

Шульга Аліна Вікторівна

кандидат технічних наук, доцент

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

ORCID: 0009-0002-8358-8657

Білоус Олександр Віталійович

аспірант

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

ORCID: 0009-0000-4221-323X

ЗАСТОСУВАННЯ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ АДАПТАЦІЇ СИСТЕМ ФОРМУВАННЯ ПЕРСОНАЛЬНИХ ЗВУКОВИХ ЗОН ДО АКУСТИКИ ПРИМІЩЕНЬ

Анотація. Проведений аналітичний огляд сучасного стану досліджень у галузі формування персональних звукових зон (Personal Sound Zones, PSZ) з акцентом на впровадженні методів штучного інтелекту (ШІ) для адаптації систем до складної акустики приміщень. Введено векторно-матричну модель системи, сформульовано цільові функції алгоритмів контролю акустичного контрасту (JACC) та узгодження тиску (JPM) з параметрами регуляризації λ та μ . Встановлено формальну аналогію між задачею PSZ і задачею beamforming у фазованих антенних решітках (ФАР). Проаналізовано, чому класичні алгоритми ACC та PM деградують в умовах реверберації, і систематизовано підходи на основі ШІ, що вирішують цю проблему: нейромережна ідентифікація ATF, глибоке навчання для компенсації нелінійних спотворень та алгоритми онлайн-адаптації з мінімальною затримкою. Огляд слугує теоретичним підґрунтям для дисертаційного дослідження, спрямованого на розробку ШІ-системи адаптивного керування PSZ. Проведений аналіз показує, що алгоритми формування персональних звукових зон (PSZ) допускають єдину векторно-матричну постановку у вигляді $p = Hw$, де цільові функції JACC та JPM формально збігаються з критеріями MVDR та LCMV beamforming у фазованих антенних решітках (ФАР). Ця математична рамка є точкою докладання методів штучного інтелекту: ШІ не замінює класичні алгоритми, а вирішує їхню головну вразливість, а саме залежність від точної матриці H , яка в реальних умовах постійно змінюється. Практичне значення огляду полягає в тому, що він формує теоретичну і методологічну основу для дисертаційного дослідження, центральною задачею якого є розробка ШІ-системи адаптивного керування PSZ.

Ключові слова: локалізовані зони звуку, аудіо-анклави, персональні звукові зони, масиви гучномовців, параметричний масив, акустичний контраст, узгодження тиску, розбірливість мовлення, ультразвук, ультразвукові випромінювачі, акустичні метаповерхні.

AlinaPhD in Technical Sciences, Associate Professor
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
ORCID: 0009-0002-8358-8657**Shulha**

Professor

OleksandrPhD Student
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
ORCID: 0009-0000-4221-323X**Bilous**

Student

APPLICATION OF ARTIFICIAL INTELLIGENCE METHODS FOR ADAPTING PERSONAL SOUND ZONE FORMATION SYSTEMS TO ROOM ACOUSTICS

Abstract. An analytical review of the current state of research in the field of Personal Sound Zones (PSZ) formation is presented, with an emphasis on the use of Artificial Intelligence (AI) methods for adapting such systems to complex room acoustics. A vector-matrix model of the system is introduced, and the objective functions of Acoustic Contrast Control (JACC) and Pressure Matching (JPM) algorithms are formulated with regularization parameters λ and μ . A formal analogy between the PSZ problem and beamforming in Phased Array Antennas (PAA) is established. The reasons why classical ACC and PM algorithms degrade under reverberant conditions are analyzed, and AI-based approaches aimed at solving this problem are systematized, including neural-network-based ATF identification, deep learning for nonlinear distortion compensation, and low-latency online adaptation algorithms. The review serves as a theoretical basis for dissertation research aimed at developing an AI-based adaptive PSZ control system. The conducted analysis shows that PSZ formation algorithms can be represented within a unified vector-matrix formulation, $p = Hw$, where the objective functions JACC and JPM formally correspond to the MVDR and LCMV beamforming criteria used in phased antenna arrays. This mathematical framework defines the point of application for artificial intelligence methods: AI does not replace classical algorithms but addresses their key vulnerability, namely their dependence on the accurate matrix H , which continuously changes under real acoustic conditions. The practical significance of the review lies in the fact that it forms the theoretical and methodological basis for dissertation research whose central task is the development of an AI-based system for adaptive PSZ control.

Keywords: localized sound zones, audio enclaves, Personal Sound Zones, loudspeaker arrays, parametric array, acoustic contrast, pressure matching, speech intelligibility, ultrasound, ultrasonic transducers, acoustic metasurfaces.

Вступ. Сьогоднішній розвиток акустичних технологій диктує перехід від загального озвучування приміщень до створення індивідуалізованих зон, де звук стає приватним інструментом, а не джерелом шуму для оточуючих. Технології формування локалізованих зон звуку (personal sound zones, PSZ), або як їх дедалі частіше називають в інженерній практиці — аудіо-анклави, представляють собою методики створення акустичного поля, у якому цільовий сигнал відтворюється з високою якістю лише в обмеженій області простору (region of interest, ROI). Водночас поза цією областю (non-region of interest, NROI) рівень сигналу повинен бути настільки низьким, щоб не заважати іншим або забезпечувати повну конфіденційність повідомлення [1, 2].

На відміну від традиційних засобів індивідуального прослуховування, таких як навушники, аудіо-анклави забезпечують кращий досвід. Людина не ізолюється фізично від середовища, що дозволяє їй вільно комунікувати з оточуючими, одночасно отримуючи персоналізовану звукову інформацію. Попит на такі рішення стрімко зростає у сферах відкритого офісного планування, де акустичний дискомфорт є основною причиною зниження продуктивності, а також у музейних експозиціях, ритейлі та транспорті, де необхідно розділити контент для водія та пасажирів [1, 2].

Основна інженерна складність полягає в тому, що звук у вільному просторі поширюється сферично, а його взаємодія з перешкодами та стінами приміщення породжує реверберацію, яка нівелює ефект локалізації. Сучасні дослідження 2024-2025 років зосереджені на пошуку компромісів між акустичним контрастом, якістю відтворення та стійкістю системи до переміщень користувача.

Метою статті є: (1) ввести векторно-матричну модель PSZ-системи та сформулювати цільові функції алгоритмів ACC і PM як задачі оптимізації з параметрами регуляризації λ та μ ; (2) встановити аналогію між задачею PSZ і задачею beamforming у фазованих антенних решітках (ФАР); (3) проаналізувати, чому класичні алгоритми деградує в умовах реальної акустики приміщень, і систематизувати методи ШІ, що усувають цю проблему; (4) окреслити напрям дисертаційного дослідження — розробку ШІ-системи адаптивного керування PSZ із онлайн-оновленням ATF [3].

Наукова новизна отриманих результатів полягає в тому, що:

- вперше систематизовано підходи до адаптації систем формування персональних звукових зон на основі методів штучного інтелекту з урахуванням впливу ревербераційного середовища;
- отримала подальший розвиток векторно-матрична модель PSZ-системи шляхом формалізації цільових функцій акустичного контрасту та узгодження тиску як задач оптимізації;
- встановлено формальну аналогію між задачами формування персональних звукових зон та методами beamforming у фазованих антенних решітках;
- розвинуто підхід до адаптивного керування PSZ-системами на основі поєднання нейромережної ідентифікації акустичних передавальних функцій та алгоритмів навчання з підкріпленням.

Основна частина

Фізичні принципи утворення локалізованих зон звуку. Фундаментальною основою створення аудіо-анклавів є принципи лінійної та нелінійної акустики. Більшість систем базуються на хвильовій моделі, де акустичний тиск у конкретній точці розглядається як результат суперпозиції внесків від множини випромінювачів. Керування цим процесом зводиться до маніпуляції фазою та амплітудою хвиль для досягнення конструктивної інтерференції в цільовій зоні та деструктивної у зоні пригнічення. З погляду теорії сигналів задача PSZ є прямим аналогом задачі формування діаграми спрямованості (beamforming) у фазованих антенних решітках (ФАР): вектор сигналів керування гучномовцями w відповідає вектору комплексних ваг антенної решітки, матриця акустичних передавальних функцій H — матриці steering vectors, а акустичний контраст AC — критерію Signal-to-Interference Ratio (SIR). Це дозволяє безпосередньо застосовувати до PSZ-задачі розвинений апарат просторової фільтрації, включаючи методи MVDR та LCMV [8, 9, 10].

Хвильова модель та інтерференційне керування. У системі з L гучномовців звуковий тиск p у точці r на частоті f описується через передавальні функції середовища (ATF). Математично це виглядає як лінійна комбінація:

$$p(r, f) = \sum_{i=1}^L h_i(r, f) \cdot q_i(f)$$

де $h_i(r, f)$ - передавальна функція від i джерела до точки r , а $q_i(f)$ - комплексний сигнал керування. Ключова фізична ідея полягає в тому, що ми можемо підібрати такі значення q_i , щоб у точках світлої зони (bright zone) амплітуди хвиль додавалися, а в темній зоні (dark zone) взаємно віднімалися.

Проте інтерференція в реальному приміщенні стає проблемою. Якщо у безлунній камері ми маємо справу лише з прямими хвилями, то в кімнаті з'являються відбиття, які створюють стоячі хвилі та моди. Це призводить до того, що оптимальні параметри керування, знайдені для однієї точки, можуть бути неефективними при зміщенні мікрофона на кілька сантиметрів. Ця фізична нестабільність особливо виражена на високих частотах, де довжина хвилі порівнянна з похибками позиціонування [3].

Дифракційні обмеження та апертура. Здатність системи фокусувати звук обмежена законами дифракції. Фундаментальний принцип стверджує, що чим вужчу зону ми хочемо отримати, тим більшою має бути апертура випромінювача відносно довжини хвилі. Кутова ширина променя $\Delta\alpha$ для лінійного масиву довжиною D приблизно дорівнює:

$$\Delta\alpha \approx 1.22 \cdot \frac{\lambda}{D}$$

На низьких частотах (наприклад, 100 Гц) створення вузької зони вимагає масивів розміром у кілька метрів. Це пояснює, чому більшість компактних персональних аудіо-систем ефективно працюють лише у середньочастотному та високочастотному діапазонах, тоді як локалізація НЧ залишається однією з найскладніших задач [1].

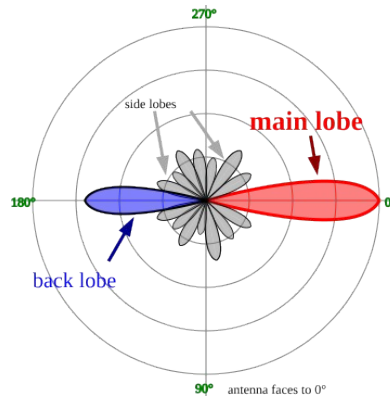


Рис. 1. Схема формування звукової зони: головна пелюстка в ROI та бічні пелюстки, що створюють витік у NROI.

Нелінійна акустика та параметричні явища. Окремим і надзвичайно перспективним напрямком є параметричні масиви (parametric arrays). Вони використовують нелінійні властивості повітря для самодемодуляції ультразвукового сигналу. Коли ультразвукова хвиля (носіє, наприклад, 40 кГц) модулюється аудіосигналом і випромінюється у простір, внаслідок нелінійної взаємодії в повітрі виникає вторинна хвиля потрібного діапазону [12].

Фізика цього явища описується моделлю Берктейя, згідно з якою тиск демодульованого сигналу $p_2(t)$ пропорційний другій похідній від квадрата обвідної ультразвукового сигналу $E(t)$:

$$p_2(t) \propto \frac{d^2}{dt^2} [E(t)]^2$$

Це створює ефект ліхтаря: оскільки довжина хвилі ультразвуку дуже мала, ми отримуємо надзвичайно вузький промінь звуку навіть при малих габаритах випромінювача. Проте ціна такої спрямованості низький ККД та високі нелінійні спотворення, що потребує складних алгоритмів налаштувань [14].

Акустичні метаповерхні. Останніми роками фізика аудіо-анклавів збагатилася концепцією метаповерхонь. Це штучні структури з субхвильовими елементами, які дозволяють керувати фазою хвилі на кожній ділянці поверхні. Метаповерхні можуть працювати як градієнтні лінзи (GRIN-лінзи), фокусуючи звук у конкретній точці простору без керування кожним каналом окремо. Особливо цікавими є «самозгинальні» промені (self-bending beams), які здатні огинати перешкоди і зустрічатися в точці прослуховування, мінімізуючи розсіювання [16, 18].

Технічні засоби створення перепаду звукового тиску. Сьогодні на ринку представлено кілька основних типів пристроїв, кожен з яких реалізує власний спосіб формування локалізованих зон звуку. Вибір між ними залежить від балансу між якістю звуку, енергоефективністю та рівнем ізоляції зони.

Масиви у чутому діапазоні (ACC та PM). Традиційний підхід створення аудіо-анклавів базується на використанні масивів звичайних електродинамічних гучномовців. Ключовим технічним засобом тут є цифровий сигнальний процесор (DSP), який розраховує фільтри для кожного динаміка.

Метод контролю акустичного контрасту (ACC) спрямований на максимізацію енергетичного відношення між ROI та NROI. У векторно-матричній формі цільова функція JACC записується як: $JACC(w) = (w^H Hb^H Hb w) / (w^H Hd^H Hd w + \lambda \|w\|^2)$, де $Hb \in C^{MB \times L}$ та $Hd \in C^{MB \times L}$ — матриці ATF для світлої та темної зони відповідно, а λ — параметр регуляризації Array Effort, що обмежує загальну потужність збудження масиву [4, 3]. Оптимальний вектор w^* є власним вектором, що відповідає максимальному власному значенню матриці $(Hd^H Hd + \lambda I)^{-1} Hb^H Hb$. Аналогічна структура задачі зустрічається в MVDR-beamformer для ФАР, де λ виконує роль навантаження (diagonal loading) для підвищення робастності до похибок steering vector. Це ідеально для сповіщень у шумних залах [4].

Метод узгодження тиску (Pressure Matching, PM) мінімізує середньоквадратичну похибку між реальним і цільовим звуковим полем. Цільова функція JPM у матричній формі: $JPM(w) = \|Hb w - d\|^2 + \mu \|Hd w\|^2$, де $d \in C^{MB}$ — вектор цільового тиску (наприклад, поле ідеального монополя), а μ — параметр регуляризації, що визначає компроміс між якістю відтворення в ROI та рівнем пригнічення в NROI [5, 11]. Аналітичне рішення має вигляд: $w^* = (Hb^H Hb + \mu Hd^H Hd)^{-1} Hb^H d$, що є прямим аналогом LCMV-beamformer у ФАР. На відміну від ACC, PM відтворює точну форму сигналу, тому забезпечує вищу якість музики, але дає нижчий акустичний контраст, зокрема у приміщеннях з $RT60 > 0,3$ с [5].

Сучасні системи, такі як розробки компаній Noveto або Audfly, використовують сотні мікро-динаміків, що дозволяє створювати дуже гнучкі діаграми спрямованості та навіть реалізовувати "звукові кишені" біля вух користувача [1].

Параметричні ультразвукові випромінювачі. Параметричні гучномовці, такі як Audio Spotlight від Holosonics технічно складаються з масиву п'єзоелектричних трансд'юсерів. Вони випромінюють

вузькоспрямований ультразвуковий пучок інтенсивністю близько 110-130 дБ SPL (на частотах понад 40 кГц). Аудіосигнал накладається на цей носій за допомогою амплітудної модуляції (АМ). Основним викликом для таких пристроїв є нелінійні спотворення. У 2024 році було запропоновано використовувати нейромережі типу WaveNet для ідентифікації та компенсації нелінійностей у реальному часі, що дозволило знизити коефіцієнт гармонік (THD) з 15-20% до прийнятних 3% [14, 15]. Це відкриває шлях до використання параметричних систем у Hi-Fi аудіо.

Метаповерхні та пасивні фокусуєчі пристрої. Технічна реалізація метаповерхонь зазвичай передбачає 3D-друк складних лабіринтоподібних структур або використання мембран з керованим натягом. Останні розробки демонструють акустичні метаповерхні (ІАМ), де фокусна відстань змінюється шляхом регулювання тиску повітря всередині мембрани. Це дозволяє динамічно переміщувати звукову зону за слухачем без складних електронних фазообертачів [16, 17].

Гібридні конфігурації. Найбільш досконалі пристрої на сьогодні це гібриди. Вони поєднують ультразвукову панель для точної спрямованості на високих частотах і малий масив звичайних динаміків для відтворення низьких частот, які УЗ-системи відтворюють неефективно. Такі системи часто доповнюються камерами для відстеження положення голови, що дозволяє системі тримати звуковий анклав точно на вухах користувача, навіть якщо він рухається [6].

Нижче наведено порівняльну таблицю характеристик основних технологічних підходів.

Таблиця 1

Порівняння технічних засобів формування звукових зон

Технологія	Метод створення перепаду тиску	Спрямованість	Якість звуку (THD)	Габарити
Масиви динаміків	Інтерференція (АСС/РМ)	Середня	Висока (< 1%)	Великі (особливо для НЧ)
Параметричні УЗ	Нелінійна демодуляція	Дуже висока	Низька (3-15%)	Компактні панелі
Метаповерхні	Фазова затримка в структурі	Висока	Висока	Дуже тонкі структури
Гібридні системи	Комбіноване керування	Висока	Середня/Висока	Помірні

Огляд методів вимірювання якості звуку. Оцінювання ефективності аудіо-анклавів потребує специфічних метрик, оскільки звичайні параметри, наприклад АЧХ, не дають повної картини просторової ізоляції.

Фізичні характеристики поля. Першим і найважливішим показником є Акустичний контраст (АС). Це відношення середньої енергії у світлій зоні до енергії в темній зоні. У наукових роботах він часто обчислюється як відношення середньоквадратичних значень тиску в контрольних точках:

$$AC = 10 \frac{\sum_{i \in ROI} |p_i|^2}{\sum_{j \in NROI} |p_j|^2}$$

Типові значення для успішних систем становлять 12-18 дБ. Якщо контраст нижче 10 дБ, ефект локалізації стає майже непомітним у реальних умовах через фоновий шум [19].

Другим параметром є Ширина зони (-3 dB / FWHM). Вона показує просторову область, де звук залишається гучним. У добре сфокусованих системах ширина зони становить близько 15-30 см, що відповідає розміру голови людини. Якщо зона ширша, виникає ризик того, що звук почують інші.

Перцептивні метрики розбірливості. Для мовних сценаріїв ключовим є Індекс передачі мовлення (STI). Він вимірює, наскільки добре зберігається модуляція сигналу після проходження через тракт. STI варіюється від 0 до 1.

STI > 0.60: Гарна розбірливість, необхідна в ROI.

STI < 0.40: Низька розбірливість, бажана в NROI для забезпечення конфіденційності [19].

Для музики використовують складніші моделі, такі як PEAQ (Perceptual Evaluation of Audio Quality), які враховують психоакустичне маскування. Оцінка ODG (Objective Difference Grade) показує, наскільки відтворений звук відрізняється від оригіналу за шкалою від 0 (непомітно) до -4 (дуже дратівливо) [13].

Таблиця 2

Ключові метрики та їхні цільові значення

Метрика	Опис	Цільове значення (ROI)	Цільове значення (NROI)
ΔSPL (Контраст)	Різниця рівнів ROI/NROI	> 12-15 дБ	< -12 дБ від ROI
STI	Розбірливість мовлення	> 0.65	< 0.35 (приватність)
THD	Гармонійні спотворення	< 1% (динаміки), < 5% (УЗ)	-
FWHM	Ширина звукової плями	20-40 см	-
C50	Чіткість (speech clarity)	> +2 дБ	-

Вплив реверберації та параметри приміщення. Вимірювання часто проводять у два етапи: у безлунній камері (для оцінки потенціалу) та в реальному офісі. Тут важливі параметри C50 (чіткість мовлення) та C80 (чіткість музики), які показують співвідношення ранньої енергії до пізньої реверберації. Для аудіо-анклавів

критично, щоб система не лише створювала перепад тиску, а й не розмивала часову структуру сигналу. [20]

Задачі та перешкоди на шляху створення кращого пристрою. Незважаючи на значні зусилля, ідеальний аудіо-анклав досі залишається викликом для інженерів. Виділимо чотири основні проблеми, які заважають масовому впровадженню технології.

Стабільність у ревербераційних середовищах. Це найголовніша перешкода і водночас головна точка застосування ШІ-методів. Алгоритми АСС та РМ надзвичайно чутливі до точності матриці ATF: навіть незначна зміна температури повітря на 1–2°C змінює швидкість звуку і спричиняє фазовий зсув, що руйнує зону. У реальних офісах з металевими та скляними перегородками контраст падає на 5–10 дБ порівняно з лабораторними умовами. Три напрями ШІ демонструють найбільший потенціал для вирішення цієї задачі. По-перше, нейромережна ідентифікація ATF в реальному часі: згорткові мережі (CNN) та рекурентні мережі (LSTM) навчаються відновлювати матрицю H за коротким вимірjuвальним сигналом (10–50 мс), що дозволяє оновлювати фільтри керування до того, як акустика встигає змінитись [21, 22]. По-друге, навчання з підкріпленням (Reinforcement Learning, RL) для онлайн-адаптації параметрів λ та μ : агент отримує зворотний зв'язок у вигляді поточного значення АС і коригує регуляризацію без повного перерахунку матриці H , що знижує обчислювальну складність з $O(L^3)$ до $O(L)$ на кожному кроці адаптації [24, 25]. По-третє, архітектури типу WaveNet та U-Net застосовуються для прекомпенсації нелінійних спотворень у параметричних УЗ-системах, знижуючи THD з 15–20% до 3% [15]. Спільним критерієм якості для всіх ШІ-підходів є збереження АС ≥ 12 дБ при затримці адаптації не більше 30 мс – вимога, що витікає з психоакустичних досліджень сприйняття просторового звуку [1].

Низька ефективність та спотворення параметричних систем. Параметричні УЗ гучномовці дають найкращу спрямованість, але вони забезпечують низьку якість звуку та високе енергоспоживання. Через нелінійну природу демодуляції виникає величезна кількість гармонік та інші завади. Крім того, УЗ пристрої майже не відтворюють низькі частоти. Дослідження 2024-2025 років показують, що вирішення лежить у площині глибокого навчання, яке дозволяє прогнозувати спотворення та компенсувати їх ще до випромінювання [14, 15].

Проблема апертури та низьких частот. Як було зазначено в розділі про фізику, для локалізації НЧ потрібні величезні випромінювачі. Це робить неможливим створення компактних настільних систем, які б повноцінно відтворювали музику. Вихід шукають у використанні акустичних метаматеріалів, які могли б використовувати ефекти ближнього поля. Але поки що це залишається на рівні експериментальних прототипів [1].

Затримки та відстеження користувача. Якщо ми хочемо, щоб аудіо-анклав слідував за головою слухача, система повинна постійно перераховувати сотні фільтрів. Сумарна затримка від моменту руху голови до адаптації поля має бути менше 30 мс. При використанні складних алгоритмів РМ це вимагає величезних обчислювальних потужностей. Розробка розподілених алгоритмів обробки, які б ділили навантаження між вузлами масиву, є одним із пріоритетних завдань [5].

Висновок. Проведений аналіз показує, що алгоритми формування персональних звукових зон (PSZ) допускають єдину векторно-матричну постановку у вигляді $p = Hw$, де цільові функції JACC та JPM формально збігаються з критеріями MVDR та LCMV beamforming у фазованих антенних решітках (ФАР). Ця математична рамка є точкою докладання методів штучного інтелекту: ШІ не замінює класичні алгоритми, а вирішує їхню головну вразливість, а саме залежність від точної матриці H , яка в реальних умовах постійно змінюється [1, 3]. Кількісний аналіз показує межі класичних підходів: АСС забезпечує контраст 15–22 дБ та STI $< 0,35$ у темній зоні, але втрачає 6–9 дБ при зростанні RT60 від 0,1 до 0,5 с; РМ стійкіший до реверберації (деградація 2–4 дБ) і дає THD $< 1\%$, але поступається АСС за рівнем ізоляції на 5–8 дБ [3, 5]. Саме ця деградація при RT60 $> 0,3$ с є кількісним обґрунтуванням необхідності ШІ-адаптації: нейромережна ідентифікація ATF та онлайн-налаштування λ і μ через навчання з підкріпленням дозволяють утримувати АС ≥ 12 дБ у реальних умовах, де статичні алгоритми не справляються [21, 24].

Практичне значення огляду полягає в тому, що він формує теоретичну і методологічну основу для дисертаційного дослідження, центральною задачею якого є розробка ШІ-системи адаптивного керування PSZ. Конкретна архітектура, що обґрунтовується цим оглядом: нейромережний блок швидкої ідентифікації ATF (CNN/LSTM, час оновлення ≤ 30 мс) у зв'язці з RL-агентом адаптації параметрів λ та μ . Ціль – забезпечити стабільний акустичний контраст АС ≥ 12 дБ у реальних офісних приміщеннях з RT60 до 0,5 с, де сьгоднішні статичні алгоритми втрачають 6–9 дБ і стають практично непридатними. Успішне вирішення цієї задачі наближає технологію персонального аудіо без навушників до масового практичного застосування.

Декларація про штучний інтелект. Штучний інтелект не використовувався.

Конфлікт інтересів. Автори заявляють про відсутність конфлікту інтересів та підтверджують, що під час підготовки цієї роботи не існувало жодних комерційних, фінансових чи інших взаємовідносин, які могли б бути розцінені як такі, що здатні вплинути на результати дослідження або їх інтерпретацію. Робота виконана відповідно до принципів академічної доброчесності, етичних норм проведення наукових досліджень та вимог редакційної політики щодо запобігання конфлікту інтересів.

Список літератури

1. Betlehem T. *Personal Sound Zones: Delivering Interface-Free Audio to Multiple Listeners* / T. Betlehem, W. Zhang, M. A. Poletti, T. D. Abhayapala // *IEEE Signal Processing Magazine*. — 2015. — Vol. 32, No. 2. — P. 81–91. — DOI: 10.1109/MSP.2014.2360707.
2. Jacobsen R. M. *Living with Sound Zones: A Long-Term Field Study of Dynamic Sound Zones in a Domestic Context* / R. M. Jacobsen, K. F. Skov, S. S. Johansen, M. B. Skov, J. Kjeldskov // *Proc. CHI Conference on Human Factors in Computing Systems (CHI'23)*. — New York : ACM, 2023. — Art. 405. — DOI: 10.1145/3544548.3581535.
3. Elliott S. J. *Robustness and Regularization of Personal Audio Systems* / S. J. Elliott, J. Cheer, J.-W. Choi, Y. Kim // *IEEE Transactions on Audio, Speech, and Language Processing*. — 2012. — Vol. 20, No. 7. — P. 2123–2133. — DOI: 10.1109/TASL.2012.2197613.
4. Chang J.-H. *A Realization of Sound Focused Personal Audio System Using Acoustic Contrast Control* / J.-H. Chang, C.-H. Lee, J.-Y. Park, Y.-H. Kim // *The Journal of the Acoustical Society of America*. — 2009. — Vol. 125, No. 4. — P. 2091–2097. — DOI: 10.1121/1.3082114.
5. Shi L. *Generation of Personal Sound Zones With Physical Meaningful Constraints and Conjugate Gradient Method* / L. Shi, T. Lee, L. Zhang, J. K. Nielsen, M. G. Christensen // *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. — 2021. — Vol. 29. — P. 823–837. — DOI: 10.1109/TASLP.2021.3052564.
6. Zhu Y. *A Robust Hybrid ACC-PM Approach for Personal Sound Zones* / Y. Zhu, L. Zhou, H. Liu, L. Shi, L. Gan // *Proc. Interspeech 2025*. — Rotterdam, 2025. — P. 3853–3857. — DOI: 10.21437/Interspeech.2025-65.
7. Yang J. *A Review of Sound Field Control* / J. Yang, M. Wu, L. Han // *Applied Sciences*. — 2022. — Vol. 12, No. 14. — Art. 7319. — DOI: 10.3390/app12147319.
8. Capon J. *High-Resolution Frequency-Wavenumber Spectrum Analysis* / J. Capon // *Proceedings of the IEEE*. — 1969. — Vol. 57, No. 8. — P. 1408–1418. — DOI: 10.1109/PROC.1969.7278.
9. Frost O. L. *An Algorithm for Linearly Constrained Adaptive Array Processing* / O. L. Frost // *Proceedings of the IEEE*. — 1972. — Vol. 60, No. 8. — P. 926–935. — DOI: 10.1109/PROC.1972.8817.
10. Van Trees H. L. *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory* / H. L. Van Trees. — New York : John Wiley & Sons, 2002. — 1472 p. — DOI: 10.1002/0471221104.
11. Souden M. *A Study of the LCMV and MVDR Noise Reduction Filters* / M. Souden, J. Benesty, S. Affes // *IEEE Transactions on Signal Processing*. — 2010. — Vol. 58, No. 9. — P. 4925–4935. — DOI: 10.1109/TSP.2010.2051803.
12. Berkta H. O. *Possible Exploitation of Non-Linear Acoustics in Underwater Transmitting Applications* / H. O. Berkta // *Journal of Sound and Vibration*. — 1965. — Vol. 2, No. 4. — P. 435–461. — DOI: 10.1016/0022-460X(65)90116-3.
13. Yoneyama M. *The Audio Spotlight: An Application of Nonlinear Interaction of Sound Waves to a New Type of Loudspeaker Design* / M. Yoneyama, J. Fujimoto, Y. Kawamo, S. Sasabe // *The Journal of the Acoustical Society of America*. — 1983. — Vol. 73, No. 5. — P. 1532–1536. — DOI: 10.1121/1.389414.
14. Pompei F. J. *The Use of Airborne Ultrasonics for Generating Audible Sound Beams* / F. J. Pompei // *Journal of the Audio Engineering Society*. — 1999. — Vol. 47, No. 9. — P. 726–731.
15. Li M. *Deep Learning-Based Approach for Identification and Compensation of Nonlinear Distortions in Parametric Array Loudspeakers* / M. Li, T. Zhuang, K. Chen, J.-X. Zhong, J. Lu // *IEEE Signal Processing Letters*. — 2025. — Vol. 32. — P. 1455–1459. — DOI: 10.1109/LSP.2025.3553434.
16. Zhang H. *Optimizing Multi-User Indoor Sound Communications with Acoustic Reconfigurable Metasurfaces* / H. Zhang, Q. Wang, M. Fink, G. Ma // *Nature Communications*. — 2024. — Vol. 15. — Art. 1270. — DOI: 10.1038/s41467-024-45435-4.
17. He Y. *Audible Enclaves Crafted by Nonlinear Self-Bending Ultrasonic Beams* / Y. He, Q. Bi, L. Zheng, Y. Tian, N. Fang, X. Liu // *Proceedings of the National Academy of Sciences*. — 2024. — Vol. 121, No. 14. — e2315513121. — DOI: 10.1073/pnas.2315513121.
18. Cummer S. A. *Controlling Sound with Acoustic Metamaterials* / S. A. Cummer, J. Christensen, A. Alu // *Nature Reviews Materials*. — 2016. — Vol. 1. — Art. 16001. — DOI: 10.1038/natrevmats.2016.1.
19. Qiao Y. *Isolation Performance Metrics for Personal Sound Zone Reproduction Systems* / Y. Qiao, L. Guadagnin, E. Choueiri // *JASA Express Letters*. — 2022. — Vol. 2, No. 10. — Art. 104801. — DOI: 10.1121/10.0014604.
20. Wang S. *A Stable Sound Field Control Method for a Personal Audio System* / S. Wang, Q. Liu, J. Tian, J. Pan, X. Jiang // *Applied Sciences*. — 2023. — Vol. 13, No. 22. — Art. 12209. — DOI: 10.3390/app132212209.
21. Yu W. *Room Acoustical Parameter Estimation from Room Impulse Responses Using Deep Neural Networks* / W. Yu, W. B. Kleijn // *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. — 2021. — Vol. 29. — P. 436–447. — DOI: 10.1109/TASLP.2020.3043115.
22. Steinmetz C. J. *Filtered Noise Shaping for Time Domain Room Impulse Response Estimation from Reverberant Speech* / C. J. Steinmetz, V. K. Ithapu, P. Calamia // *Proc. 2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*. — 2021. — P. 221–225. — DOI: 10.1109/WASPAA52581.2021.9632680.
23. Ratnarajah A. *Towards Improved Room Impulse Response Estimation for Speech Recognition* / A. Ratnarajah, I. Ananthabhotla, V. K. Ithapu, P. Hoffmann, D. Manocha, P. Calamia // *Proc. ICASSP 2023*. — 2023. — P. 1–5. — DOI: 10.1109/ICASSP49357.2023.10094770.
24. Luo Z. *GFANC-RL: Reinforcement Learning-Based Generative Fixed-Filter Active Noise Control* / Z. Luo, H. Ma, D. Shi, W.-S. Gan // *Neural Networks*. — 2024. — Vol. 180. — Art. 106687. — DOI: 10.1016/j.neunet.2024.106687.

25. *Mismar F. B. Deep Reinforcement Learning for 5G Networks: Joint Beamforming, Power Control, and Interference Coordination / F. B. Mismar, B. L. Evans, A. Alkhateeb // IEEE Transactions on Communications. — 2020. — Vol. 68, No. 3. — P. 1581–1592. — DOI: 10.1109/TCOMM.2019.2961332.*

Надійшла до редакції: 10.04.26

Прийнята до друку: 12.06.26

Опубліковано: 30.06.26