилю постійної частоти, якщо він працює без змін параметрів руху.

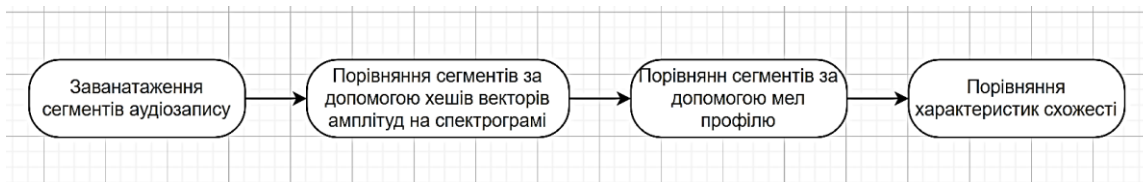


Рис. 3. Узагальнена послідовність етапів порівняння аудіофрагментів

Джерело: розроблено авторами.

Узагальнена послідовність програмної реалізації, яка використовувалася для розрахунків коефіцієнтів схожості, представлена на рис. 3. За базове значення розрахунку коефіцієнта схожості використаний метод MelSpectrogram [6; 7], який постачається в пакеті librosa. Порівняння на основі хешу алгоритм зображений на рис. 4, яке розглядається в [4; 5].

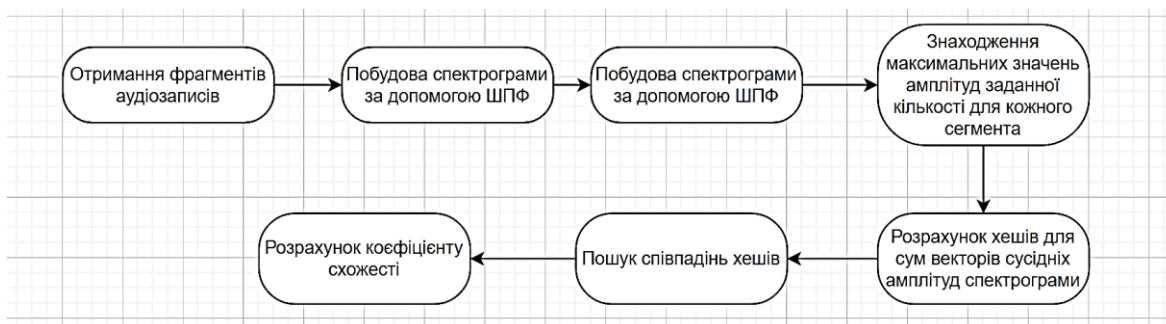


Рис. 4. Узагальнена послідовність етапів алгоритму порівняння аудіофрагментів на основі хешу

Джерело: розроблено авторами.

Алгоритм побудови аудіовідбитків реалізує підхід так званого landmark-based audio fingerprinting, запропонований [5]. Його основна ідея полягає у використанні локальних максимумів спектрограми для формування компактного представлення звукового сигналу. На першому етапі аудіосигнал перетворюється у частотно-часове представлення за допомогою короткочасного перетворення Фур'є.

У програмній реалізації для обчислення спектрограми використовується короткочасне перетворення Фур'є (STFT) з параметрами:  $n\_fft = 4096$  та  $hop\_length = 256$ . Як віконна функція застосовується вікно Ханна. Виділення спектральних піків здійснюється шляхом вибору фіксованої кількості найбільш енергетичних точок (TOP-K), що фактично виконує роль порогу та дозволяє відфільтрувати шумові компоненти сигналу.

На основі отриманої спектрограми обираються  $K$  точок з найбільшими значеннями амплітуди. Кожна така точка характеризується трійкою параметрів:

$$p_i = (t_i, f_i, a_i)$$

де  $t_i$  – часова координата піку;

$f_i$  – частота;

$a_i$  – амплітуда.

Далі формуються пари піків, які лежать у певному часовому інтервалі:

$$\Delta t = t_j - t_i, t_{min} < \Delta t < t_{max}.$$

Для розрахунку хешів частоти піків переводяться у логарифмічну шкалу та округлюються, щоб згладити невеликі зміни у сигналі. Аналогічно, різниця часу між піками також округлюється. Отримані значення використовуються для створення хешів, що дозволяє знаходити схожі звуки навіть при наявності шуму або невеликих спотворень. Для кожного хешу другого сигналу виконується пошук відповідних хешів у першому з урахуванням допустимої похибки. Для кожного знайденого збігу обчислюється часовий зсув між сигналами. Отримані значення накопичуються у вигляді гістограми. У випадку подібних сигналів гістограма має виражений максимум, що відповідає спільному часовому зсуву.

З отриманої гістограми часових зсувів визначається найбільш ймовірний зсув між сигналами як той, що має найбільшу кількість збігів. Відношення цієї кількості до числа сформованих хешів використовується як міра подібності сигналів. Чим більше значення цього показника, тим більша ймовірність того, що аудіофрагменти є подібними.

Для побудови Mel-представлення використовується  $n\_mels = 64$  фільтрів у діапазоні частот  $f_{min} = 50$  Гц,  $f_{max} = 8000$  Гц. Після обчислення Mel-спектрограми значення переводяться у логарифмічну шкалу, щоб зменшити різницю між слабкими і сильними компонентами. Далі по часу обчислюється медіана, що дозволяє отримати стабільний спектральний профіль сигналу. Отриманий вектор нормалізується, щоб порівняння не залежало від гучності запису.

Як тестові дані використаний власний набір файлів, записаних за допомогою мікрофону «Sony F-V120» на вулиці при прольоті БПЛА типу «Shahed-136» при візуальному спостереженні та «Xiaomi Mavic 3» на вулиці на відстані 10 м. Запис зроблений з частотою дискретизації 22 кГц та бітрейтом 353 кб/с у файли формату WAV з використанням кодування аудіопотоку PCM. Для перевірки гіпотези сформовані 24 пари з вирізаних аудіофрагментів однакової довжини з отриманих файлів довжиною 10-30 с (таблиця 1). Пари сформовані для отримання коефіцієнтів схожості: два фрагменти з одного файлу, з різних файлів запису звуку БПЛА типу «Shahed-136» та пара записаного звуку БПЛА типу «Shahed-136» та «Xiaomi Mavic 3».

Загальна послідовність етапів порівняння аудіо-фрагментів для хеш-орієнтованого підходу та методу Mel-спектрального профілю наведена на рисунку 3. Діаграма відображає спільні етапи попередньої обробки сигналів та подальше паралельне виконання двох алгоритмів із формуванням відповідних критеріїв подібності.

Таблиця 1 – Результати порівняння критеріїв подібності для хеш-орієнтованого та Mel-спектрального підходів

Кількість макс. значень амплітуди, які обираються зі спектрограми	Довжина запису	З допомогою хеш (критерій дає результат >1 – схожі, <0.2 – різні)			Mel-профіль (критерій лежить в проміжку 0...1)		
		аудіо блоки з одного файлу	аудіо блоки де чути БПЛА «Shahed-136»	звуки записані БПЛА «Shahed-136» та «Xaomi Mavic 3»	аудіо блоки з одного файлу	аудіо блоки де чути БПЛА «Shahed-136»	звуки записані БПЛА «Shahed-136» та «Xaomi Mavic 3»
10	1	0,333	0	0	0,887	0,345	0,233
20	1	1,904	0	0	0,887	0,345	0,233
30	1	3,116	0	0	0,887	0,345	0,233
10	2	2,9	0	0	0,907	0,287	0,415
20	2	4,759	0	0	0,907	0,287	0,415
30	2	3,436	0	0	0,907	0,287	0,415
30	3	2,163	0	0	0,923	0,283	0,565
30	4	2,163	0	0	0,916	0,205	0,604

Дані табл. 1 дають підстави вважати хеш-орієнтований підхід більш придатним для задачі зіставлення аудіофрагментів БПЛА, ніж Mel-спектральний профіль. Перевага хеш-підходу полягає в кращому розділенні схожих і різних сигналів: для аудіоблоків з одного файлу його score зростає від 0,333 до 1,904 і 3,116 при збільшенні кількості відібраних піків від 10 до 20 і 30 відповідно, тоді як для різних записів (Shahed-136 та Xaomi Mavic 3) значення залишаються на рівні 0. Натомість Mel-профіль хоча й дає високі значення для фрагментів з одного файлу (0,887-0,923), одночасно показує і відносно високі значення для різних записів (0,233-0,604), що зменшує контраст між класами «схожі» і «різні». Параметр «кількість максимальних значень амплітуди» означає число найбільш виразних спектральних піків, які відбираються зі спектрограми для побудови аудіовідбитка. Збільшення цього числа до 20-30, як правило, підвищує score для споріднених фрагментів, оскільки опис сигналу стає більш інформативним; однак надмірне збільшення може частково залучати нестійкі або шумові піки, через що ріст показника не завжди є монотонним.

У ході проведеного дослідження було проаналізовано можливість застосування алгоритмів аудіовідбитків для поєднання розрізаних подій у гетерогенних системах мультимедійного спостереження. Отримані результати підтверджують доцільність використання звукової інформації як додаткового джерела для встановлення зв'язку між подіями, зафіксованими різними сенсорами.

Порівняльний аналіз амплітудно-частотних характеристик показав суттєву відмінність між музичними записами та звуками роботи двигуна БПЛА. Музичні композиції характеризуються складною структурою спектра, наявністю великої кількості короткочасних піків та широким частотним діапазоном, що ускладнює виділення стабільних характерних ознак. Натомість гул двигуна має виражену квазістаціонарну природу та проявляється у вигляді стійкої смуги на спектрограмі, що відповідає майже постійній основній частоті обертання двигуна.

Саме ця особливість дозволяє ефективно використовувати алгоритми аудіовідбитків і хешування для зіставлення записів, отриманих у різних часових точках та з різних мікрофонів. Експериментальні результати підтвердили, що для шумів техногенного походження (зокрема гулу мотору) критерії подібності демонструють стабільні значення, тоді як для суб'єктивно різних джерел звуку показники різко зменшуються, що знижує ймовірність хибних збігів.

**Висновки.** Порівняння з підходом на основі Mel-спектрограм показало, що хеш-орієнтовані методи є більш придатними для задачі ідентифікації подій, пов'язаних із роботою двигунів, оскільки вони менш чутливі до фонового шуму та змін амплітуди сигналу. Це робить запропонований підхід перспективним для практичного застосування в системах раннього виявлення зовнішнього втручання.

Таким чином, отримані результати підтверджують, що різниця в спектральній структурі музичних сигналів і гулу двигуна може бути використана як надійна ознака для автоматичного поєднання аудіоподій у системах спостереження. Подальші дослідження доцільно спрямувати на адаптацію алгоритмів до змін режимів роботи двигуна та інтеграцію аудіо-аналізу з відеоданими в реальному часі.

### Список використаних джерел

1. Sun, Y., & Zhi, X. (2023). Enhancing UAV detection in surveillance camera videos through spatiotemporal information and optical flow. *Sensors*, 23, 6037.
2. Lee, D. R., & Kim, S. (2024). A multi-stage approach to UAV detection, identification, and tracking using region-of-interest management and rate-adaptive video coding. *Applied Sciences*, 14, 5559.
3. Kumar, C. A., & Giridhar, O. (2024). UAV detection multi-sensor data fusion. *Journal of Research in Science and Engineering*, 6(7).
4. Веремєнко, В. В., & Роговенко, А. І. (2025, 2 жовтня). Застосування зліпку аудіозаписів для пошуку збігів джерел звуку, знятих з різних пристроїв запису в системі відеоспостереження. У «Вільне програмне забезпечення у освіті, науці та бізнесі». Чернігів: НУ «Чернігівська політехніка».
5. Wang, A. L.-C. (2003). An industrial-strength audio search algorithm. У *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2003)*. Baltimore, MD, USA.
6. librosa. (n.d.). *librosa.feature.melspectrogram*. <https://librosa.org/doc/main/generated/librosa.feature.melspectrogram.html>.
7. Chou, S.-Y., & Cheng, K.-H. (2019, February 18). Learning to match transient sound events using attentional similarity for few-shot sound recognition. National Taiwan University, Taipei, Taiwan.
8. Strauss, M. (n.d.). *How Shazam works: An explanation in Python*. <https://michaelstrauss.dev/shazam-in-python>.

### References

1. Yu Sun, Xiyang Zhi “Enhancing UAV Detection in Surveillance Camera Videos through Spatiotemporal Information and Optical Flow”, *Sensors* 2023, 23, 6037.
2. Dongkyu ‘Roy’ Lee, Sanghong Kim “A Multi-Stage Approach to UAV Detection, Identification, and Tracking Using Region-of-Interest Management and Rate-Adaptive Video Coding”, *Applied sciences* 2024, 14, 5559.
3. Chiranjeevi Amit Kumar, Ozkan Giridhar “UAV Detection Multi-sensor Data Fusion”, *Journal of Research in Science and Engineering* Volume-6, Issue-7, July 2024.
4. Veremeienko V.V., Rogovenko A.I. “Application of Audio Recording Fingerprints for Matching Sound Sources Captured by Different Recording Devices in a Video Surveillance System”, “Free Software in Education, Science, and Business”, Chernihiv Polytechnic National University – Chernihiv, October 2, 2025.
5. Avery Li-Chun Wang «An Industrial-Strength Audio Search Algorithm» ISMIR 2003, 4th International Conference on Music Information Retrieval, Baltimore, Maryland, USA, October 27-30, 2003, Proceedings.
6. librosa.feature.melspectrogram – Librosa documentation <https://librosa.org/doc/main/generated/librosa.feature.melspectrogram.html>.

7. Szu-Yu Chou, Kai-Hsiang Cheng «Learning to match transient sound events using attentional similarity for few-shot sound recognition» Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei, Taiwan 18.02.2019.

8. Michael Strauss. “How Shazam Works - An explanation in Python” <https://michaelstrauss.dev/shazam-in-python>.

Дата першого надходження статті до видання: 12.02.2026  
Дата прийняття статті до друку після рецензування: 05.03.2026

UDC 004.622

**Vasyl Veremeienko<sup>1</sup>, Andrii Rohovenko<sup>2</sup>**

<sup>1</sup>PhD student, Department of Information and Computer Systems  
Chernihiv Polytechnic National University (Chernihiv, Ukraine)

E-mail: [vasiliy.veremeenko@gmail.com](mailto:vasiliy.veremeenko@gmail.com). ORCID: <https://orcid.org/0009-0004-2742-6367>

<sup>2</sup>PhD in Technical Sciences, Associate Professor of the Department of Information and Computer Systems  
Chernihiv Polytechnic National University (Chernihiv, Ukraine)

E-mail: [arogovenko@stu.cn.ua](mailto:arogovenko@stu.cn.ua). ORCID: <https://orcid.org/0000-0003-4594-5692>

ResearcherID: [G-3926-2014](https://orcid.org/0000-0003-4594-5692). Scopus Author ID: [57484900000](https://orcid.org/0000-0003-4594-5692)

## APPLICATION OF AUDIO ANALYSIS FOR LINKING DISPARATE EVENTS IN UAV SURVEILLANCE SYSTEMS

*Modern surveillance systems operate under conditions of increasing threats associated with the use of unmanned aerial vehicles (UAVs), which necessitates improving the efficiency of event detection and interpretation. Of particular relevance is the problem of linking disparate events in multi-camera systems where there is no overlap of fields of view, and where video data may be noisy or of insufficient quality. Under such conditions, relying solely on visual information is insufficient, which justifies the use of additional information channels, particularly audio signals.*

*The research problem lies in developing methods for establishing connections between audio fragments obtained from different sources and at different points in time, with the aim of forming a unified incident chain in heterogeneous multimedia surveillance systems. Particular attention is paid to the analysis of acoustic signals produced by UAV engines, which exhibit characteristic spectral properties and can be used as an identifying feature.*

*The objective of this study is to analyze the possibilities of applying audio fingerprinting algorithms to link disparate events in multimedia surveillance systems based on UAV engine audio signals. To achieve this goal, spectral analysis methods were used, particularly the short-time Fourier transform (STFT), audio fingerprinting algorithms based on local maxima of the spectrogram, as well as the Mel-spectral profile method using cosine similarity. The implementation of the algorithms was carried out in the Python programming language using specialized audio signal processing libraries.*

*As a result of the study, it was established that UAV engine sound signals have a quasi-stationary spectral structure, which ensures the stability of similarity criteria when comparing audio fragments. Experimental analysis showed that the hash-based approach to audio fingerprinting demonstrates a better ability to distinguish between similar and different signals compared to the Mel-spectral method. In particular, increasing the number of spectral peaks (local amplitude maxima) enhances the informativeness of the signal representation and increases the similarity score for related fragments, while for different sound sources the value remains low. In contrast, the Mel-based approach is characterized by lower contrast between classes, which increases the likelihood of false matches.*

*The obtained results confirm the feasibility of using audio analysis as an additional information channel for event correlation in heterogeneous surveillance systems. The proposed approach can be used to improve threat detection efficiency and support real-time decision-making. Prospects for further research include extending the analysis methods to different UAV operating modes and integrating audio and video data into a unified monitoring system.*

**Keywords:** audio analysis; audio fingerprinting; UAV; spectral analysis; STFT; Mel spectrogram; event correlation; multimedia surveillance.

Fig.: 4. Tables: 1. References: 8.