

*М.А. ОМАРОВ, д-р техн. наук, Р.І. ЦЕХМІСТРО, канд. фіз.-мат. наук,
С.В. ШАПОВАЛОВ, канд. техн. наук, В.В. УСИК, канд. техн. наук,*

ОСОБЛИВОСТІ ВПЛИВУ РОЗПОДІЛУ АКУСТИЧНОГО ПОЛЯ КІЛЬЦЕВОЇ ДВОСЕКЦІЙНОЇ МІКРОФОННОЇ РЕШІТКИ З ЕЛЕКТРОННИМ УПРАВЛІННЯМ НА СТВОРЕННЯ АЛГОРИТМІВ ПОШУКУ І ЛОКАЛІЗАЦІЮ ЗВУКУ

Вступ

Запропоновано рекомендації, що дозволяють врахувати вплив параметрів розташування джерел звуку мікрофонними двосекційними ґратами з метою спрощення та уточнення розв'язання задачі на пошук, виявлення та локалізацію максимального джерела звуку. Проведено оцінку кореляції Пірсона між сусідніми джерелами (приймачами) звуку з урахуванням впливу між ними. Показано можливості вдосконалення існуючих алгоритмів пошуку максимуму звуку та його локалізації.

Особливості побудови та обрання алгоритму досліджень

Локалізація джерела звуку відіграє важливу роль у деяких методах просторової фільтрації, що використовуються у багатьох телекомунікаційних системах. Мовні сигнали, що надходять з різних напрямків, можуть заважати цільовому сигналу, але можуть маскувати його. Отже, необхідна обробка та розпізнавання сигналів з кожного мікрофона, що є нелегким програмно-апаратним завданням. Людський слух відомий як найефективніший детектор напряму прибуття звуків. Тому, щоб зробити алгоритми обробки сигналів ефективнішими, їх принцип роботи має відповідати деяким властивостям людського слуху. Незважаючи на великий розвиток науки в галузі людського сприйняття, питання, пов'язані з локалізацією звуку, остаточно не визнані, тому явища, що лежать у їх основі, як і раніше, є предметом інтенсивних досліджень [1–5].

Відомо, що алгоритми з нейро-грубого обчислювального підходу застосовуються до цифрового отримання аудіосигналів з урахуванням автоматичної локалізації джерел звуку за наявності шуму і паразитного ехо [5].

Зазвичай тестуються нейронні мережі для автоматичного визначення напрямку прибуття звуку, були також деякі експерименти з використанням грубих наборів. Результати роботи вказаних алгоритмів порівнюються з експериментальними даними. У ряді робіт розглянуто нейро-грубий підхід до вирішення цієї проблеми. Вихід кожного алгоритму, що тестується, повинен був надавати інформацію про напрям прибуття звуку. У разі нейро-грубого алгоритму результат його дії може бути доступний у вигляді слів, що визначають напрям прибуття звуку.

Відповідно до поточного стану знань, сприйняття спрямованості звуку бінауральною системою людини ґрунтується на наступних двох основних сутностях [5]. Міжмежова різниця рівнів (ILD), різниця інтенсивностей хвильових форм у лівому та правому вухах. Міжмежна різниця часу (ITD): різниця часу прибуття відповідних хвильових форм в обидва вуха, що еквівалентно різниці фаз хвильових форм. В області цифрової обробки сигналів визначення локалізації джерела звуку може здійснюватися за допомогою масиву мікрофонів, який може бути як лінійний, так і нелінійний [7, 8].

Більш складні фазовані мікрофонні решітки складаються з чотирьох і більше мікрофонів. Їхня кількість може досягати 64 і навіть 512 штук. Головне призначення фазованих мікрофонних решіток – створити потрібну діаграму спрямованості в заданому напрямку, оперативно змінювати цей напрям у просторі і тим самим збільшити співвідношення сигнал/шум від обраного акустичного джерела, на який спрямована фазована антена решітка. Велика

кількість мікрофонів значно впливає на вартість решітки, проте вона краще впорається із завданням локалізації звуку.

Для спрощення завдання обробки сигналів решітки ведуться дослідження, в роботах [7, 8] показано, як при меншій кількості включених мікрофонів можна отримати еквівалентний або більший рівень сигналу на конкретній відстані від решітки, тобто від джерела звуку.

Мета роботи – пошук можливих шляхів спрощення завдань інтелектуального отримання аудіосигналу за допомогою нейронних мереж, тобто автоматичне відстеження джерела звуку. Подібні дослідження застосовуються до деяких передових систем відеоконференцій, що використовують відеокамеру, яка автоматично повертається.

Використання двосекційної антенної решітки дає можливість краще визначити джерело звуку та економити апаратні ресурси в обробці сигналів. Різні дослідники досліджували численні методи локалізації джерела [3–6]. Більшість таких досліджень засновано на технології цифрової обробки сигналів і вимагають великих обчислювальних витрат. Існують деякі методи, що використовують просторову фільтрацію, кореляційну обробку і т.д.

Постановка задачі

У роботі [8] описано аналіз просторового розподілу амплітуди акустичного поля кільцевої мікрофонної решітки (рис. 1). Кожен мікрофон був представлений моделлю – випромінювачем сферичної хвилі (ізотропний випромінювач). Проведено аналіз впливу числа випромінювачів, радіусу ґрат на концентрацію акустичного поля в центрі ґрат і на довільній відстані. Складено алгоритм, що дозволяє враховувати довільне парне та непарне число випромінювачів, розташоване по довжині дуги при рівномірному та нерівномірному розташуванні [4, 7].

Схема системи реєстрації звуку представлена на рис. 1. У роботі [1] використовувалися 8 мікрофонів, розташованих симетрично на обідку діаметром 30 см. На рис. 2 зображено двухсекційну мікрофонну решітку, яка досліджувалась у [7].

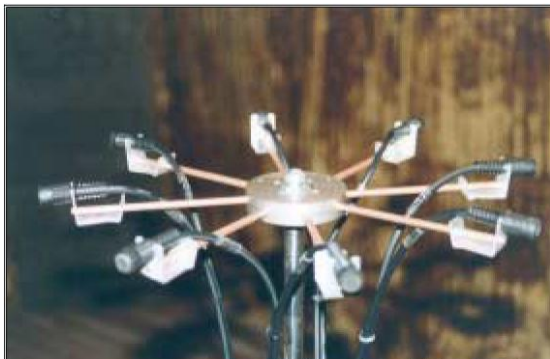


Рис. 1. Акустична система реєстрації звуку з 8 мікрофонів

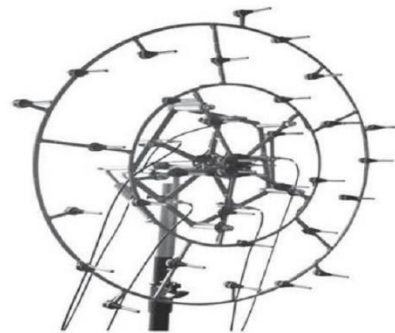


Рис. 2. Двосекційна мікрофонна решітка

Оцінка розташування джерела на основі обробки акустичних сигналів за допомогою мікрофонної решітки являє собою детерміноване завдання. Однак у реальних умовах виникають різні спотворення і перешкоди, такі як фоновий шум, відбиті звуки та інші. Отже, сигнали, що приймаються кільцевою мікрофонною решіткою, виражаються наступним співвідношенням [1]:

$$x_i(t) = \alpha_i U[t - (i-1) \cdot \tau]; \quad (1)$$

де α_i – коефіцієнт згасання i -го мікрофона, замість співвідношення $U(t-(i-1)\tau)$ (τ -часова затримка акустичної хвилі між сусідніми мікрофонами) використовуємо співвідношення [7], $n_i(t)$ – шум, створюваний i -м мікрофоном. Замість ґрат, представлених на рис. 1, ми розглянемо двосекційну решітку (рис. 2), яка має великі практичні можливості.

Замість сигналу $s(t)$ введемо на розгляд сигнал $U_n(t)$ аналогічно [7, 8]. Замість значення тимчасової затримки між сусідніми мікрофонами введемо на розгляд геометричний еквівалент R_n :

$$U_n(t) = \alpha_i h_i(t) U_0 \frac{\cos(\omega t - kR_n)}{R_n} + n_i(t); \quad (2)$$

$$\text{де } R_n = \sqrt{Z^2 + (Y_n \pm Y)^2 + (X_n \pm X)^2}; \quad (3)$$

X_n, Y_n – координати n -го випромінювача, X, Y – координати точки спостереження в декартових координатах, пов'язаних з ґратами, Z – відстань від площини решітки, як показано на рис. 2.

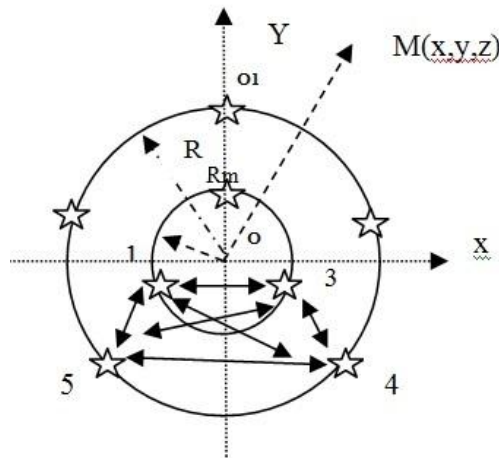


Рис. 3. Схема розташування мікрофонів у кільцевій двосекційній решітці (8 мікрофонів) з номерами окремих мікрофонів.

Розглянемо лінійні ґрати з $N+N_m$ мікрофонів, розташованих за великим і малим радіусом відповідно. Введемо декартову систему координат, як показано на рис. 3. Для знаходження амплітуди поля поблизу ґрат мікрофонів розташуємо їх уздовж осі OX , і відстань R_n до точки спостереження будемо визначати в декартових координатах аналогічно [8].

Основна частина

Облік алгоритму (1) ускладнюють завдання локалізації джерела звуку, і тому для її вирішення існує ряд різних методів [1, 2]. Більшість з них засновано на оцінці положення джерела звуку на основі сигналів, що приймаються мікрофонами в матриці, включаючи методи крос-кореляції [9], адаптивну фільтрацію [10] або обчислення відповідних векторів та матриць власних значень. У свою чергу, у разі відстеження або локалізації ряду джерел використовуються методи, засновані на максимальній правдоподібності. Більш детальну інформацію можна знайти у великій літературі з локалізації акустичних джерел мультимедійних додатків.

Таким чином, оцінка розташування джерела на основі обробки акустичних сигналів за допомогою мікрофонної решітки – це детерміноване завдання. Даний похід поклав основу задачі про пошук можливих шляхів спрощення задач інтелектуального отримання аудіосигналу з використанням нейронних мереж. Це автоматичне відстеження джерела звуку, яке застосовується до деяких передових систем відеоконференцій, що використовує автоматичне повертання відеокамери [1].

У [6, 7] сигнал заданий функцією сферичної хвилі (ізотропним випромінювачем). Ці роботи заслуговують на увагу, оскільки завдання інтелектуального розпізнавання складні як у теорії, так і на практиці, і будь-яке допустиме спрощення значно економить програмні та апаратні ресурси. Наприклад, у [8] показано, що на відстанях, що відповідають проміжній зоні, помітно меншим числом мікрофонів можна отримати еквівалентний рівень сигналу у двосекційній решітці у напрямку перпендикулярному центру решітки.

Даний результат дає підставу розглянути вплив параметрів решітки, врахованих в алгоритмі (1) – (3). Застосуємо вказаний метод до завдань визначення положення джерела звуку.

Більшість таких завдань засновано на оцінці положення джерела звуку на основі сигналів, що приймаються мікрофонами в матриці, включаючи методи крос-кореляції.

Максимум кореляції повинен відповідати змінам у часі сигналам між мікрофонами, що становлять інтерес в даний момент. Оскільки мовний сигнал може включати значні зміни енергії, максимумами кореляційної функції можуть відповідати пікам енергії.

Тому дана робота використовує аналогічний похід до оцінки кореляційних залежностей рівнів сигналів між сусідніми мікрофонами на фіксованих відстанях від ґрат при однаковому числі точок спостереження на чотирьох інтервалах $zd(0-5\lambda)$, $(0-10\lambda)$, $(0-15\lambda)$, $(0-20\lambda)$. Точки спостереження мали координати $(x=0.5, y=0, zd)$, $(x=2.5, y=1.5, zd)$,

З використанням формул (1)–(3) розроблено алгоритм та програму комп'ютерного аналізу, що дозволяє досліджувати просторовий розподіл звуку всередині решітки з урахуванням та без урахування особливостей полів ближньої та проміжної зони елемента решітки аналогічно [7–9].

Аналіз розрахунків кореляційних співвідношень з двохсекційним розміщенням мікрофонів

У [1, 2] параметри кореляції розраховуються для кожної пари мікрофонів в межах послідовних частотних піддіапазонів. Беручи до уваги даний факт, проведемо розрахунки кореляційних залежностей рівнів сигналів між мікрофонами, позначеними на рис. 3. В [1, 2] кореляційний аналіз виконується в межах октавних піддіапазонів, тобто автокореляційні залежності і залежності на підставі співвідношення Пірсона були розраховані на основі зміни частот мікрофонів і затримки сигналів.

Алгоритм, представлений за співвідношенням (1), дозволяє враховувати як тимчасові, так і частотні залежності. У моделі мікрофона у вигляді ізотропного випромінювача є можливість обліку часу та частоти $\cos(wt-kr)$. Однак не менш цікавим є вплив параметрів решітки на оцінки кореляційних співвідношень між сусідніми та протилежними мікрофонами у двосекційній еквідистантній решітці. У роботі [7] показано, що при меншому підсумковому числі мікрофонів можна досягти еквівалентного рівня сигналу, що приймається, на відстанях, відповідних проміжній зоні, особливо в точці рівновіддаленої від усіх мікрофонів.

Розрахунки кореляції Пірсона (4), наведені в табл. 1, проводилися для рівня сигналу для відстаней, відповідних трьом інтервалам zd : $(0-10\lambda)$, $(0-15\lambda)$, $(0-20\lambda)$, від площини решітки в точках з координатами $M(0.5, 0, zd)$ і $M(2.5\lambda, 1.5, zd)$ при масиві $Nd=20$ значень. Коефіцієнт Пірсона розраховувався за парними комбінаціями: $\text{corr}(U5, U4)$, $\text{corr}(U1, U3)$, $\text{corr}(U5, U3)$, $\text{corr}(U1, U4)$, $\text{corr}(U5, U1)$, $\text{corr}(U4, U3)$. Вибрані вихідні параметри дозволяють схилити рівновіддаленість точки спостереження від вибраних мікрофонів, що важливо для обчислювального експерименту. Як видно з рис. 3, три мікрофони розташовані на секції з малим діаметром, та 5 мікрофонів – на секції з великим радіусом при співвідношеннях $R=2R_m$, радіус великого кола становить дві довжини хвилі.

На рис. 4–6 подано кореляційні поля між сусідніми мікрофонами відповідно до схеми, яка представлена на рис. 1. Розрахунки U_n представлені відповідно до алгоритму 1–5. Параметр zd – відстань точки спостереження, тобто джерела звуку змінювалося в інтервалі 0–5 довжин хвиль із кроком 0.25λ ($Nd=20$). На відміну від 8-мікрофонних решіток (рис. 1), де 5 мікрофонів розташовано на великій дузі і 3 – на малій, у ґратах із 6 мікрофонів троє пере-

бувають у малій і трое – у великий дузі, тобто кут між мікрофонами складає 60 градусів. Радіус великої дуги становить дві довжини хвилі, а малої – довжину хвилі. На цих рисунках координати точки спостереження ($x = 0,5\lambda$, $y = 0$, $z = zd$).

Формула для розрахунків, результати яких наведено у табл. 1:

$$\text{corr}(U1, U2) = \frac{\sum_{i=0}^{Nd} (U1 - \bar{U}1)(U2 - \bar{U}2)}{\sqrt{\sum_{i=0}^{Nd} (U2 - \bar{U}2)^2} \sqrt{\sum_{i=0}^{Nd} (U1 - \bar{U}1)^2}} \quad (4)$$

Таблиця 1

Коефіцієнти кореляції для решітки N=6 (3 – по малому радіусу, 3 – по максимальному) M ($x=0.5, y=0, zd$).				
Інтервал розрахунків (0-zd)	0-5λ	0-10λ	0-15λ	0-20λ
corr(U5,U4)	-0.07	-0.033	0.01	0.039
corr(U1,U3)	-0.097	-0.023	0.019	0.05
corr(U5,U3)	0.03	-0.008	0.014	0.038
corr(U1,U4)	0.872	0.879	0.883	0.086
corr(U5,U1)	0.279	0.266	0.29	0.309
corr(U4,U3)	0.013	0.072	0.11	0.137
corr(U5,U01)	-0.07	-0.033	0.096	0.039
Коефіцієнти кореляції для решітки N=6 (3 – по малому радіусу, 3 – по максимальному) M ($x=2.5, y=1.5, zd$).				
Інтервал розрахунків (0-zd)	0-5λ	0-10λ	0-15λ	0-20λ
corr(U5,U4)	-0.11	0.181	0.037	-0.11
corr(U1,U3)	0.14	-0.071	-0.115	-0.018
corr(U5,U3)	-0.079	0.011	-0.026	-0.071
corr(U1,U4)	0.245	0.006	-0.041	-0.044
corr(U5,U1)	0.645	0.295	0.236	0.233
corr(U4,U3)	0.985	0.987	0.987	0.988
corr(U5,U01)	-0.011	0.081	0.037	-0.011

Далі розглянемо властивості коефіцієнта кореляції як вимірювача ступеня лінійності стохастичної залежності (рис. 4–6).

Кореляційне співвідношення (4) може мати як позитивне, так і негативне значення. Знак + означає, що зі зростанням значень U збільшується умовне математичне очікування (розглянемо Y і X замість U2 і U1). Негативне значення свідчить про протилежну тенденцію. Якщо кореляційна залежність між двома величинами відсутня, то $\text{corr}_{xy}=0$. Дійсно, при відсутності кореляційної залежності величини Y від X відсутня. Проте зворотне твердження, яке завжди правильне, – з рівності нулю генерального коефіцієнта кореляції, який завжди впливає у разі, коли функція регресії лінійна. Коли $\text{corr}_{xy}=0$, відсутність кореляційної залежності Y від X. За умови $\text{corr}_{xy}=1$ спостерігається лінійна функціональна залежність. Чим ближче corr_{xy} до 1, тим ближче стохастична залежність між величинами, що корелюються (ближче до лінійної). Чим вищий рівень лінійної стохастичної залежності X і Y, тим corr_{xy} ближче до 1.

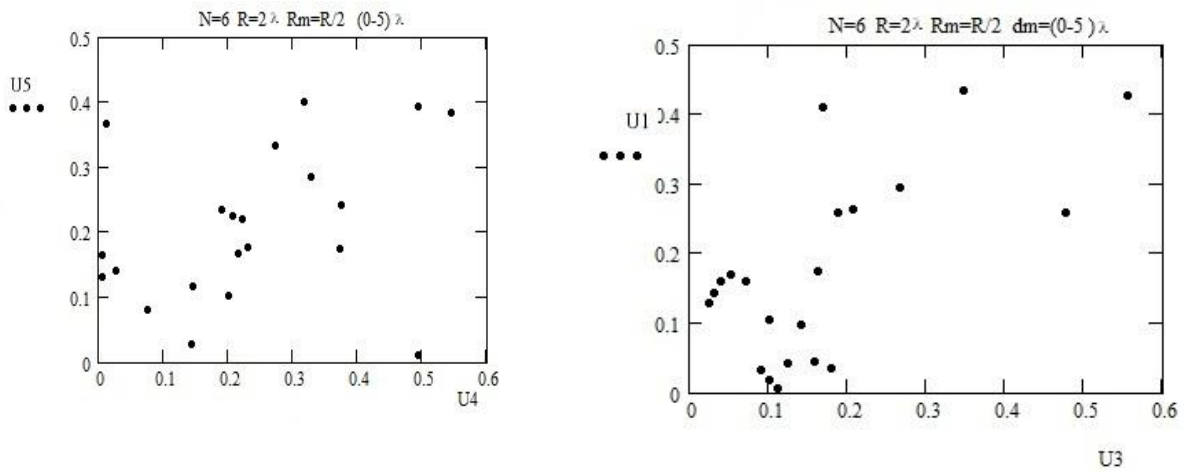


Рис. 4. Кореляційні поля для решітки з 6 мікрофонів (3+3) між мікрофонами 5-4 і 1-3 відповідно до схеми, представленої на рис. 3

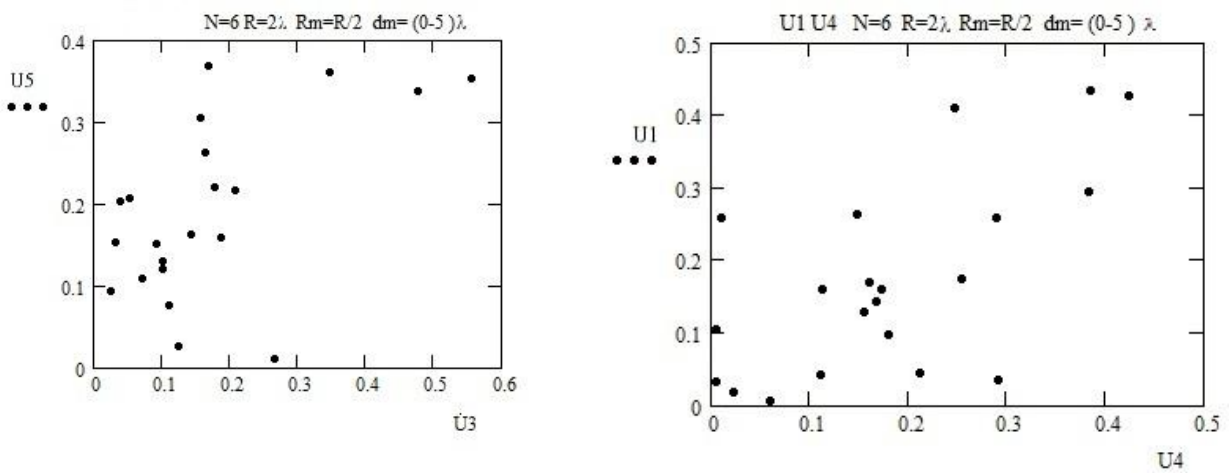


Рис. 5. Кореляційні поля для ґрат з 6 мікрофонів (3+3) між мікрофонами 5-3 і 1-4 відповідно до схеми, представленої на рис. 3

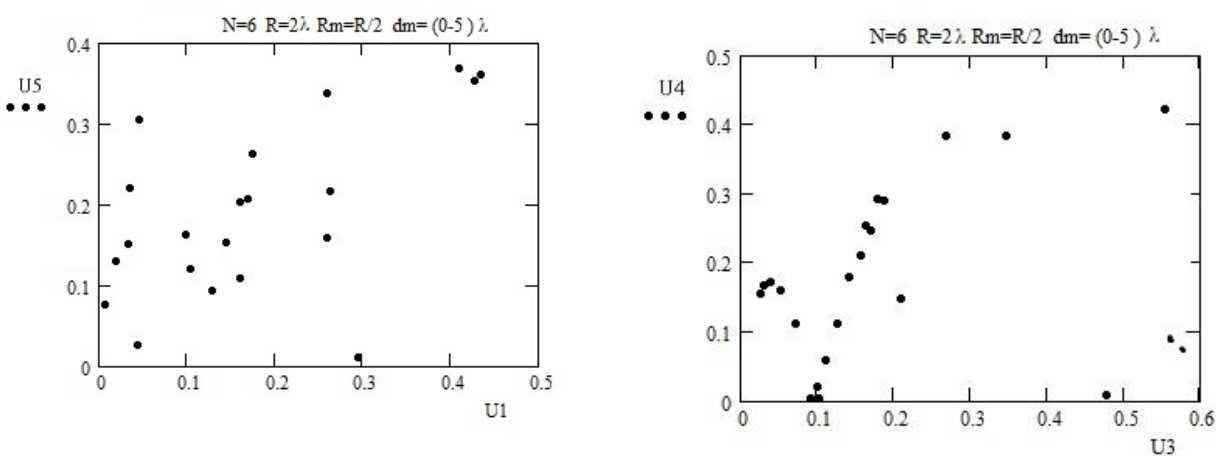


Рис. 6. Кореляційні поля для решітки з 6 мікрофонів(3+3) між мікрофонами 5-1 і 4-3 відповідно до схеми, представленої на рис. 3

Висновки

Аналіз кореляційних залежностей показав наявність впливу сигналів мікрофонів, який узгоджується із залежностями, побудованими на експериментальних залежностях, наведених у [1]. Кореляційні поля та табличні результати також вказують на можливість отримання прийняттого рівня звуку при меншій кількості задіяних мікрофонів. Аналіз кореляційних залежностей для різного числа мікрофонів на масиві дистанцій, відповідних середній та дальній зонам, сприяють пошуку удосконалення алгоритмів нейромереж. Кореляційний аналіз має перевагу порівняно з виробленням рекомендацій щодо вдосконалення аналізу на основі врахування взаємного впливу звукового поля мікрофонів.

Список літератури:

1. Andrzej Czyzewski. Intelligent Acquisition of Audio Signal Employing Neural Network and Rough Set Algorithms // Lecture Notes in Computer Science Technical University of Gdansk, Sound & Vision Engineering Department Narutowicza 11/12. Gdansk, Poland, 1999. P.1–22.
2. Andrzej Czyzewski. Automatic identification of sound source position employing neural networks and rough sets // Sound and Vision Engineering Department, Technical University of Gdansk, ul. Narutowicza, 80-952 Gdansk, Poland, Pattern Recognition Letters 24, 2003. P. 921–933.
3. Bodden M. Modeling Human Sound-Source Localization and the Cocktail-Party-Effect // Acta Acustica. 1993. № 1. P. 43–55.
4. Czyzewski A., Krolikowski R. Neuro-Rough Control of Masking Thresholds for Audio Signal Enhancement // Neurocomputing. 2001. № 36. P. 5–27.
5. Czerniawski J. The verification of new algorithms for identifying sound source position and sound acquisition methods based on spatial filtering. M.Sc. Diploma Thesis. Technical University of Gdansk, Gdansk, Poland, 2001. (in Polish).
6. Czyzewski A., Krolikowski R. Neuro-Rough Control of Masking Thresholds for Audio Signal Enhancement // Neurocomputing. 2001. № 36. P. 5–27.
7. Omarov M., Tsekhmistro R., Shapovalov S. Analysis of Acoustic Field Distribution of Circular Microphone Array in Free Space // Proceedings 16th International Conference on Advanced Trends in Radioelectronics // Telecommunications and Computer Engineering, TCSET 2022. P. 134–138.
8. Gorobets N.N., Gorobets Yu.N., Tsekhmistro R.I. Near-Field Effects in the Electromagnetic Power Distribution in the Vicinity of an Lattice of Hertzian Dipoles // Telecommunications and Radioengineering. 1999. Vol. 53, № 3. P.24–29.
9. Цехмістро Р.І. Особливості випромінювання телекомунікаційних імпедансних дровових антен // Радіотехніка. 2004. № 139. С. 28–32.
10. Bolotina I.O., Kroening H.-M., Kvasnikov K.G., Sednev D.A., Sumtsova O.V. Acoustic Field Simulation of an Antenna Array at Scanning by the SPA Method for Modern Ultrasonic Testing Technologies // J. Advanced Materials Research. 2014. Vol. 1040. P. 959–964.

Надійшла до редколегії 17.01.2026

Прийнята до друку після рецензування 23.04.2026

Публікація (оприлюднення) 30.04.2026

Відомості про авторів:

Омаров Мурад Анверович – д-р техн. наук, професор, Харківський національний університет радіоелектроніки, проректор з науково-педагогічної роботи, Україна; e-mail: murad.omarov@nure.ua; ORCID: <https://orcid.org/0000-0003-4842-4972>

Цехмістро Роман Іванович – канд. фіз.-мат. наук, Національний технічний університет "ХПІ", доцент кафедри мультимедійних та інтернет-технологій і систем, Україна; e-mail: tsekhmistorroman@gmail.com; ORCID: <https://orcid.org/0000-0003-3628-3658>

Шаповалов Сергій Вікторович – канд. техн. наук, Харківський національний університет радіоелектроніки, доцент кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; e-mail: serhii.shapovalov@nure.ua; ORCID: <https://orcid.org/0000-0003-0066-5291>

Усик Вікторія Валерівна – канд. техн. наук, Харківський національний університет радіоелектроніки, професор кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; e-mail: Victoria.Usik@nure.ua; ORCID: <https://orcid.org/0000-0002-3515-4849>