

МЕТОД ІНТЕЛЕКТУАЛЬНОЇ ІДЕНТИФІКАЦІЇ РУХОМИХ ОБ'ЄКТІВ НА ОСНОВІ КЛАСТЕРИЗАЦІЇ ОЗНАК У МЕЛ-КЕПСТРАЛЬНОМУ ПРОСТОРИ

Анотація. Впровадження інтелектуальних систем моніторингу в промислово автоматизацію та новітні транспортні системи потребує розробки методів ідентифікації рухомих об'єктів, здатних стабільно функціонувати в умовах високого рівня завад та обмеженої видимості. Традиційні амплітудні методи обробки є чутливими до нестаціонарності акустичних сигналів та зміни відстані до джерела, що ускладнює точну сегментацію динамічних станів об'єкта. Метою роботи є розробка методу ідентифікації фаз руху на основі мел-кепстрального аналізу, інтелектуальної кластеризації методом К-середніх, віконному перетворенні Фур'є, психоакустичній фільтрації та декореляції ознак у кепстральному просторі. В ході дослідження отримано матрицю ознак, яка дозволяє виділити цифровий відбиток об'єкта, інваріантний до коливань гучності. Експериментально підтверджено ефективність сегментації фаз наближення, проїзду та віддалення з високою швидкістю обчислень, що дозволяє інтегрувати метод у системи автоматизованого керування та технічної діагностики в режимі реального часу.

Ключові слова: ідентифікація рухомих об'єктів, мел-кепстральні коефіцієнти, кластеризація, автоматизація моніторингу, акустичний сигнал, цифрова обробка сигналів, промислова автоматизація, сегментація фаз руху.

Постановка проблеми. Сучасний стан розвитку інтелектуальних транспортних систем та систем автоматизованого моніторингу висуває високі вимоги до точності та швидкодії алгоритмів розпізнавання рухомих об'єктів. Традиційні методи ідентифікації, що базуються на візуальному аналізі даних, а саме застосуванні відеокамер, часто виявляються малоефективними в умовах обмеженої видимості, несприятливих погодних умов або складного освітлення. У зв'язку з цим, дедалі більшої ваги набувають акустичні методи моніторингу, які є пасивними, економічно доцільними та здатними функціонувати в безперервному режимі.

Аналіз акустичних сигналів рухомих об'єктів супроводжується низкою проблем, зумовлених їх нестаціонарністю, високим рівнем фонових завад та амплітудними спотвореннями внаслідок ефекту Доплера та зміни відстані до мікрофона. Існуючі підходи часто фокусуються лише на класифікації типу об'єкта, ігноруючи необхідність точної сегментації фаз його руху (наближення, проїзд, віддалення), що ускладнює автоматизацію процесів прийняття рішень у реальному часі, тому виникає необхідність у розробці

стійкого методу ідентифікації, який би поєднував ефективні способи виділення інформативних ознак, що інваріантні до енергетичних коливань сигналу, з алгоритмами інтелектуальної кластеризації. Використання кепстрального аналізу в комбінації з ітераційним перерахунком центроїдів станів об'єкта дозволить автоматизувати процес розпізнавання динаміки руху, забезпечуючи високу вірогідність ідентифікації в складних акустичних середовищах.

Особливої актуальності дана задача набуває в умовах промислових підприємств та автоматизованих складських комплексів, де рухомі об'єкти, такі як внутрішньозаводський транспорт, робототехнічні пристрої, конвеєрні лінії, функціонують у середовищі з високим рівнем техногенних завад, а стандартні амплітудні методи ідентифікації стають недієздатними через постійні відлуння та накладання шумів від працюючих агрегатів. Тому існує потреба у впровадженні методів інтелектуальної ідентифікації, здатних виділяти корисний сигнал на основі його спектральної структури, а не лише інтенсивності, що є необхідною умовою для забезпечення безпеки та безперервності автоматизованих виробничих процесів.

Аналіз останніх досліджень і публікацій. Проблема автоматизованої ідентифікації рухомих об'єктів за їх акустичними сигналами займає важливе місце у сучасних дослідженнях, що присвячені розвитку інтелектуальних транспортних систем та систем технічної діагностики. Питання аналізу нестационарних шумів пов'язані з переходом від класичних спектральних методів до часово-частотних перетворень [1]. Доведено, що використання стандартного швидкого перетворення Фур'є не завжди забезпечує достатню інваріантність до шумів навколишнього середовища [2] і стає актуальним використання мел-кепстрального аналізу [3], який був розроблений для розпізнавання мовлення, але є надзвичайно ефективним у технічній акустиці [4]. Наприклад, при розпізнаванні типів двигунів мел-кепстральний аналіз [5] дозволяє ефективно декорелювати спектральні компоненти та виділяти огинаючу спектра, яка є унікальною для різних режимів роботи об'єкта.

У сфері інтелектуального аналізу даних використовуються алгоритми навчання без учителя для сегментації фаз руху, при цьому одним з найбільш продуктивних підходів використовується метод К-середніх через його низьку обчислювальну складність. Кластеризація у кепстральному просторі забезпечує вищу точність розділення стану руху об'єкта порівняно з аналізом у часовій області, оскільки кепстральні коефіцієнти менш чутливі до амплітудних флуктуацій, викликаних зміною відстані [6].

Проте залишається недостатньо вивченим питання стабільності ідентифікації об'єктів в умовах високої інтенсивності фонових завад та специфічних імпульсних перешкод, що виникають у реальних промислових та міських локаціях. Більшість існуючих методів фокусуються на класифікації типів об'єктів, наприклад, легкове чи вантажне авто, залишаючи поза увагою задачу точної часової сегментації фаз руху в автоматичному режимі без попереднього навчання системи. Саме інтеграція кепстрального аналізу з ітераційною кластеризацією для вирішення задачі розпізнавання динамічних станів рухомих об'єктів потребує подальшого дослідження та вдосконалення.

Метою дослідження є розробка та дослідження методу автоматизованої ідентифікації фаз руху рухомих об'єктів на основі інтелектуального аналізу мел-кепстральних ознак та ітераційної кластеризації, що забезпечує стабільну сегментацію динамічних станів об'єкта в умовах нестационарних акустичних завад.

Викладення основного матеріалу дослідження. Процес отримання ознак для ідентифікації транспортних засобів базується на комбінації класичного спектрального аналізу та психоакустичного моделювання. Математична модель алгоритму складається з наступних кроків:

1. Попередня обробка та сегментація [2]. Оскільки акустичний сигнал автомобіля є нестационарним, він розділяється на короткі кадри (фрагменти) тривалістю N відліків. Для мінімізації ефекту розтікання спектра для вхідного дискретного сигналу виконується віконна обробка за методом Геммінга, результатом якої є зважена послідовність $s_w(n)$:

$$s_w(n) = s(n) \left[0.54 - 0.46 \cos \frac{2\pi \cdot n}{N-1} \right], \quad (1)$$

де $s(n)$ – дискретні значення амплітуди звукового сигналу в межах одного аналізованого кадру, n – дискретний час.

2. Швидке перетворення Фур'є [1]. Для переходу з часової області у частотну для кожного вікна обчислюється дискретне перетворення Фур'є:

$$X(k) = \sum_{n=0}^{N-1} s_w(n) \cdot e^{-i \frac{2\pi}{N} nk}, \quad (2)$$

де k – індекс частотної компоненти, ціле число, яке вказує на конкретну дискретну частоту після проведення швидкого перетворення Фур'є.

На основі отриманого спектра розраховується спектральна щільність потужності $P(k)$, яка відображає розподіл енергії сигналу за частотами:

$$P(k) = \frac{1}{N} |X(k)|^2. \quad (3)$$

3. Мел-фільтрація [6]. Перетворення лінійної частотної шкали f (Гц) у логарифмічну шкалу мелів m , яка імітує нелінійне сприйняття звуку людським вухом. Зв'язок між частотами визначається як:

$$m = 2595 \cdot \lg \left(1 + \frac{f}{700} \right). \quad (4)$$

Енергія в кожній мел-смузі $E(m)$ обчислюється шляхом згортки $P(k)$ з набором трикутних фільтрів $H_m(k)$:

$$E(m) = \sum_{k=0}^{N/2} P(k) \cdot H_m(k), \quad m = 1, \dots, M, \quad (5)$$

де M – кількість фільтрів.

4. Логарифмування енергій. Для імітації логарифмічної чутливості слуху до інтенсивності звуку та для нормалізації динамічного діапазону обчислюється логарифм

енергії кожного фільтра $S(m)$, що надає алгоритму стійкості до значних змін амплітуди при зміні відстані від рухомого об'єкта до мікрофона, $S(m) = \ln(E(m))$.

5. Дискретне косинусне перетворення. Обчислення кепстральних коефіцієнтів шляхом застосування дискретного косинусного перетворення до логарифмічних енергій, що дає можливість декорелювати ознаки та виділити огинаючу спектра (спектральний контур):

$$C(n) = \sum_{m=1}^M S(m) \cdot \cos\left(\frac{\pi \cdot n}{M}(m - 0.5)\right), \quad n = 0, \dots, L - 1, \quad (6)$$

де L – кількість вихідних коефіцієнтів.

Коефіцієнт $C(0)$ несе інформацію про середню енергію кадру, а $C(1) \dots C(L-1)$ описують форму спектральних піків, що є унікальними для різних типів двигунів та режимів їх роботи.

Перевага використання мел-кепстрального аналізу у задачах ідентифікації транспортних засобів полягає в тому, що кепстральні коефіцієнти дозволяють розділити інформацію про джерело звуку (вібрації двигуна) та передавальну характеристику середовища. Це робить систему розпізнавання інваріантною до акустичних особливостей місцевості та погодних умов, фокусуючись на фундаментальних характеристиках об'єкта.

Після етапу формування матриці ознак, наступним кроком є автоматизований розподіл отриманих даних на групи, що відповідають фізичним станам руху транспортного засобу. Для цього найчастіше використовується метод K -середніх. Метою кластеризації є розбиття множини векторів ознак $X = \{x_1, x_2, \dots, x_M\}$ на K заздалегідь визначених кластерів $S = \{S_1, S_2, \dots, S_K\}$. У даному дослідженні $K = 4$, що відповідає фазам: фоновий шум, наближення, проїзд та віддалення.

Суть метода K -середніх полягає в мінімізації сумарного квадратичного відхилення точок кластерів від їх центрів (центроїдів), що математично виражається як

$$J = \sum_{i=0}^{K-1} \sum_{x \in S_i} \|x - \mu_i\|^2, \quad (7)$$

де x – вектор кепстральних коефіцієнтів, μ_i – центроїд i -го кластера (еталонний спектральний портрет фази руху), $\|x - \mu_i\|$ – евклідова відстань між вектором ознак та центром кластера.

Процес розрахунку виконується ітераційно і полягає в наступній послідовності дій. Випадковим чином або за певною стратегією обираються початкові координати K центроїдів μ_i . Кожен вектор x_j відносять до того кластера, відстань до центроїда якого є мінімальною:

$$S_i^{(l)} = \left\{ x_j : \|x_j - \mu_i^{(l)}\|^2 \leq \|x_j - \mu_i^{(t)}\|^2 \right\} \text{ для всіх } l = 0, \dots, K - 1. \quad (8)$$

Обчислюються нові координати центроїдів як середнє арифметичне всіх векторів, що потрапили до відповідного кластера

$$\mu_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x \in S_i^{(t)}} x. \quad (9)$$

Цикл повторюється до повної стабілізації центроїдів (збіжності алгоритму). Результатом кластеризації є набір з K центроїдів. Оскільки кожен центроїд μ_i є вектором у кепстральному просторі, для візуалізації та фізичного аналізу проводиться зворотне перетворення або аналіз їхньої енергії у частотній області.

Доцільність застосування методу K – середніх для ідентифікації фаз руху об'єкта ґрунтується на сукупності технічних та аналітичних переваг, що відповідають специфіці акустичного моніторингу. Алгоритм характеризується високою швидкістю обчислень, що дозволяє ефективно обробляти великі масиви багатовимірних даних у режимі реального часу, забезпечуючи миттєву реакцію системи на зміну дорожньої обстановки. Метод забезпечує чітку фізичну інтерпретацію результатів, оскільки кожен сформований кластер безпосередньо корелює з рівнем акустичної енергії та спектральним складом звуку на різних відстанях від точки реєстрації. Застосування евклідової відстані в декорельованому кепстральному просторі гарантує високу стійкість класифікації, дозволяючи алгоритму стабільно розділяти фази руху об'єкта навіть за умов варіативності його швидкості або зміни інтенсивності фонового шуму.

Для перевірки ефективності розробленого методу було проведено аналіз натурального акустичного сигналу, що виникає під час руху об'єкта повз точку реєстрації. На рис. 1 представлена осцилограма отриманого сигналу, яка демонструє характерну амплітудну модуляцію. Форма сигналу зумовлена нестационарністю процесу та зміною відстані між джерелом звуку та мікрофоном, що є класичним прикладом для відпрацювання алгоритмів сегментації та ідентифікації станів руху

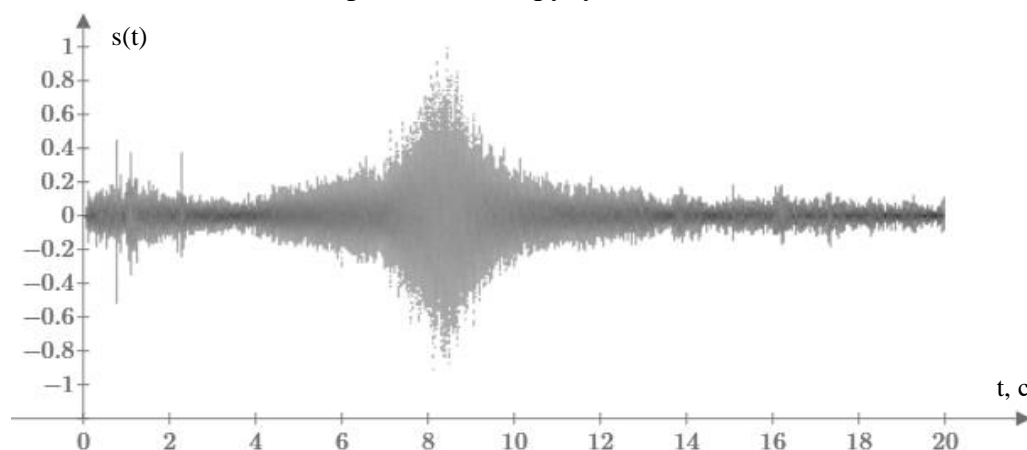


Рисунок 1 – Осцилограма акустичного сигналу транспортного засобу

Отриманий акустичний сигнал має характерну для рухомих об'єктів форму, що відображає динаміку зміни звукового тиску в часі протягом 20-секундного інтервалу запису. Осцилограма сигналу, нормована в діапазоні $[-1; 1]$, що дозволяє провести фазовий аналіз процесу ідентифікації. Початковий етап (0–6 сек) характеризується низькою та стабільною амплітудою (до 0.2), де переважає фоновий шум та віддалені низь-

кочастотні складові роботи двигуна. На інтервалі 6–8.5 сек спостерігається експоненціальне зростання енергії, що відповідає фазі активного наближення об'єкта. Точка максимального наближення фіксується приблизно на 8.5 сек, де пік амплітуди досягає значень 0.8...1, що корелює з нульовим рівнем децибел при спектральному аналізі. Подальше віддалення об'єкта (8.5–12 сек) супроводжується асиметричним спаданням амплітуди, зумовленим спрямованістю звукового випромінювання вихлопної системи та проявом ефекту Доплера, після чого сигнал поступово згасає до рівня фонового шуму (12–20 сек).

Для перетворення вхідного сигналу $s(n)$ у матрицю ознак були послідовно виконані операції, що поєднують часовий та спектральний аналіз. На початковому етапі неперервний акустичний сигнал був розділений на кадри довжиною $N = 1024$ відліків. Оскільки кожен фрагмент обробляється окремо, виникає ефект розриву сигналу на межах кадрів, для усунення якого застосовується вікно Геммінга (1), що надає можливість знизити амплітуду сигналу до мінімуму на краях вікна та отримати спектр без паразитної високочастотної складової. Для кожного згладженого кадру обчислено швидке перетворення Фур'є (2), яке переводить дані з часової області у частотну. Отримані комплексні амплітуди перетворені на спектр потужності $P(k)$ згідно (3), що дозволяє ідентифікувати розподіл енергії рухомого об'єкта за гармоніками. Спектр проходить через набір з 26 трикутних фільтрів, що рівномірно розподілені за шкалою мелів (4) для інтегрування енергії у специфічних частотних смугах (5). Використання мел-шкали є важливим, оскільки вона дозволяє виділити низькочастотні компоненти, до яких система ідентифікації має бути найбільш чутливою. Отримані значення енергій логарифмуються, що наближає цифрову модель до логарифмічного характеру слуху людини та зменшує вплив значних коливань амплітуди, які виникають при зміні відстані до об'єкта, при цьому ознаки стають більш стабільними. До логарифмічних енергій застосовано дискретне косинусне перетворення (6), що дозволило трансформувати спектральні дані у кепстральні коефіцієнти, тобто стиснути основну інформацію про тембр звуку у перші 13 коефіцієнтів, відсікаючи при цьому дрібні спектральні шуми.

В результаті виконання процедур обробки акустичного сигналу отримана матриця S , де кожен стовпець представляє собою вектор з 13 елементів, що є унікальним «цифровим відбитком» акустичного стану об'єкта в конкретний момент часу. Тривимірною поверхнею матриці ознак S , яка містить набір кепстральних коефіцієнтів для кожного часового кадру досліджуваного аудіосигналу наведена на рис.2. Така структура даних є оптимальною для подальшої кластеризації, оскільки вона поєднує в собі високу інформативність та мінімальний обчислювальний об'єм, що дозволяє реалізувати систему ідентифікації в режимі реального часу.

На тривимірній моделі спектрально-часового рельєфу акустичного сигналу вісь ординат відображає номери мел-кепстральних коефіцієнтів, які описують тембр звуку та форму огинаючої спектра, ігноруючи дрібномасштабні шумові складові. Вісь абсцис представляє часову розгортку процесу, де кожен стовпець матриці є окремим зрізом

звукового поля в конкретний момент часу, кількість яких визначається загальною тривалістю запису та кроком сегментації. Вісь аплікат демонструє амплітуду кепстральних енергій. Характерні піки та підняття на поверхні C чітко локалізують фазу максимального наближення, де концентрація енергії у відповідних кепстральних смугах досягає максимуму. Така візуалізація підтверджує, що алгоритм успішно порівнює вектори ознак (стовпці матриці) між собою. Подібність структури стовпців у межах певних часових інтервалів створює умови для їх подальшого об'єднання алгоритмом K – середніх у гомогенні кластери («Шум», «Наближення», «Проїзд», «Віддалення»), забезпечуючи високу достовірність автоматизованої сегментації фаз руху.

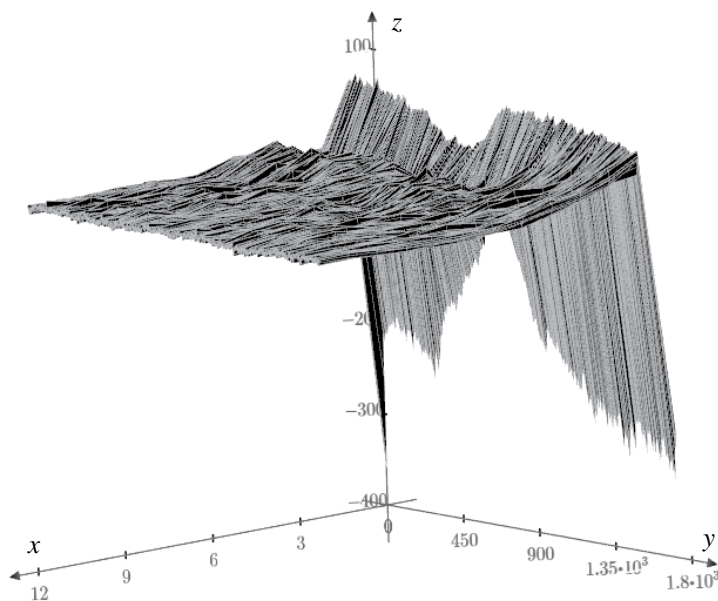


Рисунок 2 – Тривимірна поверхня матриці ознак C

Візуалізація матриці ознак C у вигляді тривимірної поверхні демонструє чітку часову динаміку акустичного портрета рухомого об'єкта. Найбільша амплітудна варіативність спостерігається у низькочастотних кепстральних коефіцієнтах (1–4), що підтверджує їх пріоритетну роль у формуванні унікального цифрового образу об'єкта. Наявність характерних рельєфних зон дозволяє автоматизовано розділяти сигнал на стадії шум, наближення, проїзд та віддалення з високим ступенем достовірності за допомогою алгоритму K – середніх (7–9). На основі обчисленої матриці мелкепстральних коефіцієнтів алгоритм розділив аудіосигнал на 4 кластери, кожен з яких має чітку фізичну інтерпретацію. Розподіл акустичних кадрів за кластерами наведено на рис.3.

Аналіз результатів кластеризації дозволяє виділити чотири ключові сегменти, що відповідають просторовому положенню рухомого об'єкта відносно точки спостереження:

Кластер «Фоновий шум» (рис. 3,а) охоплює початкову ділянку сигналу, що в кепстральному просторі характеризується низькою дисперсією енергій та відсутністю ви-

ражених доміантних частот та відповідає навколишньому звуковому фону та знаходженні об'єкта на великій відстані.

Кластери «Наближення» (рис. 3,б) та «Віддалення» (рис. 3,г) включають вектори ознак, що описують динамічну зміну спектра. Спостерігається характерний дрейф центроїдів кластера, викликаний наростанням потужності та зміною спектрального складу внаслідок ефекту Доплера.

Кластер «Проїзд» (рис. 3,в) характеризується групою векторів з найвищими значеннями кепстральних коефіцієнтів та відповідає моменту проїзду об'єкта поблизу мікрофона, де акустичний портрет є найбільш насиченим та стабільним.

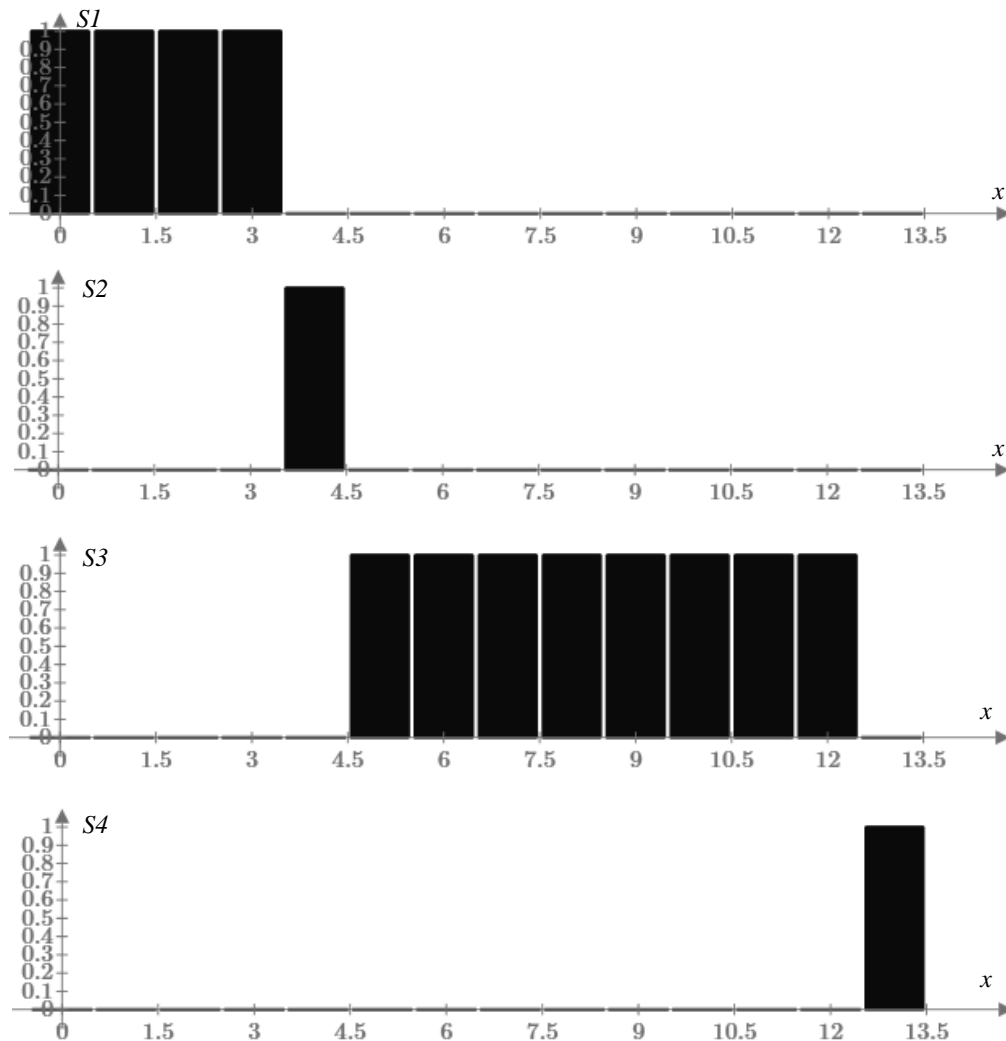


Рисунок 3 – Розподіл акустичних кадрів за кластерами: а) кластер «Фоновий шум», б) кластер «Наближення», в) кластер «Проїзд», г) кластер «Віддалення»

Алгоритм продемонстрував високу чутливість до моментів зміни станів. На відміну від класичних порогових методів, що базуються лише на амплітуді, метод K – середніх у мел-кепстральному просторі дозволив чітко зафіксувати часові межі фаз навіть за умов асиметричного згасання звуку. Точки переходу між кластерами на часовій осі збігаються з результатами візуального аналізу осцилограми, що підтверджує достовірність ідентифікації. Завдяки використанню евклідової відстані в декорельова-

ному просторі ознак, алгоритм ігнорує вплив короткочасних імпульсних завад, які були виявлені на початкових етапах аналізу. Такі перешкоди або ігноруються як статистичні викиди, або поглинаються найбільш наближеним кластером, не створюючи помилкових сегментів руху. Отримані результати підтверджують, що поєднання мел-кепстрального аналізу з ітераційною кластеризацією забезпечує стійку автоматизовану ідентифікацію станів рухомого об'єкта.

Важливим результатом проведеного дослідження є отримані спектральні портрети центроїдів, що представляють еталонні образи чотирифазної моделі руху об'єкта (рис. 4). Вісь ординат представлена у відносних децибелах, а вісь абсцис відображає центральні частоти мел-фільтрів. Візуалізація центроїдів кластерів у кепстральному просторі підтверджує високу сепарабельність станів та ефективність обраної комбінації методів мел-кепстрального аналізу та K – середніх для задач ідентифікації. Графічний аналіз свідчить про те, що сформовані кластери є лінійно роздільними, що є критично важливим для стабільної роботи автоматизованих систем класифікації.

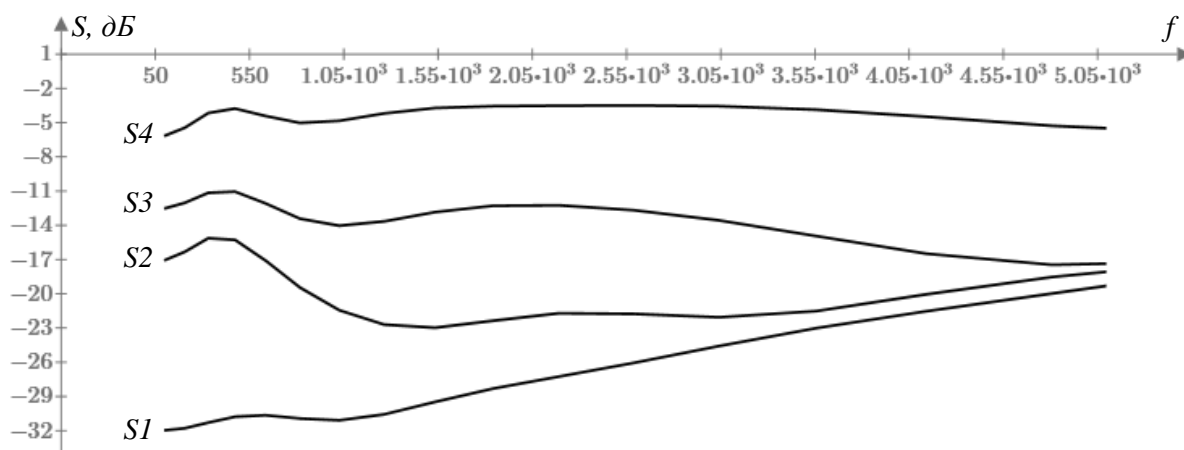


Рисунок 4 – Спектральні портрети центроїдів

Між граничними станами $S1$ (шум) та $S4$ (максимальне наближення) спостерігається амплітудна різниця понад 25 дБ. Це гарантує високу надійність розрізнення фаз руху навіть в умовах зашумленого промислового середовища. Перетин характеристик кластерів $S2$ (наближення) та $S3$ (віддалення) у високочастотній області (понад 5 кГц) вказує на схожість акустичних властивостей об'єкта на значних відстанях. Проте ідентифікація цих станів залишається достовірною завдяки суттєвим відмінностям у низькочастотному діапазоні, де зосереджена основна енергія механічних вузлів.

Кожен центроїд відображає унікальний спектральний відбиток конкретної фази руху. Кластер $S4$ відповідає точці максимального наближення, характеризується найвищим енергетичним рівнем ($-4 \dots -6$) дБ та рівномірним широкосмуговим спектром. Це свідчить про повну реєстрацію акустичного випромінювання: від низькочастотних вібрацій двигуна до високочастотного шуму тертя шин об поверхню. Кластер $S1$ відображає пасивний стан системи або рівень фонового шуму ($-20 \dots -32$) дБ відносно піку. Спектральний нахил у бік високих частот підтверджує домінування білого шуму

навколишнього середовища за відсутності корисного сигналу. Кластери S_2 та S_3 ідентифікують фази віддаленого руху.

Висновки. Запропонований підхід базується на припущенні, що кожен рухомий об'єкт у процесі функціонування генерує специфічне акустичне поле. Динаміка зміни цього поля дозволяє не лише ідентифікувати тип об'єкта, а й класифікувати стадії його просторового переміщення відносно точки спостереження за допомогою аналізу нестационарних спектральних характеристик. Високе співвідношення сигналу та шуму забезпечує стабільну роботу алгоритму виділення мел-кепстральних ознак. Явна нестационарність сигналу повністю обґрунтовує доцільність використання віконного швидкого перетворення Фур'є для коректного відображення спектральної динаміки. При цьому наявність поодиноких імпульсних завад на початкових секундах запису вимагає особливої уваги під час кластеризації методом K – середніх, оскільки такі сплески можуть формувати окремі мікрокластери перешкод або потребувати поглинання кластером фоновому шуму для підвищення достовірності ідентифікації. Отримані спектральні портрети центроїдів доводять, що запропонований метод дозволяє не просто фіксувати наявність рухомого об'єкта, а й детально сегментувати процес його переміщення. Це відкриває можливості для створення адаптивних систем автоматизації, здатних прогнозувати час проїзду та ідентифікувати режими роботи об'єкта за його цифровим відбитком.

ЛІТЕРАТУРА

1. O. Polyvoda, H. Rudakova, I. Kondratieva et al. Digital Acoustic Signal Processing Methods for Diagnosing Electromechanical Systems. *Lecture Notes in Computational Intelligence and Decision Making. ISDMCI 2019. Advances in Intelligent Systems and Computing*. Cham: Springer, 2020. Vol. 1020. P. 97–109.
DOI: https://doi.org/10.1007/978-3-030-26474-1_7.
2. Бондаренко М. Е., Іващенко Г. С. Використання послідовності методів попередньої обробки в системах голосової ідентифікації. *Системи управління, навігації та зв'язку. Збірник наукових праць*. 2025. № 2. С. 90–96.
DOI: <https://doi.org/10.26906/SUNZ.2025.2.090>.
3. Rudresh M. D., Latha A. S., Suganya J., Nayana C. G. Performance analysis of speech digit recognition using cepstrum and vector quantization. *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*. Mysuru, India, 2017. P. 1–6.
DOI: <https://doi.org/10.1109/ICEECCOT.2017.8284580>.
4. Савула А. А., Коротинський А. П. Застосування звукових ознак для моніторингу техніко-експлуатаційного стану промислового обладнання. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки*. 2024. Т. 35 (74), № 6. Ч. 2. С. 215 – 220. DOI: <https://doi.org/10.32782/2663-5941/2024.6.2/30>
5. Alkhatib B., Kamal Eddin M. M. W. Voice Identification Using MFCC and Vector Quantization. *Baghdad Science Journal*. 2020. Vol. 17, Iss. 3 (Suppl.). P. 1019–1026.
DOI: [https://doi.org/10.21123/bsj.2020.17.3\(Suppl.\).1019](https://doi.org/10.21123/bsj.2020.17.3(Suppl.).1019).

6. Abdul Z., Al-Talabani A. Mel Frequency Cepstral Coefficient and its Applications: A Review. *IEEE Access*. 2022. Vol. 10. P. 122136–122158.

DOI: <https://doi.org/10.1109/ACCESS.2022.3223444>.

REFERENCES

1. Polyvoda, O., Rudakova, H., Kondratieva, I., Rozov, Y., & Lebedenko, Y. (2020). Digital Acoustic Signal Processing Methods for Diagnosing Electromechanical Systems. In V. Lytvynenko, S. Babichev, W. Wójcik, O. Vynokurova, S. Vyshemyrskaya, & S. Radetskaya (Eds.), *Lecture Notes in Computational Intelligence and Decision Making. ISDMCI 2019. Advances in Intelligent Systems and Computing* (Vol. 1020, pp. 97–109). Springer, Cham. https://doi.org/10.1007/978-3-030-26474-1_7

2. Bondarenko, M. E., & Ivashchenko, H. S. (2025). Vykorystannia poslidovnosti metodiv poperednoi obrobky v systemakh holosovoi identyfikatsii [Using a sequence of preprocessing methods in voice identification systems]. *Systems of Control, Navigation and Communication. Collection of Scientific Works*, (2), 90–96. <https://doi.org/10.26906/SUNZ.2025.2.090>

3. Rudresh, M. D., Latha, A. S., Suganya, J., & Nayana, C. G. (2017). Performance analysis of speech digit recognition using cepstrum and vector quantization. *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, 1–6. <https://doi.org/10.1109/ICEECCOT.2017.8284580>

4. Savula, A. A., & Korotynskyi, A. P. (2024). Zastosuvannia zvukovykh oznak dlia monitorynhu tekhniko-eksploatatsiinoho stanu promyslovoho obladdannia [Application of sound features for monitoring the technical and operational state of industrial equipment]. *Scientific Notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences*, 35(74), 6(2), pp 215 – 220. <https://doi.org/10.32782/2663-5941/2024.6.2/30>

5. Alkhatib, B., & Kamal Eddin, M. M. W. (2020). Voice Identification Using MFCC and Vector Quantization. *Baghdad Science Journal*, 17(3 Suppl.), 1019–1026. [https://doi.org/10.21123/bsj.2020.17.3\(Suppl.\).1019](https://doi.org/10.21123/bsj.2020.17.3(Suppl.).1019)

6. Abdul, Z., & Al-Talabani, A. (2022). Mel Frequency Cepstral Coefficient and its Applications: A Review. *IEEE Access*, 10, 122136–122158.

<https://doi.org/10.1109/ACCESS.2022.3223444>

Received 23.04.2026.

Accepted 24.04.2026.

Published 30.04.2026

Intelligent method for moving object identification based on feature clustering in the mel-frequency cepstral domain

The implementation of intelligent monitoring systems in modern industrial automation and advanced transportation networks requires the development of highly reliable methods for moving object identification. The relevance of this study is driven by the necessity to create systems capable of stable operation under conditions of intense acoustic noise, limited optical visibility, and dynamic environmental parameters. Traditional amplitude-based sound processing methods prove to be ineffective due to their high sensitivity to signal non-stationarity and the Doppler effect, which prevents accurate segmentation of an object's dynamic states as the distance to the sensor changes.

The objective of this work is to develop and scientifically substantiate a method for identifying motion phases based on the complex application of Mel-frequency cepstral analysis and intelligent clustering. The proposed approach is based on the synthesis of Short-Time Fourier Transform) algorithms using the Hamming window to minimize spectral leakage, psychoacoustic filtering based on Mel-scale filters, and Discrete Cosine Transform for feature disintegration and decorrelation in the cepstral domain.

The research methodology is based on converting a non-stationary acoustic signal into a time-frequency representation followed by the calculation of Mel-Frequency Cepstral Coefficients. This allowed for the formation of a multidimensional feature matrix that acts as a unique "digital fingerprint" of the object. A significant advantage of this method is its invariance to amplitude (loudness) fluctuations, achieved through logarithmic energy scaling within each critical frequency band. For automatic state recognition, an unsupervised machine learning method – K-means clustering – was applied, allowing for the grouping of feature vectors into compact centroids corresponding to specific motion phases.

The scientific novelty of the results lies in the improvement of the dynamic state segmentation process for moving objects by transitioning from time-domain characteristic analysis to analysis within the cepstral feature space. It has been experimentally confirmed that distinct differentiation is possible between the phases of approach, transit (point of closest proximity), and departure of the object. The results demonstrate high computational stability and speed, ensuring the possibility of integrating the method into local microprocessor-based automated control systems, PLCs, and real-time technical diagnostic systems. The practical value of the work lies in the potential to create autonomous acoustic detectors for monitoring industrial equipment and traffic flows without the need for expensive video surveillance systems.

Keywords: moving object identification, mel-frequency cepstral coefficients, clustering, monitoring automation, acoustic signal, digital signal processing, industrial automation, motion phase segmentation.

Поливода Оксана Валеріївна – к.т.н., доцент, доцент кафедри автоматизації, робототехніки і мехатроніки, Херсонський національний технічний університет.

ORCID: <http://orcid.org/0000-0002-6323-3739>

Назмєєв Ілля Олегович – аспірант кафедри автоматизації, робототехніки і мехатроніки, Херсонський національний технічний університет.

ORCID: <http://orcid.org/0009-0008-2360-0347>

Polyvoda Oksana Valeriivna – Candidate of Technical Sciences, Associate Professor, Associate Professor at the Department of Automation, Robotics and Mechatronics, Kherson National Technical University.

ORCID: <http://orcid.org/0000-0002-6323-3739>

Nazmieiev Illia Olehovych – Postgraduate Student at the Department of Automation, Robotics and Mechatronics, Kherson National Technical University.

ORCID: <http://orcid.org/0009-0008-2360-0347>