

КОТ Максим

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

<https://orcid.org/0009-0006-7178-5871>

e-mail: [redkotyara@yahoo.com](mailto:redkotyara@yahoo.com)

СТЕПАНОВ Михайло

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

<https://orcid.org/0000-0001-6376-4268>

e-mail: [2m.stepanov@gmail.com](mailto:2m.stepanov@gmail.com)

## МЕТОДИ ОБРОБКИ КОРИСНИХ СИГНАЛІВ У ЗВУКОВОМУ ДІАПАЗОНІ

У роботі розглядаються основні методи аналізу акустичних сигналів, які застосовуються в сучасних дослідженнях та практичних розробках. Особлива увага приділяється трьом ключовим підходам: спектральному аналізу, вейвлет-перетворенню та методам машинного навчання, зокрема нейронним мережам. Описуються принципи спектрального аналізу, який базується на перетворенні Фур'є та його модифікаціях (дискретне перетворення Фур'є, швидке перетворення Фур'є). Підкреслюється, що спектральний аналіз є особливо ефективним для вивчення стаціонарних процесів, дозволяючи отримати точну характеристику енергетичного розподілу за частотами. Проте для аналізу нестаціонарних сигналів (таких як шум автомобілів, музичні та мовні сигнали) застосовується часово-залежне перетворення Фур'є (STFT), яке має певні обмеження щодо одночасної роздільної здатності у часовій та частотній областях.

Вейвлет-перетворення представлено як альтернативний математичний інструмент, що забезпечує одночасне представлення сигналу як у часовій, так і в частотній областях. На відміну від класичного перетворення Фур'є, цей метод дозволяє локалізувати спектральні компоненти в часі, що особливо важливо для нестаціонарних акустичних сигналів. Принцип вейвлет-перетворення полягає у розкладі сигналу на базисні функції – вейвлети, отримані шляхом масштабування та зсуву материнського вейвлету. Такий підхід дозволяє виявляти особливості сигналу на різних масштабах та у різні моменти часу, а також ефективно знижувати рівень шуму без істотної втрати корисної інформації.

У статті також розглядаються сучасні методи машинного навчання та нейронних мереж для аналізу акустичних сигналів. Підкреслюється, що за останні два десятиліття використання машинного навчання для обробки аудіо сигналів суттєво зросло, і сьогодні ці методи домінують у нових підходах до обробки звукових сигналів. Особлива увага приділяється глибоким нейронним мережам, які часто перевершують традиційні методи обробки сигналів. Автор зазначає, що, незважаючи на запозичення багатьох методів глибокого навчання з обробки зображень, існують важливі відмінності між цими областями, які вимагають спеціальних підходів до аналізу аудіо.

Ключові слова: звук, сигнал, обробка.

KOT Maksym, STEPANOV Mikhailo

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

## METHODS OF PROCESSING AUDIO SIGNALS

The article examines the principal methods of acoustic signal analysis employed in contemporary research and practical applications. Special attention is devoted to three key approaches: spectral analysis, wavelet transformation, and machine learning methods, particularly neural networks. The author provides a detailed description of spectral analysis principles, which are based on Fourier transformation and its modifications (Discrete Fourier Transform, Fast Fourier Transform). The article emphasizes that spectral analysis is particularly effective for studying stationary processes, enabling precise characterization of energy distribution across frequencies. However, for analyzing non-stationary signals (such as automobile noise, musical and speech signals), time-dependent Fourier transform (STFT) is applied, which has certain limitations regarding simultaneous resolution in time and frequency domains.

Wavelet transformation is presented as an alternative mathematical tool that provides simultaneous representation of signals in both time and frequency domains. Unlike classical Fourier transformation, this method allows for the localization of spectral components in time, which is especially important for non-stationary acoustic signals. The principle of wavelet transformation involves decomposing signals into basis functions—wavelets—obtained through scaling and shifting of a mother wavelet. This approach enables the detection of signal features at different scales and at different moments in time, and also effectively reduces noise levels without significant loss of useful information.

The article also explores contemporary machine learning methods and neural networks for acoustic signal analysis. It emphasizes that over the past two decades, the use of machine learning for audio signal processing has grown substantially, and today these methods dominate new approaches to sound signal processing. Particular attention is paid to deep neural networks, which often outperform traditional signal processing methods. The author notes that despite borrowing many deep learning methods from image processing, there are important differences between these fields that require specialized approaches to audio analysis. Audio signals form one-dimensional time series that fundamentally differ from two-dimensional images and must be studied sequentially in chronological order. These properties have given rise to audio-specific solutions in the field of signal processing. The article concludes that the integration of these diverse methods allows for more comprehensive analysis of complex acoustic phenomena in various applications.

Keywords: sound, signal, processing.

Стаття надійшла до редакції / Received 16.04.2025

Прийнята до друку / Accepted 11.05.2025

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Акустичні сигнали є важливим джерелом інформації про навколишнє середовище та процеси, що в ньому відбуваються. Виявлення та обробка корисних акустичних сигналів на фоні шумів та завад представляє собою складну науково-технічну задачу, що має безпосереднє відношення до багатьох практичних застосувань. З розвитком технологій та збільшенням шумового забруднення середовища, важливість ефективних методів обробки акустичних сигналів постійно зростає.

Проблема виділення корисних акустичних сигналів є актуальною в таких областях як системи зв'язку та безпеки. У кожній з цих сфер існують специфічні вимоги до якості сигналу та швидкості обробки алгоритмів в умовах реального середовища.

Незважаючи на значний прогрес у розробці методів обробки акустичних сигналів, залишається ряд невирішених проблем, пов'язаних із адаптацією до змінних умов середовища, забезпеченням надійності розпізнавання в умовах низького співвідношення сигнал/шум, а також зниженням обчислювальної складності алгоритмів для забезпечення можливості їх реалізації в мобільних та вбудованих системах. Це зумовлює необхідність комплексного аналізу існуючих підходів та визначення перспективних напрямків подальших досліджень.

## АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

У праці "Discrete-Time Signal Processing" наведено детальні дані про перетворення Фур'є, включаючи математичні формули для прямого та зворотнього перетворення, його основні властивості та основні теоретичні положення. У подальших розділах міститься розгорнута інформація про алгоритм, який називається швидке перетворення Фур'є (FFT) і часово-залежне перетворення Фур'є (time-dependent Fourier transform), яке може проаналізувати нестационарні сигнали, динамічні характеристики якого змінюються з часом. Однак, важливо зазначити, що застосування цього методу потребує суттєво більших обчислювальних ресурсів, що може становити критичне обмеження для систем із низькою продуктивністю чи енергоефективних обчислювальних пристроїв з обмеженою потужністю.

Стаття "Spectral Analysis for Nonstationary Audio" пропонує формальне визначення нестационарного процесу.

Паралельно з цим, робота "Analysis of sound patterns through wavelet transforms" містить детальний теоретичний опис Вейвлет перетворення разом з наочними графічними ілюстраціями результатів самого перетворення.

Інформація, описана в патенті US6990453B2 "System and methods for recognizing sound and music signals in high noise and distortion", представляє собою систему та метод класифікації звукових сигналів за допомогою порівняльного аналізу ключових ознак вхідного сигналу з еталонними характеристиками, які зберігаються у попередньо сформованій базі даних аудіо зразків, дозволяючи ефективно ідентифікувати зафіксований аудіодетекторною системою звук в умовах значного зашумлення.

У книзі "Wavelet Tour of Signal Processing" є опис про вейвлет перетворення, його властивості, та інформацію про материнські вейвлет функції.

Стаття "Noise Reduction of Speech Signal using Wavelet Transform with Modified Universal Threshold" присвячена зниженню рівня шумів у сигналів за допомогою вейвлет перетворення.

У статті "Machine Learning Applied to Music/Audio Signal Processing" йдеться мова про збільшення використання машинного навчання для обробки аудіо сигналів, потім наведені приклади застосування машинного навчання на практичних завданнях обробки сигналів.

У роботі "Deep Learning for Audio Signal Processing" описується історія використання глибокого навчання, а потім наведені дані про різні методи та моделі для обробки аудіосигналів.

## ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою роботи є: дослідження та сформування огляду сучасних методів обробки сигналів у звуковому діапазоні, виявлення обмежень кожного методу для подальшого використання під час виконання процедури виділення та аналізу корисної інформації, що реєструється за допомогою аудіодетекторних систем.

## Спектральний аналіз

Спектральний аналіз дозволяє детально вивчати частотний склад сигналів, що є особливо корисним при аналізі стаціонарних процесів, коли необхідно отримати точну характеристику енергетичного розподілу за частотами. Цей метод є ефективним для аналізу стаціонарних процесів. Основним інструментом спектрального аналізу є перетворення Фур'є, дискретне перетворення Фур'є (DFT), швидке перетворення Фур'є (FFT) та інші модифікації.

Швидке перетворення Фур'є (FFT) є алгоритмічно оптимізованою реалізацією дискретного перетворення Фур'є, що значно зменшує обчислювальну складність, дозволяючи обробляти великі об'єми даних за короткий час. Цей метод має численні переваги, зокрема ефективність та точність при аналізі

стаціонарних сигналів, але його застосування може бути обмеженим для сигналів, що змінюються в часі, оскільки класичне FFT не враховує часову динаміку, що не підходить для аналізу нестационарних сигналів, як, наприклад, шум автомобілів, музичні та мовні сигнали, шум вітру [1]. Щоб розв'язати це питання, застосовують часово-залежне перетворення Фур'є (time-dependent Fourier transform або STFT), яке полягає у послідовному застосуванні FFT до малих сегментів сигналу, що отримуються шляхом ковзання вікна по часовій області. Цей підхід дозволяє отримати часово-частотну характеристику сигналу [2-4].

Проте для STFT є недолік, який полягає в обмеженні на одночасну роздільну здатність у часовій та частотній областях. Це означає, що при зменшенні вікна аналізу покращується часова роздільна здатність, а для частотна роздільна здатність погіршується. Під час збільшення вікна у частотній області роздільна здатність стає кращою, а у часовій – погіршується. Цей недолік також може мати вплив на продуктивність малопотужних електронно-обчислювальних машинах, у випадках, коли треба отримати високу роздільну здатність в обох областях.

### Вейвлет перетворення

Вейвлет перетворення – це математичний інструмент для аналізу акустичних сигналів, який забезпечує одночасне представлення як у часовій, так і в частотній областях. На відміну від класичного перетворення Фур'є, вейвлет-аналіз дозволяє локалізувати спектральні компоненти в часі, що особливо важливо для нестационарних акустичних сигналів.

Основний принцип вейвлет перетворення полягає у розкладі сигналу на базисні функції – вейвлети, які отримуються шляхом масштабування та зсуву материнського вейвлету [5]. Така декомпозиція дозволяє виявляти особливості сигналу на різних масштабах та у різні моменти часу. Вейвлет перетворення дозволяє проаналізувати високочастотні і низькочастотні компоненти сигналу. Вейвлет-аналіз дозволяє значно знизити рівень шуму в сигналі без істотної втрати корисної інформації [6]. Але існує складність вибору оптимального материнського вейвлету для конкретної задачі, що може вимагати попереднього знання про характеристики сигналу.

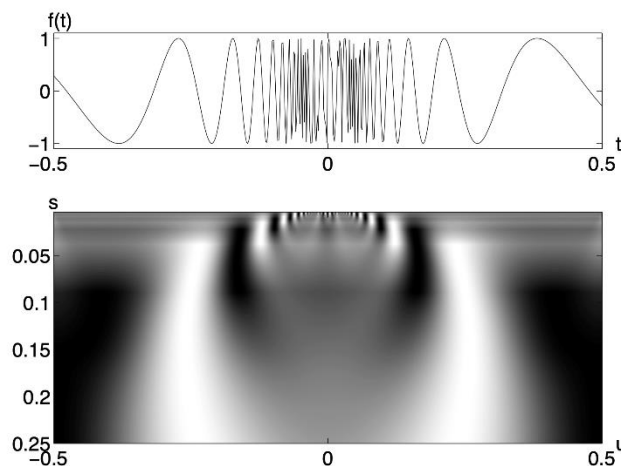


Рис. 1. Графік функції та результат вейвлет перетворення

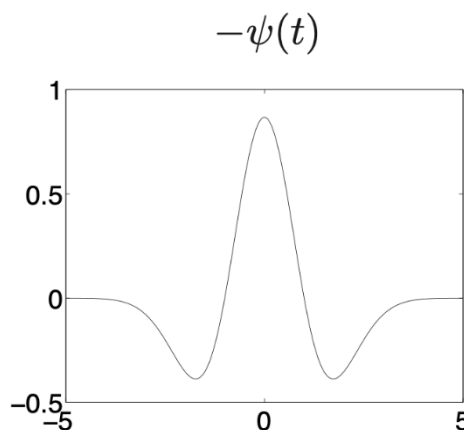


Рис. 2. Графік вейвлету Мексиканська шляпа у часовій області

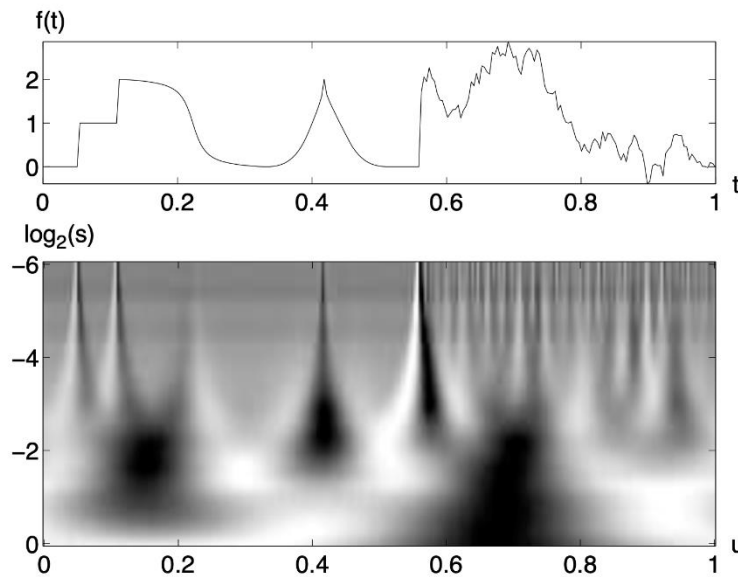


Рис. 4. Графік функції та результат вейвлет перетворення за допомогою вейвлету Мексиканська шляпа. Чорні, сірі та білі точки означають додатні, нульові та від'ємні вейвлет коефіцієнти

### Машинне навчання та нейронні мережі

Методи машинного навчання створюють новий підхід до аналізу та класифікації акустичних сигналів. На відміну від традиційних методів, які вимагають ручного проектування ознак, алгоритми машинного навчання можуть автоматично виявляти важливі характеристики сигналів безпосередньо з даних.

За останні два десятиліття використання машинного навчання для обробки аудіо сигналів різко зросло. Сьогодні у нових підходах у сферах обробки звукових сигналів домінують рішення машинного навчання. Існує широкий спектр завдань, які необхідно вирішити в аналізі та обробці аудіосигналів, більшість з яких вимагають спеціально адаптованих підходів до машинного навчання[7].

Серед методів машинного навчання, які активно застосовуються для обробки акустичних сигналів, слід виділити глибокі нейронні мережі. Сплеск інтересу до глибокого навчання дозволив застосувати його в багатьох областях обробки сигналів, часто перевершуючи традиційну обробку сигналів у великих масштабах. У цій останній хвилі глибоке навчання спочатку набуло поширення в обробці зображень, але потім було широко застосовано в обробці мови, музики та звуку навколишнього середовища, а також у численних додаткових галузях, таких як геноміка, квантова хімія, обробка природної мови та системи рекомендацій. Як наслідок, раніше використовувані методи обробки аудіосигналів, такі як моделі суміші Гауса, приховані моделі Маркова та факторізація невід'ємної матриці, часто перевершували моделі глибокого навчання в програмах, де є достатньо даних [8, 9].

Незважаючи на те, що багато методів глибокого навчання були запозичені з обробки зображень, існують важливі відмінності між областями, які вимагають особливого погляду на аудіо. Необроблені аудіосигнали утворюють одновимірний сигнал часового ряду, який принципово відрізняється від двовимірних зображень. Аудіосигнали зазвичай перетворюються на двовимірне частотно-часове представлення для обробки, але дві осі, час і частота, не є однорідними, як горизонтальна та вертикальна осі в зображенні [10]. Зображення — це миттєві знімки цілі, які часто аналізуються в цілому або частинами з невеликими обмеженнями порядку; однак звукові сигнали потрібно вивчати послідовно в хронологічному порядку. Ці властивості породили специфічні для аудіо рішення.

Також, слід пам'ятати, що використання машинного навчання для обробки акустичних сигналів вимагають значних обсягів навчальних даних.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ

#### І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У даній роботі був проведений огляд сучасних методів обробки сигналів у звуковому діапазоні, виявлення обмежень кожного методу. Аналіз отриманих результатів свідчить про те, що немає ідеального способу обробки аудіо сигналів, кожен з проаналізованих методів має свої недоліки та переваги. Результати даної роботи знадобляться на наступних етапах під час розробки алгоритму обробки звукових сигналів, який пов'язаний із адаптацією до змінних умов середовища, забезпеченням надійності розпізнавання в умовах низького співвідношення сигнал/шум, а також зниженням обчислювальної складності алгоритмів для забезпечення можливості їх реалізації в мобільних та вбудованих системах

### Література

1. Meynard A. Spectral Analysis for Nonstationary Audio [Електронний ресурс] / Adrien Meynard, Bruno Torresani // IEEE/ACM Transactions on Audio, Speech, and Language Processing. – 2018. – Т. 26, № 12. – С. 2371–2380. – Режим доступу: <https://doi.org/10.1109/taslp.2018.2862353> (дата звернення: 22.03.2025). – Назва з екрана.
2. Oppenheim A. Discrete-Time Signal Processing / Alan Oppenheim, Ronald Schaffer. – 2-ге вид. – Upper Saddle River, New Jersey 07458 : Prentice Hall, 1998. – 893 с.
3. Kronland-Martinet R. Analysis of sound patterns through wavelet transforms / R. Kronland-Martinet, J. Morlet, A. Grossmann // International journal of pattern recognition and artificial intelligence. – 1987. – Т. 1, № 2. – С. 273–302.
4. System and methods for recognizing sound and music signals in high noise and distortion : пат. US6990453B2 Сполучені Штати Америки (США) / A. Wang, J. Smith. – Опубл. 24.01.2001. – 30 с.
5. Mallat S. Wavelet Tour of Signal Processing / Stephane Mallat. – [Б. м. : б. в.], 2008.
6. Noise Reduction of Speech Signal using Wavelet Transform with Modified Universal Threshold / Rajeev Aggarwal [та ін.] // International Journal of Computer Applications. – 2011. – Т. 20, № 5. – С. 14–19.
7. Lerch A. Machine Learning Applied to Music/Audio Signal Processing [Електронний ресурс] / Alexander Lerch, Peter Knees // Electronics. – 2021. – Т. 10, № 24. – С. 3077. – Режим доступу: <https://doi.org/10.3390/electronics10243077> (дата звернення: 24.03.2025). – Назва з екрана.
8. Deep Learning for Audio Signal Processing [Електронний ресурс] / Hendrik Purwins [та ін.] // IEEE Journal of Selected Topics in Signal Processing. – 2019. – Т. 13, № 2. – С. 206–219. – Режим доступу: <https://doi.org/10.1109/jstsp.2019.2908700> (дата звернення: 24.03.2025). – Назва з екрана.
9. Семенко, А. І., Бойко, Ю. М., Шпур, О. М., Стрелковська, І. В., Корчинський, В. В., & Яровий, Р. О. (2024). Сучасні технології інфокомунікаційних та комп'ютерних мереж: Монографія (А. І. Семенко, Ред.). Європейський університет, ФО-П Білецький Р. Г., 557 с. <https://elar.khmnmu.edu.ua/handle/123456789/17381>.
10. Буняк, О. В., & Бойко, Ю. М. (2009). Використання вейвлет-перетворення для задач обробки сигналу. Вісник Хмельницького національного університету, (2), 153-154.

### References

1. Meynard A. Spectral Analysis for Nonstationary Audio [Elektronnyi resurs] / Adrien Meynard, Bruno Torresani // IEEE/ACM Transactions on Audio, Speech, and Language Processing. – 2018. – Т. 26, № 12. – С. 2371–2380. – Rezhym dostupu: <https://doi.org/10.1109/taslp.2018.2862353> (data zvernennia: 22.03.2025). – Nazva z ekrana.
2. Oppenheim A. Discrete-Time Signal Processing / Alan Oppenheim, Ronald Schaffer. – 2-he vyd. – Upper Saddle River, New Jersey 07458 : Prentice Hall, 1998. – 893 s.
3. Kronland-Martinet R. Analysis of sound patterns through wavelet transforms / R. Kronland-Martinet, J. Morlet, A. Grossmann // International journal of pattern recognition and artificial intelligence. – 1987. – Т. 1, № 2. – С. 273–302.
4. System and methods for recognizing sound and music signals in high noise and distortion : pat. US6990453B2 Spolucheni Shtaty Ameriky (SShA) / A. Wang, J. Smith. – Opubl. 24.01.2001. – 30 s.
5. Mallat S. Wavelet Tour of Signal Processing / Stephane Mallat. – [B. m. : b. v.], 2008.
6. Noise Reduction of Speech Signal using Wavelet Transform with Modified Universal Threshold / Rajeev Aggarwal [ta in.] // International Journal of Computer Applications. – 2011. – Т. 20, № 5. – С. 14–19.
7. Lerch A. Machine Learning Applied to Music/Audio Signal Processing [Elektronnyi resurs] / Alexander Lerch, Peter Knees // Electronics. – 2021. – Т. 10, № 24. – С. 3077. – Rezhym dostupu: <https://doi.org/10.3390/electronics10243077> (data zvernennia: 24.03.2025). – Nazva z ekrana.
8. Deep Learning for Audio Signal Processing [Elektronnyi resurs] / Hendrik Purwins [ta in.] // IEEE Journal of Selected Topics in Signal Processing. – 2019. – Т. 13, № 2. – С. 206–219. – Rezhym dostupu: <https://doi.org/10.1109/jstsp.2019.2908700> (data zvernennia: 24.03.2025). – Nazva z ekrana.
9. Semenکو, A. I., Boiko, J. M., et al. (2024). Suchasni tekhnolohii infokomunikatsiinykh ta kompiuternykh merezh: Monohrafiia (A. I. Semenکو, Red.). European University, Ahrar Media Print, 557 s. <https://elar.khmnmu.edu.ua/handle/123456789/17381>.
10. Buniak, O. V., & Boiko, J. M. (2009). Vykorystannia veivlet-peretvorennia dlia zadach obrobky syhnalu. Visnyk Khmelnytskoho natsionalnoho universytetu, (2), 153-154.