

Визначення області стереофонічного звучання джерел інформаційних сигналів

Вдовенко М. В., ORCID [0000-0003-1666-3389](https://orcid.org/0000-0003-1666-3389)

e-mail mv.vdovenko@aae.kpi.ua

Луньова С. А., к.ф.-м.н. доц., ORCID [0000-0003-0683-1211](https://orcid.org/0000-0003-0683-1211)

e-mail svetlana_lunyova@yahoo.com

Кафедра акустики та акустoeлектроніки acoustic.kpi.ua

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського» kpi.ua

Київ, Україна

Анотація—В даній статті розглядається проблема локалізації людиною джерел звуку. Запропонована формула для уточнення розрахунків зони стереоефекту для сигналів з мовленнєвою компонентою. Результати, одержані за допомогою аналітичних розрахунків, перевіряються експериментальним шляхом.

В якості критеріїв оцінки стереофонічного звучання використовувались параметри функції міжвушної кореляції: коефіцієнт крос-кореляції, інтервал кореляції, ширина головного максимуму та загальний вигляд функції кореляції. Вимірювання виконувались за допомогою штучної голови на глядацьких місцях в залі середніх розмірів.

Практично наявність зони стереофонії для інформаційних сигналів можна визначати за значеннями коефіцієнту крос-кореляції $IACF > 0,2$. Для різних типів сигналів це значення відповідає різному часу міжвушної затримки сигналів, що зумовлює різні розміри області стереофонії для музичних і мовленнєвих сигналів.

Бібл. 14, рис. 3, табл. 3.

Ключові слова — бінауральний ефект; локалізація звуку; уявне джерело звуку; крос-кореляційна функція; міжвушна кореляція; коефіцієнт міжвушної кореляції; бінауральні моделі; локалізаційні моделі; зона стереофонії; стереофонічний ефект.

I. ВСТУП

Виникнення стереофонічного ефекту, або, так званої сумарної локалізації звуків пов'язане з особливостями просторового слуху людини, а саме, зі здатністю створення уявних джерел звуку та локалізацією їх у просторі.

Якщо слухач знаходиться на осі симетрії системи двох однакових джерел звуку, то між сигналами, що надходять до правого і лівого вуха слухача, відсутні часова затримка (ITD — interaural arrival-time differences) та різниця звукових тисків (ILD — interaural level differences). При цьому сигнали об'єднуються в спільний звуковий образ, який локалізується посередині бази джерел звуку (так зване "уявне джерело звуку" або УДЗ) [1].

Коли сигнал приходить від системи джерел, зміщеної відносно центра голови, то звук надходить в одне вухо раніше. Це дозволяє мозку визначати надходження звукового сигналу зліва чи справа і приблизно оцінювати кут розташування джерела відносно медіальної площини голови. Чисельна різниця в часі надходження сигналу до правого та лівого вуха, що складає ± 1 мс, зміщує УДЗ в бік вуха, що сприймає сигнал раніше.

Отже, розміри області існування і локалізації УДЗ обмежені у просторі. При цьому межі цієї області

визначають, виходячи з умови, що УДЗ поєднується з крайнім джерелом звуку (тобто від центру переміщується на половину розміру бази двох джерел).

Оскільки ця характеристика просторового слуху є перш за все суб'єктивною, вона є недостатньо вивченою, незважаючи на значну кількість присвячених цьому питанню робіт [1]–[10]. Проаналізувавши відомі на сьогодні об'єктивні параметри, що впливають на сприйняття стереозвучання, ми дійшли висновку, що вони не є вичерпними і не завжди співпадають з суб'єктивною оцінкою людини.

II. АНАЛІТИЧНИЙ ОГЛЯД ЛІТЕРАТУРИ ТА ПОСТАНОВКА ПРОБЛЕМИ

Для моделювання алгоритму локалізації джерела звуку використовують бінауральні моделі слухового апарату [1], [2], [4]–[9], які класифікують за наступними ознаками:

а) за складністю ситуацій, у яких вони використовуються:

- локалізація одного джерела в безлуновій камері;
- локалізація джерела в контексті декількох конкуруючих джерел звуку за наявності реверберації [3];



б) в залежності від аспектів просторових сцен, які вони досліджують [2]:

- просторове направлення і властивості приміщення;
- відстань до уявного джерела;
- охоплення слухача;
- виявлення джерела або руху слухача.

Також розрізняють моделювання, що зосереджується або на копіюванні людського сприйняття (базується на цифровій обробці сигналу) [4], або на фізіології та модульних моделях периферійної слухової системи [5].

В математичних моделях локалізації звукового джерела використовують сигнали:

- бінауральні (інтерауральні), що потребують локалізації двома вухами;
- монофонічні, які можна вилучити лише одним вухом.

Бінауральні сигнали, як правило, більш надійні в сенсі локалізації джерела звуку. Проте вони зменшують можливість визначення напрямку фронт-тил та кута у вертикальній площині.

Бінауральні сигнали зазвичай характеризують часовою затримкою ITD та різницею рівнів ILD.

Першою спробою використання ITD була модель Джеферса [1]. В цій моделі вуха працюють окремо, і для кожного вводиться своя лінія затримки. Сигнали, що надходять до різних вух, поширюються в протилежних напрямках і в деякий момент часу перетинаються на ділянці збігу. В результаті, обидва сигнали потрапляють у зміщену в боковому напрямку ділянку відносно медіальної площини голови, положення якої відповідає боковому положенню джерела.

На сьогодні, появу УДЗ пояснюють тим, що звукові сигнали, які надходять на різні вуха, статистично пов'язані. З ослабленням цього зв'язку УДЗ спочатку гірше локалізується, а далі розпадається на два незалежних джерела звуку.

В ряді сучасних моделей локалізації джерела звуку для аналізу використовують крос-кореляційну функцію (функцію міжвушної кореляції) або метод ІАСС (interaural cross-correlation), що початково був введений Чері та Сеерсом [6], як міра подібності двох ідентичних сигналів в залежності від зміщення одного відносно іншого. В теорії обробки сигналів така функція відома як взаємно-кореляційна. Така функція широко застосовується в електронній томографії, криптоаналізі та нейрофізіології.

В задачах локалізації вона визначається за результатами вимірювань імпульсних характеристик приміщення на лівому і правому вусі $p_l(t)$ та $p_r(t)$ в різних точках приміщення¹ і оцінюється як кореляційна

функція нормована на взаємну енергію відповідних сигналів:

$$IACF(\tau) = \frac{\int_{-\infty}^{\infty} p_l(t) p_r(t + \tau) dt}{\sqrt{\int_{-\infty}^{\infty} p_l^2(t) dt \cdot \int_{-\infty}^{\infty} p_r^2(t) dt}}, \quad (1)$$

де τ – часова затримка, $p_l(t)$ та $p_r(t)$ – звуковий тиск на лівому та правому вухах.

На сьогодні є достатньо поширеним використання бінаурального запису для оцінки можливості локалізації джерела звуку людиною, а саме, визначення розмірів області, в якій діє ефект сумарної локалізації.

Традиційно, розрахунок зони стереофонії проводиться, виходячи з припущення, що час запізнення сигналів, які надходять на різні вуха, складає ± 1 мс [2], [11], [12].

Така думка пов'язана, по-перше, з тим, що цей час затримки відповідає знаходженню джерела звуку в боковому напрямі слухової системи. По-друге, вважаємо, що такий підхід виник, водночас, під впливом «ефекту передування» або ефекту Хааса, який відтворює один з механізмів локалізації звуку в приміщенні [10]. Це явище зводиться до наступного: якщо звуки надходять з різних напрямків з короткими інтервалами затримки, то локалізація відбувається за правилом “першої хвилі”, тобто за першим звуком, що надійшов. Якщо інтервал між двома короткими звуками стає меншим за 1 мс, то цей ефект не виявляється і відбувається так звана “сумарна локалізація”. Але також відомо, що для коротких імпульсних сигналів цей інтервал сумарної локалізації знаходиться в межах 5 мс. Для мовних та музичних сигналів інтервал часової затримки для сумарної локалізації ще більший, і за даними різних авторів складає від 40 мс до 60 мс [2], [11], [12]. Крім того, за результатами досліджень, виконаних Ю. А. Ковалгінім [11], при часі затримки сигналів в межах $0,8 \div 1,2$ мс переміщення УДЗ залежить від типу сигналу. Визначити область існування УДЗ можливо тільки для джерел, які не мають виражених нерівномірностей розподілення енергетичного спектру (наприклад, арфа, рояль, кастаньети, труба).

Для сигналів, спектри потужності яких мають енергетичні піки, переміщення УДЗ носить індивідуальний характер (жіночий голос, флейта, скрипка), точно розрахувати межі області стереофонічного звучання неможливо, оскільки зміщення УДЗ в залежності від часової затримки носить індивідуальний характер.

Величина коливання місцезнаходження УДЗ стає незначною тільки при $\tau = 7$ мс. При цьому час затримки УДЗ практично досягає положення крайнього джерела звуку, незалежно від типу сигналу. Подальша поведінка УДЗ залишається невизначеною, але остаточне розпадання його на два джерела звуку настає при значно більшому часі затримки $\tau > 30$ мс.

¹ ISO 3382-1:2009 “Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces”



Такі дані одержані на основі суб'єктивних оцінок експериментальних випробувань для приміщень малого об'єму (до 300 м^3) і малих розмірів бази стереосистеми (до 2,8 м) [11].

Таким чином, виходячи з огляду наведених результатів досліджень, постають питання:

- за якими об'єктивними критеріями можна оцінювати зону існування УДЗ в умовах реального приміщення;
- як залежить можливість локалізації звукового джерела при стереовідтворенні сигналів більш складних, ніж імпульсний, а саме, мовних і музичних інформаційних сигналів.

Метою дослідження є удосконалення аналітичного розрахунку області стереофонічного звучання для різних типів сигналів та експериментальна перевірка одержаних результатів за допомогою об'єктивних показників, зокрема, параметрів функції міжвушної кореляції.

III. МЕТОДИКА ПРОВЕДЕННЯ РОЗРАХУНКІВ ТА ЕКСПЕРИМЕНТУ

Для розрахунку області стереофонічного звучання складних сигналів авторами запропонована формула знаходження меж зони, виходячи з даних роботи [11]. Формула враховує відносне розташування слухача і стереопари. Час затримки сигналів, які надходять на ліве і праве вухо слухача, знаходиться через різницю відстаней від джерел звуку до відповідного вуха. Для спрощення перехресні шляхи надходження звуку не

враховуються. Для визначення крайнього положення УДЗ приймається час міжвушної затримки таким, що дорівнює 7 мс [11].

Таким чином, розрахована область стереофонічного звучання як функція віддаленості слухача від лінії бази стереопари обмежується координатою $x=S$ по горизонталі, де значення $x=0$ відповідає центру бази:

$$S = 7b \sqrt{1 + \frac{y^2}{\left(\frac{B-b}{2}\right)^2 - 49b^2}}, \quad (2)$$

де S – положення уявного джерела відносно осі x , м; y – координата приміщення, перпендикулярна базі стереопари; b – міжвушна база ($b=0,18\text{ м}$); B – ширина бази між джерелами звуку в стереопарі, м.

Експериментальний аналіз характеристик просторового слуху людини виконувався на основі функцій міжвушної кореляції (1), одержаних шляхом обробки в середовищі MATLAB® звукових сигналів, записаних за допомогою двох мікрофонів, вмонтованих у слухові проходи лівого та правого вуха штучної голови (рис. 1).

Експеримент проводився в залі з розмірами $14,3 \times 18,5 \times 6,25$ м, об'ємом $V \approx 1653\text{ м}^3$ (конференц-зал факультету електроніки Київського політехнічного інституту ім. Ігоря Сікорського), який відповідає залу середніх розмірів².

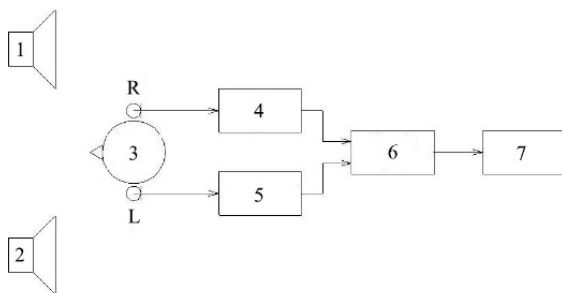


Рис. 1 Схема підключення стенду для досліджень: 1, 2 – гучномовці Radiotekhnika S-30B, розташовані на відстані 4,5 м і на висоті 1,5 м; 3 – манекен голови (R, L – вимірювальні мікрофони Behringer ECM8000); 4, 5 – підсилювачі потужності ParkAudio V4-900 MkII; 6 – звукова карта M-Audio Fast Track Pro; 7 – персональний комп'ютер.

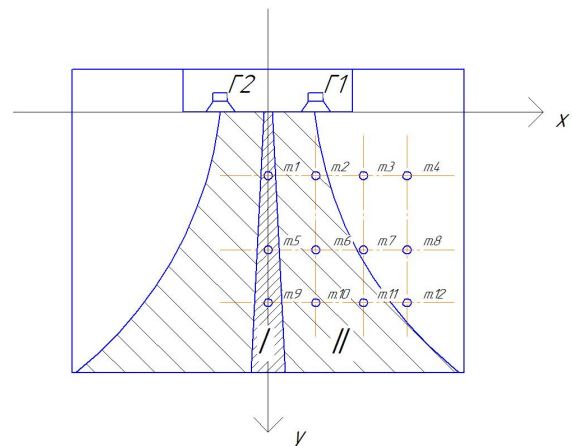


Рис. 2 Схема розташування точок вимірювань у залі: I – область стереофонії, розрахована за рекомендаціями [12]; II – область стереофонічного сприйняття звуку для складних сигналів [11]

² ISO 3382-1:2009 "Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces"

Таблиця 1 Показники статистичного зв'язку бінауральних імпульсних сигналів у відповідних точках вимірювань

№ точки вимірювань	1	2	3	4	5	6	7	8	9	10	11	12
Максимальна часова затримка, мс	0,3	7,3	11	13	0,2	4,3	7,6	9,9	0,3	2	6	8,3
IACC	0,88	0,8	0,82	0,8	0,84	0,78	0,63	0,83	0,8	0,72	0,69	0,78
Інтервал кореляції, мс	1,8	1,6	1,8	1,7	1,8	2	1,6	1,9	1,8	2	1,7	1,6
Ширина максимуму на рівні (-3)дБ, мс	0,7	0,5	0,9	1	0,7	0,5	0,8	0,9	0,5	0,8	0,8	0,8

Штучна голова почергово розташовувалась в точках звукового поля, що відповідають розміщенню слухачів в кріслах на рівні їх голів (висота над рівнем підлоги $h \approx 1,2$ м). Схема відповідного розміщення гучномовців та слухачів у залі наведена на рис. 2.

Ширина бази стереопари – 4,5м. Відстань від лінії бази до першого ряду точок вимірювання - 3м; відповідно, до другого ряду – 6,5м і до третього - 9м Відстані між точками вимірювання по горизонталі: 0; 2,25; 4,5; 7м.

Ділянка I на рис.2, що позначена як зона стереофонії, розрахована за рекомендаціями [12], виходячи з положення, що різниця в часі надходження сигналів до лівого та правого вуха не повинна перевищувати ± 1 мс. Розміри області II обчислені за формулою (2).

Вимірювання проводились для наступних видів звукових сигналів: прямокутний імпульс тривалістю 5 мс; фрагментів мовленнєвого повідомлення, хорошого співу та симфонічної музики — 20 с. В якості мовленнєвого повідомлення використано фрагмент роману “Майстер та Маргарита” Михайла Булгакова; хорошого співу — пісня «Венеціанська ніч» (музика Михайла Глінки, слова Івана Козлова); симфонічної музики — запис симфонії №9 Людвіга Ван Бетховена.

Експериментальна оцінка області стереофонічного сприйняття звуку в приміщенні виконувалась на основі таких характеристичних параметрів функції крос-кореляції, як коефіцієнт $IACC$, що відповідає максимальному значенню функції крос-кореляції, та інтервал кореляції. Інтервал кореляції визначився, як ширина центрального піку функції крос-кореляції за розмірами першого перетину нульової лінії. Чим менший інтервал кореляції (тобто швидше спадає кореляційна функція), тим слабший зв'язок між двома сигналами.

Додатково оцінювалась ширина головного максимуму $IACF_{max}$, виміряного на рівні $0,707 \cdot IACF_{max}$ або -3 дБ, що відповідає половині енергії взаємодії сигналів, які надходять у ліве та праве вуха слухача. За результатами досліджень, виконаних у роботі [14], ця величина надає додаткову інформацію щодо якості звучання: чим більша ширина піку функції кореляції, тим більше просторовість звучання. Оскільки відчуття просторовості пов'язане зі ступенем несхожості сигналів, то зміна ширини максимуму функції разом зі значенням коефіцієнту кореляції характеризує ступінь локалізації слухачем уявного джерела.

IV. РЕЗУЛЬТАТИ ТА ЇХ ОБГОВОРЕННЯ

A. Експеримент 1. Імпульсний сигнал.

Попередньо, для одержання додаткових даних щодо оцінки області стереофонії у залі розраховано часові затримки між сигналами, що надходять на ліве та праве вуха слухача від гучномовців 1 та 2 для вказаних на схемі (рис. 1) місць слухачів. З урахуванням перехресних шляхів надходження звуку, значення максимального часу затримки для відповідних точок приміщення наведені у табл. 1.

З числових даних випливає, що час міжвушної затримки (табл. 1) у точках вимірювань змінюється від 0,2 мс до 13 мс.

Функція міжвушної крос-кореляції, обчислена за алгоритмом, вказаним у формулі (1), представлена у табл. 1 відповідними числовими показниками (а саме: коефіцієнтом кореляції, інтервалом кореляції та шириною максимуму функції на рівні 0,707 (або -3 дБ).

Коефіцієнт кореляції бінауральної пари сигналів в межах області I (т.1,5,9) має значення 0,88-0,8.

У всіх точках області I інтервал кореляції дорівнює 1,8 мс.

Ширина піку функції крос-кореляції на рівні -3 дБ від її максимального значення для зони I становить $0,7 \div 0,5$ мс, тобто відрізняється несуттєво. Незначна зміна ширини піку функції крос-кореляції відповідає стабільному відчуттю звучання і впевненій локалізації УДЗ.

В області стереофонії II до області I додаються точки 2, 6, 10, 11, в яких коефіцієнт кореляції змінюється в межах 0,89-0,69; інтервал кореляції розширюється до 2 мс, а ширина центрального піку функції кореляції на рівні -3 дБ збільшується до значення 0,8 мс.

Такі зміни параметрів є цілком закономірними, враховуючи ослаблення взаємозв'язку сигналів, що надходять на ліве та праве вуха. Проте, проаналізовані характеристики за своїми значеннями мало відрізняються від показників області I.

Аналіз бінауральних характеристик, виміряних в решті точок залу, що не входять до області II (точки 3, 4, 7, 8, 12), дає наступні результати: коефіцієнт кореляції знижується до значення 0,63; інтервал кореляції знаходиться в тих самих межах, ширина піку на рівні -3 дБ досягає 1 мс.



За суб'єктивною оцінкою якості звучання в усьому приміщенні спостерігається зона повної або часткової стерефонії (тобто уявне джерело не розпадається на два джерела звуку. Максимальна міжвушна часова затримка при цьому складає 13 мс.

Таким чином, для імпульсних сигналів стереофонічна зона не обмежена ні часом міжвушної затримки ± 1 мс, ні, навіть, часом затримки у 7 мс, а є значно ширшою. Показником наявності стереозони можуть слугувати значення коефіцієнту функції крос-кореляції, оскільки цей параметр функції змінюється найбільш суттєво в залежності від положення слухача.

За даними Ю.А.Ковалгіна [11], локалізація УДЗ можлива, поки час міжвушної затримки не перевищує 15 мс, якщо різниця рівнів відповідних імпульсних сигналів не перевищує 10 дБ. Для вказаних параметрів виміряні значення коефіцієнту міжвушної кореляції становлять $IACC = 0,6$. Виходячи з цього, значення $IACC = 0,6$ може бути цілком достатнім для задовільного прослуховування стереозвучання.

В. Експеримент 2. Мовленнєве повідомлення, хоровий спів та симфонічна музика.

Запис звукових сигналів виконувався в точках 1, 4, 12 (рис. 1). Графіки одержаних функцій крос-кореляції сигналів, що надходили на ліве та праве вухо, представлені на рис. 3.

Основні показники функції міжвушної кореляції (коефіцієнт кореляції та інтервал кореляції) для досліджуваних типів сигналів вказані в таблиці 2.

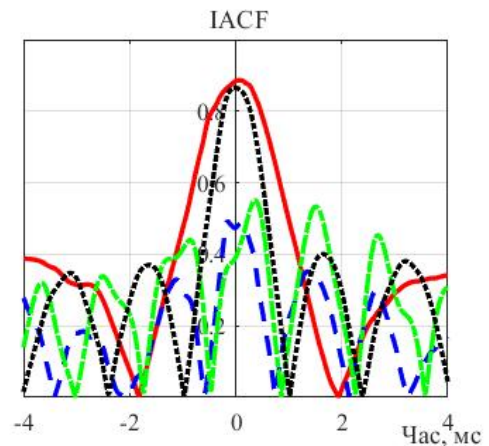
Як випливає з графіків, зображених на рис. 3а (для точки 1, яка знаходиться в зоні чіткої стерефонії) найбільший коефіцієнт кореляції 0,88 має імпульсний сигнал. До нього наближається відповідний показник симфонічної музики — коефіцієнт кореляції 0,86. Що стосується хорового співу і мовленнєвого повідомлення, то коефіцієнт кореляції спадає, відповідно, до значень 0,55 і 0,49. При цьому, інтервал кореляції зменшується від 1,8 мс для імпульсного сигналу до 0,7 мс для мовленнєвого та хорового. Такі дані свідчать про суттєве зменшення взаємозв'язку бінауральних сигналів, що мають мовленнєвий характер. Але за суб'єктивною оцінкою розбірливості мови цю зону можна вважати задовільною.

Графіки, представлені на рис. 3б,в, вказують на подальше зниження статистичного зв'язку між бінауральними сигналами з мовленнєвою складовою (мова та хоровий спів). Коефіцієнт кореляції в т. 4 зменшується до значень 0,22...0,18, а в т. 12 знаходиться в межах 0,32...0,22; причому інтервал кореляції спадає до значень 0,4...0,2 мс. При цих значеннях коефіцієнта кореляції мова стає вкрай нерозбірливою.

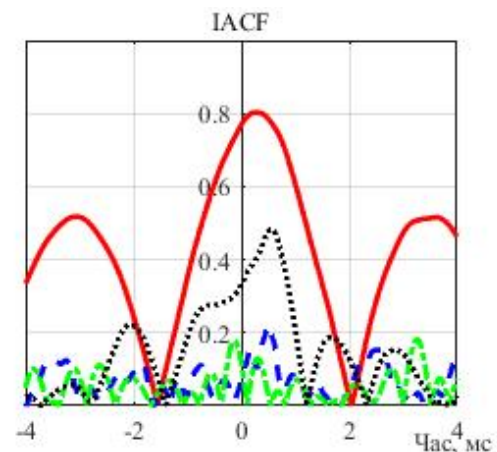
Крос-кореляційні характеристики сприйняття симфонічної музики мають показники, близькі до імпульсного сигналу. Коефіцієнт міжвушної кореляції $\geq 0,48$.

Отже, водночас з тим, як при поширенні області І для імпульсного сигналу та для симфонічної музики все ще утворюється стереополе з достатньо чіткою локалізацією уявного джерела, якість сприйняття

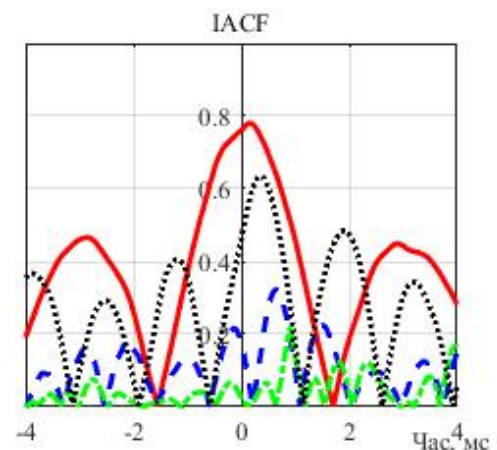
сигналів з мовленнєвою компонентою значно погіршується.



а



б



в

Рис. 3 Графіки функції міжвушної крос-кореляції, виміряні на відповідних слухачьких місцях: а – т.1; б – т.4; в – т.12; суцільна лінія — імпульсний сигнал; пунктирна — мова; штрихпунктирна — хоровий спів; точкова — симфонічний оркестр

Таблиця 2 Значення числових показників функції міжвушної кореляції в точках вимірювання для різних типів сигналів

Точки вимірювань	1		4		12	
	<i>IACC</i>	<i>Інтервал кореляції, мс</i>	<i>IACC</i>	<i>Інтервал кореляції, мс</i>	<i>IACC</i>	<i>Інтервал кореляції, мс</i>
Характеристики функції міжвушної кореляції						
Імпульсний сигнал	0,88	1,8	0,8	1,7	0,78	1,6
Симфонічна музика	0,86	1	0,48	1,3	0,63	1
Хоровий спів	0,55	0,7	0,18	0,3	0,2	0,2
Мовленнєвий сигнал	0,49	0,7	0,2	0,4	0,32	0,7

Таблиця 3 Зміщення інтервалу кореляції в різних точках приміщення для різних типів сигналів, мс

Тип сигналу	Точка 1	Точка 4	Точка 12
Імпульс	0,1	0,4	0,2
Мова	0,08	0,7	-
Хоровий спів	0,18	0,9	-
Симфонічний оркестр	0,05	-	0,3

Такий ефект пояснюється тим, що в мовному сигналі окремі склади в своєму поєднанні несуть певний інформаційний зміст. При поширенні звуку в приміщенні звучання нових складів накладається на звучання попередніх, маскуючи їх і погіршуючи чіткість і розбірливість мовлення. Значна часова міжвушна затримка призводить, по суті, до дихотичного сприйняття звуку, коли на різні вуха надходить різна інформація. За даними Хааса [7] при міжвушній затримці мовного сигналу більше, ніж на 5 мс, вже спостерігається ефект блукання віртуального джерела вздовж бази.

За даними експерименту, виконаного в роботі, при збільшенні часу затримки мовленнєвих сигналів до $t > 7$ мс уявне джерело звуку взагалі знаходиться на межі розпадання, що суб'єктивно підтверджується дуже поганою розбірливістю звучання. Цей факт відповідає висновкам [11], що УДЗ розпадається при значеннях коефіцієнту кореляції 0,2...0,05.

Для додаткового аналізу зазначимо, що точка 1 знаходиться праворуч відносно вертикальної осі симетрії приміщення і відповідає першому кріслу від середини першого ряду. Таке положення місця запису сигналів фіксується зміщенням максимуму функції кореляції відносно осі ординат (рис. 3а). Проте, більш доцільно (а іноді єдино можливо) аналізувати зміщення не піку, а всього інтервалу кореляції вздовж часової осі. Факт зміщення інтервалу кореляції вказує на несиметричне положення слухача відносно стереосистеми випромінювання. Величина зміщення інтервалу кореляції в часі дозволяє розрахувати положення УДЗ для конкретного слухача (табл.3).

Як бачимо, якщо в т. 1 (зоні чіткої стереофонії I) значення часової затримки для різних сигналів з певним наближенням можна вважати приблизно однаковими, то в т. 4 вони сильно відрізняються, що свідчить про непевну локалізацію уявного джерела слухачем. В т. 12 можливість локалізації джерел звуку з мовленнєвою компонентою взагалі відсутня, в той

час як визначення положення джерела імпульсного та музичного сигналів зберігається. Це свідчить про обмеженість зони локалізації мовленнєвих сигналів.

Крім того, сам вигляд функції кореляції може надати інформацію щодо існування області стереозвучання. Так в точці 4, яка за значеннями коефіцієнту кореляції та на основі суб'єктивної оцінки розбірливості мови, не входить у зону стереосприйняття інформаційних сигналів (рис. 3б), пік функції починає видозмінюватися, розділяючись на два максимуми, що особливо помітно для запису симфонічної музики. Саме тому визначення інтервалу кореляції ускладнюється та стає неоднозначним. Цей факт вказує на зникання УДЗ, тобто слухач починає розрізняти джерела звуку окремо. При цьому для мовленнєвих сигналів різко звукується інтервал кореляції, різниця між головним і боковими максимумами стає несуттєвою (або значення кореляційної функції взагалі не досягають нуля).

З проведеного аналізу випливає, що практично область стереофонічного звучання доцільно обраховувати за значенням коефіцієнту кореляції бінауральної стереопари, тобто числової характеристики, яка враховує всі зазначені закономірності стереофонічного сприйняття.

За експериментальними даними на зовнішній межі області стереофонії II значення коефіцієнту міжвушної кореляції для сигналів з мовленнєвою компонентою $IACF > 0,2$, що є граничним значенням існування УДЗ [11]

Для симфонічних творів ця область виявляється набагато ширшою і наближається до стереозони імпульсних сигналів. За оцінками, наведеними в роботі [11], для імпульсних сигналів УДЗ практично розпадається при значно більшому часі міжвушної затримки, який може досягати 30 мс при різниці рівнів звукового тиску на різних вухах у 6 дБ, що відповідає випадку, коли відстань до одного джерела вдвічі більша, ніж до іншого, а різниця відстаней до відповідного вуха складає приблизно 10 м.

ВИСНОВКИ

Розміри області стереофонічного звучання можна визначити об'єктивним способом на основі параметрів міжвушної крос-кореляційної функції бінауральної пари сигналів.

Музичні сигнали, зокрема оркестрова музика, за її сприйняттям наближаються до імпульсних сигналів. Зона стереофонії для даних сигналів виходить за межі



досліджуваного простору. Для задовільного стереопрослуховування музичних творів максимальний час затримки сигналів бінауральної пари може перевищувати 15мс. При більшому часі міжвушної затримки область існування УДЗ залежить від співвідношення рівнів звукових тисків сигналів, що надходять на ліве та праве вуха.

При випромінюванні стереосистемою сигналів, які мають мовленнєві складові, можливість локалізації УДЗ в приміщенні значно погіршується, що призводить до суттєвого зменшення розмірів стереозони.

Зону стереофонії для інформаційних сигналів практично можна визначити за значенням коефіцієнту кореляції $IACF > 0,2$, що відповідає суб'єктивному відчуттю задовільної розбірливості мови. Для мовленнєвих сигналів це значення відповідає часу міжвушної затримки сигналів 7 мс. Отже розміри стереофонічної зони для таких сигналів можна визначити за запропонованою авторами формулою (3). З її використанням зона стереофонії значно розширюється на відміну від початкової формули, що враховує замале значення різниці в надходженні сигналів на праве та ліве вуха.

Таким чином, область стереофонії в приміщенні слід визначити для конкретного типу звукового сигналу або сигналів, що переважно звучать у даному залі.

ПЕРЕЛІК ПОСИЛАНЬ

- [1] L. A. Jeffress, "A place theory of sound localization.," *J. Comp. Physiol. Psychol.*, vol. 41, no. 1, pp. 35–39, 1948, DOI: [10.1037/h0061495](https://doi.org/10.1037/h0061495).
- [2] A. Kohlrausch, J. Braasch, D. Kolossa, and J. Blauert, "An Introduction to Binaural Processing," in *The Technology of Binaural Listening*, J. Blauert, Ed. Springer-Verlag Berlin Heidelberg, 2013, pp. 1–32.
- [3] T. May, S. van de Par, and A. Kohlrausch, "Binaural Localization and Detection of Speakers in Complex Acoustic Scenes," in *The Technology of Binaural Listening*, J. Blauert, Ed. Springer-Verlag Berlin Heidelberg, 2013, pp. 397–425.
- [4] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 14, no. 4, pp. 1462–1469, 2006, DOI: [10.1109/TSA.2005.858005](https://doi.org/10.1109/TSA.2005.858005).
- [5] M. L. Jepsen, S. D. Ewert, and T. Dau, "A computational model of human auditory signal processing and perception," *J. Acoust. Soc. Am.*, vol. 124, no. 1, pp. 422–438, 2008, DOI: [10.1121/1.2924135](https://doi.org/10.1121/1.2924135).
- [6] E. C. Cherry and B. M. A. Sayers, "Human "Cross-Correlator" — A Technique for Measuring Certain Parameters of Speech Perception," *J. Acoust. Soc. Am.*, vol. 28, no. 5, pp. 889–895, 1956, DOI: [10.1121/1.1908506](https://doi.org/10.1121/1.1908506).
- [7] V. V. Furduev Akusticheskiye osnovy veshchaniya. Moscow, USSR: Gosudarstvennoye izdatel'stvo literatury po voprosam svyazi i radio, 1980, p.318.
- [8] J. Blauert and W. Cobben, "Some Consideration of Binaural Cross Correlation Analysis," *Acta Acust. united with Acust.*, vol. 39, no. 2, pp. 96–104, 1978.
- [9] R. M. Stern and H. S. Colburn, "Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position," *J. Acoust. Soc. Am.*, vol. 64, no. 1, pp. 127–140, 1978, DOI: [10.1121/1.381978](https://doi.org/10.1121/1.381978).
- [10] H. Haas, "The Influence of a Single Echo on the Audibility of Speech," *J. Audio Eng. Soc.*, vol. 20, no. 2, pp. 146–159, 1972.
- [11] Y. A. Kovalgin, *Stereofonia [Stereophony]*. Moscow, USSR: Radio i svyaz, 1989.
- [12] Y. S. Vakhitov, *Teoreticheskiye osnovy elektroakustiki i elektroakusticheskaya apparatura [Theoretical fundamentals of electroacoustics and electroacoustic equipment]*. Moscow, USSR: Iskusstvo, 1982.
- [13] R. Otnes and L. Enokson, *Prikladnoy analiz vremennykh ryadov [Applied time series analysis]*. Moscow, USSR, 1982.
- [14] J. Blauert, *Prostranstvennyy slukh [Spatial hearing]*. Moscow, USSR: Enrgiya, 1979.

Надійшла до редакції 29 листопада 2018 р.

УДК 534.134

Определение области стереофонического звучания источников информационных сигналов

Вдовенко М. В., ORCID [0000-0003-1666-3389](https://orcid.org/0000-0003-1666-3389)

e-mail mv.vdovenko@aac.kpi.ua

Лунёва С. А., к.ф.-м.н. доц., ORCID [0000-0003-0683-1211](https://orcid.org/0000-0003-0683-1211)

e-mail svetlana_lunyova@yahoo.com

Кафедра акустики и акустоэлектроники acoustic.kpi.ua

Национальный технический университет Украины

«Киевский политехнический институт имени Игоря Сикорского» kpi.ua

Киев, Украина



Аннотация—В данной статье рассматривается проблема локализации человеком источников звука. Обосновывается идея корректировки расчета зоны стереофонического эффекта и вводятся некоторые уточнения (обосновывается идея уточнения расчета зоны стереофонического эффекта). Особое внимание уделено появлению и распаду воображаемого источника звука.

В качестве критериев оценки стереофонического звучания использовались параметры функции междушной корреляции: коэффициент кросс-корреляции, интервал корреляции, ширина главного максимума и общий вид функции корреляции. Измерения выполнялись с помощью искусственной головы на зрительских местах в зале средних размеров.

Практически наличие зоны стереофонии для информационных сигналов можно определять по значениям коэффициента кросс-корреляции $IACF > 0,2$. Для разных типов сигналов это значение соответствует разному времени между ушной задержки сигналов, что приводит различные размеры области стереофонии для музыкальных и речевых сигналов.

Библ. 14, рис. 3, табл. 3.

Ключевые слова — бинауральный эффект; локализация звука; воображаемый источник звука; кросс-корреляционная функция; междушная корреляция; коэффициент междушной корреляции; бинауральные модели; локализационные модели; зона стереофонии; стереофонический эффект.

UDC 534.134

Determination of the Area of Stereo Sound of Sources of Information Signals

M. V. Vdovenko, ORCID [0000-0003-1666-3389](https://orcid.org/0000-0003-1666-3389)

e-mail mv.vdovenko@aae.kpi.ua

S. A. Luniova, PhD Assoc.Prof., ORCID [0000-0003-0683-1211](https://orcid.org/0000-0003-0683-1211)

e-mail svetlana_lunyova@yahoo.com

Department of Acoustics and Acoustoelectronics acoustic.kpi.ua

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" kpi.ua

Kyiv, Ukraine

Abstract—The purpose of the work done was to estimate the human binaural capability of locating stereo sound sources of information signals, especially speech, choir singing, and symphonic music as compared to pulse signals. The measurements were conducted in an average-sized hall where sound was emitted by a stereo system consisting of two loudspeakers with the base width of 4.5 m.

Based on the results of analysis of existing methods of binaural sound perception modelling, the estimation was performed based on correlational processing of signals recorded using a head dummy placed in various points of the room.

For the recording points selected, the maximum time delays between signals arriving at a listener's left and right ears were calculated, and the interaural cross-correlation functions were obtained. The general functions, especially the peak shift, correlation interval, and peak sharpness (including the conditions when it splits into two individual ones), were analysed. The correlation factor values and levels were calculated. As a result, the conclusions regarding the capability of locating an imaginary sound source were made based on the cross-correlation function factors values, which simplified the application of this method in practice.

Based on the results of experiments conducted and on the subjective sensations of perception, the authors have come to the conclusion that a stereophony area has sizes much wider than those assumed earlier from the signal time delay of 1 ms at a receiving stereo pair (representing a shift of an imaginary source towards the ear perceiving the signal earlier).

At the same time, it was found that signals with a speech component are much harder to locate than pulse signals. While music signals, especially symphonic music, are close to pulse signals in terms of human locating capabilities.

The found patterns have allowed us to introduce adjustments to the stereophony area calculations. Based on the research results, we suggest defining the stereophony zone border by the correlation factor value of 0.5. Given that, the interaural cross-correlation function properties and subjective perception provide for acceptable speech legibility and music transparency. The key conclusion is that this area is quite narrow for speech signals and is actually limited to the time delay of signals arriving at left and right ears – around 1 ms. Music signals have a wider stereophony area defined by the time delay between perception binaural pair components of around 10 ms. Therefore, sizes of a stereophonic sound area ought to be defined with regard to an information signal's type.

Ref. 14, fig. 3, tabl. 3.

Keywords — binaural effect; sound localization; imaginary sound source; cross-correlation function; inter ear correlation; inter-ear correlation coefficient; binaural models; localization models; stereo zone; stereo effect.

