

МАТЕМАТИЧНА МОДЕЛЬ ІДЕНТИФІКАЦІЇ АКУСТИЧНИХ СИГНАЛІВ У МЕЛ-КЕПСТРАЛЬНОМУ ПРОСТОРИ

Поливода О.В.¹ [ORCID], Назмєєв І.О.² [ORCID]

¹Херсонський національний технічний університет, к.т.н., доцент, Україна

²Херсонський національний технічний університет, аспірант, Україна

Анотація. Впровадження інтелектуальних систем моніторингу в промислову 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 нестационарності акустичних сигналів та зміни відстані до джерела, що ускладнює точну сегментацію динамічних станів об'єкта. Метою роботи є розробка математичної моделі 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. 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.

Сучасний стан розвитку інтелектуальних транспортних систем та систем автоматизованого моніторингу потребує впровадження алгоритмів ідентифікації, здатних з високою точністю та швидкодією розпізнавати рухомі об'єкти. Традиційні методи ідентифікації з використанням комп'ютерного зору, часто виявляються малоефективними в умовах обмеженої видимості, несприятливих погодних умов або складного освітлення, що зумовлює перехід до методів моніторингу заснованих на аналізі акустичних полів [1]. Тому актуальним завданням є математична формалізація процесу розпізнавання рухомих об'єктів через перетворення часової реалізації сигналу у багатовимірний простір кепстральних ознак, що дозволяє абстрагуватися від енергетичних параметрів та зосередитися на ідентифікації спектральної

огиноючої, що відображає фундаментальні фізичні властивості джерела звуку. Впровадження таких математичних моделей у системи промислової автоматизації забезпечує інваріантність процесу ідентифікації до зовнішніх збурень та гарантує високу швидкодію в режимі реального часу.

Метою дослідження є розробка математичної моделі ідентифікації фаз руху об'єктів шляхом перетворення часових реалізацій акустичних сигналів у простір мел-кепстральних коефіцієнтів з наступною сегментацією станів на основі ітераційних алгоритмів кластеризації.

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]. Акустичний сигнал розділяється на короткі кадри тривалістю N відліків та виконується віконна обробка за методом Геммінга, результатом якої є зважена послідовність $s_{\omega}(n)$:

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

де n – дискретний час.

2. Швидке перетворення Фур'є [3]. Для переходу з часової області у частотну для кожного вікна обчислюється дискретне перетворення Фур'є:

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

де k – індекс частотної компоненти.

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 [4]. 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.

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.

Наступний крок – 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 $S = \{S_1, S_2, \dots, S_K\}$ за рахунок 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, $\|x - \mu_i\|$ – евклідова відстань між вектором ознак та центром кластера.

Важливим результатом проведеного дослідження є отримані спектральні портрети центроїдів, що представляють еталонні образи чотирифазної моделі руху об'єкта (рис. 1). Вісь ординат представлена у відносних децибелах, а вісь абсцис відображає центральні частоти мел-фільтрів. Графічний аналіз свідчить про те, що сформовані кластери є лінійно роздільними, що є критично важливим для стабільної роботи автоматизованих систем класифікації.

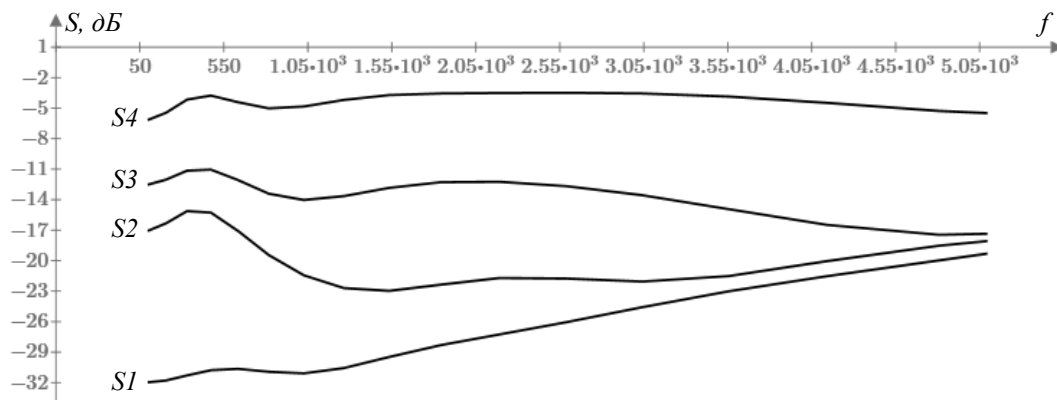


Рисунок 1 – Спектральні портрети центроїдів

Між граничними станами $S1$ (шум) та $S4$ (максимальне наближення) спостерігається амплітудна різниця понад 25 дБ, що гарантує високу надійність розрізнення фаз руху навіть в умовах зашумленого 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.

Висновки

Запропонована математична модель ідентифікації акустичних сигналів у мел-кепстральному просторі характеризується високою швидкістю обчислень, що дозволяє ефективно обробляти великі масиви багатовимірних даних у режимі реального часу. 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.

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 (ICECCOT). Mysuru, India, 2017. P. 1–6. DOI: <https://doi.org/10.1109/ICECCOT.2017.8284580>.
4. 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>.

MATHEMATICAL MODEL FOR ACOUSTIC SIGNAL IDENTIFICATION IN THE MEL-FREQUENCY CEPSTRAL DOMAIN

Oksana Polyvoda, Illia Nazmieiev

Abstract. *The implementation of intelligent monitoring systems in industrial automation and advanced transportation networks requires the development of moving object identification methods capable of stable operation under high noise levels and limited visibility. Traditional amplitude-based processing methods are sensitive to the non-stationarity of acoustic signals and changes in distance to the source, which complicates the accurate segmentation of the object's dynamic states. The objective of this work is to develop a mathematical model for motion phase identification based on Mel-frequency cepstral analysis, intelligent K-means clustering, Short-Time Fourier Transform, psychoacoustic filtering, and feature decorrelation in the cepstral domain. The effectiveness of phase segmentation for approach, transit, and departure has been experimentally confirmed with high computational speed, enabling the integration of the method into automated control and technical diagnostic systems in real-time.*

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

REFERENCE

1. Polyvoda, O., Rudakova, H., Kondratieva, I., Rozov, Y., & Lebedenko, Y. (2020). 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 (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 (ICECCOT), 1–6. <https://doi.org/10.1109/ICECCOT.2017.8284580>
4. 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>