

Апаратно-програмний комплекс «Штучна голова». Частина 1

Коригування частотної характеристики тракту

Дворник^f О. О., ORCID [0000-0003-4735-2225](https://orcid.org/0000-0003-4735-2225)

Моторнюк Д. Є., ORCID [0000-0001-9027-5259](https://orcid.org/0000-0001-9027-5259)

Дідковська М. В., к.т.н. доц, ORCID [0000-0003-0818-2008](https://orcid.org/0000-0003-0818-2008)

Продеус^s А. М., д.т.н. проф., ORCID [0000-0001-7640-0850](https://orcid.org/0000-0001-7640-0850)

Національний технічний університет України

"Київський політехнічний інститут імені Ігоря Сікорського"

Київ, Україна

Анотація—Розроблено технологію коригування частотної характеристики вимірювального тракту апаратно-програмного комплексу «Штучна голова», призначеного для акустичної експертизи приміщень. Показано, що таке коригування може бути виконано шляхом контрольованого ділення частотної характеристики системи «гучномовець-приміщення-мікрофон» на попередньо отриману оцінку амплітудної частотної характеристики підсистеми «гучномовець-мікрофон». Виконано аналіз характеру та ступеня впливу такого коригування на точність оцінювання імпульсної характеристики приміщення. Показано необхідність застосування методу регуляризації для досягнення задовільної точності обчислень, вказано можливий варіант множника регуляризації та значення параметра регуляризації.

Ключові слова — штучна голова; частотна характеристика; коригування; імпульсна характеристика приміщення; оцінювання

I. Вступ

Штучну голову (ШГ), що містить модель зовнішнього вуха, використовують в акустиці досить давно при розв'язанні наступних задач: запис звуку в концертній залі, запис відгуку приміщення на звуковий стимул для подальшого обчислення імпульсної характеристики (ІХ) приміщення, контроль якості навушників, контроль якості та розбірливості мови в лініях зв'язку [1]–[3].

Прикладом ефективного використання штучної голови є база даних, що містить записи бінауральних імпульсних характеристик кількох приміщень різного призначення [4]–[7]. Науково-практична цінність таких записів полягає в можливості моделювання звуку в різних точках приміщення (ефект «акустичне дзеркало» [8]), оцінювання якості та розбірливості мовних презентацій, якості музичних презентацій, дослідження алгоритмів дереверберації [9]–[14].

На жаль, згадана база бінауральних ІХ [4], [5] є досить обмеженою. Наприклад, вона містить записи бінауральних ІХ лише однієї аудиторії, тоді як архітекторам та звукоінженерам корисно було б мати бінауральні ІХ для низки лекційних та офісних приміщень різних розмірів та із різними архітектурними особливостями. Для усунення вказаного недоліку авторами створено апаратно-програмний комплекс «Штучна голова», із використанням якого планується отримати низку акустичних портретів у вигляді оцінок бінауральних ІХ для приміщень різного призначення.

II. ПОСТАНОВКА ЗАДАЧІ

Спрощену структурну схему процедури оцінювання ІХ приміщення за допомогою ШГ представлено на Рис. 1.

Тестовий звуковий сигнал $x(t)$ випромінюють за допомогою гучномовця, розміщеного в точці приміщення, де зазвичай знаходиться диктор. Відгук приміщення $y(t)$ на стимул $x(t)$ сприймають парою мікрофонів, притиснутих до вušних раковин штучної голови, розміщеної в одній із точок приміщення, де зазвичай знаходяться слухачі.

Оскільки амплітудно-частотні характеристики гучномовця та мікрофона не є ідеально рівномірними в смузі частот тестового сигналу, замість ІХ приміщення $h_r(t)$ насправді буде оцінено ІХ $h_{\Sigma}(t)$ системи «гучномовець-приміщення-мікрофон»

$$h_{\Sigma}(t) = h_l(t) \otimes h_r(t) \otimes h_m(t) = h_r(t) \otimes h_{lm}(t), \quad (1)$$

де \otimes – символ згортки; $h_l(t)$ – ІХ гучномовця; $h_m(t)$ – ІХ одного із мікрофонів; $h_{lm}(t) = h_l(t) \otimes h_m(t)$ – ІХ підсистеми «гучномовець-мікрофон» (ГМ).

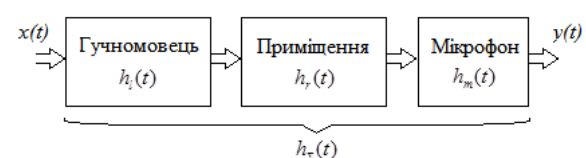


Рис. 1 Структурна схема процедури оцінювання ІХ приміщення



Спосіб обчислення ІХ приміщення $h_r(t)$ є досить очевидним. Дійсно, в частотній області співвідношенню (1) відповідає добуток відповідних частотних характеристик

$$\begin{aligned} H_{\Sigma}(f) &= H_l(f) \cdot H_r(f) \cdot H_m(f) = \\ &= H_r(f) \cdot H_{lm}(f), \end{aligned} \quad (2)$$

де $H(f) = \mathbb{F}\{h(t)\}$, \mathbb{F} – символ прямого перетворення Фур'є. Із (2) отримуємо

$$\begin{aligned} h_r(t) &= \mathbb{F}^{-1}\{H_r(f)\} = \\ &= \mathbb{F}^{-1}\left\{\frac{|H_{\Sigma}(f)|}{|H_{lm}(f)|} \cdot \exp[\theta_{\Sigma}(f) - \theta_{lm}(f)]\right\}, \end{aligned} \quad (3)$$

де \mathbb{F}^{-1} – символ оберненого перетворення Фур'є, $|\cdot|$ – символ модуля, $\theta_{\Sigma}(f)$ та $\theta_{lm}(f)$ – фазочастотні характеристики (ФЧХ) функцій $H_{\Sigma}(f)$ та $H_{lm}(f)$, відповідно.

Проблемним місцем для реалізації обчислень за виразом (3) є операція ділення, оскільки модуль $|H_{lm}(f)|$ (амплітудно частотна характеристика (АЧХ) підсистеми ГМ) може містити малі числові значення, що призведе до переповнення розрядної сітки обчислювальної системи та аварійної зупинки комп'ютера. Разом із тим, зрозуміло, що якщо забезпечити належний контроль за властивостями функції $|H_{lm}(f)|$, рівняння (3) може використовуватися для обчислень.

Що стосується ФЧХ $\theta_{\Sigma}(f)$ та $\theta_{lm}(f)$, тут є два варіанти: 1) враховувати інформацію про обидві ФЧХ; 2) враховувати інформацію лише про $\theta_{\Sigma}(f)$, припускаючи, що $\theta_{lm}(f) = 0$.

Таким чином, метою даної роботи є пошук методів обчислення АЧХ підсистеми «гучномовець-мікрофон» достатньо точних та відносно простих с точки зору їх технічної реалізації.

III. ОРГАНІЗАЦІЯ ДОСЛІДЖЕНЬ

Доцільно скористатися інформацією, викладеною в [4] щодо властивостей апаратно-програмного комплексу, що був використаний при оцінюванні бінауральних ІХ різних приміщень.

Як вказано в [4], вимірювання проводилися за допомогою професійної аудіоапаратури. Це акустичний двоканальний активний студійний монітор Genelec 8130, що має цифровий інтерфейс AES/EUBU, всенаправлені конденсаторні вимірювальні мікрофони Beyerdynamic MM1 та мікрофонний підсилювач RME Octamic II у поєднанні з аудіоінтерфейсом RME Multiface II. Для бінауральних записів використано штучну голову HMS2 фірми «HEAD acoustics». Вимірювання проводилися з частотою дискретизації 48 кГц з 24-бітовою точністю. Базовим елементом при побудові тестового сигналу в [4] слугував mls-сигнал із ступенем 16 (тобто сигнал містив 2^{16}

вибірок), що відповідає тривалості сигналу 1,36 с при частоті дискретизації 48 кГц. Тестовий сигнал створювався шляхом повторювання цього базового елементу 17 разів. Імпульсну характеристику приміщення обчислювали шляхом усереднення 16 останніх сплесків оцінки взаємної кореляційної функції (ВКФ) між сигналом з виходу мікрофона та тестовим сигналом, що дозволило підвищити відношення сигнал-шум на 12 дБ.

Авторами даної статті використано такі ж тестові сигнали, як і в [4], з тією лише відмінністю, що частота дискретизації становила 44,1 кГц замість 48 кГц. Але більш суттєвою відмінністю розробленого програмно-апаратного комплексу було використання побутових та напівпрофесійних гучномовців та мікрофонів:

- 1) малогабаритний MP3-плеєр Optima MK-4 (вихідна потужність 5 Вт, частотний діапазон 150-18000 Гц, відношення сигнал-шум 86 дБ);
- 2) активна колонка Genius SP-HF 2.0 500 (14 Вт, 65-20000 Гц, 73 дБ);
- 3) активні акустичні монітори Yamaha MSP5 (67 Вт, 50-40000 Гц, 102 дБ);
- 4) вимірювальні конденсаторні мікрофони Superlux ESM-999 (20 Гц – 20 кГц, динамічний діапазон 106 дБ, відношення сигнал-шум 70 дБ).

Враховуючи порівняно невисоку якість використаних гучномовців та мікрофонів, стає більш зрозумілою важливість коригування результатів вимірювань ІХ приміщення на основі виразу (3). Інформація про подібне коригування при отриманні результатів роботи [4], на жаль, відсутня. Можливо, в [4] таке коригування вважалося зайвим з огляду на професійний рівень аудіоапаратури.

Як відомо [5], ІХ $h_{\Sigma}(t)$ ланцюжка «гучномовець-приміщення-мікрофон» може бути обчисленою за виразом

$$\begin{aligned} K_{XY}(\tau) &= \langle x(t)y(t+\tau) \rangle = \\ &= \int_{-\infty}^{\infty} h_{\Sigma}(v) \langle x(t)x(t+\tau-v) \rangle dv = \quad (4) \\ &= \int_{-\infty}^{\infty} h_{\Sigma}(v) K_X(\tau-v) dv, \end{aligned}$$

де $\langle \cdot \rangle$ – символ математичного очікування; $K_X(\tau)$ – кореляційна функція тестового сигналу $x(t)$; $K_{XY}(\tau)$ – ВКФ тестового сигналу $x(t)$ та сигналу $y(t)$ на виході мікрофона. Дійсно, за умови

$$K_X(\tau) = D_X \cdot \delta(\tau),$$

де $\delta(\tau)$ – дельта-функція Дірака, D_X – дисперсія тестового сигналу $x(t)$, із (2) отримуємо

$$K_{XY}(\tau) = D_X \cdot h_{\Sigma}(\tau). \quad (5)$$

Фур'є-перетворення від (5) приводить до виразу



$$S_{XY}(f) = H_{\Sigma}(f) \frac{N_0}{2}, \quad (6)$$

де $S_{XY}(f)$ – взаємний спектр сигналів $x(t)$ та $y(t)$; N_0 – спектральна щільність білого шуму $x(t)$.

Якщо вимірювання виконувати в заглушеній кімнаті на малій відстані між мікрофоном та гучномовцем, тоді, згідно (5) та (6), можна отримати інформацію про ІХ та частотну характеристику підсистеми ГМ:

$$\begin{aligned} K_{XYa}(\tau) &= D_X \cdot h_{lm}(\tau), \\ S_{XYa}(f) &= H_{lm}(f) \frac{N_0}{2}; \end{aligned} \quad (7)$$

оскільки форма вихідного сигналу $y(t)$ буде визначатися властивостями підсистеми ГМ.

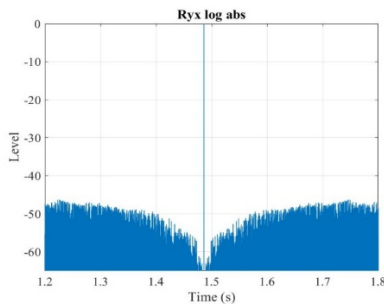


Рис. 2 Автокореляційна функція mls-сигналу

Оскільки при практичних вимірюваннях запис та обробка сигналів виконуються у цифровій формі тоді при використанні виразів (7) достатньо забезпечити щоб спектр потужності тестового сигналу $x(t)$ був рівномірним у смузі частот $F_s/2$, де F_s – частота дискретизації. Для обраного тестового mls-сигналу можна вважати, що ця вимога виконується з достатньою для практичної діяльності точністю, оскільки ширина центрального сплеску є мінімально можливою та становить $1/F_s$, а рівень бокових пелюстівок є близьким до мінус 47 дБ (Рис. 2).

Приклади форми сигналів $x(t)$ та $y(t)$ на вході та виході, відповідно, підсистеми ГМ із гучномовцем Genius наведено на Рис. 3, а відповідних спектрів потужності — на Рис. 4. Як бачимо з Рис. 3, вибірки сигналу $y(t)$, на відміну від вибірок сигналу $x(t)$, мають різну висоту, що свідчить про спотворення вхідного сигналу $x(t)$ підсистемою ГМ. Наведені на Рис. 4 оцінки спектрів потужності більш інформативні, оскільки дозволяють визначити, на яких частотах відбувається найбільш сильне спотворення вхідного сигналу (Рис. 4а) не перевищує $\pm 0,5$ дБ, то нерівномірність спектру вихідного сигналу в смузі частот від 300 Гц до 10 кГц становить ± 3 дБ, в смузі частот від 10 кГц до 16 кГц — є близькою до ± 6 дБ, а в смузі від 50 Гц до 300 Гц спостерігається провал глибиною мінус 15 дБ.

Згідно із (7), форму АЧХ підсистеми ГМ можна також визначити шляхом взаємного кореляційного аналізу сигналів $x(t)$ та $y(t)$. На Рис. 5 наведено форму такої оцінки ВКФ для підсистеми ГМ із гучномовцем Genius.

Модуль Фур'є-перетворення усередненого сплеску ВКФ, згідно (7), є оцінкою АЧХ підсистеми ГМ.

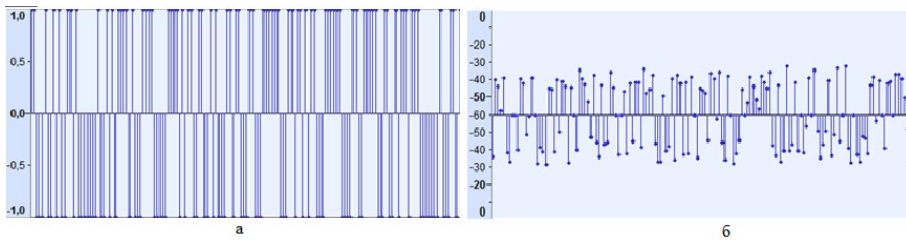


Рис. 3 Сигнали $x(t)$ (а) та $y(t)$ (б) на вході та виході підсистеми ГМ

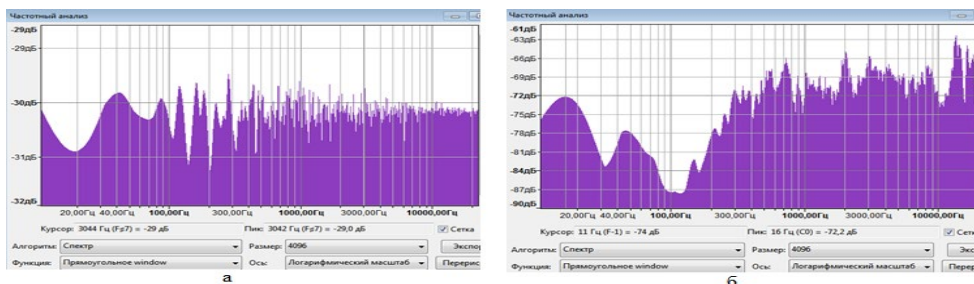


Рис. 4 Спектри потужності сигналів $x(t)$ (а) та $y(t)$ (б) на вході та виході підсистеми ГМ



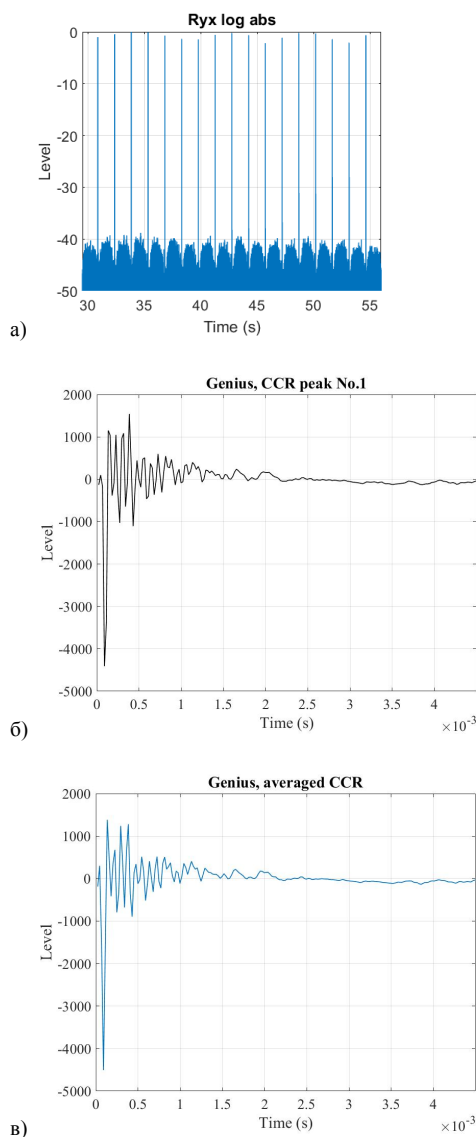


Рис. 5 ВКФ сигналів $x(t)$ та $y(t)$ (а), форма першого (б) та усередненого (в) сплесків ВКФ

IV. РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ

На Рис. 6 показано оцінки АЧХ $|H_{lm}(f)|$ підсистеми ГМ для 3-х видів гучномовців.

На кожному із рисунків наведено графіки оцінок АЧХ $|H_{lm}(f)|$ чотирьох типів:

- 1) Туре 0 – оцінка спектра потужності Велча як усереднення спектрів потужності 17 сегментів сигналу $y(t)$, кожен сегмент має довжину 65536 вибірок, вагове вікно при оцінюванні спектру не застосовується: $[Py, fy] = pwelch(y, rectwin(65536), 0, 65536, Fs)$;
- 2) Туре 1 – згладжений за частотою, трикутним вікном шириною 50 Гц, модуль дискретного

перетворення Фур'є (ДПФ) від результату усереднення 17 піків ВКФ;

- 3) Туре 2 – згладжений за частотою, трикутним вікном шириною 50 Гц, результат усереднення (17 разів) модулів ДПФ від 17 сплесків ВКФ;
- 4) Туре 3 – оцінка спектра потужності Велча із використанням вагового вікна Ханна, що забезпечує роздільну здатність 50 Гц: $[Py, fy] = pwelch(y, hann(delw), round(delw/2), 65536, Fs)$, $delw = (1,5F_s)/df$, $F_s = 44100$ Гц, $df = 50$ Гц.

Процедуру згладжування трикутним вікном, використану в оцінках 1-го та 2-го типів, було реалізовано шляхом цифрової фільтрації незгладженої оцінки спектру, для чого було застосовано рекурсивний фільтр першого порядку. Це двоетапна процедура, на першому етапі якої виконується згладжування прямокутним вікном зліва направо:

$$y_n = y_{n-1} + \frac{1}{N}(x_n - x_{n-N}), \quad n = N+1, \dots, Nfft/2 + 1,$$

а на другому етапі виконується згладжування прямокутним вікном справа наліво:

$$y_n = y_{n+1} + \frac{1}{N}(x_n - x_{n+N}), \quad n = Nfft/2 - N + 1, \dots, 1,$$

де $N = \frac{Nfft}{F_s} \cdot \frac{df}{2}$, $Nfft = 65536$, $F_s = 44100$ Гц, $df = 50$ Гц.

Як бачимо, значення параметру N обчислювалося, виходячи із ширини прямокутного вікна в 25 Гц. Це зумовлене тим, що подвійне згладжування прямокутним вікном еквівалентне одиночному згладжуванню трикутним вікном подвоєної ширини. Необхідність в застосуванні двох напрямків фільтрації зумовлена тим, що при цьому не треба дописувати якісь вибірки поза краями первинної оцінки спектру, що підлягає згладжуванню.

Порівнюючи між собою наведені на Рис. 6 оцінки $|H_{lm}(f)|$, доходимо наступних висновків:

- по-перше, виходячи із ступеня фактичної нерівномірності та фактичному частотному діапазону оцінки АЧХ $|H_{lm}(f)|$, слід віддати перевагу активній колонці Genius. Як це не дивно, але акустичний монітор Yamaha виявився неконкурентним, що можна пояснити його фізичним старінням;
- по-друге, як видно із наведених графіків, незгладжену оцінку $|H_{lm}(f)|$ нульового типу недоцільно використовувати для подальшої корекції, з огляду на неприйнятно велику дисперсію в області середніх та високих частот;



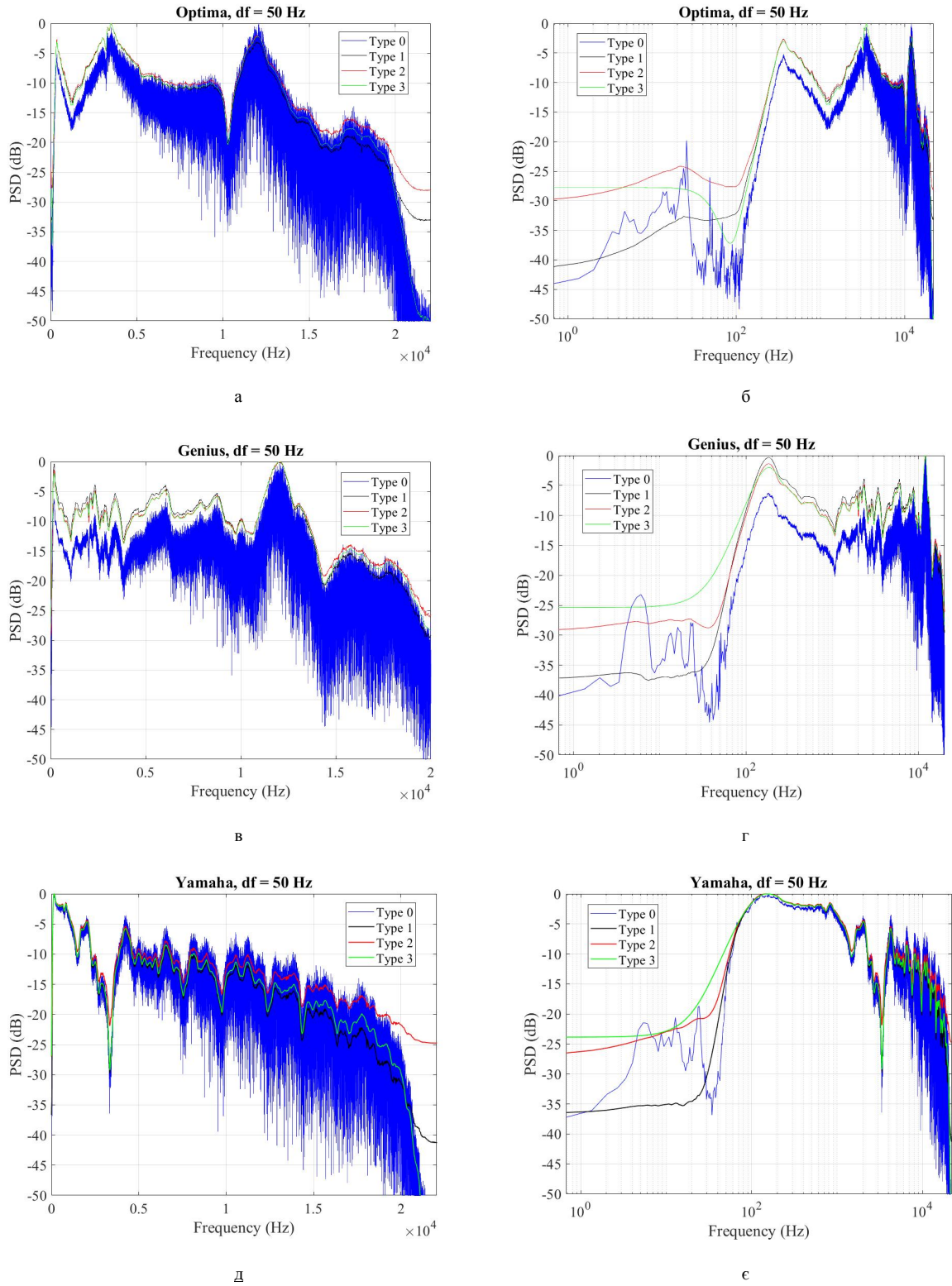


Рис. 6 Оцінки спектру сигналу гучномовців (заглушене приміщення)

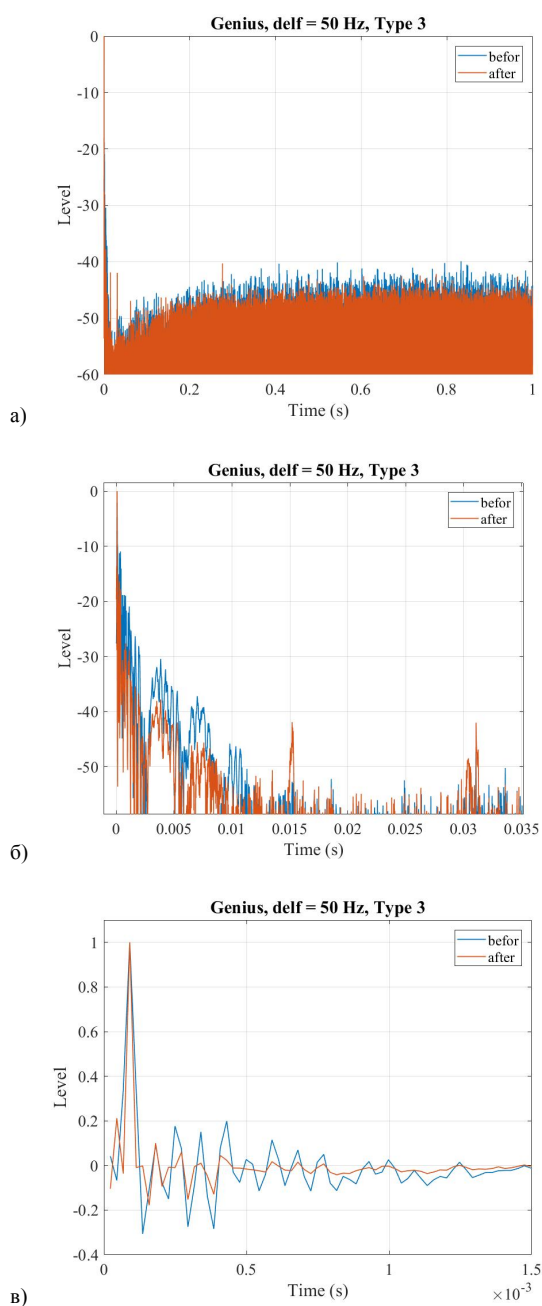


Рис. 7 Скориговані оцінки ІХ для гучномовця Genius

- по-третє, серед згладжених оцінок найбільш перспективною є оцінка третього типу як така, що найкращим чином забезпечує зниження дисперсії оцінки без надмірного підвищення її зміщення. Зауважимо, що певна доля суб'єктивності при виборі типу оцінки є неминучою, оскільки істинна АЧХ підсистеми ГМ є невідомою.

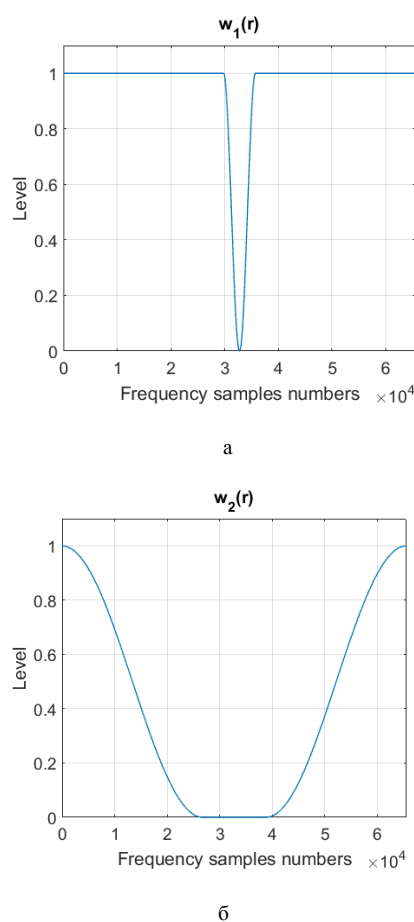
Результати обчислень ІХ підсистеми ГМ згідно рівняння (3) із використанням згладженої оцінки $|H_{lm}(f)|$ для гучномовця Genius показані на Рис. 7 для різних значень часового масштабу. Зазначимо,

що ІХ підсистеми ГМ можна називати «апаратною функцією» розробленого комплексу, яка характеризує потенційні можливості зазначеного комплексу в сенсі точності вимірювань ІХ приміщень.

Результат коригування, як бачимо, має як позитивні, так і негативні якості. Негативною є поява окремих дискретних сплесків достатньо високого рівня - 40 дБ (Рис. 7а). Позитивні:

- 1) рівень бокових пелюстків знизився приблизно на 3-4 дБ (Рис. 5а);
- 2) головний пік ІХ став вужчим в 2 рази, а рівень найближчих до нього сплесків став нижчим майже на 6 дБ (Рис. 5б,в).

З огляду на гостроту зазначених вище сплесків доречно припустити, що вони спричинені збільшенням дисперсії оцінки АЧХ, отриманої із використанням виразу (6), на високих частотах, через ділення на малі значення згладженої оцінки $|H_{lm}(f)|$. Очевидним та простим способом пригнічення таких сплесків є множення відкоригованої оцінки АЧХ на спектральне вагове вікно. В часовій області це відповідає процедурі згладжування оцінки ІХ шляхом ковзного усереднення.

Рис. 8 Вагове вікно $w_1(r)$ для $f_1=19$ кГц (а) та вікно $w_2(r)$ для $f_2=18$ кГц (б)

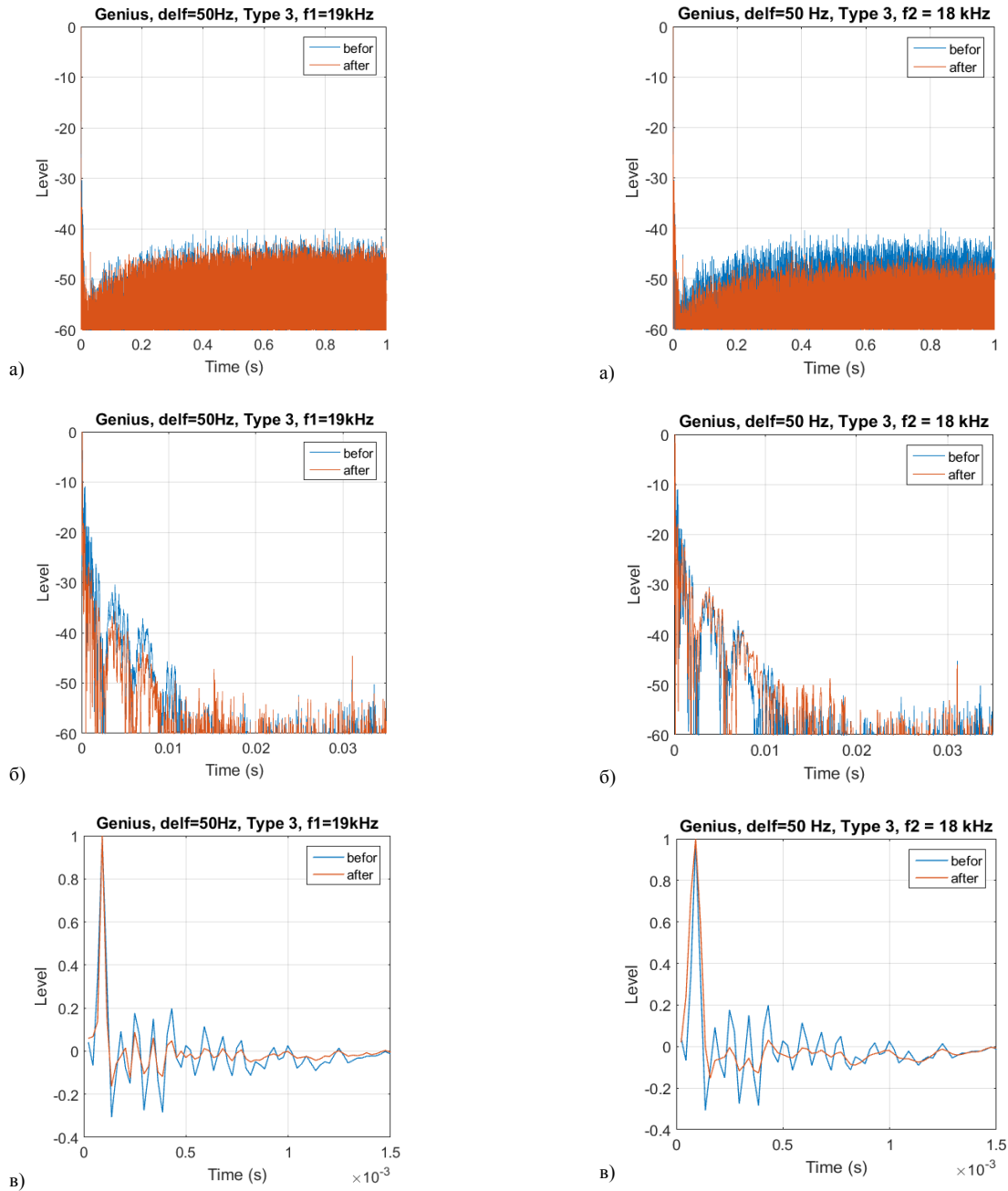


Рис. 9 Скориговані ІХ для гучномовця Genius, із вікном $w_1(r)$ для $f_1=19$ кГц

Рис. 10 Скориговані ІХ для гучномовця Genius, із вікном $w_2(r)$ для $f_2=18$ кГц

В рамках даної роботи розглянуто два види такого вагового вікна. В першому випадку це вікно описується виразом:

$$w_1(r) = \begin{cases} 1, & r = 1, \dots, \frac{N_{fft}}{2} - k, \quad r = \frac{N_{fft}}{2} + 2 + k, \dots, N_{fft}, \\ 0,5 \left[1 - \cos \frac{\pi}{k} \left(r - \frac{N_{fft}}{2} - 1 \right) \right], & r = \frac{N_{fft}}{2} + 1 - k, \dots, \frac{N_{fft}}{2} + 1 + k; \end{cases}$$

де k – визначає інтервал, який відповідає половині періоду косинусоїди, що описує кінцеву частину вікна, яка відповідає за пригнічення високих частот:

$$k = (F_s/2 - f_1) / (F_s/N_{fft}).$$

Друге вагове вікно – це вікно Ханна

$$w_2(r) = 0,5 \left[1 - \cos \left(\pi(r - N_{fft}/2 - 1) / \Delta N \right) \right],$$

$$r = 1, \dots, N_{fft},$$

де ΔN – половина ширини вікна Ханна у вибірках.

Графіки вагових вікон для $f_1 = 19$ кГц та $f_2 = 18$ кГц ($f_2 = \Delta N \cdot F_s / N_{fft}$) наведено на Рис. 8. Неважко помітити, що в окремому випадку, коли $f_1 = 0$, $f_2 = F_s/2$, ці вікна співпадають: $w_1(r) = w_2(r)$.

Результати застосування вікон $w_1(r)$ та $w_2(r)$ наведено на Рис. 9 та Рис. 10.

Порівнюючи Рис. 9 та Рис. 10, неважко зробити висновок, що вікно $w_2(r)$ з параметром $f_2 = 18$ кГц є більш перспективним, оскільки дозволяє зменшити рівень бокових пелюстків оцінки ІХ підсистеми ГМ на 4 дБ, тоді як вікно $w_1(r)$ з параметром $f_1 = 19$ кГц дозволяє зменшити рівень бокових пелюстків лише на 1 дБ (Рис. 9а та Рис. 10а). В інтервалі 0-10 мс вікно $w_1(r)$ є кращим, оскільки забезпечує рівень бокових пелюстків на 6 дБ нижчий, ніж вікно $w_2(r)$. Проте ця перевага не є принциповою, оскільки у випадку вікна $w_2(r)$ рівень цих бокових пелюстків не перевищує -30 дБ, а ширина головного сплеску ІХ на рівні -30 дБ не перевищує 2 мс, що забезпечує достатню для інженерних застосувань точність оцінювання ІХ приміщення.

Зазначимо, що використання вагового вікна $w_2(r)$ можна розглядати як регуляризацію, з множителем регуляризації $w_2(r)$ та параметром регуляризації f_2 , рішення нестійкої оберненої задачі [15].

Оскільки в даній роботі розглянуто варіант коригування частотної характеристики тракту ШГ, де враховується лише нерівномірність АЧХ підсистеми ГМ, в подальшому слід розглянути доцільність додаткового врахування оцінки фазової характеристики підсистеми ГМ, а також оцінити доцільність використання адаптивних фільтрів.

ВИСНОВКИ

Розроблено технологію коригування частотної характеристики апаратно-програмного комплексу «Штучна голова», що містить елементи побутової та напівпрофесійної аудіотехніки. Показано, що таке коригування можна виконати на етапі обробки сигналів шляхом контрольованого ділення оцінки частотної характеристики системи «гучномовець-приміщення-мікрофон» на попередньо отриману оцінку АЧХ підсистеми «гучномовець-мікрофон». При

цьому задовільну для практичного застосування точність оцінювання можна забезпечити за умови регуляризації рішення. В якості множника регуляризації запропоновано спектральне вікно Ханна із значенням параметру регуляризації 18 кГц при частоті дискретизації сигналу 44,1 кГц, що дозволяє знизити рівень бокових пелюстків апаратної функції вимірювального комплексу на 4 дБ, із мінус 40 дБ до мінус 44 дБ. При цьому ширина головного сплеску апаратної функції на рівні мінус 30 дБ не перевищує 2 мс, що забезпечує достатню для інженерних застосувань точність оцінювання ІХ приміщення.

ПЕРЕЛІК ПОСИЛАНЬ

- [1] J. Blauert, Ed., *The Technology of Binaural Listening*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, ISBN: 978-3-642-37761-7. DOI: [10.1007/978-3-642-37762-4](https://doi.org/10.1007/978-3-642-37762-4)
- [2] "HEAD acoustics - Telecom Division - Binaural Recording Systems - Artificial Head Measurement System HMS II.3 - Overview." [Online]. Available: https://www.head-acoustics.com/eng/telecom_hms_II_3.htm. [Accessed: 03-Mar-2020].
- [3] P. L. Søndergaard and P. Majdak, "The auditory modeling toolbox," in *The Technology of Binaural Listening*, Springer Berlin Heidelberg, 2013, pp. 33–56.
- [4] M. Jeub, M. Schäfer, and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in *2009 16th International Conference on Digital Signal Processing*, 2009, pp. 1–5, DOI: [10.1109/ICDSP.2009.5201259](https://doi.org/10.1109/ICDSP.2009.5201259).
- [5] "IKS: Aachen Impulse Response Database." [Online]. Available: <https://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/aachen-impulse-response-database/>. [Accessed: 03-Mar-2020].
- [6] G. A. Soulodre, N. Popplewell, and J. S. Bradley, "Combined effects of early reflections and background noise on speech intelligibility," *J. Sound Vib.*, vol. 135, no. 1, pp. 123–133, Nov. 1989, DOI: [10.1016/0022-460X\(89\)90759-1](https://doi.org/10.1016/0022-460X(89)90759-1).
- [7] M. Jeub, M. Schäfer, H. Krüger, C. Nelke, C. Beaugeant, and P. Vary, "Do We Need Dereverberation for Hand-Held Telephony?," in *20th International Congress on Acoustics, ICA 2010*, 2010, URL: www.ind.rwth-aachen.de/www.infineon.com.
- [8] "SOUND FORGE Audio Studio 14 – Домашня студія редакування звука." [Online]. Available: <https://www.magix.com/ru/muzyka/sound-forge/sound-forge-audio-studio/>. [Accessed: 03-Mar-2020].
- [9] E. A. P. Habets, N. D. Gaubitch, and P. A. Naylor, "Temporal selective dereverberation of noisy speech using one microphone," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 4577–4580, DOI: [10.1109/ICASSP.2008.4518675](https://doi.org/10.1109/ICASSP.2008.4518675).
- [10] J. S. Bradley, H. Sato, and M. Picard, "On the importance of early reflections for speech in rooms," *J. Acoust. Soc. Am.*, vol. 113, no. 6, p. 3233, 2003, DOI: [10.1121/1.1570439](https://doi.org/10.1121/1.1570439).
- [11] I. (Technical U. of D. Arweiler, J. (Technical U. of D. Buchholz, and T. (Technical U. of D. Dau, "Speech intelligibility enhancement by early reflections | Proceedings of the International Symposium on Auditory and Audiological Research," in *Speech processing and perception under adverse conditions*, 2009, pp. 289–298, URL: <https://proceedings.isaar.eu/index.php/isaarproc/article/view/2009-29>.
- [12] H. Sato and J. S. Bradley, "Evaluation of acoustical conditions for speech communication in working elementary school classrooms," *J. Acoust. Soc. Am.*, vol. 123, no. 4, pp. 2064–2077, 2008, DOI: [10.1121/1.2839283](https://doi.org/10.1121/1.2839283).
- [13] A. Prodeus, V. Didkovskiy, M. Didkovska, I. Kotvitskiy, D. Motomiuk, and A. Khrapachevskiy, "Objective and Subjective Assessment of the Quality and Intelligibility of Noised Speech,"



in *2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, 2018, pp. 71–74, DOI: [10.1109/INFOCOMMST.2018.8632125](https://doi.org/10.1109/INFOCOMMST.2018.8632125).

- [14] A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, and D. Motorniuk, "Automated Subjective Assessment of Speech Intelligibility under Diotic and Dichotic Listening," in *Data-Centric Business and Applications*, Springer, 2020 (in print).

- [15] A. N. Tikhonov, "O nekorrektnykh zadachakh lineynoy algebry i ustoychivom metode ikh resheniya [On ill-posed problems of linear algebra and a stable method for solving them]," *DAN USSR*, vol. 163, no. 3, pp. 591–594, 1965.

Надійшла до редакції 13 березня 2020 року

UDC 534.84

Artificial Software Complex "Artificial Head". Part 1

Adjusting the Frequency Response of the Path

O. O. Dvornyk^f, ORCID [0000-0003-4735-2225](https://orcid.org/0000-0003-4735-2225)

D. Motorniuk, ORCID [0000-0001-9027-5259](https://orcid.org/0000-0001-9027-5259)

M. V. Didkovska, PhD Assoc.Prof., ORCID [0000-0003-0818-2008](https://orcid.org/0000-0003-0818-2008)

A. M. Prodeus^s, Dr.Sc.(Eng.) Prof., ORCID [0000-0001-7640-0850](https://orcid.org/0000-0001-7640-0850)

National technical university of Ukraine "Igor Sikorsky Kyiv polytechnic institute"
Kyiv, Ukraine

Abstract—In this paper, the technology of correction of the frequency response of the measuring path of the hardware-software complex "Artificial head", intended for acoustic examination of the rooms, has been developed. This correction is necessary because the amplitude-frequency response of the loudspeaker and the microphone are not perfectly uniform in the frequency band of analysis. Therefore, instead of the room impulse response, the convolution of the room impulse response with the impulse responses of the loudspeaker (test signal source) and the microphone will actually be evaluated. It is shown that such correction can be made by controlled dividing the frequency response of the loudspeaker-room-microphone system by the previously obtained estimate of the amplitude-frequency response of the loudspeaker-microphone subsystem.

The problem with such calculations is the division operation, because the amplitude-frequency response of the loudspeaker-microphone subsystem may contain small numerical values, which will overflow the discharge grid of the computing system and crash the computer. However, it is clear that if proper control over the amplitude-frequency characteristics of the loudspeaker-microphone subsystem is ensured, such a division can be practically implemented.

Since the amplitude-frequency response of the subsystem "loudspeaker-microphone" takes the smallest values at the edges of the frequency range, and the variance of the evaluation of the mutual spectrum of the system "loudspeaker-room-microphone" is the largest on the right edge of the frequency range, it is advisable to apply the method of regularization to achieve necessary calculation accuracy. In this case, the role of the regularizing factor can be played by a spectral weighting window whose values are close to one at low and medium frequencies, and with the approaching to the right edge of the frequency range, the values of this window are close to zero. The role of the regularization parameter is thus the width of such a weighting window.

The nature and value of such correction effect on the accuracy of the room impulse response estimate are analyzed in this paper. It is shown that the Hann (Hanning) window can be used as a regularizing factor, while the width value of this window was experimentally found, which provides satisfactory properties for evaluating the room impulse response. It has been also shown that the width of the Hann window, which is close to 80% of the total frequency range analyzed, is satisfactory for practical applications.

Keywords — artificial head; frequency response; correction; room impulse response; estimation.

