

Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Кваліфікаційна наукова
праця на правах рукопису

Дворник Олександр Олександрович

УДК 621.391.83


ДИСЕРТАЦІЯ
МЕТОДИ ТА СИСТЕМИ ОЦІНКИ ТА КОРЕКЦІЇ АКУСТИЧНИХ
ХАРАКТЕРИСТИК ПРИМІЩЕНЬ ДЛЯ ПУБЛІЧНИХ ВИСТУПІВ

17 - Електроніка та телекомунікації

171 - Електроніка

Подається на здобуття наукового ступеня доктора філософії.

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

 / Дворник О.О.

Науковий керівник Продеус Аркадій Миколайович, доктор технічних наук,
професор

Київ - 2024

АНОТАЦІЯ

Дворник О.О. Методи та системи оцінки та корекції акустичних характеристик приміщень для публічних виступів. – Кваліфікаційна робота на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 171 «Електроніка». - Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», МОН України, Київ, 2024.

Дисертаційна робота спрямована на дослідження шляхів підвищення точності та швидкості оцінювання акустичних характеристик приміщень в умовах негативної дії шуму та реверберації, а також на дослідження шляхів корекції цих характеристик.

Зміст дисертаційного дослідження викладено в чотирьох розділах, де представлено та обґрунтовано основні результати роботи.

Актуальність дисертаційної роботи обґрунтовано у вступі, де сформульовано мету та задачі дослідження, описано методи дослідження, надано інформацію про наукову новизну та практичне значення одержаних результатів.

У першому розділі виконано огляд існуючих методів і підходів оцінювання розбірливості мови у приміщенні, аналізу ранніх відбиттів у приміщенні, а також методам корекції акустики приміщення. Показано, що в існуючих наукових працях, присвячених дослідженню розбірливості мови, недостатньо вивчене питання, який з методів оцінювання розбірливості мови є оптимальним, за критеріями точності та швидкості вимірювань, для вирішення поставленої задачі, із врахуванням апріорної інформації про заводові умови. Дослідження методів виявлення ранніх відбиттів є актуальним, оскільки аналіз моментів появи ранніх відбиттів дозволяє робити висновки стосовно їх впливу на розбірливість мовлення, а також локалізувати напрями приходу відбиттів до слухача. Аналіз існуючих наукових праць показав, що ґрунтовної відповіді на дане питання поки немає. Ще одним

перспективним напрямом дослідження є частотна корекція акустики приміщення. Хоч дане питання досить активно досліджується у світі, однак наразі важко знайти рекомендації стосовно вибору методу корекції в реальному часі з урахуванням зміни наповненості приміщення.

Другий розділ присвячений порівнянню різновидів модуляційного методу оцінювання розбірливості мовлення за точністю вимірювань. Результатом порівняння є рекомендації щодо застосування методів в залежності від задачі. Запропоновано використання багатокomпонентного тестового сигналу для порівняння прямого та непрямого методу оцінки індексу передачі мовлення, також проведено порівняння прямого та непрямого методів оцінки індексу передачі мовлення.

У третьому розділі представлено методи акустичної експертизи приміщення з використанням штучної голови, показано, що за допомогою використання побутового акустичного обладнання можливо провести вимірювання, що за точністю результатів не поступаються отриманим на професійному обладнанні. Показано, що досягнення прийнятної точності вимірювань є можливим за умови використання запропонованого методу корегування частотної характеристики вимірювального тракту.

У четвертому розділі представлені результати інструментальної оцінки якості та корекції якості приміщення. Представлено результати суб'єктивних та об'єктивних оцінок розбірливості мовлення з урахуванням впливу шуму, ранніх та пізніх відбиттів. Розроблено апаратно-програмний комплекс для аналізу напрямків відбиттів звуку у приміщенні, а також продемонстровано працездатність запропонованого методу корекції частотної характеристики приміщення у реальному часі із застосуванням багатоканального компресора.

Представлені в дисертації нові теоретичні та практичні результати можна рекомендувати до використання при розробці та експлуатації систем, що

призначені для акустичної експертизи приміщення й, зокрема, вимірювання розбірливості мовлення, та частотної корекції приміщення.

В дисертаційній роботі отримано наступні наукові результати:

1. Вперше виконано порівняння точності оцінювання індексу розбірливості мовлення в умовах дії шуму та реверберації із використанням повного традиційного методу та повного формантно-модуляційного методу для широких діапазонів значень відношення сигнал-шум та тривалості тестових сигналів.
2. Вперше виконано порівняння швидкого та повного варіантів формантно-модуляційного методу оцінювання індексу розбірливості мовлення в умовах дії шуму та реверберації для широких діапазонів значень відношення сигнал-шум та тривалості тестових сигналів.
3. Вперше запропоновано багатокomпонентний тестовий сигнал та алгоритми його обробки, на основі яких виконано порівняння двох видів непрямого методу оцінки індексу передачі мовлення із прямим методом.
4. Уточнено попередні висновки про те, що формантний метод оцінювання індексу передачі мовлення є найбільш точним та швидким в умовах переважної дії шумової завади.
5. Уточнено умови використання апаратури непрофесійного рівня в апаратно-програмному комплексі «Штучна голова», призначеного для оцінки бінауральних імпульсних характеристик приміщення.
6. Вдосконалено та експериментально перевірено дієвість методу багатоканальної компресії сигналу, спотвореного реверберацією, в задачі послаблення нерівномірності амплітудно-частотної характеристики приміщення.

7. Показано можливість визначення часу реверберації та напрямків приходу відбиттів до слухача шляхом просторового акустичного сканування простору.

Практичне значення отриманих результатів полягає у наступному:

1. Встановлено принципову можливість використання повного та швидкого формантно-модуляційних методів оцінювання STI замість традиційного повного модуляційного методу.
2. Встановлена залежність точності формантного, повного модуляційного, повного формантно-модуляційного та швидкого формантно-модуляційного методів оцінювання STI від відношення сигнал-шум та від тривалості тестового сигналу.
3. Виконане коректне порівняння, за точністю вимірювань, прямого та непрямого методів оцінки STI.
4. Отримано практичні рекомендації стосовно корекції частотної характеристики вимірювального тракту системи «Штучна голова», що містить обладнання непрофесійного рівня, для підвищення точності оцінювання бінауральних імпульсних характеристик приміщення.
5. Встановлено працездатність та ефективність методу корекції сигналу, спотвореного реверберацією, на основі методу багатоканальної компресії, що дозволяє послабити нерівномірність амплітудно-частотної характеристики приміщення
6. Встановлено працездатність запропонованого акустичного модуля для вимірювання просторових акустичних характеристик приміщення.

Ключові слова: розбірливість мовлення, оцінка розбірливості мовлення, штучна голова, акустична експертиза приміщення, корекція амплітудно-частотної характеристики приміщення, просторовий аналіз відбиттів у приміщенні.

Список публікацій здобувача

1. Dvornyk, Oleksandr Oleksandrovych. 2023. "Real-Time Correction of the Frequency Response of a Public Address System". *Microsystems, Electronics and Acoustics* 28 (2). Kyiv, Ukraine:280467.1-67. <https://doi.org/10.20535/2523-4455.me.280467>.
2. Didkovska, Maryna & Dvornyk, Oleksandr & Motorniuk, Daria & Prodeus, Arkadiy. (2020). Artificial Software Complex "Artificial Head". Part 1. Adjusting the Frequency Response of the Path. *Microsystems Electronics and Acoustics*. 22. 56-64.
3. Prodeus, Arkadiy & Dvornyk, Oleksandr & Didkovska, Maryna & Motorniuk, Daria. (2020). Hardware and Software System "Artificial Head". Part 2. Evaluation of Speech Intelligibility in Classrooms. *Microsystems Electronics and Acoustics*. 22. 48-55.
4. Kozeruk, Sergei, & Dvornyk Oleksandr. 2023. "Device for Operational Assessment of Acoustic Characteristics of Room". *Visnyk NTUU KPI Seriya - Radiotekhnika Radioaparobuduvannia*, no. 94 (December):24-31. <https://doi.org/10.20535/RADAP.2023.94.24-31>.
5. A. Prodeus, O. Dvornyk, A. Naida, M. Didkovska, O. Grebin. (2023). Multicomponent Signal for Comparing Direct and Indirect Methods of Speech Transmission Index Measurement. *Electronics and Control Systems*. 1. 27-35. 10.18372/1990-5548.75.17546.
6. А. Продеус, А. Вітик, О. Дворник, І. Котвицький, О. Чайка, М. Ярошенко, Суб'єктивне оцінювання розбірливості мови на тлі шуму та реверберації. "Мікросистеми, електроніка та акустика", том 23, № 2, 2018, с. 66-73, ISSN 2523-4447, DOI: 10.20535/2523-4455.2018.23.2.128820
7. A. Prodeus, K. Bukhta, P. Morozko, O. Serhiienko, I. Kotvytskyi, O. Dvornyk, "Automated Subjective Assessment of Speech Intelligibility in Various Listening Modes," ISSN 2523-4447. *Microsystems, Electronics and Acoustics*, 2018, vol. 23, no. 3, pp.49-57. DOI: <https://doi.org/10.20535/2523-4455.2018.23.3.130367>

8. Продеус, А. М., Дворник, О. О., & Найда, А. С. (2024). Точність оцінки індексу передачі мовлення формантно-модуляційним методом. *Вісті вищих учбових закладів. Радіoeлектроніка* (прийнято до друку) <https://doi.org/10.20535/S002134702306002X>
9. А. М. Продеус, І. В. Котвицький, М. В. Дідковська, В. С. Дідковський, К. А. Кухарічева, Д. Є. Моторнюк, О. О. Дворник Спосіб виявлення кліпування мовного та музичного сигналів // Патент UA 144291 U, МПК G01R 23/20, опубл. 25.09.2020. Prodeus, Arkadiy & Dvornyk, Oleksandr & Naida, Anton & Didkovska, Maryna.

SUMMARY

Dvornyk O.O. Methods and systems for evaluation and correction of acoustic characteristics of premises for public performances. - Qualifying thesis as a manuscript.

Dissertation for the Doctor of Philosophy degree in Electronics. - National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ministry of Education and Science of Ukraine, Kyiv, 2024.

The dissertation is aimed at researching ways to improve the accuracy and speed of evaluating the acoustic characteristics of premises under the influence of noise and reverberation, as well as investigating methods for correcting these characteristics.

The content of the dissertation research is presented in four chapters, where the main results of the work are outlined and justified.

The relevance of the dissertation is substantiated in the introduction, where the purpose and objectives of the research are formulated, research methods are described, and information is provided on the scientific novelty and practical significance of the obtained results.

In the first chapter, a review of existing methods and approaches for assessing speech intelligibility in a room, analyzing early reflections in the room, and methods for correcting room acoustics is conducted. It is shown that in existing scientific works dedicated to studying speech intelligibility, the question of which method of speech intelligibility assessment is optimal, according to the criteria of accuracy and measurement speed, remains insufficiently studied to address the set task, considering the prior information about the adverse conditions. The investigation of early reflection detection methods is relevant because analyzing the moments of appearance of early reflections allows conclusions to be drawn regarding their impact on speech intelligibility and to localize the directions of reflection arrival to the listener. The analysis of existing scientific works has shown that there is no comprehensive answer to this question yet. Another promising direction for research is frequency correction of room acoustics. Although this issue is quite actively

researched worldwide, it is currently difficult to find recommendations regarding the choice of correction method in real-time considering the changing occupancy of the room.

The second chapter is dedicated to comparing various types of modulation methods for assessing speech intelligibility in terms of measurement accuracy. The comparison results in recommendations for the application of methods depending on the task. The use of a multi-component test signal is proposed for comparing direct and indirect methods of speech transmission index assessment, and a comparison of direct and indirect methods of speech transmission index assessment is also conducted.

The third chapter presents methods of acoustic expertise of premises using an artificial head. It is shown that household acoustic equipment can be used to conduct measurements that are not inferior in accuracy to those obtained with professional equipment. It is demonstrated that achieving acceptable measurement accuracy is possible provided that the proposed method of correcting the frequency response of the measurement path is used.

The fourth chapter presents the results of instrumental assessment of the quality and correction of the premises. The results of subjective and objective assessments of speech intelligibility, taking into account the influence of noise, early and late reflections, are presented. A hardware and software complex for analyzing sound reflection directions in the premises has been developed, and the effectiveness of the proposed method of correcting the frequency response of the premises in real time using a multichannel compressor has been demonstrated.

The new theoretical and practical results presented in the dissertation can be recommended for use in the development and operation of systems designed for acoustic assessment of premises, including speech intelligibility measurement, and frequency correction of premises.

In the dissertation research, the following scientific results have been obtained:

8. For the first time, a comparison of the accuracy of speech intelligibility index estimation in conditions of noise and reverberation using the full traditional method and the full formant-modulation method has been performed across wide ranges of signal-to-noise ratios and test signal durations.
9. A comparison of the rapid and full versions of the formant-modulation method for speech intelligibility index estimation in conditions of noise and reverberation has been conducted for wide ranges of signal-to-noise ratios and test signal durations.
10. A multi-component test signal and processing algorithms have been proposed for comparison of two types of indirect speech intelligibility index estimation methods with the direct method.
11. It has been clarified that the formant method for speech intelligibility index estimation is the most accurate and rapid under predominantly noise conditions.
12. Conditions for using non-professional equipment in the hardware-software complex "Artificial Head," designed for binaural impulse response assessment of premises, have been refined.
13. The effectiveness of the multi-channel signal compression method in mitigating amplitude-frequency characteristic non-uniformities of the room has been improved and experimentally verified.
14. The possibility of determining the reverberation time and directions of reflections arrival at the listener by spatial acoustic scanning of the space has been demonstrated.

The practical significance of the obtained results lies in the following:

7. The possibility of using both full and rapid formant-modulation methods for STI assessment instead of the traditional full modulation method has been established.

8. The dependence of the accuracy of formant, full modulation, full formant-modulation, and rapid formant-modulation methods of STI assessment on the signal-to-noise ratio and the duration of the test signal has been determined.
9. A comparative analysis, based on measurement accuracy, of direct and indirect methods of STI assessment has been conducted.
10. Practical recommendations for correcting the frequency response of the measurement system's "Artificial Head," which includes non-professional-grade equipment, have been developed to improve the accuracy of binaural impulse response measurements in room acoustics analysis.
11. The effectiveness of the proposed signal correction method, based on multichannel compression, for mitigating irregularities in the amplitude-frequency characteristics of the room distorted by reverberation, has been demonstrated.
12. The functionality of the proposed acoustic module for measuring spatial acoustic characteristics of a room has been verified.

Key words: speech intelligibility, speech intelligibility assessment, artificial head, room acoustics analysis, correction of room amplitude-frequency characteristics, spatial reflections analysis in rooms.

List of applicant's publications

1. Dvornyk, Oleksandr Oleksandrovych. 2023. "Real-Time Correction of the Frequency Response of a Public Address System". *Microsystems, Electronics and Acoustics* 28 (2). Kyiv, Ukraine:280467.1-67. <https://doi.org/10.20535/2523-4455.mea.280467>.
2. Didkovska, Maryna & Dvornyk, Oleksandr & Motorniuk, Daria & Prodeus, Arkadiy. (2020). Artificial Software Complex "Artificial Head". Part 1. Adjusting the Frequency Response of the Path. *Microsystems Electronics and Acoustics*. 22. 56-64.

3. Prodeus, Arkadiy & Dvornyk, Oleksandr & Didkovska, Maryna & Motorniuk, Daria. (2020). Hardware and Software System "Artificial Head". Part 2. Evaluation of Speech Intelligibility in Classrooms. *Microsystems Electronics and Acoustics*. 22. 48-55.
4. Kozeruk, Sergei, & Dvornyk Oleksandr. 2023. "Device for Operational Assessment of Acoustic Characteristics of Room". *Visnyk NTUU KPI Seriiia - Radiotekhnika Radioaparotobuduvannia*, no. 94 (December):24-31. <https://doi.org/10.20535/RADAP.2023.94.24-31>.
5. A. Prodeus, O. Dvornyk, A. Naida, M. Didkovska, O. Grebin. (2023). Multicomponent Signal for Comparing Direct and Indirect Methods of Speech Transmission Index Measurement. *Electronics and Control Systems*. 1. 27-35. 10.18372/1990-5548.75.17546.
6. А. Продеус, А. Вітик, О. Дворник, І. Котвицький, О. Чайка, М. Ярошенко, Суб'єктивне оцінювання розбірливості мови на тлі шуму та реверберації. "Мікросистеми, електроніка та акустика", том 23, № 2, 2018, с. 66-73, ISSN 2523-4447, DOI: 10.20535/2523-4455.2018.23.2.128820
7. A. Prodeus, K. Bukhta, P. Morozko, O. Serhiienko, I. Kotvytskyi, O. Dvornyk, "Automated Subjective Assessment of Speech Intelligibility in Various Listening Modes," ISSN 2523-4447. *Microsystems, Electronics and Acoustics*, 2018, vol. 23, no. 3, pp.49-57. DOI: <https://doi.org/10.20535/2523-4455.2018.23.3.130367>
8. Продеус, А. М., Дворник, О. О., & Найда, А. С. (2024). Точність оцінки індексу передачі мовлення формантно-модуляційним методом. *Вісті вищих навчальних закладів. Радіoeлектроніка* (прийнято до друку) <https://doi.org/10.20535/S002134702306002X>
9. А. М. Продеус, І. В. Котвицький, М. В. Дідковська, В. С. Дідковський, К. А. Кухарічева, Д. Є. Моторнюк, О. О. Дворник Спосіб виявлення кліпування мовного та музичного сигналів // Патент UA 144291 U, МПК G01R 23/20, опубл. 25.09.2020. Prodeus, Arkadiy & Dvornyk, Oleksandr & Naida, Anton & Didkovska, Maryna.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ.....	15
ВСТУП.....	17
1. СУЧАСНИЙ СТАН ОЦІНЮВАННЯ РОЗБІРЛИВОСТІ МОВЛЕННЯ У ПРИМІЩЕННІ ТА КОРЕКЦІЇ АКУСТИКИ ПРИМІЩЕННЯ.....	24
1.1.Методи оцінювання розбірливості мовлення у приміщеннях.....	24
1.2.Аналіз ранніх відбиттів звуку у приміщенні.....	30
1.3.Корекція акустики приміщення	35
Висновки до розділу 1.....	40
2. ПОРІВНЯННЯ РІЗНОВИДІВ МОДУЛЯЦІЙНОГО МЕТОДУ ОЦІНЮВАННЯ РОЗБІРЛИВОСТІ МОВЛЕННЯ ЗА ТОЧНІСТЮ ТА ШВИДКІСТЮ ВИМІРЮВАНЬ.....	41
2.1. Порівняння повного модуляційного методу оцінювання розбірливості мовлення із повним формантно-модуляційним методом.....	41
2.2.Порівняння швидкого формантно-модуляційного методу із повним формантно-модуляційним методом.....	48
2.3.Багатокомпонентний тестовий сигнал для порівняння прямого та непрямого методів оцінки індексу передачі мовлення.....	59
2.4.Порівняння прямого та непрямого методів оцінки індексу передачі мовлення.....	70
Висновки до розділу 2.....	78
3. АКУСТИЧНА ЕКСПЕРТИЗА ПРИМІЩЕННЯ ІЗ ВИКОРИСТАННЯМ ШТУЧНОЇ ГОЛОВИ.....	79
3.1. Коригування частотної характеристики вимірювального тракту.....	79
3.2.Вимірювання бінауральних імпульсних характеристик приміщення та розбірливості мовлення, спотвореного реверберацією.....	93
Висновки до розділу 3.....	101
4. ІНСТРУМЕНТАЛЬНА ОЦІНКА ТА КОРЕКЦІЯ ЯКОСТІ АКУСТИКИ ПРИМІЩЕННЯ.....	102

4.1.Вплив шуму, ранніх та пізніх відбиттів на розбірливість мовлення...	102
4.2.Аналіз відбиттів у приміщенні.....	109
4.3.Корекція акустики приміщення шляхом адаптації амплітудно-частотної характеристики акустичної системи під зміну наповненості приміщення.....	115
Висновки до розділу 4.....	122
ВИСНОВКИ.....	123
Список літератури.....	126

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

DSP, Digital Signal Processor – процесор цифрової обробки сигналів

EDT, Early Decay Time – час приходу ранніх відбиттів

EQ, Equalization – еквалізація

ERB, Early Reflection Benefit - користь ранніх відбиттів

JND, Just Noticeable Difference – ледь помітна різниця

MLS, Maximum-Length Sequence – послідовність максимальної довжини,
M-послідовність

RIR, Room Impulse Response - імпульсна характеристика кімнати

RRE, Room Response Equalization – вирівнювання відгуку приміщення

RTF, Room Transfer Function – відгук кімнати

SNR, Signal to Noise Ratio – співвідношення сигнал-шум

STI, Speech Transmission Index – індекс передачі мовлення

VST, Virtual Studio Technology – віртуальна студійна технологія

АС – активна акустична система

АЦП – аналого-цифровий перетворювач

АЧХ - амплітудно-частотна характеристика

БК – багатоканальний компресор

ВКФ - взаємно-кореляційна функція

ГМ - підсистема «гучномовець-мікрофон»

ДПФ – дискретне перетворення Фур'є

КД – кутова діаграма

КПКХ – керована просторова кореляційна характеристика

МП – мікшерний пульт

ПК – персональний комп'ютер

ПООАХ – пристрій для оперативної оцінки акустичних характеристик

ПФМ – повний формантний метод

ФМ – формантний метод

ЦАП – цифроаналоговий перетворювач

ШГ - штучна голова

ШФМ – швидкий формантний метод

ІРП – імпульсна реакція приміщення

ІХ – імпульсна характеристика

ВСТУП

Актуальність роботи. Дисертаційна робота направлена на дослідження методів та систем оцінки та корекції акустичних характеристик приміщень для проведення мовленнєвих заходів, а також на автоматизацію налаштувань систем звукопідсилення у приміщеннях, що призначені для мовленнєвих та музичних заходів. На сьогоднішній день на території України існує досить велика кількість мультифункціональних технічно оснащених приміщень для заходів будь-якого формату, де на першому місці стоїть візуальна складова, тобто те, як ці зали виглядають. Недоліком цих залів є те, що досить мало уваги приділяється акустичній складовій. Між тим, акустична складова будь-якого залу першою чергою потрібна для передачі інформації зі сцени до слухачів. Для цього необхідно забезпечити відповідну якість акустичних сигналів (музики, мовлення) в місцях, де знаходяться слухачі. Одним із важливих показників такої якості є розбірливість мовлення та чіткість звучання музики. Також важливо забезпечити інші акустичні характеристики приміщення, такі як час реверберації, частотна характеристика, просторовість звучання тощо. Тому питання оцінки розбірливості мови у приміщенні та аналізу акустичних характеристик приміщення є актуальним. Однак після визначення акустичних характеристик приміщення майже завжди постає питання про їх корекцію, тому ще одним важливим напрямом досліджень є корекція акустичних характеристик приміщень.

Значний внесок у дослідження та розв'язання задач аналізу характеристик приміщення та оцінки розбірливості мови, а також методів корекції акустичних характеристик приміщення зробили такі іноземні та вітчизняні вчені, як Хоутгаст Т, Стінкен Г, Кулард Дж, Белоу Г, Марино К., Бредлі Дж, Чеччі С, Банк Б, Патеро Т, Покровський Н.Б, Дідковський В.С., Гавриленко А.В. Проте аналіз праць, що стосується вибору методу оцінки розбірливості мови, аналізу та корекції акустичних характеристик приміщення свідчить про існування низки питань стосовно потенційних можливостей методів оцінки

розбірливості мови в приміщеннях. Зокрема, це питання вибору кращого методу із врахуванням наявності певної апіорної інформації про завадові умови у приміщенні та про вимоги до точності та швидкості акустичних вимірювань. Крім того, наявні методи не дають чіткої відповіді стосовно вибору кращого методу локалізації відбиттів, також слід зазначити, що недостатньо уваги приділяється частотній корекції акустики приміщень у реальному часі.

Таким чином, розробка нових та вдосконалення існуючих методів підвищення точності систем оцінювання та корекції акустичних характеристик приміщень в умовах негативної дії шуму та реверберації, відносно простих для практичної реалізації, є **актуальною** науково-технічною задачею, що має важливе прикладне значення. Цим визначаються актуальність та практичне значення теми дисертаційного дослідження.

Мета та задачі дослідження. *Метою дисертаційної роботи* є пошук шляхів підвищення точності та швидкості оцінювання акустичних характеристик приміщень в умовах негативної дії шуму та реверберації, а також шляхів корекції цих характеристик, що забезпечують можливість подальшого створення автоматизованих систем акустичної експертизи та корекції приміщень.

Об'єктом дослідження є процеси оцінювання та корекції акустичних параметрів приміщень, призначених для мовленнєвих та музичних презентацій.

Предметом дослідження є методи оцінки розбірливості мовлення в приміщеннях в умовах дії шумової та ревербераційної завад, а також методи корекції акустичних характеристик приміщень.

Для досягнення поставленої мети необхідно вирішити наступні завдання:

1. Виконати аналітичний огляд сучасних методів оцінки розбірливості мовлення в приміщеннях, а також методів корекції акустичних характеристик приміщення.
2. Виконати порівняння об'єктивних методів оцінки розбірливості мовлення за точністю та швидкістю вимірювань, з метою подальшої автоматизації акустичних вимірювань в приміщеннях.
3. Визначити принципову можливість створення апаратно-програмної системи «Штучна голова» із використанням акустичної апаратури непрофесійного рівня для подальшого використання при оцінюванні бінауральних імпульсних характеристик приміщення.
4. Виконати дослідження алгоритмів програмно-апаратного комплексу для визначення напрямку прибуття ранніх відбиттів звуку у приміщенні.
5. Виконати дослідження алгоритмів програмно-апаратного комплексу для корекції амплітудно-частотної характеристики приміщення в реальному часі.

Методи дослідження. Для досягнення поставленої мети використано методи аналітичного та комп'ютерного моделювання, а також методи експериментального оцінювання та корекції акустичних характеристик приміщення.

Наукова новизна отриманих результатів.

У дисертації представлено наступні наукові результати:

1. Вперше виконано порівняння точності оцінювання індексу розбірливості мовлення в умовах дії шуму та реверберації із використанням повного традиційного методу та повного формантно-модуляційного методу для широких діапазонів значень відношення сигнал-шум та тривалості тестових сигналів.
2. Вперше виконано порівняння швидкого та повного варіантів формантно-модуляційного методу оцінювання індексу

розбірливості мовлення в умовах дії шуму та реверберації для широких діапазонів значень відношення сигнал-шум та тривалості тестових сигналів.

3. Вперше запропоновано багатокомпонентний тестовий сигнал та алгоритми його обробки, на основі яких виконано порівняння двох видів непрямого методу оцінки індексу передачі мовлення із прямим методом.
4. Уточнено попередні висновки про те, що формантний метод оцінювання індексу передачі мовлення є найбільш точним та швидким в умовах переважної дії шумової завади.
5. Уточнено умови використання апаратури непрофесійного рівня в апаратно-програмному комплексі «Штучна голова», призначеного для оцінки бінауральних імпульсних характеристик приміщення.
6. Вдосконалено та експериментально перевірено дієвість методу багатоканальної компресії сигналу, спотвореного реверберацією, в задачі послаблення нерівномірності амплітудно-частотної характеристики приміщення.
7. Показано можливість визначення часу реверберації та локалізації напрямків ранніх відбиттів звуку шляхом просторового акустичного сканування простору.

Особистий внесок здобувача. Усі результати, наведені у дисертаційній роботі і винесені на захист, отримано за активної участі здобувача та опубліковано у спеціалізованих фахових виданнях.

У роботі [1] здобувачем особисто виконано наступне: розроблений алгоритм корекції частотної характеристики приміщення з використанням багатоканального компресору, організація та проведення експериментальних досліджень, обробка результатів. У роботі [2], опублікованій в співавторстві, здобувач особисто приймав участь у проведенні вимірювань та експериментів, також виконав аналіз результатів експериментальних досліджень. У роботі [3],

опублікований в співавторстві, здобувач особисто приймав участь у організації та проведенні експерименту двоканальної оцінки розбірливості мови, спотвореної реверберацією. У роботі [4], опублікований в співавторстві, здобувач особисто розробив методику запису сигналів та організував проведення лабораторних та натурних експериментів, а також приймав активну участь в обробці результатів. У роботі [5], опублікований у співавторстві, здобувач особисто взяв участь у створенні багатоконпонентного сигналу для порівняння прямого та непрямого методів вимірювання індексу передачі мовлення. У роботі [6] здобувач особисто взяв участь суб'єктивному оцінюванні розбірливості мови на тлі шуму та реверберації. У роботі [7] здобувач особисто приймав участь у проведенні записів звукосполучень, а також обробці результатів дослідження. У роботі [8] здобувач особисто приймав участь в обробці та трактуванні результатів дослідження. У роботі [9], опублікований в співавторстві, здобувачем особисто виконано наступне: взято участь в оформленні патенту на корисну модель.

Практичне значення одержаних результатів. Практичне значення отриманих результатів полягає у встановленні принципової можливості використання повного та швидкого формантно-модуляційних методів оцінювання STI замість традиційного повного модуляційного методу; встановленні залежності точності формантного, повного модуляційного, повного формантно-модуляційного та швидкого формантно-модуляційного методів оцінювання STI від відношення сигнал-шум та від тривалості тестового сигналу; виконанні коректного порівняння, за точністю вимірювань, прямого та непрямого методів оцінки STI; встановленні принципової можливості використання, з метою акустичної експертизи приміщень, апаратури непрофесійного рівня в апаратно-програмному комплексі «Штучна голова»; встановленні працездатності системи корекції сигналу, спотвореного реверберацією, на основі методу багатоканальної компресії, що дозволяє

послабити нерівномірність амплітудно-частотної характеристики приміщення.

Результати досліджень дозволяють створити апаратно-програмні комплекси, що за своєю вартістю є дешевшими за аналоги та дозволяють користувачу більш точно та якісно виконувати поставлені задачі, шляхом використання сучасних методів обробки сигналу та підвищення рівня автоматизації акустичної експертизи та корекції.

Зв'язок роботи з науковими програмами, планами, темами.

Викладені у дисертації нові теоретичні та практичні результати досліджень можуть використовуватися при розробці та експлуатації систем акустичної експертизи та корекції приміщень, а також у освітньому процесі кафедри акустичних та мультимедійних електронних систем за спеціальністю 171 Електроніка, в освітній програмі “Акустичні електронні системи та технології обробки акустичної інформації” Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”.

Апробація результатів дисертації. Матеріали дисертаційних досліджень обговорювалися на міжнародних конференціях:

1. IEEE 7th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), 24-27 жовтня 2023, Київ, Україна.
2. IEEE 13th International Conference on Electronics and Information Technologies (ELIT), 26-28 вересня 2023, Львів, Україна.
3. IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO), 22-24 квітня 2020, Київ, Україна.
4. XIV Міжнародна науково-технічна конференція молодих вчених «ЕЛЕКТРОНІКА-2021» 21-22 квітня 2021, Київ, Україна.
5. IEEE 42th International Conference on Electronics and Nanotechnology (ELNANO), 13-15 травня 2024, Київ, Україна.

Публікації. За результатами досліджень опубліковано 8 наукових публікацій за спеціальністю 171 Електроніка, галузь знань Електроніка та телекомунікації, з них 7 статей в наукових фахових виданнях України з категорії Б, 0 стаття в науковому фаховому виданні України з категорії А, 1 стаття прийнята до друку в періодичному науковому виданні, що входить до WoS або Scopus), 1 патент України на корисну модель, 4 тези доповідей у збірниках матеріалів міжнародних наукових та науково-практичних конференцій, організованих Українською секцією IEEE, 1 теза доповіді у збірнику матеріалів міжнародної науково-технічної конференції молодих вчених, організованою факультетом електроніки НТУУ «КПІ ім. Ігоря Сікорського».

Структура та обсяг дисертації. Робота складається зі вступу, чотирьох розділів, списку використаних джерел із 115 найменувань. Робота містить 60 рисунків та 10 таблиць. Загальний обсяг дисертаційної роботи складає 138 сторінок.

РОЗДІЛ 1. СУЧАСНИЙ СТАН ОЦІНЮВАННЯ РОЗБІРЛИВОСТІ МОВЛЕННЯ У ПРИМІЩЕННІ ТА КОРЕКЦІЇ АКУСТИКИ ПРИМІЩЕННЯ

1.1 Методи оцінювання розбірливості мовлення у приміщеннях

Розбірливістю мови, згідно міжнародного стандарту ISO/TR 4870 [10], вважається ступінь, з якою мова може бути зрозуміла (розшифрована) слухачами. Мірою розбірливості є ступінь правильного сприйняття елементів мови – величина, яка визначається як співвідношення числа правильно прийнятих по тракту, який досліджується, елементів мови (звуків, складів, слів і фраз) до загальної кількості переданих елементів та виражається у відсотках [11]. Відповідно до відсоткового значення розрізняють наступні види розбірливості: фонемна, слогова, словесна і фразова, які пов'язані одна з одною, і можуть бути переведені одна в іншу. Слід зазначити, що оцінювання розбірливості мовлення може виконуватися суб'єктивними або об'єктивними методами.

Для регламентації суб'єктивних артикуляційних випробувань існують стандарти: ГОСТ 25902-83 [12], ГОСТ 51061-97 [13], міжнародні стандарти ISO/TR4870 [10], IEC 60268-16:2020 [14]. Для реалізації вимірювань у всіх вище зазначених стандартах необхідно забезпечити досить складну підготовку та задіяти досить велику кількість людей і обладнання. Така процедура є досить вартісною з точки зору бюджету та часу, тому в останні роки велика увага приділялася автоматизації суб'єктивного оцінювання розбірливості, що дозволило ввести в практику цілий ряд нових досить ефективних комп'ютерних методів розрахунку розбірливості мови в різних заводських умовах.

Об'єктивним (інструментально-розрахунковим) методам властива висока оперативність, але ці методи мають обмеження принципового характеру (вплив реверберації, нелінійні спотворення сигналу), які звужують область їх застосування. На рис. 1.1 показана загальна схема класифікації методів оцінки розбірливості мови [15].

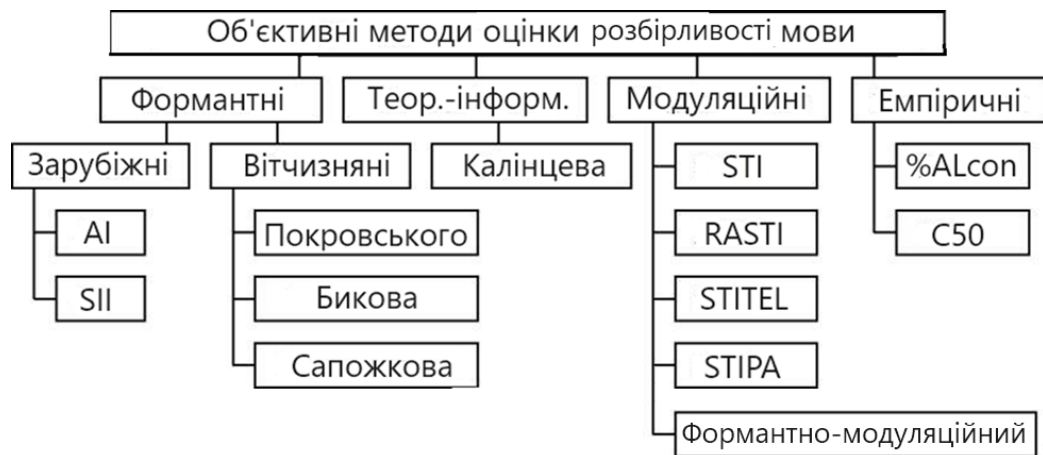


Рис. 1.1. Класифікація об'єктивних методів оцінки розбірливості мови

На сьогоднішній день на практиці частіше всього використовують формантний та модуляційний методи оцінювання розбірливості. В роботі [16] показано, що співвідношення сигнал-шум може використовуватися для оцінки розбірливості мови. Однак формантний метод вигідно відрізняється від методу парціальних відношень сигнал/шум тим, що в ролі показника якості використовується скалярна величина – формантна розбірливість [15], що представляє собою середню ймовірність сприйняття формант на тлі шуму. Важливою особливістю формантного методу є те, що показником якості є ймовірність. Використовуючи інформацію про статистичні властивості мови, середню ймовірність сприйняття формант потім можна перерахувати в ймовірність правильного розпізнавання звуків, складів, слів і фраз [17].

Це зручно, оскільки при використанні суб'єктивних методів оцінки розбірливості мови показником якості також служить ймовірність правильного розпізнавання звукосполучень, складів, слів або фраз. Отже, для зіставлення якості вирішення однієї і тієї ж задачі об'єктивним та суб'єктивним методами досить порівняти дві оцінки ймовірностей.

До модуляційного методу оцінки розбірливості мови відносять інструментальний метод оцінки розбірливості STI (speech transmission index) – індекс передачі мовлення [18]. Основою повного модуляційного методу є алгоритм, викладений у роботі [19]. Перед вимірюваннями STI формується

14 тестових сигналів:

$$x_i(t) = \xi(t) \sqrt{f_i(t)}, \quad f_i(t) = 1 + \sin 2\pi F_i t, \quad i = \overline{1, 14},$$

$$F_i = 0,63; 0,8; 1; 1,25; 1,6; 2; 2,5; 3,15; 4; 5; 6,3; 8; 10; 12,5 \text{ Гц.}$$

де $\xi(t)$ - стаціонарний шум із спектром мовлення, $f_i(t)$ - функція модуляції, F_i - частота модуляції. При вимірюваннях STI сигнали $x_i(t)$ по черзі випромінюються джерелом звуку, розташованим в точці, де зазвичай знаходиться диктор.

В точці, де знаходиться слухач, для кожного $x_i(t)$ формується сигнал $y_i(t) = x_i(t) \otimes h(t) + n(t)$, де \otimes - символ згортки, $h(t)$ - імпульсна характеристика приміщення, $n(t)$ - шумова завада. Далі сигнали $y_i(t)$ піддаються фільтрації гребінкою із семи октавних фільтрів, в результаті чого формується набір із 98 сигналів $y_{k,i}(t)$, $k = \overline{1, 7}$, $i = \overline{1, 14}$.

Індекс розбірливості мовлення STI обчислюється згідно виразу:

$$STI = \sum_{k=1}^7 \alpha_k \cdot MTI_k - \sum_{k=1}^6 \beta_k \cdot \sqrt{MTI_k \cdot MTI_{k+1}},$$

де MTI_k - індекс передачі модуляції в k -й смузї частот, α_k та β_k - вагові коефіцієнти.

Значення MTI_k обчислюють згідно виразів

$$MTI_k = \frac{1}{14} \sum_{i=1}^{14} TI_{k,i},$$

$$TI_{k,i} = \begin{cases} \frac{SNR_{eff\ k,i} + 15}{30}, & -15 < SNR_{eff\ k,i} < 15; \\ 0, & SNR_{eff\ k,i} \leq -15; \\ 1, & SNR_{eff\ k,i} \geq 15; \end{cases},$$

де $TI_{k,i}$ - індекс передачі, $SNR_{eff\ k,i}$ - ефективне відношення сигнал-шум,

$$SNR_{eff\ k,i} = 10 \lg \frac{m_{k,i}}{1 - m_{k,i}},$$

$$m_{k,i} = \frac{2|A_{k,i}(F_i)|}{|A_{k,i}(0)|}, \quad A_{k,i}(F_i) = \frac{1}{T} \int_0^T y_{k,i}^2(t) e^{-j2\pi F_i t} dt$$

де $|\cdot|$ - символ модуля, \square - символ оцінки.

Слід зазначити, що модуляційний метод оцінювання розбірливості мовлення [19] є сьогодні найбільш поширеним завдяки своїй універсальності. На відміну від формантного методу [20], модуляційний метод дозволяє враховувати не тільки дію шуму, але й дію реверберації на розбірливість мовлення. Запропонований в [21] формантно-модуляційний (ФМ) метод є варіантом модуляційного методу, тому він також є універсальним. Перевагою ФМ методу є можливість обчислення як індексу STI, так і індексу артикуляції.

Оскільки суттєвим недоліком повного модуляційного методу оцінювання індексу передачі мовлення STI є значна тривалість процедури вимірювань, на практиці замість повного модуляційного методу частіше використовують спрощені «швидкі» методи, такі як RASTI (STI для приміщень), STIPA (STI для систем звукопідсилення) та STITEL (STI для телекомунікаційних систем) [22].

ФМ метод також існує у двох варіантах, це повний та швидкий методи. Запропонований в [23] алгоритм ФМ методу відрізняється від наведеного вище алгоритму модуляційного методу лише тим, що, після обчислення $SNR_{eff\ k,i}$ згідно вище зазначених формул, індекс передачі модуляції MTI_k обчислюють згідно виразам:

$$MTI_k = \begin{cases} \frac{E_k + 15}{30}, & -15 < E_k < 15; \\ 0, & E_k \leq -15; \\ 1, & E_k \geq 15; \end{cases},$$

$$E_k = \frac{1}{14} \sum_{i=1}^{14} SNR_{eff\ k,i}, \text{ після чого обчислюють } STI.$$

Можна бачити, що відмінність між ФМ методом та модуляційним методом полягає у переміні місцями лінійної та нелінійної операцій. Дійсно, в

модуляційному методі значення $SNR_{eff\ k,i}$ спершу піддають нелінійному перетворенню при обчисленні $TI_{k,i}$, а потім усереднюють згідно MTI_k . В ФМ методі робиться навпаки: значення $SNR_{eff\ k,i}$ спершу усереднюють, обчислюючи E_k , а потім піддають нелінійному перетворенню при обчисленні MTI_k .

Корисність такої зміни порядку операцій полягає в можливості обчислити не тільки STI, але й індекс артикуляції:

$$A = \sum_{k=1}^7 p_k \cdot P_k(E_k),$$

де p_k - імовірність знаходження формант в k -й смузі частот; $P_k(E_k)$ - коефіцієнт сприйняття, $E_k = 10 \lg \frac{D_{sk}}{D_{nk}}$ - відношення сигнал-шум в k -й смузі частот, D_{sk} та D_{nk} - дисперсії сигналу та шуму в k -й смузі частот, відповідно. Зауважимо, що свою назву ФМ метод отримав саме через можливість обчислювати як STI, так і індекс артикуляції.

Враховуючи кусочно-лінійний характер залежностей $TI_{k,i}$ та MTI_k , можна очікувати, що в певній околиці $SNR_{eff\ k,i} = 0$ дБ середні значення та дисперсії оцінок STI, отриманих повними модуляційним та ФМ методами, будуть близькими. Разом із тим, за межами цього діапазону відмінності можуть бути значними. Оскільки на сьогодні це питання не є вивченим, одна із задач даної роботи полягає в заповненні цієї прогалини.

Розплатою за використання швидких методів є зменшення точності оцінювання STI, що показано в роботі [23]. З огляду на це, в окремих сучасних апаратно-програмних вимірювальних системах, таких як NTi Audio [24], існує можливість виконати вимірювання STI повним модуляційним методом.

Зазвичай при вимірюваннях STI повним модуляційним методом обирають тривалість тестового сигналу близькою до 10-15 секунд [25]. Проте в літературних джерелах дуже важко знайти обґрунтування такого вибору у

вигляді залежності похибки вимірювань STI від тривалості T тестових сигналів. Для ФМ методу такі залежності також є невідомими. Крім того, невідомим є співвідношення похибок оцінювання STI повними модуляційним та формантно-модуляційним методами.

Повний ФМ (ПФМ) метод, як і повний модуляційний метод [23], потребує досить багато часу для вимірювань STI . Наприклад, при використанні тестового сигналу тривалістю $T=10$ с процедура вимірювань модуляційним методом потребує 15 хвилин [22]. Такий значний час вимірювань пояснюється використанням набору із 14 тестових сигналів, що випромінюються, реєструються та обробляються послідовно у часі.

Використання швидкого ФМ (ШФМ) методу потребує набагато менше часу, оскільки при вимірюваннях STI випромінюється лише один тестовий полігармонічний сигнал [23].

Порівняння точності оцінок розбірливості мовлення для формантного, ПФМ та ШФМ методів, виконані в [26] для випадку переважної дії шумової завади, показали, що при фіксованій тривалості T тестового сигналу формантний метод є найбільш точним, тоді як ШФМ метод є найменш точним. Недоліком представлених в [18] результатів досліджень, виконаних шляхом комп'ютерного моделювання, є обмеженість випадком $T = 16$ с.

Іншим недоліком є відсутність інформації стосовно зміщення оцінок розбірливості мовлення, отриманих ПФМ та ШФМ методами. Цей недолік частково усунено в [20], де показано, що при малих відношеннях сигнал-шум ($SNR < -15$ дБ) та нетривалих тестових сигналах ($T \leq 16$ с) оцінка розбірливості мовлення, отримана ШФМ методом, є помітно завищеною. Зменшити таке зміщення можна шляхом збільшення T , проте залежність величини зміщення від T не була простежена у [20]. Що стосується ПФМ методу, для нього зміщення не оцінювалося зовсім.

Нарешті, при порівнянні ПФМ та ШФМ методів в [27] та [23] не було розглянуто ситуацію сумісної дії шуму та реверберації на розбірливість мовлення.

1.2 Аналіз ранніх відбиттів звуку у приміщенні

Глен Белоу у роботі [30] описав методи оцінки акустичних властивостей приміщення як суб'єктивно, за допомогою слухової системи людини, так і об'єктивно, за допомогою певних пристроїв.

Наразі відомо, що акустичні характеристики приміщення впливають на сприйняття звуку людиною досить суттєво [31 - 35], однак в даних роботах досліджується саме мовленнєвий вплив, що не скажеш про [36, 37] у яких ще сказано про вплив музичної інформації. Автори праць [41, 42, 35] досліджували вплив шуму і реверберації у шкільних аудиторіях на розбірливість мови, вони експериментально показали, що шумова завада сильніше впливає на розбірливість ніж ревербераційна. Це пов'язано із тим, що шум від людей, що сидять поряд із слухачем, впливає на сприйняття інформації слухачем більше, оскільки джерело шуму розташовано близько до слухача.

Також у даних роботах представлено оцінки залежності розбірливості мови від віку слухачів, які свідчать, що для забезпечення словесної розбірливості у 95% для школярів віком 6 років відношення сигнал-шум має бути не нижче 16 дБ, тоді як для студентів - не нижче 1 дБ.

К. Марино у роботі [32] оцінювала характеристики великого лекційного залу шляхом вимірювань часу реверберації T_{60} , часу приходу ранніх відбиттів – EDT (early decay time), індексу розбірливості мови STI та коефіцієнту ясності C_{80} . Ці дослідження виконувалися шляхом моделювання даного залу у програмі ODEON [50]. Оцінка якості приміщення показала, що час реверберації є нижчим, ніж належить для приміщень таких розмірів, тому автором було запропоновано використати системи звукопідсилення для підвищення рівня прямого звуку. В роботі Р. Юсефа [31] показано, що використання сучасних технічних аудіо та візуальних засобів у приміщеннях, не призначених для цього, може привести до погіршеного сприйняття інформації.

Недоліком даних робіт можна зазначити те, що чітко не визначено питання корисності акустичних систем для покращення розбірливості мови. Імпульсна характеристика приміщення є однією із важливих характеристик приміщення, що містить інформацію стосовно якості звуку та розбірливості мовлення.

Імпульсна характеристика (ІХ) приміщення залежить від геометричних параметрів та акустичного оформлення. В роботі [51] зазначено, що для одного і того самого приміщення існує безліч ІХ, форма яких залежить від розташування у просторі приймача та джерела звуку. За словами авторів роботи [51], приміщення є лінійною системою, тому імпульсною характеристикою можна повністю описати, як змінюється звуковий сигнал при проходженні від джерела до слухача. Припустимо, що шлях передачі звуку є лінійною системою, тоді співвідношення спектрів вхідного та вихідного сигналу буде наступним:

$$Y(j\omega) = H(j\omega)X(j\omega), \text{ де}$$

$Y(j\omega)$ - спектр вихідного сигналу, $H(j\omega)$ - передавальна функція лінійної системи, $X(j\omega)$ - спектр вхідного сигналу. Імпульсна характеристика системи пов'язана із частотною характеристикою перетворенням Фур'є:

$$h(\tau) = F^{-1}[H(j\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega) e^{j\omega\tau} d\omega.$$

Павло Малекі у роботі [52] визначив поняття точкової імпульсної характеристики приміщення. Дане поняття може описати часово-просторові властивості приміщення, а саме:

$$h_{\theta\phi}(t) = \left[\sum_{p=0}^p h_{(p\theta\phi)}(t) \right], \text{ де}$$

p - номер відбиття звуку, ϕ - нахил траєкторії відбиття, θ - азимут.

У роботах [53, 54] була розглянута так звана модальна імпульсна характеристика, суть якої полягає у тому, що нормальними модами є власні

резонансні частоти кімнати. Коли у приміщенні починає звучати джерело звуку, то воно збуджує n -ну кількість нормальних мод, при цьому, коли джерело вимикається, моди продовжують резонувати накопиченою енергією, й кожна з утворених мод розпадається зі швидкістю, що визначена певною константою.

У дисертації Себастьян Гарі [55] схематично зображена імпульсна характеристика, яка має вигляд як на рис. 1.2.1.

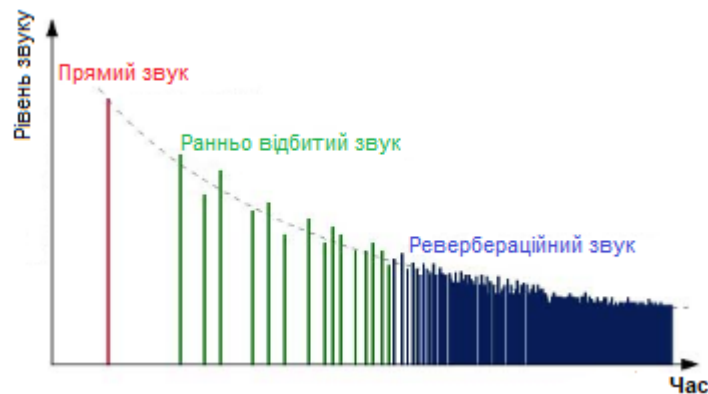


Рис. 1.2.1. Представлення імпульсної характеристики [55]

Із рис 1.2.1 добре видно, що перші відбиття після прямого звуку є дискретні і вони чітко розділені між собою у часі. У роботі [56], 1951-го року, Хельмут Хаас зазначив, що ранні відбиття сприймаються як одне ціле за умови, що вони надходять до слухача не пізніше 50 мс після приходу прямого звуку. Тоді дані відбиття є корисними, оскільки вони підсилюють прямий звук, і слухачем це сприймається як одне ціле. Також на даному рисунку видно, що імпульси з часом частішають і формують ревербераційний хвіст («пізня реверберація»).

Використовуючи оцінку імпульсної характеристики приміщення, можливо оцінити чіткість звуку для мовленнєвих та музичних сигналів у різних точках приміщення. Співвідношенням енергії ранньої та пізньої реверберації можна описати індекс чіткості мовлення C_{50} та індекс прозорості музичного звуку C_{80} :

$$C_{50} = 10 \lg \frac{E_{50}}{L_{50}}, \quad C_{80} = 10 \lg \frac{E_{80}}{L_{80}}.$$

Людські вуха можна ототожнити з мікрофонами, оскільки мікрофон перетворює механічні коливання мембрани у електричні сигнали так само як і барабанна перетинка, яка передає їх до мозку. Ідея імітувати людський слух досить широко вивчена [57 - 59]. Однак зараз широку розповсюдженість отримали імітації німецької фірми Neumann [60]. Дана мікрофонна система досить точно імітує людське вухо. Дані мікрофонні системи широко використовуються для різного роду вимірювань [61 - 64]. Зазначені мікрофонні системи не мають широкого використання у сфері шоубізнесу, а використовуються саме для оцінки імпульсних характеристик приміщення.

Як було сказано вище, одним із основних показників акустичних властивостей приміщення вважається час його реверберації [65, 66]. Слід зазначити, що для більш повної оцінки якості акустики необхідно знати характеристики формування звукового процесу в приміщенні (відносні рівні звукового тиску, відношення енергії раннього та пізнього відбиття (відлуння), вторинні енергетичні складові та рівні фонового шуму).

У стандарті [67] подано визначення та опис методів оцінювання часу реверберації – це метод переривання шуму (interrupted noise method), метод інтегрованої імпульсної перехідної характеристики (integrated impulse response method) та метод випромінювання короткого звукового імпульсу.

Зауважимо, що «на практиці неможливо створити та випромінювати справжню дельта-функцію Дірака, проте з метою вимірювання короткочасні нестационарні звуки (наприклад, звуки пострілів) можуть виявитися її прийнятною апроксимацією. Альтернативним методом вимірювань є застосування сигналу типу періодичної псевдовипадкової послідовності максимальної довжини або іншого детермінованого сигналу з плоским спектром, як у синусоїдального сигналу з частотою, що лінійно змінюється, і зворотним перетворенням вимірюного відгуку на імпульсну перехідну характеристику» [67]. На початку процесу реверберації криві спаду несуть корисну інформацію про енергію ранніх, а це, у свою чергу, наштовхує на думку про можливість локалізації відбиваючих поверхонь та можливість їх

подальшої реконструкції для покращення акустичних характеристик приміщення.

У роботі [68] представлено результати вимірювання ступеня звукопоглинання різних матеріалів, які свідчать, що різне розміщення акустичного матеріалу в ревербераційній кімнаті дійсно впливає на звукопоглинання на частотах, вищих за 500 Гц, проте питання, звідки надходить відбитий звук на мікрофон, залишається відкритим.

У вищезазначеній роботі [68] вплив ранніх відбиттів на розбірливість мовлення досліджувався на основі двох спрощених моделей раннього відбиття звуку в кімнаті. Перша модель дозволяє оцінити вплив детермінованих відбиттів, розташованих на інтервалі часу (0-50) мс, на розбірливість мовлення. Друга модель дозволяє дослідити вплив сукупності випадкових відбиттів на одному і тому ж інтервалі часу. Відмічена суттєва залежність розбірливості мовлення від місцезнаходження відбивачів та їх звукопоглинання. Автори [69], вимірюючи звукопоглинання конструктивних елементів, зазначають, що оцінки акустичних характеристик приміщення залежать не тільки від розташування конструктивних елементів, але й від місця установки мікрофонів. Рекомендується враховувати розташування конструктивних елементів та вносити відповідні поправки в оцінку звукопоглинання.

Звукоізоляція між сусідніми житловими приміщеннями [70] є однією з важливих характеристик внутрішньої екологічної ефективності квартири. В зазначеній роботі моделюються процеси передачі звуку по акустичним та вібраційним каналам в низькочастотному звуковому діапазоні. Вказується на актуальність розробки методів і пристроїв їх виявлення з метою підвищення звукоізоляції об'єктів на етапі підготовки до експлуатації.

Недоліком вищезазначених досліджень є те, що автори не піднімають питання локалізації відбиттів, а рекомендації по звукоізоляції можуть привести до хибних висновків, що потягне за собою зайві фінансові витрати.

1.3 Корекція акустики приміщення

Наразі велика кількість вітчизняних звукорежисерів, які забезпечують проведення концертів, налаштовують систему суб'єктивно, «під свій смак». Пов'язано це з тим, що їх досвід у даній сфері досить великий і вони знають, як система має «звучати», однак майже очевидно, що такий суб'єктивний підхід не є єдиним можливим та найкращим. Г. Белоу у роботі [30] зазначає, що корекція частотної характеристики (еквалізація) акустичної системи є вирішенням задачі забезпечення якісного звучання, однак вплив приміщення описаний досить дотично, що не дозволяє отримати відповідь, як діяти в нетипових ситуаціях (на кшталт великого часу реверберації). Розглянемо більш детально функцію вирівнювання (еквалізації) відгуку (частотної характеристики) приміщення (RRE - Room Response Equalization), оскільки така частотна корекція наявна у більшості акустичних систем. Стефанія Чеччі у роботі [71] зробила ґрунтовний огляд методів вирівнювання амплітудно-частотної характеристики приміщення шляхом частотної корекції. Автор досить детально розкрила, які методи ефективніше використовувати в обмеженій зоні навколо точки вимірювання, а які є кращими в більш широкому об'ємі простору. Також вона зазначила, що еквалайзери, які адаптуються до змін у середовищі, більш придатні до реальних практичних задач, бо дозволяють використовувати даний тип систем корекції звуку в реальному часі. Слід зазначити, що важливим є виділення завад звукового діапазону частот, які викликані особливостями побудови цифрових еквалайзерів. Стівен Нелі та Джонт Аллен у роботі [72] вивчали можливість обернення амплітудно-частотної характеристики (Room Transfer Function - RTF) та його наслідки. Розглядаючи передатну функцію модельованого приміщення, було показано, що якщо стіни мають низьку відбивну здатність, тоді передатна функція є мінімально-фазною. Однак для приміщень з більшими коефіцієнтами відбиття від стін імпульсна характеристика не є мінімально-фазною. Авторами було запропоновано вирівняти амплітудну

характеристику та мінімально-фазну характеристику приміщення шляхом добутку мінімальної-фазної системи $Hm(\omega)$ і стабільного всепропускаючого фільтра $A(\omega)$:

$$H(\omega) = Hm(\omega) \cdot A(\omega)$$

Коефіцієнти фільтра еквалізації можуть бути обчислені шляхом зворотного Фур'є-перетворення від спектру, оберненого до $Hm(\omega)$. Слухаючи результат вирівнювання, Нелі та Аллен зазначили, що "Ефект кімнати був видалений, але у фоновому режимі лунав тон, схожий на дзвінок". У роботі [73] розглянуто прямий та непрямий способи реалізації еквалайзера (Рис. 1.3.1).

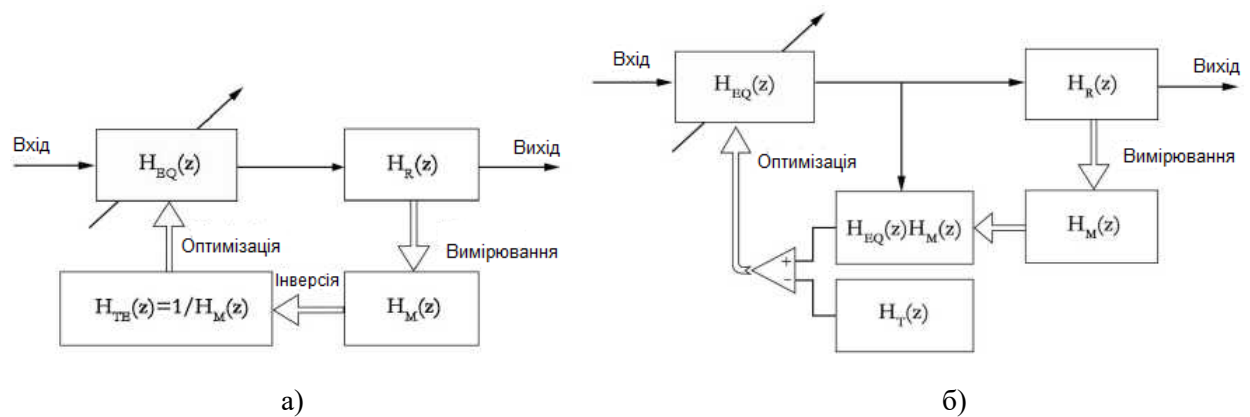


Рис. 1.3.1. Непрямий (а) та прямий (б) способи реалізації еквалайзера, де $H_{EQ}(z)$ являє собою фільтр еквалізації, $H_R(z)$ - канал відтворення, H_M - виміряна імпульсна характеристика, а $H_T(z)$ і $H_{TE}(z)$ - цільові функції [73]

Як раніше вказувалося, еквалайзери можуть бути класифіковані як мінімально-фазові та змішано-фазові. Інша цікава класифікація методів проектування еквалайзера була також подана в [73], де поділено методи проектування еквалайзера на «непрямі» та «прямі». Для еквалізації відгуку приміщення використовувались фільтри Кауца, застосовуючи неоднорідну частотну роздільну здатність цих фільтрів. Вони застосовувались як для мінімально-фазової, так і для змішано-фазової еквалізації, використовуючи як нерухомі полюси, так і налаштовані полюси. У роботах [74 - 75] автор використовував фільтри Кауца при фіксованому полюсному підході, дані фільтри були реалізовані у вигляді гребінки фільтрів із секціями другого

порядку. Даний вид фільтрації має переваги, оскільки навантаження на процесор, що обробляє дані, стає меншим. У роботі [76] було показано еквівалентність паралельних фільтрів і фільтрів Кауца, а також наведено формули для взаємного перерахунку параметрів цих структур.

У роботі [77] запропоновано модель передатної характеристики кімнати із використанням загальних акустичних полюсів, що відповідають резонансам кімнати. Загальні акустичні полюси оцінюються як загальні значення полюсів багатьох низькочастотних передатних характеристик, оцінених при ситуаціях, коли джерело і приймач були розміщені у різних місцях. Полюси обчислювались за допомогою лінійно прогностичної моделі відгуку приміщення та оцінені двома методами. Перший метод полягав у використанні методу найменших квадратів, припускаючи, що виміряні передатні характеристики мають однакові коефіцієнти для лінійно прогностичної моделі. Другий метод полягав в усередненні коефіцієнтів лінійного прогнозування, оцінених з кожної виміряної переданої характеристики. Така модель названа авторами моделлю із загальними акустичними полюсами та нулями, оскільки це нульово-полюсна модель, утворена загальними акустичними полюсами та нулями, що забезпечуються коефіцієнтами ковзного середнього. У роботі [78] запропоновано багатоточковий фільтр вирівнювання, що використовує загальні акустичні полюси. Спільні акустичні полюси знову оцінюються як загальні коефіцієнти лінійно прогностичної моделі при багаторазовому вимірюванні відгуку приміщення. Вирівнювання досягається за допомогою нерекурсивних фільтрів, що мають зворотні характеристики загальної акустичної полюсної функції. До недоліків даного методу можна віднести те, що фільтр вирівнювання є мінімально-фазним еквалайзером, який не може компенсувати провали частотної характеристики. Проте, такий фільтр може придушити загальні піки, спричинені резонансами в різних місцях, а також має низьку чутливість до змін положення приймача. Робота [79] виділяється тим, що до звичайного еквалайзера полюсів було

додано попередню обробку, яка полягала у пригніченні низькодобротних резонансів у всьому спектрі.

Б. Банк у роботах [74 - 76] зазначав, що компенсація загальних акустичних полюсів є досить вигідною при використанні гребінок фільтрів другого порядку, зокрема, таких, що мають роздільну здатність в логарифмічному масштабі частоти та можливість налаштування положень полюсів. Розглянуті вище методи еквалізації відгуку приміщення зарекомендували себе з кращої сторони, однак для їх реалізації необхідно використовувати зв'язку з декількох фільтрів, що негативно впливає на продуктивність системи. Сучасні апаратні DSP процесори у своєму арсеналі мають максимум 15 смуг, що достатньо для виправлення всіх недоліків приміщення при тому, що кінцева якість звучання буде мати мінімальні спотворення. Однак такі процесори досить вартісні і розглядати їх як інструмент системного інженера не раціонально. Також слід зазначити, що автори вищезазначених робіт розробляли методики корекції, недоліками яких є те, що вони нездатні працювати у реальному масштабі часу та не враховують змін акустичних властивостей приміщення в результаті змін наповненості приміщення.

В [80] запропоновано еквалайзер, що містив гребінку лінійно-фазних октавних фільтрів, а в [81] запропонований квазілінійний фазовий октавний графічний еквалайзер з низькою затримкою. Недоліком даних робіт є те, що ці еквалайзери не використовувались для корекції акустичних властивостей приміщення, а також те, що зміна амплітуди сигналу не є динамічною. Еквалайзер, частотна характеристика якого змінюється у часі, прийнято називати динамічним еквалайзером [82]. Еквалайзер, який поділяє вхідний аудіосигнал за допомогою кількох фільтрів, і сигнал від кожного фільтра проходить через власний компресор, називають багатоканальним компресором [82]. Реалізацію такого компресору описано у патенті [83], а структурну схему зображено на рис. 1.3.2. У роботах [85 - 87] описано

прилади, схожі до приладу патенту [83], однак вони використовувались для задач частотного прикрашення звуку. Динамічний компресор, параметри якого можуть керуватись зовнішніми джерелами (sidechain), широко застосовується у аудіо індустрії [85 - 87]. Структурну схему такого компресору зображено на рис. 1.3.3.

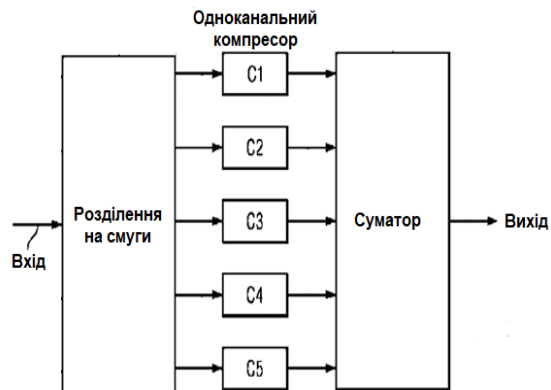


Рис. 1.3.2. Структурна схема динамічного компресору [83]

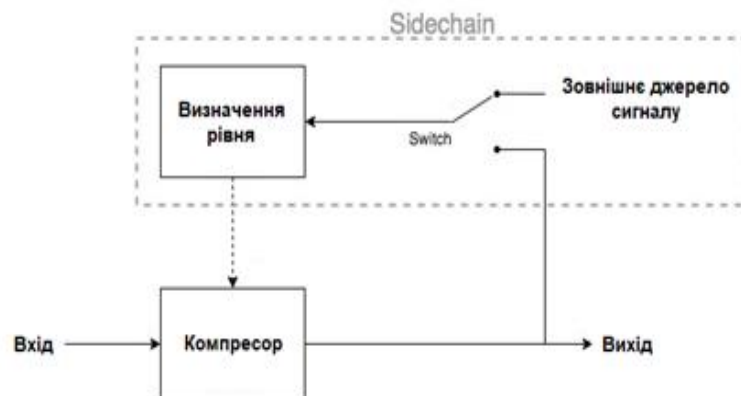


Рис. 1.3.3. Структурна схема компресору з керуючим сигналом

Дані компресори, як сказано у [87], зазвичай використовуються у студійній роботі. Також компресори можуть бути корисними при вирішенні проблеми частотних конфліктів у певній смузі частот [88]. Прикладом такого приладу є DBX DriveRack PA2 [89], що має функцію автоеквалайзера, яка дозволяє в автоматичному режимі підлаштовувати акустичну систему під приміщення. Принцип роботи такого еквалайзера схожий до принципу роботи динамічного еквалайзера. Проте недоліком такого підходу є те, що автоеквалайзер не змінює своїх параметрів в результаті зміни наповненості приміщення.

Висновки до розділу 1

1. На сьогодні в літературі надано недостатньо інформації стосовно залежності точності вимірювань індексу передачі мовлення від тривалості тестових сигналів для різних видів модуляційного методу, що утруднює обґрунтування вибору кращого методу оцінювання розбірливості мовлення в різних завадових умовах.
2. Формантний метод оцінки розбірливості мовлення переважає над модуляційними методами за точністю та швидкістю вимірювань в умовах дії шумової завади, проте ступінь такої переваги до останнього часу не було оцінено для різних тривалостей тестового сигналу та для різних значень відношення сигнал-шум.
3. Хоча непрямий метод оцінювання індексу передачі мовлення є швидшим за повний прямий метод, проте на сьогодні недостатньо досліджені два можливих різновиди непрямого методу.
4. Оцінка бінауральних імпульсних характеристик приміщення за допомогою системи "Штучна голова", що містить акустичну апаратуру непрофесійного рівня, є актуальним завданням, оскільки дозволяє суттєво зменшити вартість вимірювальної системи для акустичної експертизи приміщень.
5. Завдання локалізації напрямів ранніх відбиттів в приміщенні є актуальним, оскільки дозволяє уточнити час приходу ранніх відбиттів до слухача, проте на сьогодні це завдання досліджено не в достатній для практичного використання мірі.
6. Завдання частотної корекції приміщення є актуальним, оскільки дає можливість відкоригувати амплітудно-частотну характеристику приміщення. Проте на сьогодні питання створення простої та ефективної системи такого коригування в реальному часі, з урахуванням зміни наповненості приміщення, є недостатньо дослідженим.

РОЗДІЛ 2. ПОРІВНЯННЯ РІЗНОВИДІВ МОДУЛЯЦІЙНОГО МЕТОДУ ОЦІНЮВАННЯ РОЗБІРЛИВОСТІ МОВЛЕННЯ ЗА ТОЧНІСТЮ ТА ШВИДКІСТЮ ВИМІРЮВАНЬ

В даному розділі наведені результати досліджень порівняння різних методів оцінювання розбірливості мовлення. Метою дослідження є встановлення особливостей використання різних методів. Результатом порівняння є рекомендації щодо застосування методів в залежності від особливостей задачі, що розв'язується, та від завадових умов. Запропоновано використання багатокomпонентного тестового сигналу для порівняння прямого та непрямого методів оцінки індексу передачі мовлення, а також виконано порівняння прямого та непрямого методів оцінки індексу передачі мовлення.

2.1. Порівняння повного модуляційного методу оцінювання розбірливості мовлення із повним формантно-модуляційним методом

В даному підрозділі наведено результати порівняння повного модуляційного та повного формантно-модуляційного методів оцінювання розбірливості мовлення за точністю вимірювань в умовах дії шумової завади.

Для порівняння повних ФМ та модуляційного методів були проведені дослідження, які виконувалися шляхом комп'ютерного моделювання. Сигнали $y_i(t) = x_i(t) + n(t)$, де $n(t)$ - стаціонарний рожевий шум, генерувалися із частотою дискретизації 22050 Гц.

Обчислення STI повним модуляційним та повним ФМ методами виконувалися за алгоритмами, наведеними в Розділі 1. Впродовж кожної сесії обчислень тривалість сигналів T змінювалася та приймала значення 4, 8, 16, 32 та 64 секунди, а відношення сигнал-шум (SNR) змінювалися від мінус 28 дБ до плюс 28 дБ із кроком 4 дБ. Для кожного сполучення параметрів SNR та T обчислювалося 30 оцінок STI, що дозволяло оцінити математичне сподівання та стандартне відхилення оцінок STI з достатньою для практичних застосувань

точністю. Фільтрація сигналів $y_i(t)$ виконувалася гребінкою із семи еліптичних октавних фільтрів, центральні частоти смуг пропускання яких приймали значення 125, 250, 500, 1000, 2000, 4000 та 8000 Гц, а коефіцієнт передачі за межами смуги пропускання становив мінус 60 дБ.

Для обчислень STI повним ФМ методом використано програмне забезпечення, запропоноване в [21]. Це програмне забезпечення було модифіковано для обчислень STI повним модуляційним методом.

Обчислення виконувалися в середовищі Matlab R2022a. Тривалість сесії обчислень для кожного із методів становила приблизно 3,5 години для ПК із тактовою частотою процесора 2 ГГц та об'ємом оперативної пам'яті 12 ГБ та приблизно 2 години для ПК із тактовою частотою процесора 3,6 ГГц та об'ємом оперативної пам'яті 16 ГБ.

Результати досліджень повного модуляційного методу

Результати оцінювання математичного сподівання, зміщення та стандартного відхилення оцінок STI для повного модуляційного методу, для умов шумової завади (реверберація відсутня), наведено на рис. 2.1.1. Оцінки математичного сподівання (рис. 2.1.1а) свідчать про наявність зміщення оцінок STI, яке зменшується із зростанням тривалості T . Для отримання кількісних значень величини зміщення, в якості еталону використано прогнозну оцінку STI для випадку шумової завади, отриману формантним методом, де відношення сигнал-шум в k -й смузі частот обчислюється за виразом $E_k = 10 \lg \frac{D_{sk}}{D_{nk}}$, де D_{sk} та D_{nk} - дисперсії сигналу та шуму в k -й смузі частот, відповідно.

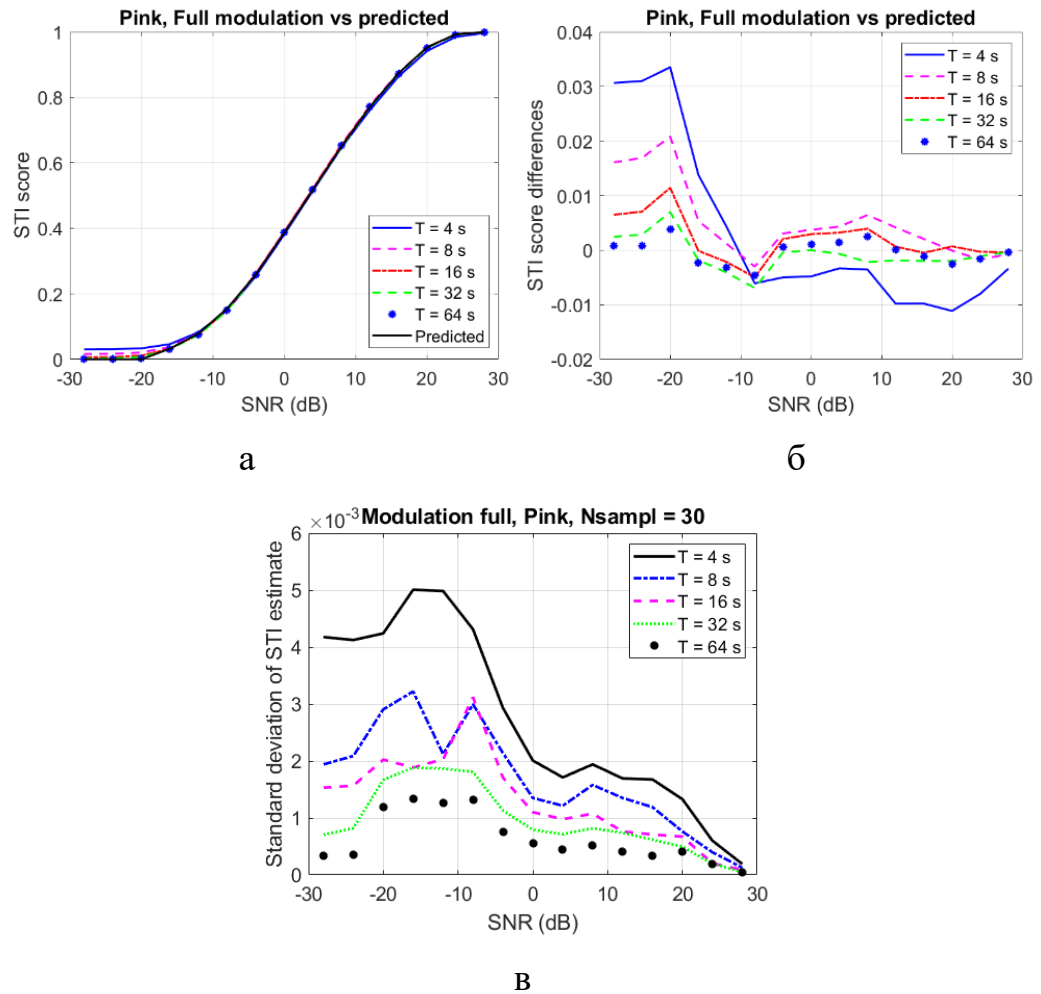


Рис. 2.1.1. Повний модуляційний метод вимірювання STI для умов дії шумової завади: оцінки математичного сподівання (а), зміщення (б) та стандартного відхилення (в)

Як можна бачити на рис. 2.1.1б, оцінка STI, отримана повним модуляційним методом, зміщена в сторону більших значень при $\text{SNR} < -15$ дБ. При $\text{SNR} > -15$ дБ оцінка STI теж дещо зміщена, але в значно меншій мірі, й величина зміщення зменшується із зростанням T та SNR. Для $T = 16$ с величина зміщення не перевищує 0.01 в діапазоні значень SNR від мінус 28 дБ до плюс 28 дБ.

Стандартне відхилення оцінки STI (рис. 2.1.1в) має максимум в діапазоні значень SNR від мінус 20 дБ до мінус 10 дБ й зменшується до дуже малих значень із наближенням SNR до 28 дБ. Незначне зменшення стандартного відхилення оцінки STI також відбувається із наближенням SNR до мінус 28

дБ. Для $T=16$ с величина стандартного відхилення не перевищує 0.003 в діапазоні значень SNR від мінус 28 дБ до плюс 28 дБ.

Результати досліджень повного ФМ методу

Результати оцінювання математичного сподівання (а), зміщення (б) та стандартного відхилення оцінок STI (в) для повного ФМ методу, для умов шумової завади (реверберація відсутня), наведено на рис. 2.1.2.

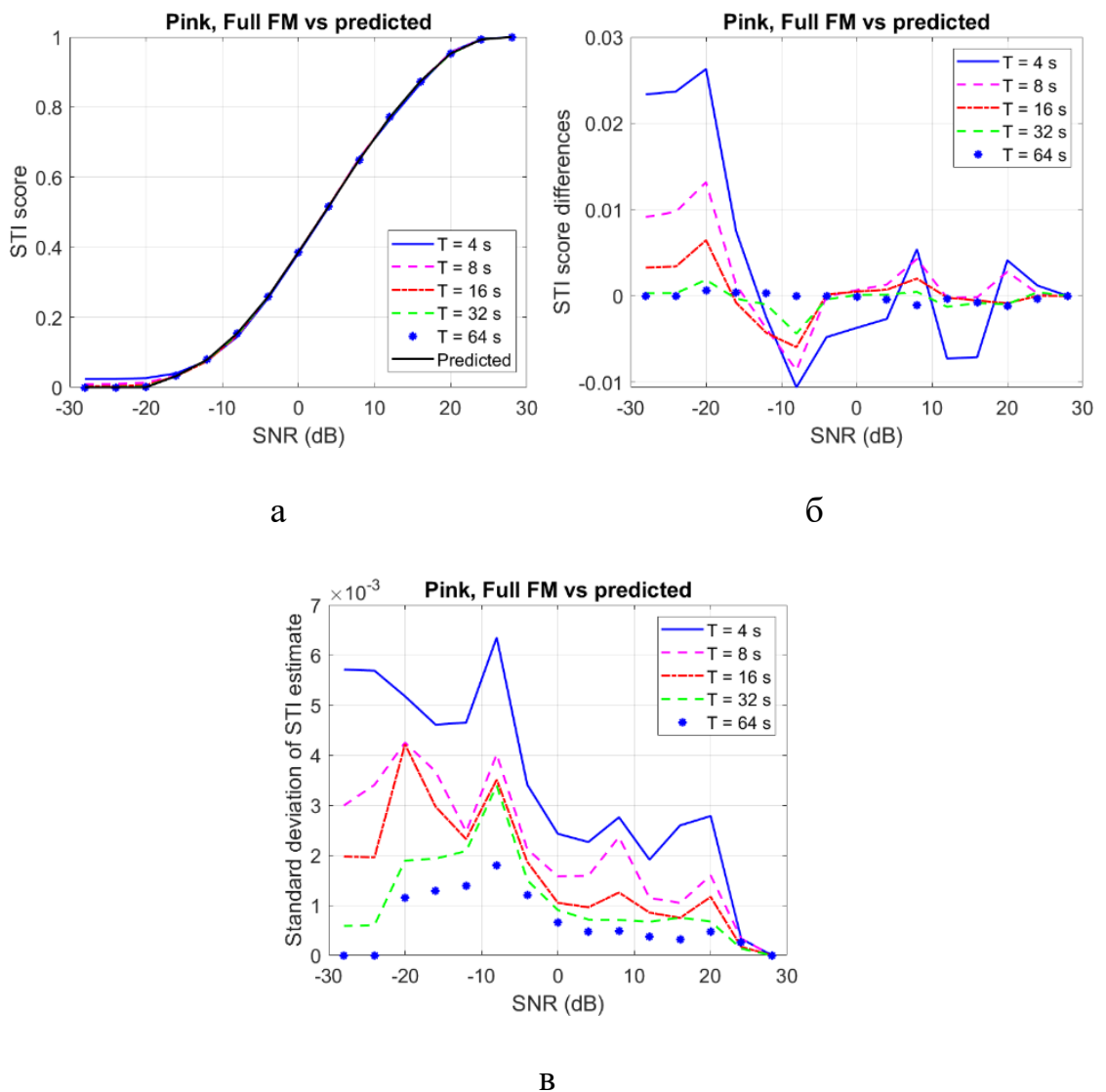


Рис. 2.1.2. Повний ФМ метод вимірювання STI для умов дії шумової завади: оцінки математичного сподівання (а), зміщення (б) та стандартного відхилення (в)

Як можна бачити на рис. 2.1.2а та 2.1.2б, оцінка STI, отримана повним ФМ методом, є помітно зміщеною в сторону більших значень при $SNR < -15$ дБ. При $SNR > -15$ дБ оцінка STI теж дещо зміщена, але величина зміщення зменшується із зростанням тривалості T та зростанням SNR. Для $T=16$ с величина зміщення не перевищує 0.006 в діапазоні значень SNR від мінус 28 дБ до плюс 28 дБ.

Стандартне відхилення оцінки STI (рис. 2.1.2в) має максимум в діапазоні значень SNR від мінус 20 дБ до мінус 8 дБ й зменшується до дуже малих значень із наближенням SNR до 28 дБ. Незначне зменшення стандартного відхилення оцінки STI також відбувається із наближенням SNR до мінус 28 дБ. Для $T=16$ с величина стандартного відхилення не перевищує 0.004 в широкому діапазоні значень SNR від мінус 28 дБ до плюс 28 дБ.

Порівняння модуляційного та ФМ методів

Наведені вище результати свідчать про значну схожість, за точністю вимірювань STI, повних модуляційного та ФМ методів. З метою більш наочного порівняння оцінок STI для цих методів, було обчислено різниці оцінок математичного сподівання:

$$\Delta_{FM,mdl} = \overline{STI}_{FM} - \overline{STI}_{mdl},$$

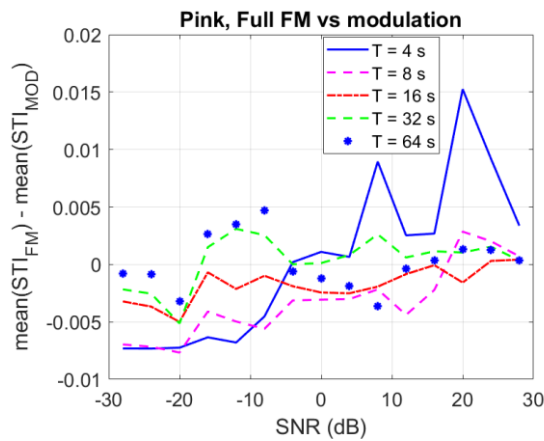
та відношення оцінок стандартних відхилень:

$$\Lambda_{FM,mdl} = \overline{\sigma STI}_{FM} / \overline{\sigma STI}_{mdl},$$

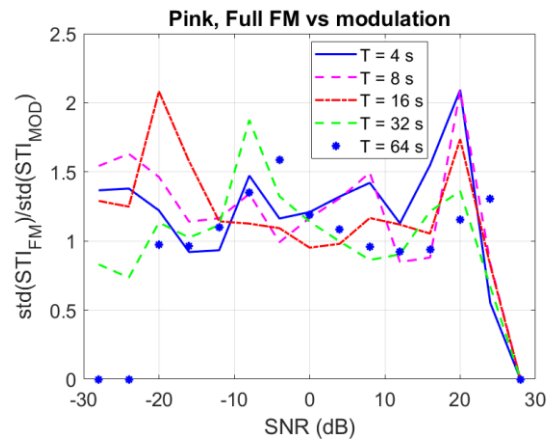
де \overline{STI}_{FM} та \overline{STI}_{mdl} - середні значення оцінок STI, отриманих повними ФМ та модуляційними методами, відповідно, $\overline{\sigma STI}_{FM}$ та $\overline{\sigma STI}_{mdl}$ - оцінки відповідних стандартних відхилень.

Наведені на рис. 2.1.3а результати свідчать, що при $T=16$ с різниця середніх значень оцінок STI не перевищує 0.005 в широкому діапазоні значень SNR від мінус 28 дБ до плюс 28 дБ. Наведені на рис. 2.1.3б результати

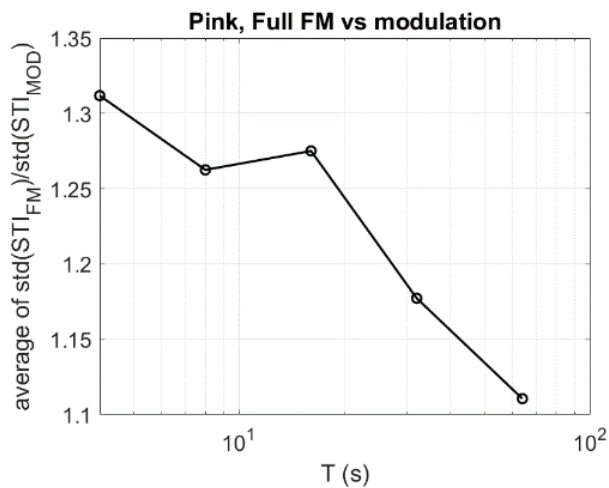
обчислень згідно $\Lambda_{FM,mdl}$ свідчать, що стандартні відхилення оцінок також мало відрізняються, їх відношення є близькими до 1. На рис. 2.1.3в наведено графік залежності усереднених, на інтервалі SNR від мінус 20 дБ до плюс 20 дБ, значень відношення $\Lambda_{FM,mdl}$ від тривалості T тестових сигналів $x_i(t)$. Можна бачити, що стандартні відхилення оцінок STI для ФМ методу лише на 10-30% вищі за такі для модуляційного методу.



а



б



в

Рис. 2.1.3. Порівняння оцінок STI повних модуляційного та ФМ методів для випадку шумової завади: різниця середніх значень (а), відношення стандартних відхилень (б), усереднене значення відношення стандартних відхилень (в)

Обговорення

Окремі отримані в даній роботі результати для $T=4$ с, $T=8$ с та $T=16$ с зведено в табл. 2.1.1, де Δ є максимальне, в межах інтервалу $-28 \text{ дБ} < \text{SNR} < 28 \text{ дБ}$, зміщення, Σ є максимальне стандартне відхилення, Ω є максимальна повна похибка вимірювань ($\Omega = \sqrt{\Delta^2 + \Sigma^2}$).

Таблиця 2.1.1.

Метод	T (с)	Δ	Σ	Ω
Модуляц.	4	0.033	0.005	0.034
	8	0.020	0.003	0.020
	16	0.010	0.003	0.011
ФМ	4	0.027	0.006	0.028
	8	0.012	0.004	0.018
	16	0.006	0.004	0.007

Оскільки прийнятною для практичного застосування похибкою оцінювання STI вважається величина «just noticeable difference» $JND=0.03$ [24], із табл. 2.1.1 можна бачити, що ця вимога практично задовольняється навіть при $T=4$ с, де максимальна повна похибка оцінювання STI є близькою до 0,03 для обох методів. При $T=8$ с та $T=16$ с максимальна повна похибка оцінювання STI для обох методів є близькою до 0.02 та 0,01, відповідно. Зазначимо, що ці результати добре узгоджуються із [22], де вказано, що при тривалості тестових сигналів 10 с похибка оцінювання STI становить 0.02.

2.2. Порівняння швидкого формантно-модуляційного методу із повним формантно-модуляційним методом

Для того, щоб оцінити STI, дослідження виконувалися комп'ютерним моделюванням тестових сигналів, спотворених шумом та реверберацією, із наступною їх обробкою. Крім того, в якості шумової завади розглянуто випадок рожевого шуму, який є досить добре наближеним до спектра реального фонового шуму. За допомогою згортки тестового сигналу з імпульсною характеристикою реального приміщення враховувався вплив реверберації, внаслідок чого модель мовленнєвого сигналу $y(t)$, спотвореного шумом та реверберацією, мала вигляд:

$$y(t) = x(t) \otimes h(t) + n(t)$$

де $x(t)$ — тестовий сигнал; $n(t)$ — шумова завада; $h(t)$ — імпульсна характеристика приміщення; \otimes — символ згортки. В якості $h(t)$ використано запис імпульсної характеристики реальної аудиторії [90].

STI оцінювалося повним та швидким ФМ методами. Опис алгоритму оцінювання STI повним ФМ методом наведено в Розділі 1. Простішим вважається алгоритм оцінювання STI швидким ФМ методом, оскільки потребує застосування лише одного тестового полігармонічного сигналу [26; 27]:

$$x(t) = \xi(t) \sqrt{f_5(t)}, \quad f_5(t) = 1 + 0,32 \cdot \sum_{i=1}^5 \sin 2\pi F_i t, \quad F_i = iF, \quad F = 0,7 \text{ Гц.}$$

При цьому

$$m_{k,i} = \frac{6.25 |A_{k,i}(F_i)|}{|A_{k,i}(0)|},$$

а замість виразу $E_k = \frac{1}{14} \sum_{i=1}^{14} SNR_{eff\ k,i}$ використовують його аналог

$$E_k = \frac{1}{5} \sum_{i=1}^5 SNR_{eff\ k,i}.$$

Результати для умов дії шуму, що переважає

Отже, розглянемо спочатку випадок переважної дії шуму, а також порівняємо за точністю вимірювань формантний метод, повний формантно-модуляційний метод (ПФМ) та швидкий формантно-модуляційний метод (ШФМ) методи.

На рис. 2.2.1 представлено графіки залежності середнього значення і стандартного відхилення оцінок STI від відношення сигнал-шум SNR та тривалості T тестового сигналу для формантного методу. При цьому для кожного сполучення значень параметрів SNR та T кількість вибірок оцінки STI становила $N_{sampler}=100$. Зауважимо, що середнє значення оцінок STI не залежить від T , у свою чергу, воно збігається із прогнозним значенням. Таким чином, оцінка STI , отримана формантним методом, є незміщеною. Стандартне відхилення оцінок STI має максимальне значення при $SNR \approx -5$ дБ й монотонно спадає із ростом T , але навіть у випадку $T=4$ с воно не перевищує 0,0015. Значення 0,03 є гранично припустимою величиною похибки вимірювань STI (у праці [91] це значення названо «ледь помітною різницею» — just noticeable difference — JND). Тому можна дійти висновку, що тривалість $T=4$ с тестового сигналу формантного методу в умовах переважної дії шумової завади забезпечує на порядок нижчу похибку вимірювань STI , ніж це визначено величиною $JND=0,03$. Такий висновок є справедливим для широкого діапазону значень SNR від -20 дБ до +20 дБ.

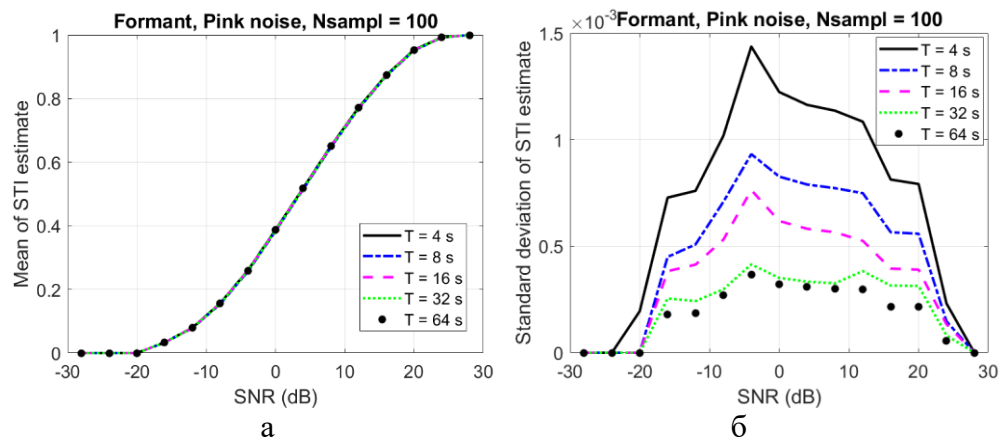


Рис. 2.2.1. Формантний метод: середнє (а) та стандартне відхилення (б) оцінок STI

Графіки залежності середнього значення та стандартного відхилення оцінок STI , отримані обробкою $N_{\text{sampl}}=30$ екземплярів оцінки STI для кожного сполучення значень параметрів SNR та T , наведено на рис. 2.2.2 для ПФМ методу.

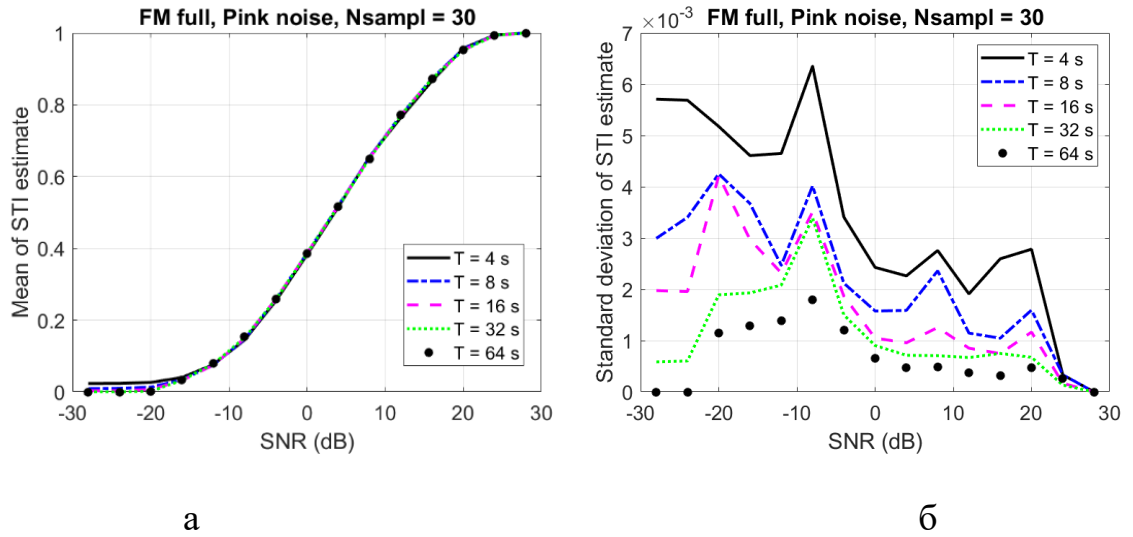


Рис. 2.2.2. Повний ФМ метод, шум: середнє (а) та стандартне відхилення (б) оцінок STI

Зіставляючи рис. 2.2.1 та 2.2.2, можна побачити, що ПФМ метод у випадку шумової завади забезпечує дуже близькі результати до результатів, отриманих формантним методом. Незважаючи на те, що $SNR < -15$ дБ оцінки є зміщеними, це зміщення зменшується зі зростанням тривалості T тестового сигналу. Звернемо увагу, що для полегшення кількісного порівняння рис. 2.2.1 та 2.2.2, на рис. 2.2.3 представлено графіки різниці середніх значень:

$$\Delta_{FM} = \overline{STI}_{FM} - \overline{STI}_{fm},$$

та відношення стандартних відхилень

$$\Lambda_{FM} = \sigma_{STI_{FM}} / \sigma_{STI_{fm}},$$

де \overline{STI}_{FM} та \overline{STI}_{fm} - усереднене значення оцінок STI для ПФМ методу та формантного методу, відповідно; $\sigma_{STI_{FM}}$ та $\sigma_{STI_{fm}}$ - стандартне відхилення оцінок STI для ПФМ методу та формантного методу, відповідно.

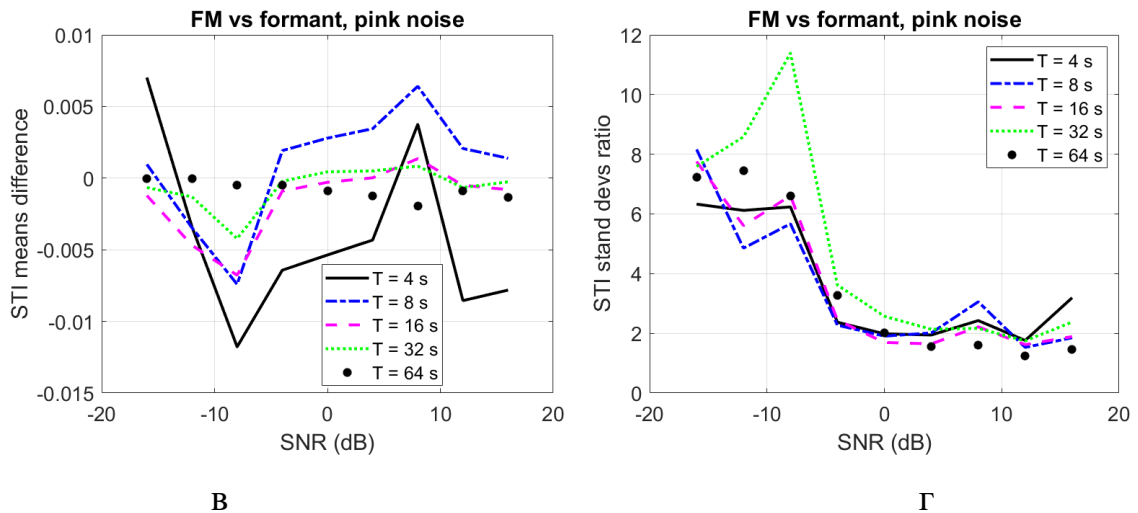


Рис. 2.2.3. Повний ФМ метод, шум: різниця Δ_{FM} (а) та відношення Λ_{FM} (б)

Зі збільшенням T різниця Δ_{FM} прямує до нуля (рис. 2.2.3а), за винятком випадку, коли $T=64$ с для $SNR>0$ дБ, де модуль різниці Δ_{FM} є більшим за такий для випадків $T=32$ с та навіть $T=16$ с. Усе це можна пояснити відносно невеликою кількістю ($N_{sampl} = 30$) екземплярів оцінки STI, яка моделювалася, а мале значення N_{sampl} — значною тривалістю часу обчислень, чого вимагає ПФМ метод.

Зауважимо, що стандартне відхилення ПФМ методу приблизно в 2 рази більше за таке для формантного методу при $SNR>0$ дБ (рис. 2.2.3б). У свою чергу, для $SNR<0$ дБ цей програш у точності вимірювань збільшується до 6 — 8 разів.

Так, при $T=16$ с зміщення не перевищує 0,007, а стандартне відхилення — 0,004 в широкому діапазоні значень SNR від -20 дБ до +20 дБ. Крім того, можна побачити, що навіть для $T=4$ с зміщення та стандартне відхилення не перевищують значень 0,012 та 0,006 відповідно. Такий варіант також є припустимим, хоча й межує із гранично припустимим значенням $JND=0,03$.

На рис. 2.2.4 представлено графіки залежності середнього значення та стандартного відхилення оцінок STI для ШФМ методу, які отримано обробкою 100 екземплярів оцінки STI.

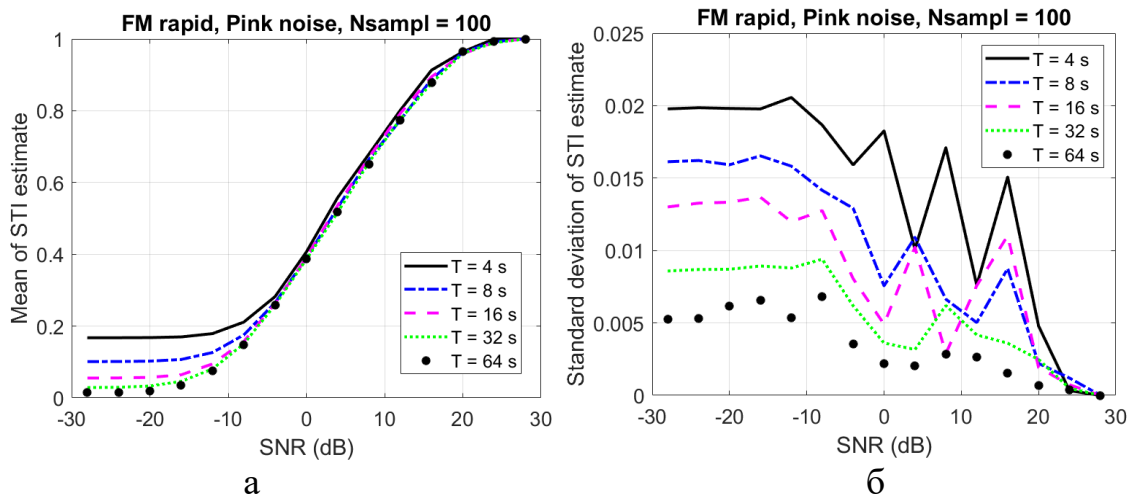


Рис. 2.2.4. Швидкий ФМ метод, шум: оцінки середнього (а) та стандартного відхилення (б) оцінок STI

Зіставляючи рис. 2.2.1 та 2.2.4, можна побачити, що швидкий ФМ метод забезпечує асимптотично спроможну оцінку STI , тому що зі зростанням тривалості T тестового сигналу при $SNR < -5$ дБ зміщення та стандартне відхилення оцінок STI прямують до нуля.

Для більш ґрунтовного порівняння рис. 2.2.1 та 2.2.3 побудуємо графіки різниці середніх значень:

$$\Delta_{RFM} = \overline{STI}_{RFM} - \overline{STI}_{frm}$$

та відношення стандартних відхилень:

$$\Lambda_{RFM} = \sigma_{STI_{RFM}} / \sigma_{STI_{frm}},$$

де \overline{STI}_{RFM} та \overline{STI}_{frm} - усереднені значення оцінок STI для ШФМ методу та формантного методу, відповідно; $\sigma_{STI_{RFM}}$ та $\sigma_{STI_{frm}}$ - стандартні відхилення оцінок STI для ШФМ методу та формантного методу, відповідно.

На рис. 2.2.5 представлено графіки залежностей Δ_{RFM} та Λ_{RFM} .

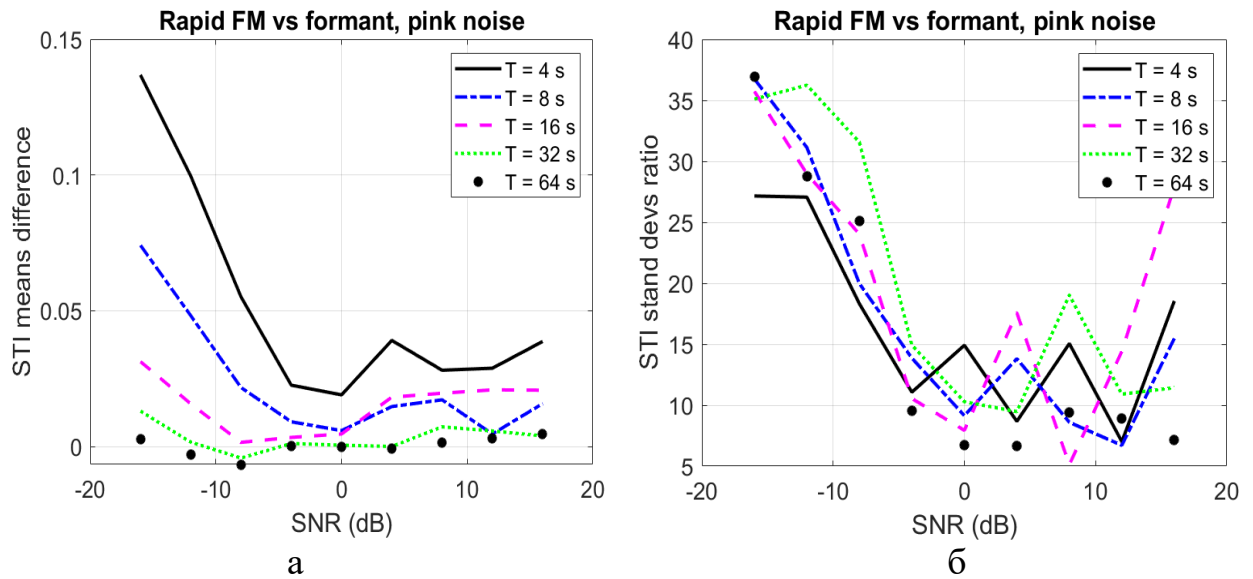


Рис. 2.2.5. Швидкий ФМ метод, шум: різниця Δ_{RFM} (а) та відношення Λ_{RFM} (б)

Слід зазначити, що при збільшенні тривалості T тестового сигналу різниця між середніми значеннями зменшується, а це, у свою чергу, можна трактувати як асимптотичну незміщеність оцінки STI (рис. 2.2.5а). Зіставляючи рис. 2.2.5а та 2.2.3а, видно, що при малих SNR зміщення оцінки STI для ШФМ методу майже на порядок перевищує таке для ПФМ методу.

Разом з тим, порівнюючи рис. 2.2.5б та 2.2.1б, можна побачити, що стандартне відхилення оцінки STI для ШФМ методу приблизно в 10 — 12 разів, за умови $SNR > 0$ дБ, перевищує таке для формантного методу, а коли $SNR < 0$ дБ, таке перевищення зростає й сягає 30 разів при $SNR \approx -15 \dots -20$ дБ.

При зіставленні стандартних відхилень ШФМ та ПФМ методів (рис. 2.2.3б та 2.2.5б), незалежно від значень SNR та T , швидкий метод приблизно в 5—6 разів програє повному методу.

Крім того, зміщення швидкої оцінки STI при $T=16$ с не перевищує $JND=0,03$ в широкому діапазоні значень SNR від -20 дБ до +20 дБ, а стандартне відхилення — 0,014 в типовому для аудиторій діапазоні $0 < SNR < 20$ дБ.

Результати для умов сумісної дії шуму та реверберації

На рис. 2.2.6 наведено графіки залежностей середнього значення та стандартного відхилення оцінок STI від SNR та T для ПФМ методу, з якого

видно, що зі зростанням T похибка оцінки STI прямує до нуля, хоча при цьому час обчислень збільшується.

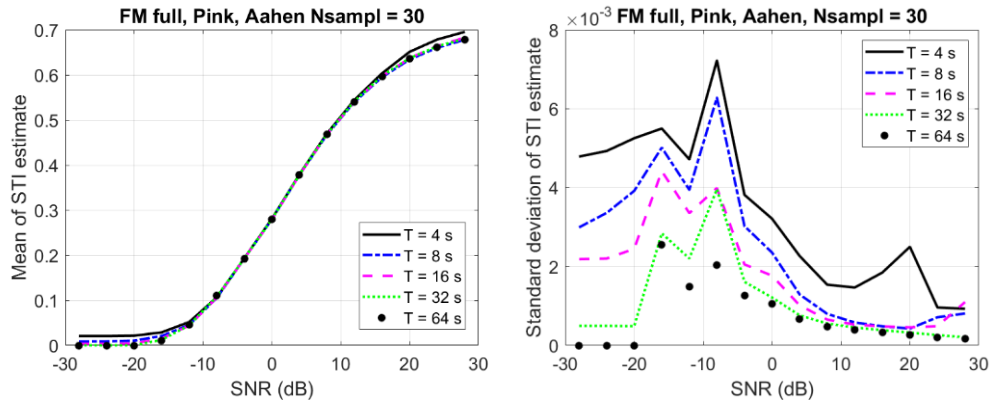


Рис. 2.2.6. Повний ФМ метод, шум та реверберація: середнє (а) та стандартне відхилення (б) оцінок STI

Таким чином, при комп'ютерному моделюванні тривалість $T=64$ с, з одного боку, можна вважати гранично припустимою через довгий час обчислень, а з іншого боку, із рис. 2.2.6а видно, що при $4 < T < 64$ с та $SNR < -15$ дБ середнє значення оцінки STI є близьким до такого при дії шуму, що переважає (рис. 2.2.1а). Графіки різниці середніх значень $\Delta_{FMnr} = \overline{STI}_{FMnr} - \overline{STI}_{FMn}$, та відношення стандартних відхилень $\Lambda_{FMnr} = \sigma_{STI_{FMnr}} / \sigma_{STI_{FMn}}$ представлено на рис. 2.2.7а та 2.2.7б, де \overline{STI}_{FMnr} та \overline{STI}_{FMn} — середні значення оцінок STI для ПФМ методу при сумісній дії шуму і реверберації та при дії лише шуму відповідно; $\sigma_{STI_{FMnr}}$ та $\sigma_{STI_{FMn}}$ — стандартні відхилення оцінок STI для ПФМ методу при сумісній дії шуму і реверберації та при дії лише шуму.

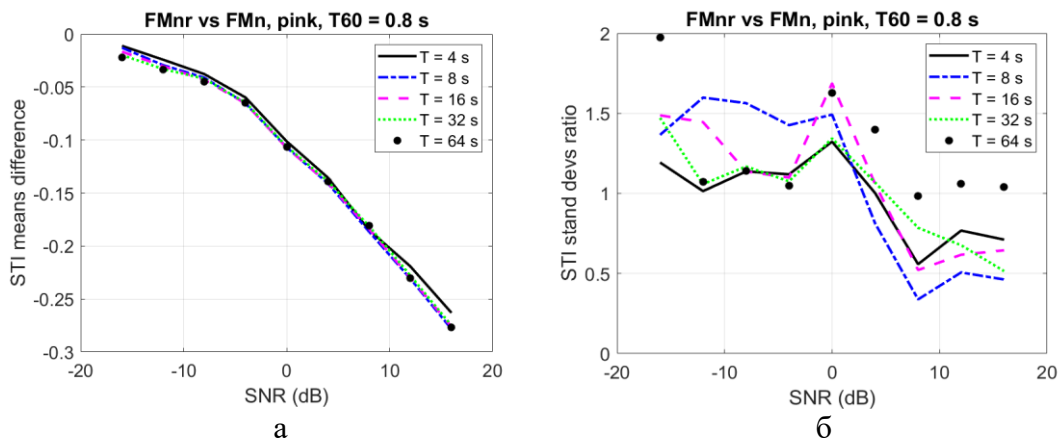


Рис. 2.2.7. Повний ФМ метод, шум і реверберація: порівняння середніх (а) та стандартних відхилень (б) оцінок STI

На рис. 2.2.7а добре видно, що модуль різниці Λ_{FMnr} не перевищує $JND=0,03$ для $SNR < -10$ дБ. При цьому дія реверберації практично не відбивається на значеннях оцінки STI, через те, що переважає вплив шумової завади. Зауважимо, що зі зростанням SNR, негативна дія реверберації на STI теж зростає, а при $SNR \approx 15-18$ дБ отримуємо $\Lambda_{FMnr} \approx -0,3$. Щодо стандартного відхилення оцінки STI, беручи до уваги показану на рис. 2.2.7б поведінку параметра Λ_{FMnr} , можна деякою мірою вважати, що воно залишилося майже таким, як у випадку дії виключно шумової завади.

Графіки залежностей середнього значення та стандартного відхилення оцінок STI від SNR та довжини тестового сигналу T для ШФМ методу представлено на рис. 2.2.8.

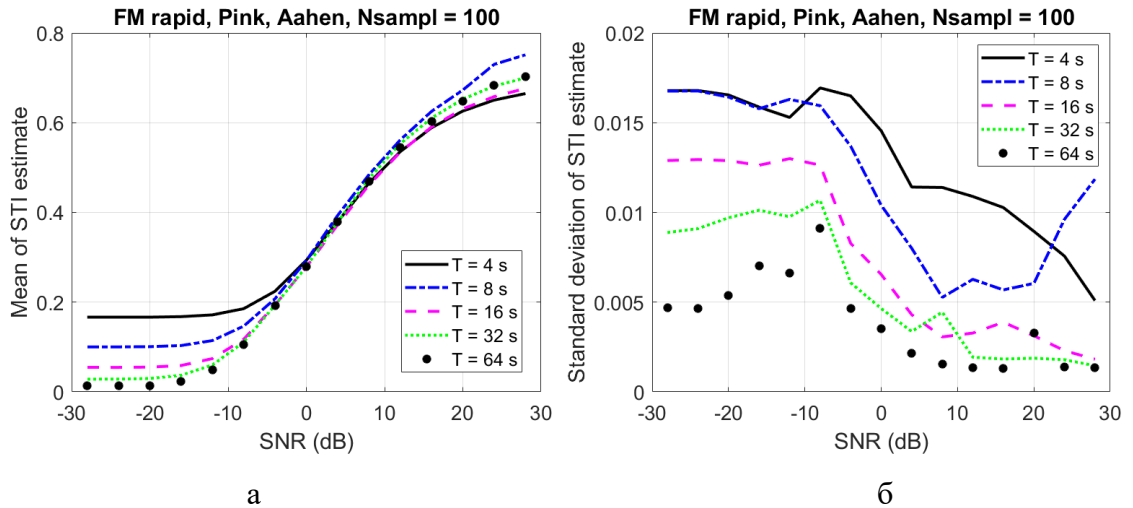


Рис. 2.2.8. Швидкий ФМ метод, шум та реверберація: середнє (а) та стандартне відхилення (б) оцінок STI

Разом з тим на рис. 2.2.9а і 2.2.9б наведено графіки різниці середніх значень $\Delta_{RFMnr} = \overline{STI}_{RFMnr} - \overline{STI}_{RFMn}$ та відношення стандартних відхилень $\Lambda_{RFMnr} = \sigma_{STI_{RFMnr}} / \sigma_{STI_{RFMn}}$, де \overline{STI}_{RFMnr} та \overline{STI}_{RFMn} - середні значення оцінок STI для ШФМ методу при сумісній дії шуму та реверберації та при дії виключно шуму, відповідно; $\sigma_{STI_{RFMnr}}$ та $\sigma_{STI_{RFMn}}$ - стандартні відхилення оцінок STI для ШФМ методу при сумісній дії шуму та реверберації та при дії виключно шуму, відповідно.

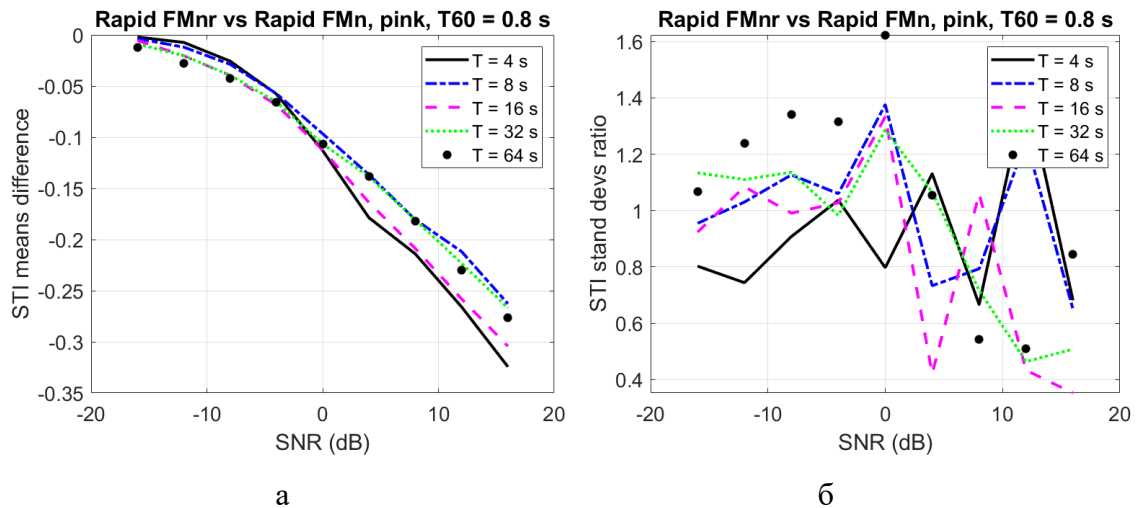


Рис. 2.2.9. Швидкий ФМ метод, шум та реверберація: порівняння середніх (а) та стандартних відхилень (б)

Зіставляючи рис. 2.2.9 та рис. 2.2.7, можна побачити, що їхня поведінка є аналогічною. Зокрема, із рис. 2.2.9а видно, що модуль різниці Δ_{RFMnr} не перевищує $JND=0,03$ для $SNR < -10$ дБ, тобто дія реверберації практично не впливає на значення оцінки STI внаслідок переваги впливу шумової завади. Проте зі зростанням SNR, негативний вплив дії реверберації на STI також зростає, тому при $SNR \approx 15-18$ дБ отримуємо $\Delta_{RFMnr} \approx -0,3$. Щодо стандартного відхилення оцінки STI (рис. 2.2.9б), то можна вважати, що воно залишається таким самим, як і у випадку дії лише шумової завади.

Обговорення результатів

У табл. 2.2.1. зведено результати досліджень для випадку дії виключно шумової завади, де взято до уваги лише мінімальне значення $T=16$ с, що забезпечує гранично припустиме значення $JND=0,03$ оцінки STI при використанні ШФМ методу. Зауважимо, що орієнтація на ШФМ метод пояснюється його практичною цінністю, яка зумовлена суттєвим скороченням тривалості процесу вимірювань, порівняно із ПФМ методом.

Таблиця 2.2.1. Показники точності для $T=16$, шумова завада

Метод	Різниця середніх значень			Стандартні відхилення			Відношення стандартних відхилень		
	Познач.	SNR<0	SNR>0	Познач.	SNR<0	SNR>0	Познач.	SNR<0	SNR>0
Формантний	-	-	-	σ_{frm}	0,0007	0,0006	-	-	-
ПФМ	Δ_{FM}	<0,007	<0,002	σ_{FM}	0,004	0,0012	Λ_{FM}	6	2
ШФМ	Δ_{RFM}	<0,03	<0,03	σ_{RFM}	0,013	0,01	Λ_{RFM}	30	12

Із таблиці 2.2.1 видно, що стандартне відхилення оцінки STI, отриманої ШФМ методом, є більшим за таке для формантного методу приблизно в 12 разів при $SNR>0$ дБ та приблизно в 30 разів при $SNR<0$ дБ.

Зменшити зміщення оцінки STI суттєво, до 0,002—0,007, дозволяє використання повної версії ФМ методу, хоча і в цьому випадку стандартне відхилення оцінки приблизно в 2—6 разів перевищує таке для формантного методу.

Результати досліджень для випадку спільної дії шуму та реверберації представлено в табл. 2.2.2.

Таблиця 2.2.2. Показники точності для $T=16$, спільна дія шуму та реверберації

Метод	Різниця середніх значень			Стандартні відхилення			Відношення стандартних відхилень		
	Познач.	SNR<0	SNR>0	Познач.	SNR<0	SNR>0	Познач.	SNR<0	SNR>0
ПФМ	Δ_{FMnr}	<0.1	0.1-0.3	σ_{FMnr}	0.004	0.001	Λ_{FMnr}	1	0.8
ШФМ	Δ_{RFMnr}	<0.1	0.1-0.3	σ_{RFMnr}	0.014	0.005	Λ_{RFMnr}	1.1	0.7

Із поданих результатів у табл. 2.2.2 видно, що при сумісній дії шуму та реверберації середні значення оцінки STI суттєвіше змінюються, коли дія реверберації більше переважає над дією шуму. Стандартне відхилення для ПФМ та ШФМ методів залишається практично таким самим, як і для випадку переважної дії шуму.

Слід зауважити, що в дослідженнях розглянуто лише одне значення часу реверберації, $T_{60}=0,8$ с, яке є типовим для аудиторій. Разом з тим для переговорних та офісних приміщень зазвичай час реверберації є меншим та близьким до 0,2—0,4 с, що є проміжним випадком між відсутністю реверберації та часом реверберації $T_{60}=0,8$ с [92], тому представлені тут результати легко розповсюдити й на ці випадки. А от щодо приміщень із часом реверберації $T_{60}>1$ с, такий випадок поки залишається недослідженим.

2.3. Багатокомпонентний тестовий сигнал для порівняння прямого та непрямого методів оцінки індексу передачі мовлення

При оцінюванні розбірливості мовлення, спотвореного шумом та реверберацією, використовують прямий або непрямий методи вимірювання індексу передачі мовлення STI. Проте залишається невідомим, наскільки суттєво різняться результати вимірювань, отримані прямим та непрямим методами. Прямі та непрямі методи оцінювання індексу розбірливості мови STI і його модифікацій (STIPA, STITEL, RASTI) описані в британському стандарті BS EN 60268-16: 2011 [22] і міжнародному стандарті IEC 60268-16 [93].

В даній дисертації представлено результати порівняння оцінок STI, отриманих прямим ШФМ методом та методами. Алгоритм обчислень ШФМ методом представлено в двох непрямим підозділі 2.2. Обчислення непрямим методами базуються на використанні співвідношення

$$m_{ki} = m_{k \text{ rev}}(F_i) \cdot m_{k \text{ noise}} = \frac{\int_0^{\infty} h_k^2(t) \exp(-j2\pi F_i t) dt}{\int_0^{\infty} h_k^2(t) dt} \cdot \left(1 + 10^{-0.1 \cdot SNR_k}\right)^{-1}.$$

При цьому у варіанті НМ1 непрямого методу оцінювання STI для обчислень STI використано MLS-сигнал $x_{MLS}(t)$ та стаціонарний шум $x_{n \text{ spch}}(t)$ зі спектром мовлення. При цьому максимальний рівень бокових пелюсток функції автокореляції сигналу $x_{MLS}(t)$ тривалістю $L = 2^{18} - 1$ вибірок становить мінус 54 дБ, що дає можливість достатньо точно оцінювати імпульсну характеристику приміщення $h(t)$.

У варіанті НМ2 непрямого методу для обчислень STI використано MLS-сигнал $x_{MLS \text{ spch}}(t)$ зі спектром мовлення. При цьому максимальний рівень бокових пелюсток функції автокореляції сигналу $x_{MLS \text{ spch}}(t)$, тривалістю

$L = 2^{20} - 1$, вибірок становить мінус 40 дБ. Незважаючи на те, що на інтервалі ± 2 с довкола максимуму функції автокореляції рівень бокових пелюсток є близьким до мінус 45 дБ, головна пелюстка функції автокореляції є розширеною та має досить складну форму. Таким чином, оцінка імпульсної характеристики приміщення $h(t)$, отримана при реалізації варіанта НМ2, має поступатися за точністю аналогічній оцінці, отриманій у варіанті НМ1.

Для пришвидшення та спрощення виконання модельних та натурних експериментальних досліджень видається доречним використання багатокomпонентних тестових сигналів, що складаються із послідовності «елементарних» тестових сигналів та пауз.

Приміром, трьохкомпонентний тестовий сигнал, представлений на рис. 2.3.1а, є зручним для порівняння результатів оцінювання STI прямим методом та непрямим методом 1 (НМ1). Справді, по обчисленню крос-кореляції між записаним відгуком приміщення та копією MLS-сигналу, на початку отриманого масиву буде отримано гострий високий сплеск, а це, у свою чергу, підходить для визначення початку першого сигналу та границь інших сигналів (стаціонарний шум і модульований шум) на тлі інтенсивної фонові завади.

Двокомпонентний сигнал, що аналогічно підходить для порівняння прямого методу із непрямим методом 2 (НМ2), представлено на рис. 2.3.1б.

Проте найбільш перспективним, з погляду економії часу та зручності досліджень, видається чотирьохкомпонентний тестовий сигнал (рис. 2.3.1в), який суттєво полегшує наскрізне зіставлення результатів оцінювання STI прямим методом та обома варіантами непрямого методу (НМ1 та НМ2). Крім того, можна говорити про проведення випробувань у практично незмінних у часі умовах завдяки тому що окремі компоненти тестового сигналу відокремлені порівняно невеликими паузами (тривалістю кілька секунд).

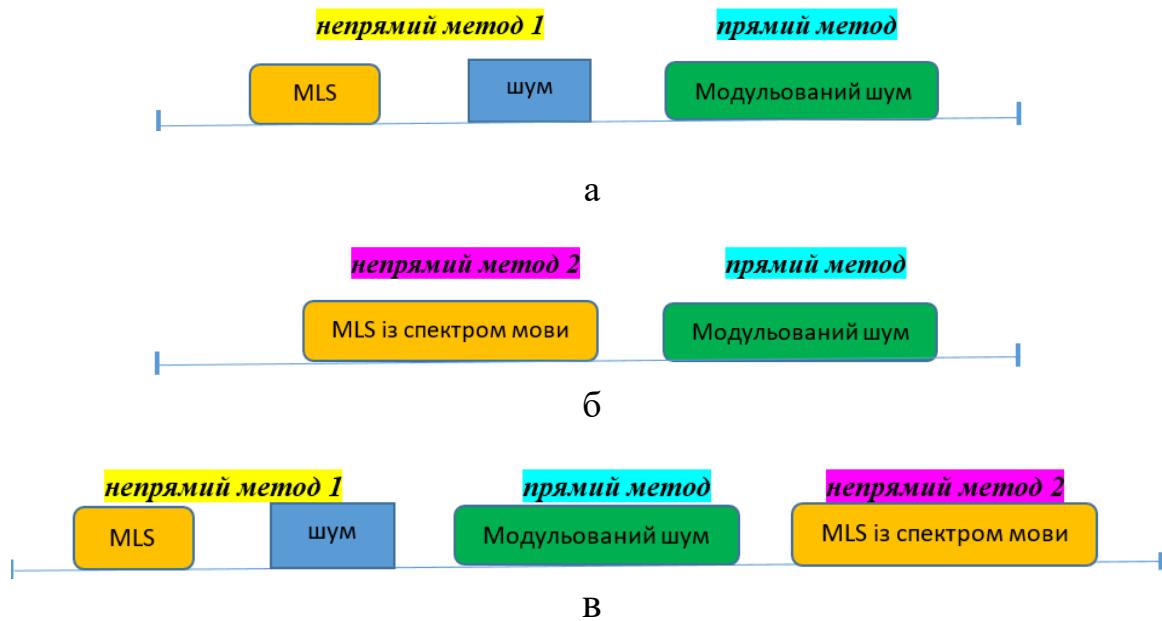


Рис. 2.3.1. Трьохкомпонентний (а), двохкомпонентний (б) та чотирьохкомпонентний (в) тестові сигнали

Зазначимо, що параметри тестового сигналу слід обирати, враховуючи очікуваний час реверберації приміщень, де планується вимірювати розбірливість мовлення. Зокрема, якщо планується проведення вимірювань в аудиторіях навчальних закладів та в офісних приміщеннях, час реверберації переважно не перевищує 1с, тому тривалість пауз між компонентами тестового сигналу може становити 2с, що, у свою чергу, забезпечить у паузах наявність інтервалів, тривалістю 1с, де фонові шумова завада переважає над реверберацією. Саме такі інтервали із фонові шумовою завадою необхідні і для оцінювання відношень сигнал — шум SNR_k , $k = 1, \dots, 7$, і для оцінювання інтегрального відношення сигнал — шум SNR.

Тривалість модульованого шуму зі спектром мовлення при використанні прямого методу STIPA має становити 15—20 с [22]. У нашому дослідженні застосовується прямий ШФМ метод вимірювання STI [23], подібний до методу STIPA, де тривалість модульованого шуму зі спектром мовлення має бути не менша ніж 16 с. Крім того, тривалість елементарного тестового сигналу у вигляді стаціонарного шуму зі спектром мовлення становить 4 с,

чого достатньо для забезпечення точності вимірювань STI, яка є близькою до точності прямого методу в умовах переважної дії шумової завади [94].

Зауважимо, що тривалість MLS-сигналів необхідно обирати, враховуючи рівень бокових пелюсток автокореляційної функції, який бажано мати не більшим за мінус 40 дБ [22]. Із графіків автокореляційних функцій mls-сигналів із рівномірним спектром, представлених на рис. 2.3.2, видно, що для кількості вибірок MLS-сигналу $L = 2^{16} - 1$ максимальний рівень бокових пелюсток є близьким до -48 дБ, а для $L = 2^{18} - 1$ цей рівень знижується до -54 дБ. Тому було обрано $L = 2^{18} - 1$, й для частоти дискретизації $F_s = 44\,100$ Гц тривалість MLS-сигналу із рівномірним спектром є близькою до 6 с.

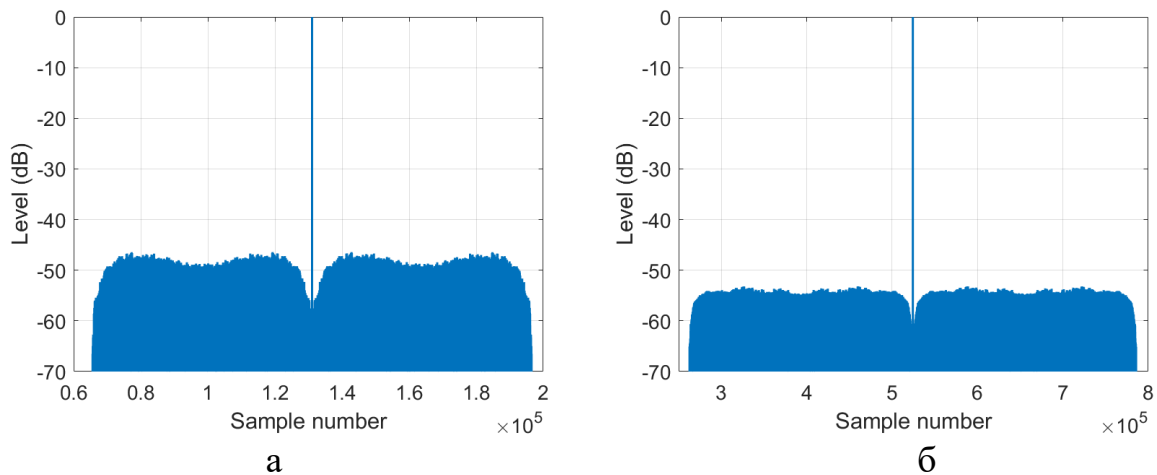


Рис. 2.3.2. Автокореляційні функція MLS сигналу для $L = 2^{16} - 1$ (а) та $L = 2^{18} - 1$ (б)

Оскільки зменшення смуги частот призводить до підвищення рівня бокових пелюсток автокореляційної функції MLS-сигналу, необхідно суттєво збільшити тривалість сигналів для таких сигналів зі спектром мовлення [22]. Із графіків автокореляційних функцій MLS-сигналів зі спектром мовлення, наведених на рис. 2.3.3, видно, що для кількості вибірок MLS-сигналу $L = 2^{19} - 1$ максимальний рівень бокових пелюсток є близьким до -37 дБ, незважаючи на те, що поблизу головного піка (в околиці ± 2 с) рівень бокових пелюсток є меншим та близьким до -40 дБ. Зауважимо, що максимальний рівень бокових пелюсток знижується до -40 дБ, коли $L = 2^{20} - 1$, а в околиці максимального

сплеску рівень бокових пелюсток на інтервалі ± 2 с є близьким до -45 дБ. Зважаючи на зазначені вище рівні бокових пелюсток, у нашому дослідженні було обрано $L = 2^{20} - 1$, де тривалість mls-сигналу зі спектром мовлення для частоти дискретизації $F_s = 44\,100$ Гц є близькою до 24 с.

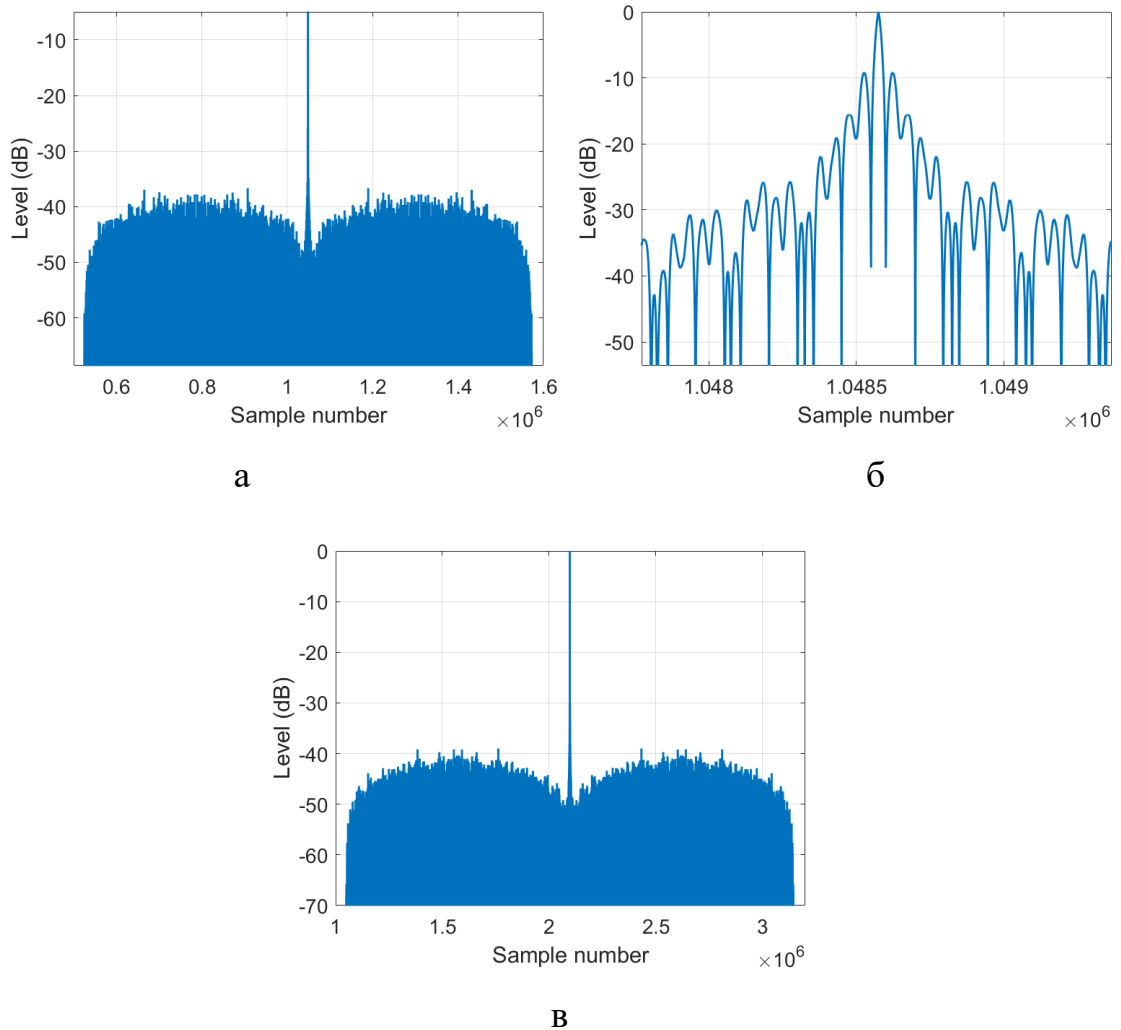


Рис. 2.3.3. Автокореляційна функція MLS-сигналу із спектром мовлення для $L = 2^{19} - 1$ (а,б) та $L = 2^{20} - 1$ (в)

Алгоритм генерування чотирьохкомпонентного тестового сигналу складається із таких етапів:

- 1) генерування MLS-сигналу із рівномірним спектром $x_{MLS}(t)$, тривалістю T_{MLS} ;

- 2) генерування MLS-сигналу із спектром мовлення $x_{MLS\ spch}(t)$, тривалістю $T_{MLS\ spch}$;
- 3) генерування відрізка гаусівського шуму із спектром мовлення $x_{n\ spch}(t)$, тривалістю $T_{n\ spch}$;
- 4) генерування амплітудно-модульованого шуму із спектром мовлення $x_{AM\ spch}(t)$, тривалістю $T_{AM\ spch}$;
- 5) вирівнювання дисперсій D компонентів тестового сигналу:

$$D[x_{MLS}(t)] = D[x_{MLS\ spch}(t)] = D[x_{n\ spch}(t)] = D[x_{AM\ spch}(t)];$$
- 6) формування остаточного тестового сигналу шляхом послідовного дописування (цю процедуру позначено символом \oplus) окремих його компонентів, із вставкою пауз $x_{sil}(t)$ тривалістю T_{sil} між цими компонентами:

$$x_{test}(t) = x_{sil}(t) \oplus x_{MLS}(t) \oplus x_{sil}(t) \oplus x_{n\ spch}(t) \oplus x_{sil}(t) \oplus x_{AM\ spch}(t) \oplus x_{sil}(t) \oplus x_{MLS\ spch}(t) \oplus x_{sil}(t);$$
- 7) нормування за максимумом сигналу $x_{test}(t)$ та запис результату нормування на диск.

Прокоментуємо окремі пункти наведених етапів.

Зокрема, пропонується формувати сигнали зі спектром мовлення за допомогою фільтрації первинних сигналів, що мають рівномірний спектр, гребінкою із 7 октавних фільтрів. Для полегшення калібровки вимірювальної системи застосовується процедура вирівнювання потужностей (п. 5 наведеного переліку). Зауважимо, що в процесі такої калібровки рієнь звуку, що випромінюється гучномовцем, має бути 60 дБА на відстані 1 м від гучномовця [22]. У подальшому результати, отримані із застосуванням різних методів вимірювання розбірливості мовлення, можуть бути коректно порівняними завдяки саме такому вирівнюванню потужностей.

На рис. 2.3.4 представлено результат синтезу чотирьохкомпонентного тестового сигналу, загальна тривалість якого є близькою до 60 с, виконаного за наведеним вище алгоритмом.

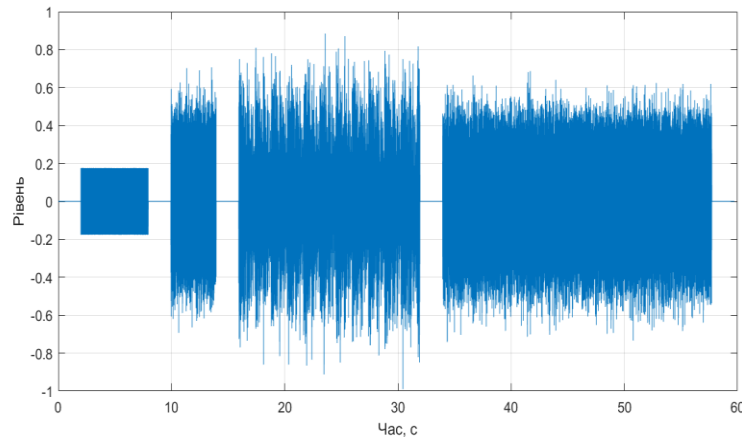


Рис. 2.3.4. Чотирьохкомпонентний тестовий сигнал

Алгоритм обробки реакції $y(t)$ приміщення на тестовий сигнал $x_{test}(t)$ складається із наступних етапів:

- 1) виконують крос-кореляційну обробку сигналу $y(t)$ та MLS сигналу $x_{MLS}(t)$, в результаті чого отримують сигнал $z(t)$;
- 2) в сигналі $z(t)$ відшуковують максимальний сплеск, положення t_0 якого на осі часу вказує на початок відгуку приміщення на MLS сигнал;
- 3) починаючи з моменту часу t_0 , із сигналу $z(t)$ вирізають масив $h(t)$ оцінки імпульсної характеристики приміщення тривалістю 1-1.5 с для подальшого обчислення $m_{k\ rev}(F_i)$;
- 4) орієнтуючись на момент часу t_0 , розраховують значення границь шумового сигналу $y_{n\ spch}(t)$, модульованого шумового сигналу $y_{AM\ spch}(t)$ та MLS-сигналу із спектром мовлення $y_{MLS\ spch}(t)$, після чого із сигналу $y(t)$ вирізають масиви відповідних даних для подальшої обробки;

- 5) оцінюють STI прямим пришвидшеним методом [23], використовуючи $y_{AM\ speech}(t)$;
- 6) оцінюють STI непрямым методом НМ1 із використанням виразу m_{ki} , при цьому для обчислення $m_{k\ rev}(F_i)$ використовують сигнал $h(t)$, а для обчислення $m_{k\ noise}$ використовують сигнал $y_{n\ speech}(t)$;
- 7) оцінюють STI непрямым методом НМ2 із використанням виразу m_{ki} , при цьому для обчислення $m_{k\ rev}(F_i)$ використовують сигнал $h_{speech}(t)$, де $h_{speech}(t)$ є результат крос-коваріації сигналів $y_{MLS\ speech}(t)$ та $x_{MLS\ speech}(t)$, а параметр $m_{k\ noise}$ обчислюють із використанням сигналу $y_{MLS\ speech}(t)$.

Для того, щоб перевірити працездатність запропонованих алгоритмів формування тестового сигналу $x_{test}(t)$ та подальшу обробку реакції $y(t)$ приміщення на цей сигнал, бралися до уваги чотири модельних приклади, де моделювання сигналів $y(t)$, спотворених шумом та реверберацією, відбувалося за виразом:

$$y(t) = x_{test} \otimes h(t) + n(t),$$

де $n(t)$ — шумова фонові завада.

У цих прикладах спостерігаються такі ситуації:

- 1) реверберація та шум практично відсутні;
- 2) реверберація є, шум практично відсутній;
- 3) реверберації практично немає, шум є;
- 4) є шум та реверберація.

Так, у першому прикладі було прийнято $y(t) = x_{test}(t)$, згідно з виразом $y(t) = x_{test} \otimes h(t) + n(t)$. У другому та четвертому прикладах використано як

$h(t)$ запис реальної імпульсної характеристики приміщення для аудиторії, об'ємом 370 м^3 та часом реверберації $0,8 \text{ с}$ [90]. У третьому та четвертому прикладах застосовано стаціонарний білий шум $n(t)$ із нормальним законом розподілу.

Разом з тим у третьому прикладі значення інтегрального відношення сигнал—шум SNR обчислювалося за виразом:

$$SNR = \frac{D_{x_{n\text{ s}p\text{c}h}}}{D_n},$$

де $D_{x_{n\text{ s}p\text{c}h}}$ - дисперсія шумового компонента із спектром мовлення, D_n - дисперсія фонові шумові завади. В четвертому прикладі значення SNR обчислювалося аналогічно із тією різницею, що замість $x_{n\text{ s}p\text{c}h}(t)$ використано згортку $x_{n\text{ s}p\text{c}h}(t) \otimes h(t)$.

У табл. 2.3.1 та на рис. 2.3.5. подано результати оцінювання STI та SNR_k для вказаних модельних прикладів.

Ситуація: T60 (s), SNR (dB)	Спосіб оцінки	STI	E ₁	E ₂	E ₃	E ₄	E ₅	E ₆	E ₇
Немає реверберації та шуму: T60 = 0; SNR = 60	Прямий метод	1,00	27,19	16,34	24,92	21,05	28,52	36,11	25,63
	HM1	0,99	11,29	13,97	16,50	18,88	21,05	22,39	22,94
	HM2	0,96	4,95	11,67	14,12	12,72	19,58	24,28	27,35
Є реверберація, шуму немає T60 = 0,8; SNR = 60	Прямий метод	0,72	6,91	1,40	3,82	4,10	6,25	11,55	11,13
	HM1	0,54	2,81	-1,95	-1,48	0,02	1,89	0,58	7,05
	HM2	0,69	-1,15	-1,76	-1,59	1,67	5,98	13,75	23,73
Реверберації немає, шум є: T60 = 0; SNR = 5	Прямий метод	0,60	21,52	15,63	23,98	7,17	0,09	-5,47	-10,72
	HM1	0,54	5,71	10,26	9,49	6,20	-0,28	-5,92	-11,14
	HM2	0,56	3,10	10,67	12,03	6,35	0,23	-5,64	-11,15
Є реверберація та шум: T60 = 0,8; SNR = 5	Прямий метод	0,41	4,19	0,81	3,18	0,91	-1,21	-11,00	-14,30
	HM1	0,33	1,33	-2,16	-2,16	-1,73	-3,11	-12,66	-16,16
	HM2	0,38	-2,02	-1,90	-1,72	-0,04	-0,71	-9,21	-15,40

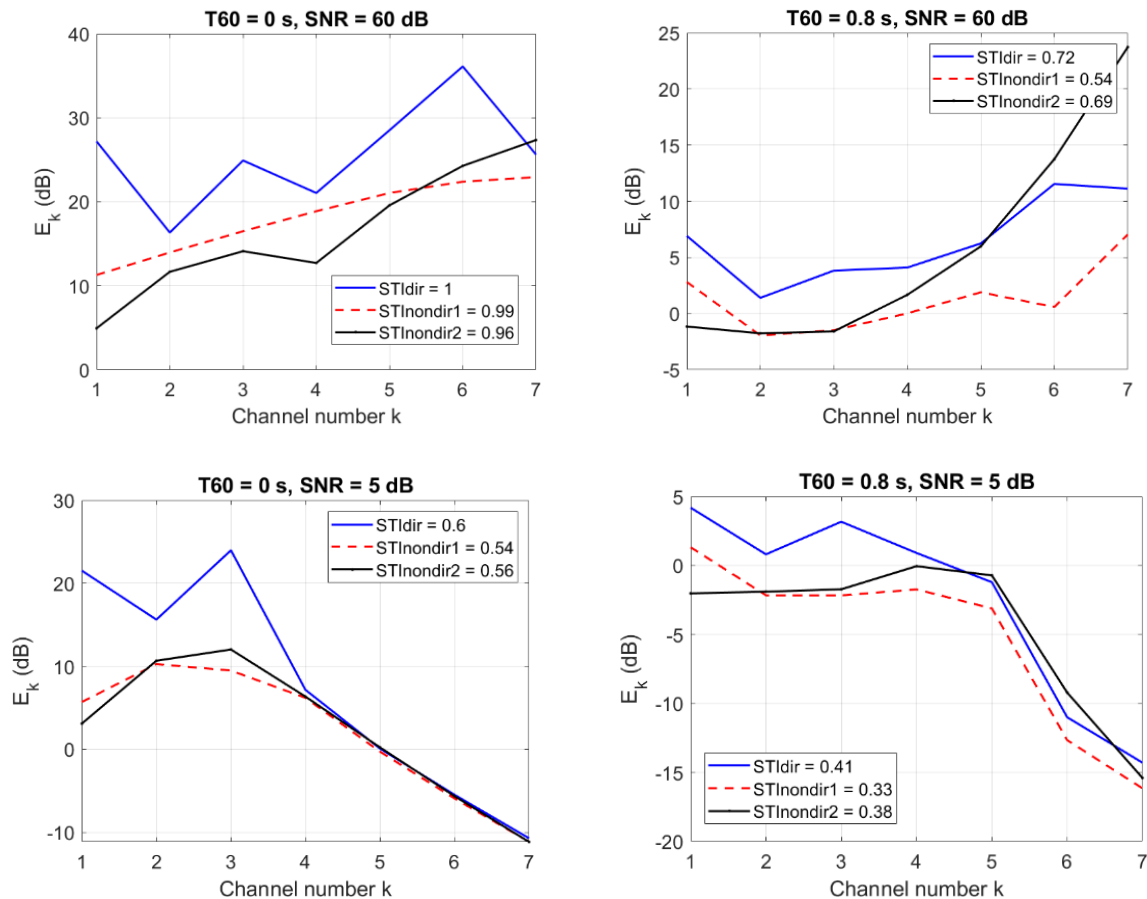


Рис. 2.3.5. STI та E_k для прикладів 1 (а), 2 (б), 3 (в) та 4 (г)

Таким чином, за отриманими результатами можна зробити висновок, що обидва варіанти непрямого методу вимірювання STI призводять до занижених результатів, на відміну від прямого методу. Причому величина такого зміщення може сягати 0,18 для НМ1 й не перевищує 0,04 для НМ2.

Зауважимо, що для малих рівнів шуму та малих значень часу реверберації (рис. 2.3.5а) спостерігається найменше зміщення оцінки НМ1, а найбільше зміщення має місце, коли спотворювальна дія реверберації переважає таку для шуму (рис. 2.3.5б). Отже, величина зміщення оцінки НМ1 суттєво залежить від завадової ситуації.

Щодо оцінки НМ2, то її суттєвою перевагою є незначна величина зміщення та практична незалежність її від завадової ситуації.

Слід зазначити, що результати проведеного моделювання добре узгоджуються із результатами досліджень [95 - 98], а це, у свою чергу, свідчить

про працездатність та правильність функціонування отриманого програмного забезпечення. Враховуючи наведені результати для модельних прикладів, у подальшому була б досить доречною спроба оптимізації параметрів запропонованого тестового сигналу. Оскільки існують різні погляди щодо форми довготривалого спектра мовлення [22], [99], [100], варто було б перевірити стійкість отриманих результатів до зміни форми довготривалого спектра мовлення. Також надалі доцільно провести експериментальні дослідження у реальних приміщеннях, оскільки модель $y(t) = x_{test} \otimes h(t) + n(t)$ є наближенням до реальної дії реверберації та шуму на мовленнєвий сигнал,

2.4 Порівняння прямого та непрямого методів оцінки індексу передачі мовлення

Для порівняння прямого та непрямого методів оцінки індексу передачі мовлення використовувався багатокомпонентний тестовий сигнал, який описаний у підрозділі 2.3 даного розділу:

$$x_{test}(t) = x_{sil}(t) \oplus x_{MLS}(t) \oplus x_{sil}(t) \oplus x_{n\ spch}(t) \oplus \\ \oplus x_{sil}(t) \oplus x_{AM\ spch}(t) \oplus x_{sil}(t) \oplus x_{MLS\ spch}(t) \oplus x_{sil}(t)$$

де $x_{MLS}(t)$ - MLS-сигнал із рівномірним спектром, тривалістю $T_{MLS} = 6$ с; $x_{n\ spch}(t)$ - відрізок гаусівського шуму із спектром мовлення, тривалістю $T_{n\ spch} = 4$ с; $x_{AM\ spch}(t)$ - амплітудно-модульований шум:

$$x_{AM\ spch}(t) = \xi(t) \sqrt{f_5(t)},$$

із спектром мовлення, тривалістю $T_{AM\ spch} = 16$ с; $x_{MLS\ spch}(t)$ MLS-сигнал із спектром мовлення, тривалістю $T_{MLS\ spch} = 24$ с; $x_{sil}(t)$ - пауза тривалістю $T_{sil} = 2$ с; \oplus - символ конкатенації (об'єднання) окремих сигналів та пауз в один багатокомпонентний тестовий сигнал. Форму тестового сигналу $x_{test}(t)$ показано на рис. 2.3.4.

Слід підкреслити, що перевагою даного тестового сигналу є можливість вимірювати STI різними методами в однакових заводових умовах, оскільки окремі компоненти тестового сигналу відокремлені невеликими (тривалістю 2 секунди) паузами, хоча така тривалість є достатньою для вимірювань STI в аудиторіях та офісних приміщеннях, де час реверберації зазвичай не перевищує 1 с. У даному випадку в паузах наявність інтервалів тривалістю не менше ніж 1 с буде фактичною, а дія реверберації не відчуватиметься. Саме такі інтервали із фоновою шумовою завадою потрібні як для оцінювання інтегрального відношення сигнал — шум SNR, так і для оцінювання парціальних відношень сигнал — шум $SNR_k, k=1, \dots, 7$. Для полегшення

калібровки вимірювальної системи та подальшого коректного зіставлення результатів вимірювань дисперсії всіх компонентів тестового сигналу зроблено однаковими.

Модельні дослідження

У модельних дослідженнях моделі записів сигналів, спотворених реверберацією та шумом, генерувалися за виразом:

$$y(t) = x(t) \otimes h(t) + n(t),$$

де $h(t)$ - імпульсна характеристика приміщення, $n(t)$ - шум (використано модель рожевого шуму), \otimes - символ операції згортки.

У даній роботі розглянуто два випадки дії завад, а саме, коли реверберація відсутня ($T_{60} = 0$ с) та реверберація наявна ($T_{60} = 0,8$ с). В якості ІХ приміщення $h(t)$ використано одну із оцінок імпульсної характеристики аудиторії університету в місті Аахен (Німеччина) [46].

Натурні дослідження

Натурні дослідження здійснювалися у Національному технічному університеті України «Київський політехнічний інститут імені Ігоря Сікорського» (Україна) в аудиторії 209 корпусу 12. На рис.2.4.1 представлено план приміщення та місця розміщення в ньому штучної голови із прикріпленими до неї двома вимірювальними мікрофонами:

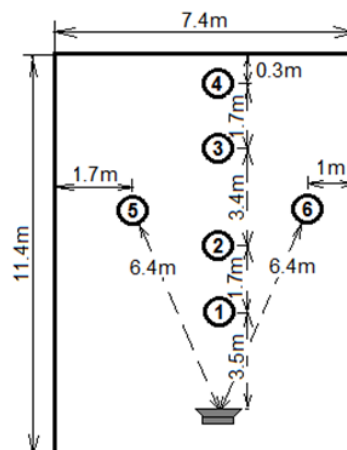


Рис. 2.4.1. План аудиторії 209 та місця розміщення в ній штучної голови

Вимірювальний комплекс складався із пари вимірювальних конденсаторних мікрофонів Superlux ECM-999 (Тайвань), які кріпилися біля вушних раковин саморобної штучної голови (ШГ), активної колонки Genius SP-HF 2,0 500 (Тайвань) та зовнішньої звукової карти Steinberg UR242 (Німеччина—Китай).

При цьому частота дискретизації тестового сигналу та сигналів, записаних із виходів мікрофонів, становила 44,1 кГц, рівень сигналу на відстані 1 м від ШГ становив 69 дБА, рівень шуму L_N в приміщенні був близьким до 40 дБ, рівень сигналу, враховуючи шум L_{S+N} , варіювався від 50 дБА (точка 4) до 60 дБ (точка 2). Разом із тим, застосовуючи вираз

$$SNR = 10 \lg(10^{0,1(L_{S+N} - L_N)} - 1),$$

неважко знайти, що при натурних дослідженнях виконувалося $SNR=10-20$ дБ.

Результати модельних досліджень

На базі моделі $y(t) = x(t) \otimes h(t) + n(t)$ для кожного значення SNR генерувалося 20 зразків тестового сигналу $x_{test}(t)$, оцінювалися середні значення STI та відповідні 95% довірчі інтервали (рис. 2.4.2). Графіки різниць Δ_1 та Δ_2 усереднених значень оцінок STI представлено на рис. 2.4.3.

$$\Delta_1 = STI_{ND1} - STI_{direct}, \quad \Delta_2 = STI_{ND2} - STI_{direct}$$

Дещо дивну поведінку графіків для прямого методу бачимо при аналізі рис. 2.4.2. Дані графіки для $SNR < -20$ дБ стабілізуються, навіть не доходячи до значення $STI=0,05$, хоча мали б наближатися до нуля. Цей факт можна пояснити зміщенням оцінки STI, притаманним прямому ШФМ методу, коли малі відношення сигнал—шум та скінчена тривалість тестового сигналу $x_{test}(t)$ [101]. На рис. 2.4.4 представлено оцінки залежності цього зміщення від тривалості сигналу та відношення сигнал—шум для випадку виключної дії шуму. Величину такого зміщення визначено відносно прогнозних оцінок m_{ki} ,

STI , MTI_k . Середні, за 100 зразками, оцінки STI , а також прогнозні значення STI наведено на рис. 2.4.4а, а на рис. 2.4.4б — різниці між усередненими оцінками та прогнозними значеннями STI :

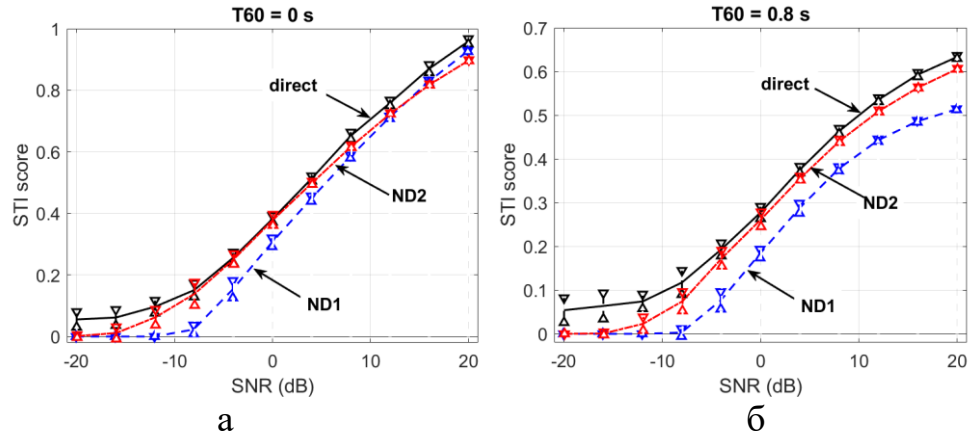


Рис. 2.4.2. Результати оцінювання STI різними методами

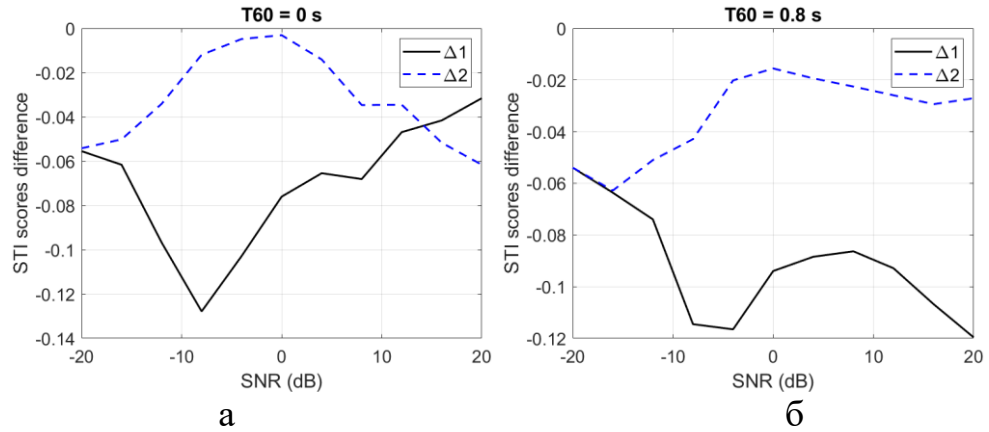


Рис. 2.4.3. Різниця оцінок STI , отриманих різними методами

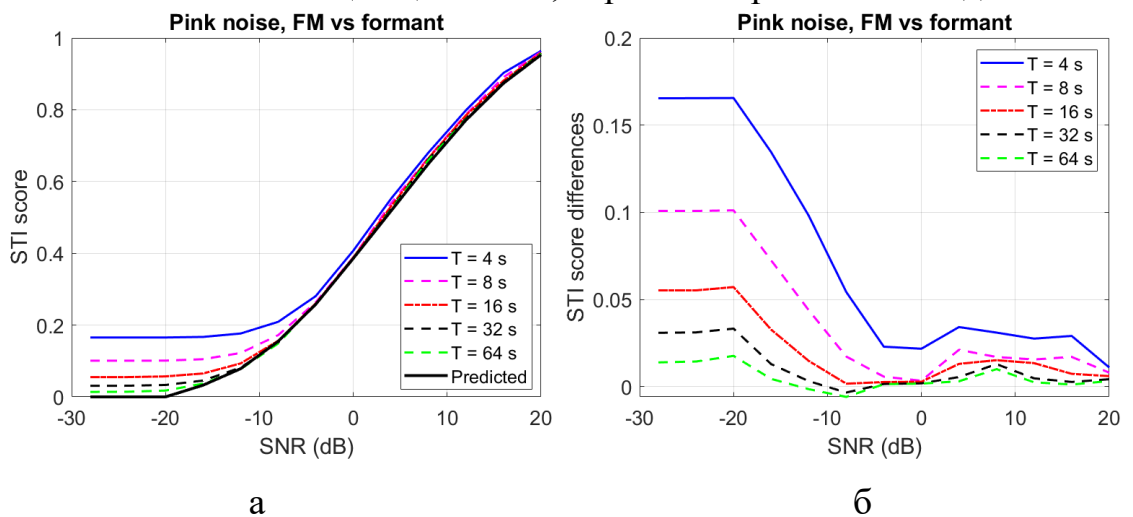


Рис. 2.4.4. Порівняння середніх оцінок STI із прогнозною оцінкою (а) та відповідна різниця (б)

Як можна побачити на рис. 2.4.4б, зміщення не перевищує 0,03 при $T \geq 32$ с у широкому діапазоні значень SNR від мінус 28 дБ до плюс 20 дБ, при $T=16$ с діапазон значень SNR, коли зміщення не перевищує 0,03, починається із мінус 15 дБ, при $T=8$ с — з мінус 10 дБ, а при $T=4$ с — з мінус 5 дБ.

На рис. 2.4.5 представлено графіки різниць між усередненими за 20 вибірками, значеннями оцінок STI, отриманих непрямыми методами, та прогнозними оцінками.

$$\delta_1 = STI_{ND1} - STI_{predict}, \quad \delta_2 = STI_{ND2} - STI_{predict}$$

Зіставлення оцінок HM1 та HM2 із прогнозними оцінками (рис. 2.4.5 та 2.4.4а), дає можливість отримати не лише точніші, але й оптимістичніші результати при дуже малих (менше ніж мінус 10 дБ) відношеннях сигнал—шум. Разом з тим для $SNR > -10$ дБ порівняння оцінок HM1 та HM2 із оцінками прямого ШФМ методу веде практично до таких самих результатів, що і порівняння із прогнозними оцінками.

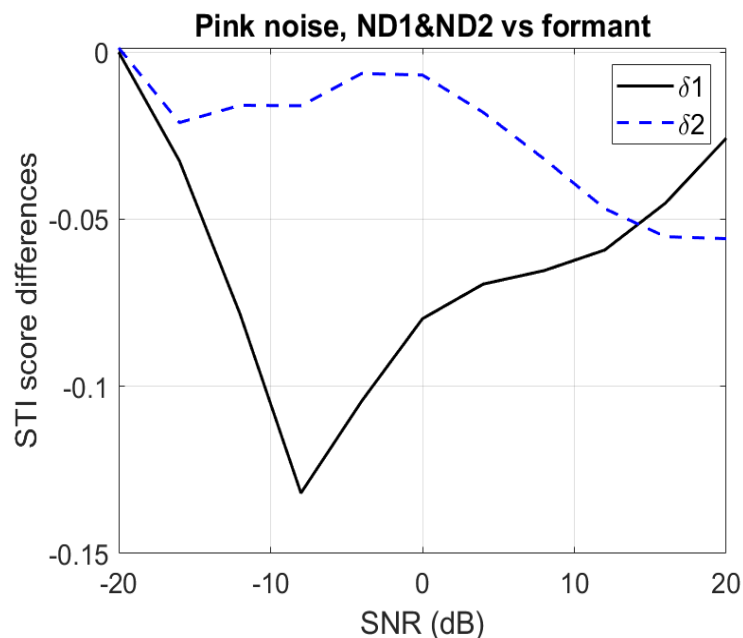


Рис. 2.4.5. Різниця між ND1, ND2 та прогножною оцінками STI

Результати натурних досліджень

Для кожної із 6 точок приміщення (рис. 2.4.1) отримано 5 двоканальних записів із використанням багатокомпонентного тестового сигналу $x_{test}(t)$, при цьому інтегральні відношення сигнал—шум в різних точках приміщення варіювалися в межах 10 —20 дБ.

На рис. 2.4.6 представлено результати обробки записів сигналів, які свідчать, що оцінки STI, отримані непрямим методом NM2, мало відрізняються від таких для прямого методу, водночас оцінки непрямого методу NM1, що добре видно, зміщені в меншу сторону.

Для усереднених, за 5 вибірок, значень оцінок STI графіки різниць Δ_1 та Δ_2 представлено на рис. 2.4.7, де для методу NM2 різниця не перевищує 0,07, у той час як для NM1 різниця може сягати 0,24.

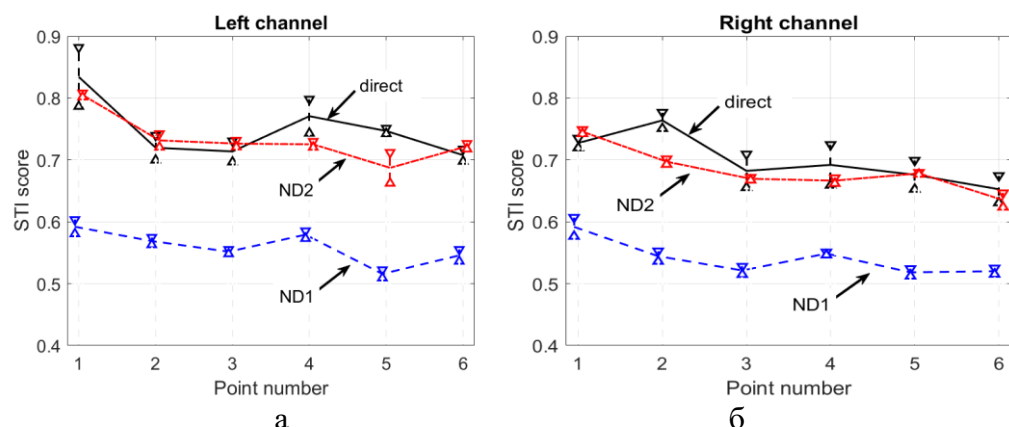


Рис. 2.4.6. Результати оцінювання STI різними методами

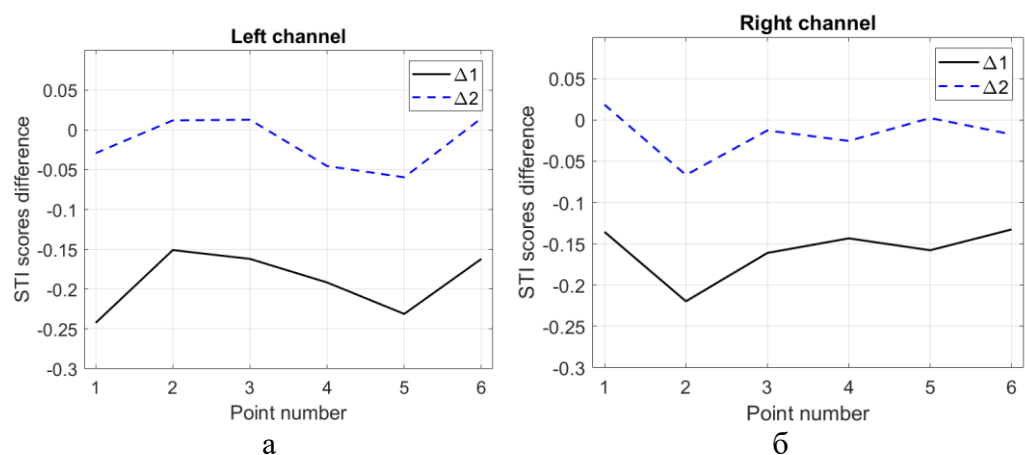


Рис. 2.4.7. Різниця оцінок STI, отриманих різними методами

Крім того, порівнюючи результати натурного та модельного експериментів, можна побачити, що у випадку НМ2 результати узгоджуються досить добре (в обох випадках різниця Δ_2 не перевищувала 0,06), а у випадку НМ1 такого узгодження не спостерігаємо. Справді, якщо в модельному експерименті різниця Δ_1 не перевищувала 0,13 (рис. 2.4.3), то в натурному експерименті ця різниця була в інтервалі 0,14 — 0,24 (рис. 2.4.7).

Оцінки STI, отримані прямим та непрямим методами, що описані у нашому дослідженні, зіставлялися в середовищі Matlab із використанням програмного забезпечення власної розробки. Перевагою такого підходу є контрольованість програмного забезпечення на рівні скриптів, що, у свою чергу, надає певні гарантії щодо правильності отриманих результатів.

Зауважимо, результати досліджень, отримані у нашій роботі, свідчать, що, на відміну від методу НМ2, НМ1 призводить до суттєво занижених результатів (на 0,13 у модельному експерименті та на 0,14—0,24 в натурному), порівняно із прямим методом. Застосуванням однакових програмних модулів можна підтверджувати відсутність помилки в обчисленнях STI непрямыми методами НМ1 та НМ2, що означає, що отримана різниця оцінок STI спричинена різницею тестових сигналів, а не різницею використаного програмного забезпечення.

Слід також зауважити, що у модельних дослідженнях ми розглянули лише одне значення часу реверберації ($T_{60}=0,8$ с), у натурному ж експерименті час реверберації був дещо нижчим. Значення оцінок T20 часу реверберації, отриманих для приміщення рис. 3 за результатами обробки MLS-сигналу та MLS-сигналу зі спектром мовлення, представлено на рис. 10. Тут середнє, за всіма точками та за обома каналами, значення оцінок T20 для MLS-сигналу є близьким до 0,6 с, водночас для MLS-сигналу зі спектром мовлення воно є суттєво нижчим та близьким до 0,53 с. Через те, що зі зростанням часу реверберації розбірливість мовлення в приміщенні спадає, ці показники оцінок

часу реверберації добре узгоджуються із результатами модельного та натурного експерименту, де для НМ1 було отримано нижчу розбірливість, ніж для НМ2. Таке узгодження можна вважати доказом правильності результатів, які було отримано у нашому дослідженні.

Висновки до розділу 2

1. В умовах дії шумової завади повні модуляційний та ФМ методи вимірювань STI забезпечують практично однакову точність вимірювань в діапазоні відношень сигнал-шум від мінус 28 дБ до плюс 28 дБ та в діапазоні тривалості тестових сигналів від 4 с до 64 с.
2. Формантний метод є найбільш точним та найбільш економним в сенсі часу вимірювань, якщо дія шумової завади переважає над дією реверберації.
3. Отримані результати залежності зміщення та стандартного відхилення оцінок STI, отриманих повним та швидким ФМ методами, від відношення сигнал-шум та від тривалості тестових сигналів, дозволяють виконувати обґрунтований вибір тривалості тестового сигналу із врахуванням особливостей завадових умов.
4. Запропоновано використовувати чотирьохкомпонентний тестовий сигнал для порівняння точності двох непрямих методів (НМ1, НМ2) оцінки STI із точністю прямого ШФМ методу.
5. Показано, що обидва варіанти (НМ1, НМ2) непрямого методу вимірювання STI призводять до занижених результатів, порівняно із прямим методом, проте різниця середніх значень оцінок STI, отриманих прямим методом та методом НМ2, суттєво менша за таку для непрямого методу НМ1. Так, для НМ2 середні значення оцінок STI відрізняються від таких для прямого методу не більше ніж на 0.06 в широкому діапазоні відношень сигнал-шум від мінус 20 дБ до плюс 20 дБ, при цьому для НМ1 ця різниця є суттєво більшою та може сягати 0.25.

РОЗДІЛ 3. АКУСТИЧНА ЕКСПЕРТИЗА ПРИМІЩЕННЯ ІЗ ВИКОРИСТАННЯМ ШТУЧНОЇ ГОЛОВИ

В даному розділі наведені результати досліджень акустичних характеристик приміщення за допомогою системи «Штучна голова», що містить акустичну апаратуру напівпрофесійного та непрофесійного рівня. Наведена методика коригування частотної характеристики вимірювального тракту. Виконано експериментальні дослідження для двоканального оцінювання розбірливості мови, спотвореної реверберацією.

3.1. Коригування частотної характеристики вимірювального тракту

У праці [46] представлено характеристики апаратно-програмного комплексу, який використовувався при оцінюванні бінауральних імпульсних характеристик (ІХ) різних приміщень університету RWTH Aachen (Германія). На рисунку 3.1.1 показана спрощена схема процедури вимірювання ІХ приміщення. Тестовий звуковий сигнал $x(t)$ випромінюється гучномовцем, розміщеним в точці приміщення, де найімовірніше знаходиться диктор. Записують відгук приміщення $y(t)$ на стимул $x(t)$ за допомогою пари мікрофонів, розташованих поблизу вушних раковин штучної голови, яка розміщена у тій точці приміщення, де можуть знаходитися слухач. Через те, що амплітудно-частотні характеристики гучномовця та мікрофона не є рівномірними в смузі частот тестового сигналу, замість ІХ приміщення $h_r(t)$ насправді буде оцінено ІХ $h_{\Sigma}(t)$ системи «гучномовець—приміщення—мікрофон»:



Рис. 3.1.1. Спрощена схема процедури оцінювання ІХ приміщення

Запис сигналів з виходів мікрофонів виконувався із 24-бітовою точністю, із частотою дискретизації 48 кГц за допомогою професійної аудіоапаратури:

- акустичний двоканальний активний студійний монітор Genelec 8130, що має цифровий інтерфейс AES/EBU;
- всенаправлені конденсаторні вимірювальні мікрофони Beyerdynamic MM1 та мікрофонний підсилювач RME Octamic II разом із заудіоінтерфейсом RME Multiface II;
- штучна голова HMS2 фірми «HEAD acoustics».

При побудові тестового сигналу базовим елементом, який повторювався 17 разів, слугував mls-сигнал із 2^{16} вибірок, що відповідає тривалості сигналу 1,36 с при частоті дискретизації 48 кГц. ІХ приміщення обчислювали шляхом усереднення 16 останніх оцінок взаємної кореляційної функції (ВКФ) між сигналом з виходу мікрофона та тестовим сигналом, що дало можливість підвищити відношення сигнал-шум на 12 дБ.

В даній дисертаційній роботі, як і в праці [46], використано такі ж тестові сигнали, лише частота дискретизації становила 44,1 кГц, а не 48 кГц. Більш суттєвою відмінністю розробленого програмно-апаратного комплексу було використання побутових та напівпрофесійних гучномовців та мікрофонів:

- 1) малогабаритний MP3-плеєр Optima МК-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, 65–20 000 Гц, 73 дБ);
- 3) активні акустичні монітори Yamaha MSP5 (67 Вт, 50–40 000 Гц, 102 дБ);
- 4) вимірювальні конденсаторні мікрофони Superlux ECM-999 (20 Гц – 20 кГц, динамічний діапазон 106 дБ, відношення сигнал—шум 70 дБ).

Зважаючи на порівняно невисоку якість використаної апаратури, стає зрозумілою важливість коригування результатів вимірювань ІХ приміщення. Щодо праці [46], то інформація про подібне коригування при отриманні

результатів роботи, на жаль, відсутня. Можливо, це пояснюється тим, що таке коригування вважалося зайвим з огляду на професійний рівень аудіотехніки.

За працею [102], ІХ $h_{\Sigma}(t)$ системи «гучномовець-приміщення-мікрофон» можна обчислити, використовуючи вираз:

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

де $\langle \cdot \rangle$ – символ математичного очікування; $K_X(\tau)$ – автокореляційна 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 $H_{\Sigma}(f) = H_r(f) \cdot H_{lm}(f)$, отримуємо:

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

Фур'є-перетворення від $K_{XY}(\tau)$ приводить до виразу

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

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

Для забезпечення точних вимірювань ІХ приміщення згідно рис. 3.1.1, потрібно знати ІХ підсистеми «гучномовець-мікрофон» (ГМ). Для цього необхідно проводити вимірювання в заглушеній кімнаті на невеликій відстані між мікрофоном та гучномовцем, тоді отримаємо:

$$K_{XYa}(\tau) = D_X \cdot h_{lm}(\tau),$$

$$S_{XYa}(f) = H_{lm}(f) \frac{N_0}{2};$$

Звідси випливає, що форма вихідного сигналу $y(t)$ буде визначатися властивостями підсистеми ГМ.

Скориговану, із врахуванням неідеальності підсистеми ГМ, оцінку ІХ приміщення можна отримати згідно із рівнянням

$$h_r(t) = \mathbb{F}^{-1} \left\{ \frac{|H_\Sigma(f)|}{|H_{lm}(f)|} \cdot \exp[\theta_\Sigma(f) - \theta_{lm}(f)] \right\}.$$

Використовуючи вищезазначені вирази, необхідно забезпечити, щоб спектр потужності тестового сигналу $x(t)$ був рівномірним у смузі частот $F_s/2$, де F_s — частота дискретизації, оскільки при практичних вимірюваннях запис та обробка сигналів виконуються у цифровій формі. Таким чином, можна вважати, що для обраного тестового MLS-сигналу ця вимога виконується досить точно, бо рівень бокових пелюсток є близьким до мінус 47 дБ, а ширина центрального сплеску є мінімально можливою та становить $1/F_s$ (рис. 3.1.2).

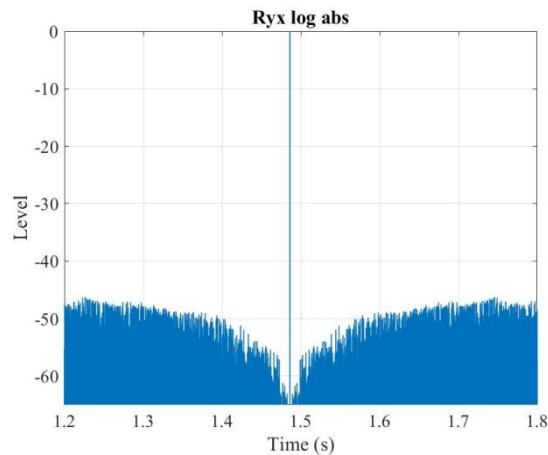


Рис 3.1.2. Автокореляційна функція MLS-сигналу

На рис. 3.1.3 наведено приклади форми сигналів $x(t)$, та $y(t)$ на вході та виході підсистеми ГМ із гучномовцем Genius, а на рис. 3.1.4 — форми відповідних спектрів потужності, до 300 Гц спостерігається провал глибиною мінус 15 дБ.

Шляхом взаємного кореляційного аналізу сигналів $x(t)$, та $y(t)$ можна визначити, згідно із $K_{xya}(\tau)$ та $S_{xya}(f)$, форму АЧХ підсистеми ГМ. Форму такої оцінки ВКФ для підсистеми ГМ із гучномовцем Genius подано на рис. 3.1.5.

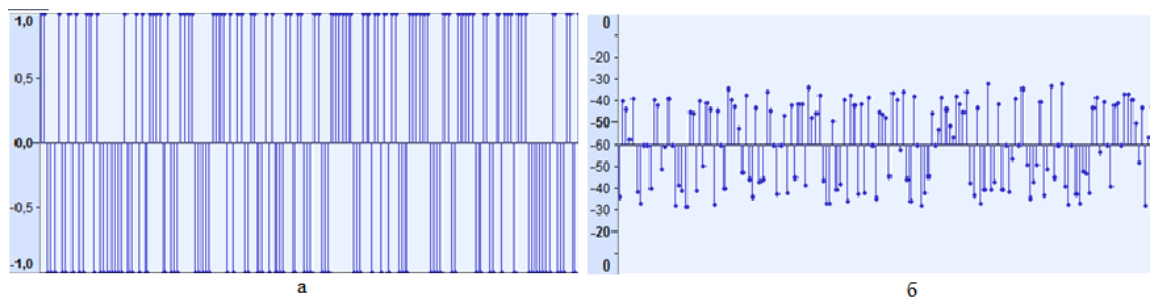


Рис. 3.1.3. Сигнали $x(t)$ (а) та $y(t)$ (б) на вході та виході підсистеми ГМ

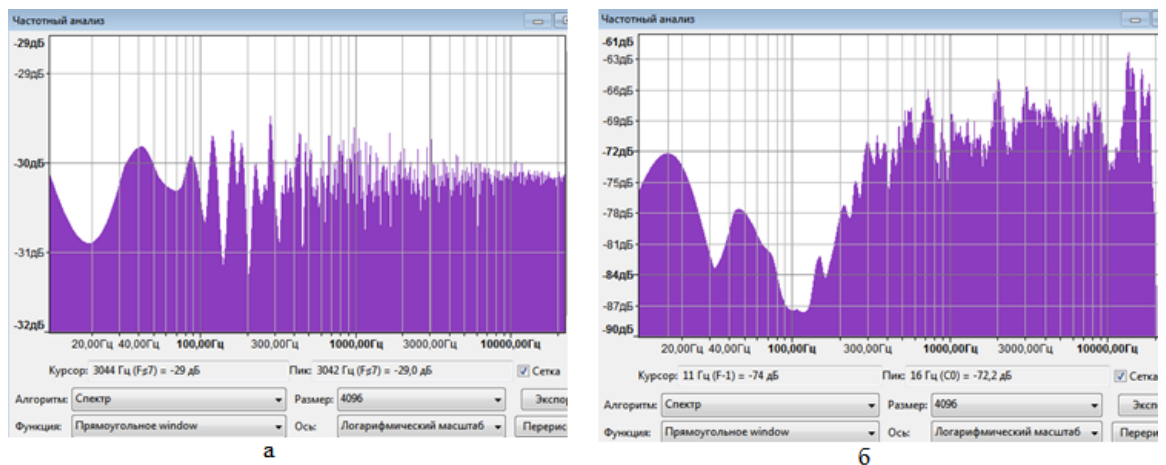


Рис. 3.1.4. Спектри потужності сигналів $x(t)$ (а) та $y(t)$ (б) на вході та виході підсистеми ГМ.

Як видно з рис. 3.1.3, вибірки сигналу $y(t)$, на відміну від вибірок сигналу $x(t)$, мають різну висоту, що вказує на спотворення входного сигналу $x(t)$ підсистемою ГМ. На рис.3.1.4 оцінки спектрів потужності більш інформативні, оскільки дають можливість визначити, на яких саме частотах

відбувається спотворення вхідного сигналу. Таким чином, дуже добре видно, що якщо нерівномірність спектра вхідного сигналу (рис. 3.1.4а) не перевищує $\pm 0,5$ дБ, то нерівномірність спектра вихідного сигналу в смузі частот від 300 Гц до 10 кГц становить ± 3 дБ, у смузі ж частот від 10 кГц до 16 кГц є близькою до ± 6 дБ.

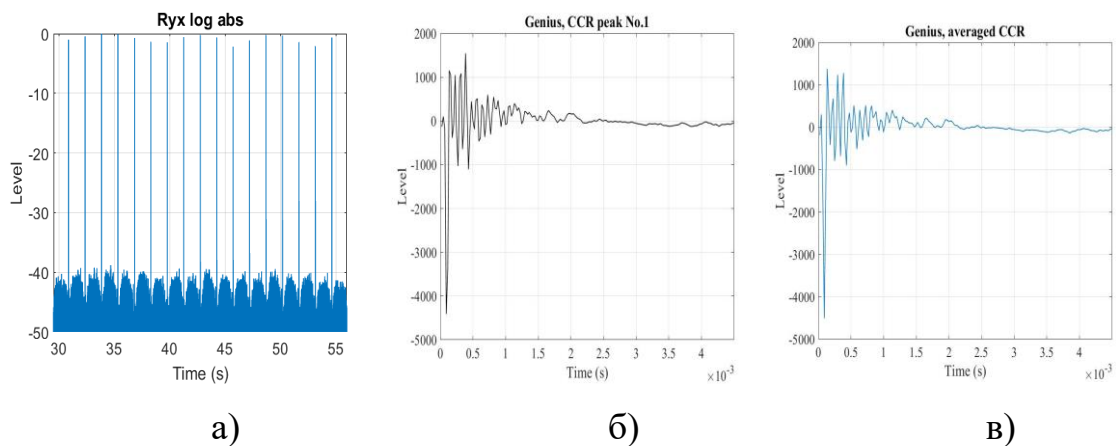


Рис. 3.1.5. ВКФ сигналів $x(t)$ та $y(t)$ (а), форма першого (б) та усередненого (в) сплесків ВКФ

За виразами $K_{xya}(\tau)$ та $S_{xya}(f)$, модуль Фур'є-перетворення усередненого сплеску ВКФ є оцінкою АЧХ підсистеми ГМ.

Оцінки АЧХ $|H_{lm}(f)|$ підсистеми ГМ для трьох видів гучномовців представлено на рис. 3.1.6, де наведено графіки оцінок АЧХ $|H_{lm}(f)|$ чотирьох типів:

- 1) *Type 0* — оцінка спектра потужності Велча (вагове вікно при оцінці не застосовується) як усереднення спектрів потужності 17 сегментів сигналу $y(t)$, при цьому кожен сегмент має довжину 65536 вибірок:
 $[Py, f_{yy}] = \text{pwelch}(y, \text{rectwin}(65536), 0, 65536, F_s);$
- 2) *Type 1* — модуль дискретного перетворення Фур'є (ДПФ) від результату усереднення 17 піків ВКФ, згладжений за частотою, трикутним вікном завширшки 50 Гц;

- 3) *Type 2* — результат усереднення (17 разів) модулів ДПФ від 17 сплесків ВКФ, згладжений за частотою трикутним вікном завширшки 50 Гц;
- 4) *Type 3* — показники оцінки спектра потужності Велча із використанням вагового вікна Ханна, що забезпечує роздільну здатність 50 Гц:
 $[P_{yy}, f_{yy}] = \text{pwelch}(y, \text{hann}(\text{delw}), \text{round}(\text{delw}/2), 65536, F_s), \text{delw} = (1,5 F_s) / df$
 $F_s = 44100 \text{ Гц}, df = 50 \text{ Гц}.$

За допомогою цифрової фільтрації незгладженої оцінки спектра було реалізовано процедуру згладжування трикутним вікном, використану в оцінках 1-го та 2-го типів (*Type 1*, *Type 2*), для чого було застосовано рекурсивний фільтр першого порядку. На першому етапі цієї двоетапної процедури виконується зліва направо згладжування прямокутним вікном, таким чином отримуємо:

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

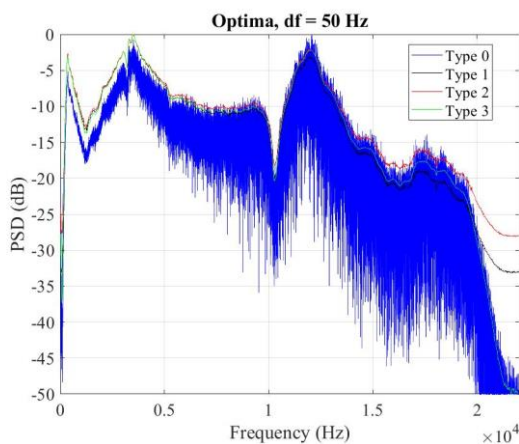
а на другому етапі навпаки — справа наліво:

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

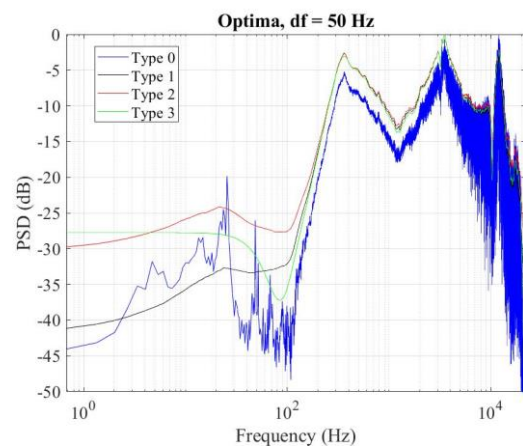
де $N = \frac{N_{fft}}{F_s} \cdot \frac{df}{2}$, $N_{fft} = 65536$, $F_s = 44100 \text{ Гц}$, $df = 50 \text{ Гц}$.

З виразу видно, що значення N обчислювалося, враховуючи ширину прямокутного вікна, що становить 25 Гц. Усе це зумовлено тим, що подвійне згладжування прямокутним вікном еквівалентне одиночному згладжуванню трикутним вікном подвоєної ширини. Така необхідність в застосуванні двох напрямків фільтрації пояснюється тим, що поза краями первинної оцінки спектра, що підлягає згладжуванню, не потрібно дописувати якісь вибірки. Отже, порівнюючи між собою наведені на рис. 3.1.5 оцінки АЧХ $|H_{lm}(f)|$ для різних гучномовців, робимо такі висновки:

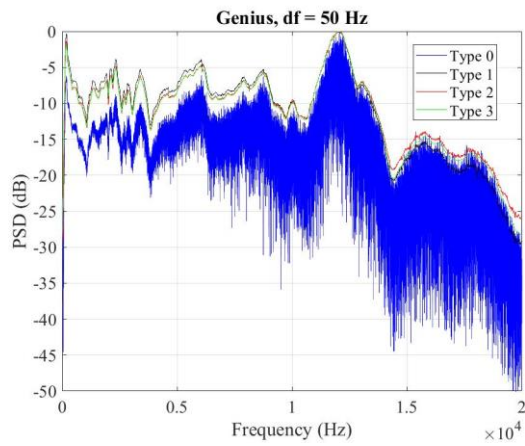
- по-перше, слід віддати перевагу активній колонці Genius, зважаючи на ступінь фактичної нерівномірності та фактичний частотний діапазон оцінки АЧХ $|H_{lm}(f)|$. Найцікавіше те, що акустичний монітор Yamaha виявився неконкурентним (що можна пояснити його фізичним старінням);
- по-друге, на графіках чітко спостерігається незгладжена оцінка $|H_{lm}(f)|$ нульового типу, яку недоцільно використовувати для подальшої корекції через неприпустимо велику дисперсію в області середніх та високих частот;
- по-третє, найбільш перспективною серед згладжених оцінок є оцінка третього типу, яка найкраще забезпечує зниження дисперсії оцінки без підвищення надміру її зміщення. Слід зауважити, що певна суб'єктивність при виборі типу оцінки є неминучою, бо істинна АЧХ підсистеми ГМ є невідомою.



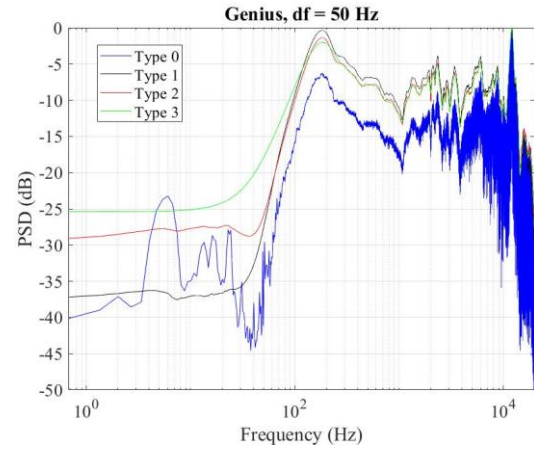
а



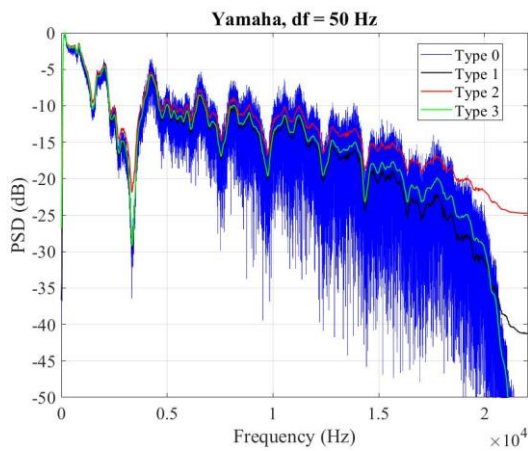
б



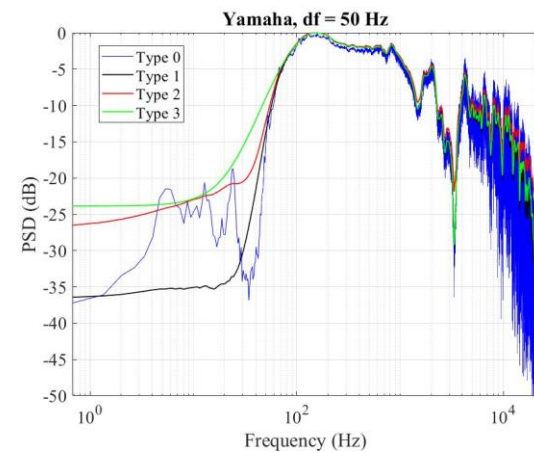
В



Г



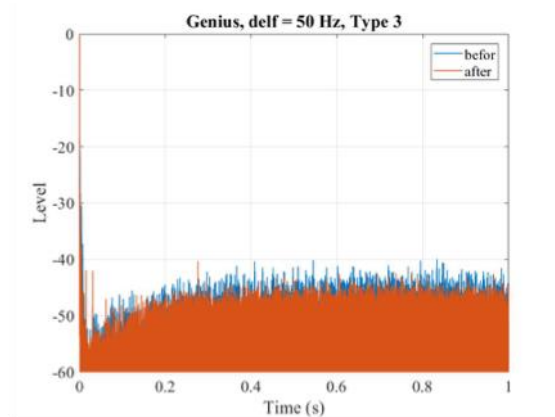
Д



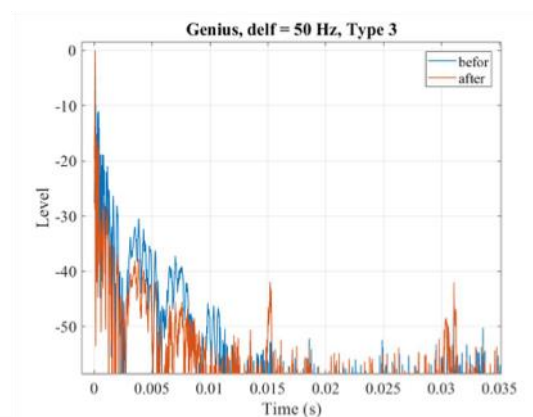
Е

Рис. 3.1.6. Оцінки АЧХ різних гучномовців (заглушене приміщення)

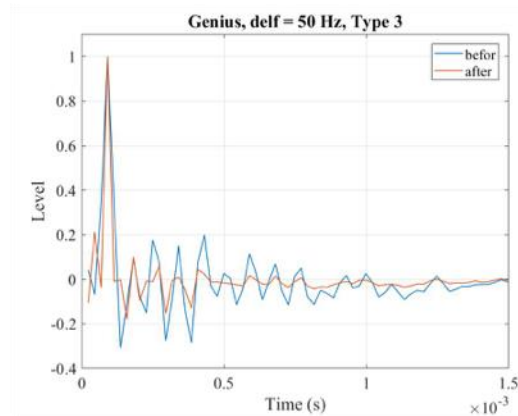
Результати обчислень ІХ підсистеми ГМ із використанням згладженої оцінки АЧХ для гучномовця Genius для різних значень часового масштабу представлено на рис. 3.1.7.



а)



б)



в)

Рис. 3.1.7. Скориговані оцінки ІХ для гучномовця Genius

Варто зауважити, що ІХ підсистеми ГМ можна називати «апаратною функцією» розробленого комплексу, яка характеризує його потенційні можливості у значенні точності вимірювань ІХ приміщень.

Як бачимо, результат коригування має як позитивні, так і негативні якості. Серед негативних можна відзначити появу окремих дискретних сплесків достатньо високого рівня -40 дБ (рис. 3.1.7а). Серед позитивних:

- 1) рівень бокових пелюсток знизився приблизно на 3–4 дБ (рис. 3.1.5а);
- 2) рівень найближчих до нього сплесків став нижчим майже на 6 дБ, а головний пік ІХ став вужчим у 2 рази (рис. 3.1.5б, в).

Зважаючи на зазначені вище сплески, можна припустити, що вони спричинені збільшенням дисперсії оцінки АЧХ на високих частотах, через ділення на малі значення згладженої оцінки $|H_{lm}(f)|$. Простим та очевидним способом пригнічення таких сплесків є множення відкоригованої оцінки АЧХ на спектральне вагове вікно, що відповідає в часовій області процедури згладжування оцінки ІХ через ковзне усереднення. В даній роботі розглянуто два види вагового вікна $w_1(r)$ та $w_2(r)$.

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

Де k – визначає інтервал, який відповідає половині 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:

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

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

де – ΔN половина ширини вікна Ханна у вибірках.

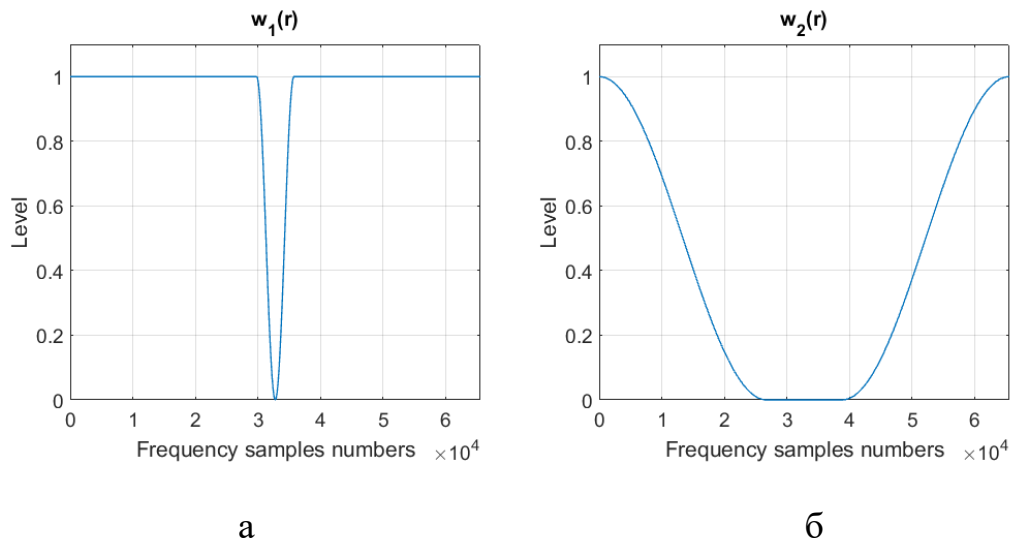


Рис. 3.1.8. Вагове вікно $w_1(r)$ для $f_1 = 19$ кГц (а) та вікно $w_2(r)$ для $f_2 = 18$ кГц (б)

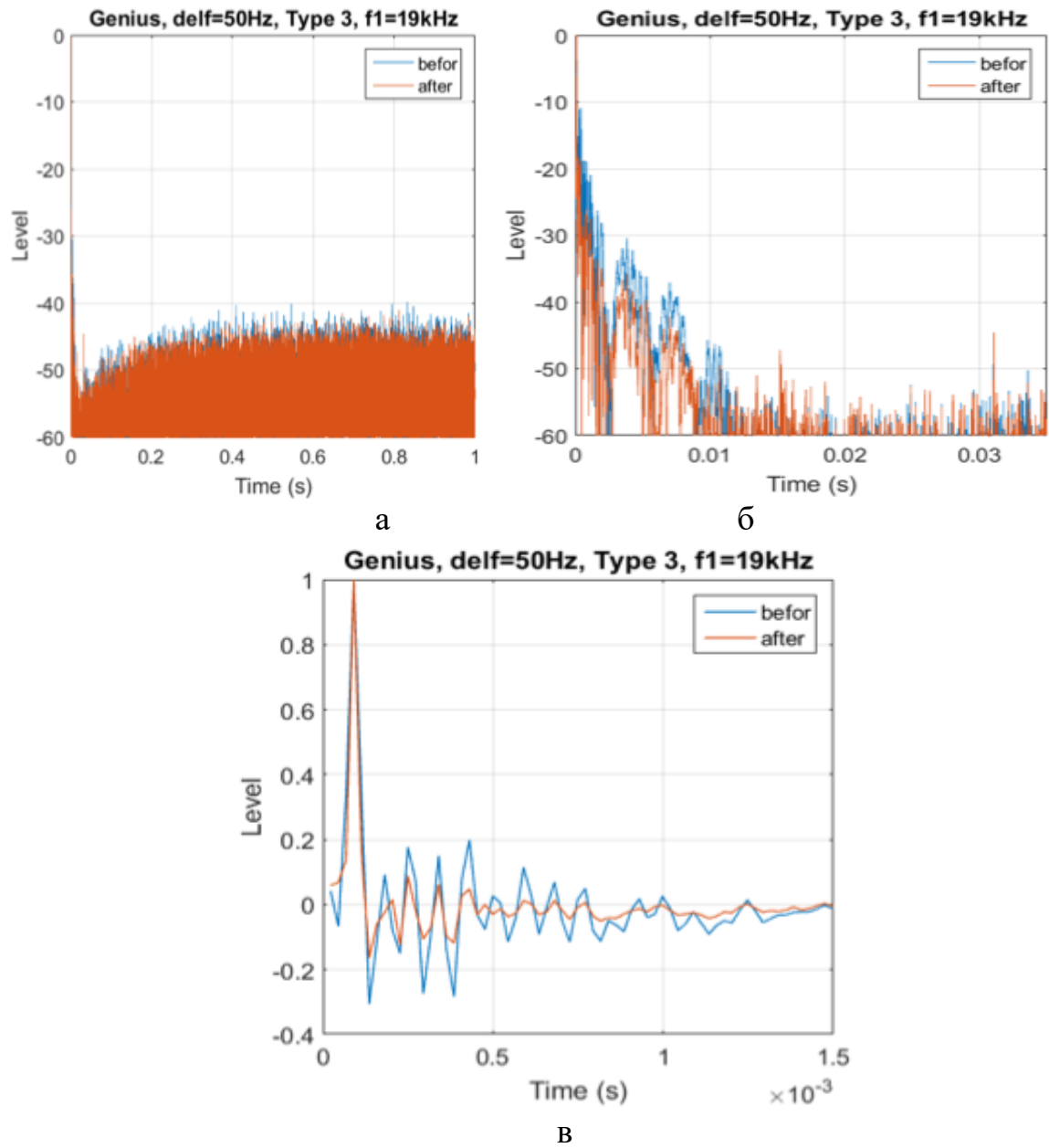


Рис. 3.1.9. Скориговані ІХ для гучномовця Genius, із вікном $w_1(r)$ для $f_1 = 19 \text{ кГц}$

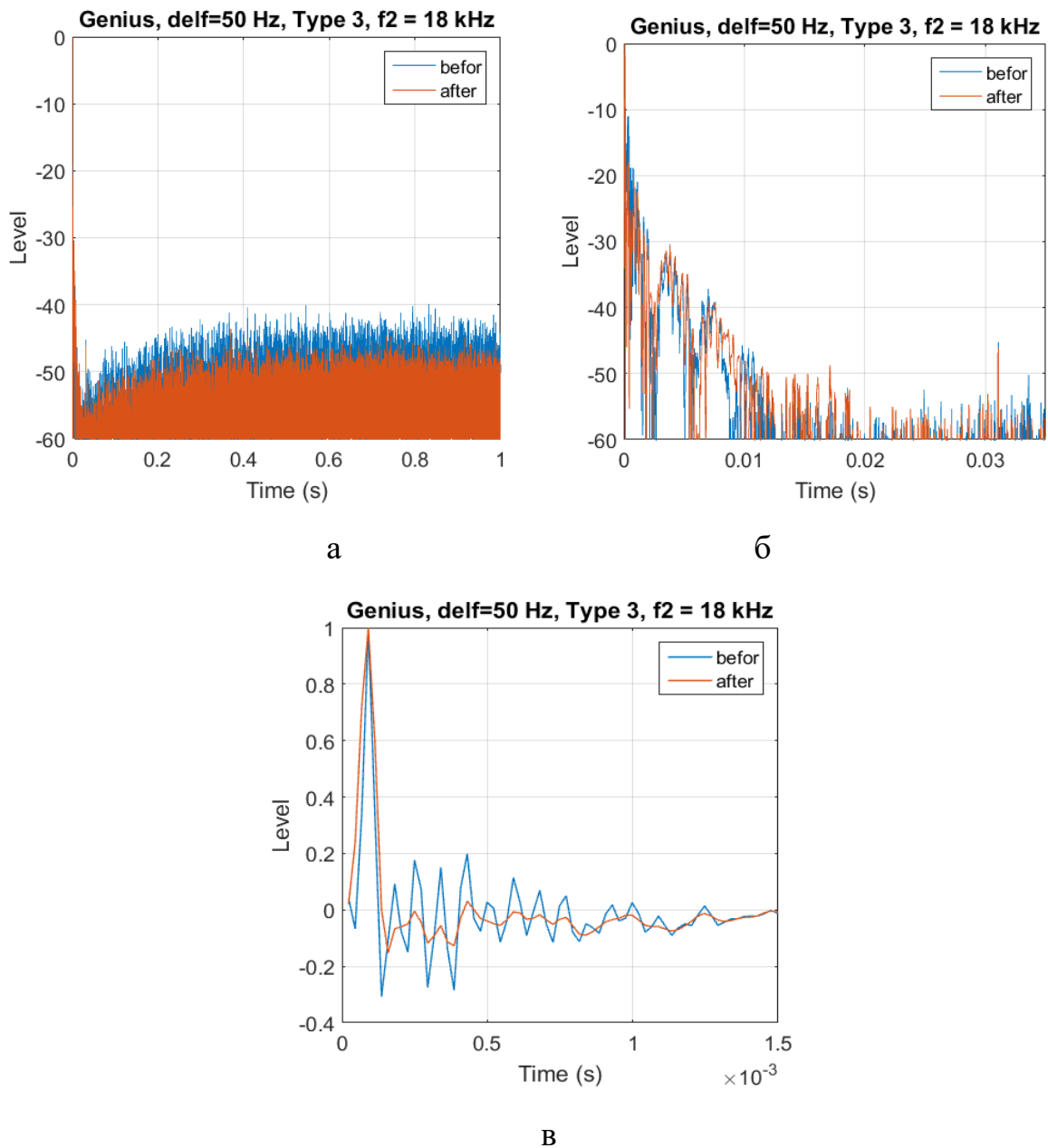


Рис. 3.1.10. Скориговані ІХ для гучномовця Genius, із вікном $w_2(r)$ для $f_2 = 18 \text{ кГц}$

На рис. 3.1.8 представлено графіки вагових вікон для $f_1 = 19 \text{ кГц}$ та $f_2 = 18 \text{ кГц}$ ($f_2 = \Delta N \cdot F_s / N_{fft}$), на яких добре видно, що в окремому випадку, коли $f_1 = 0, f_2 = F_s / 2$, ці вікна збігаються: $w_1(r) = w_2(r)$.

На рис. 3.1.9 та на рис. 3.1.10 наведено результати застосування вікон $w_1(r)$ і $w_2(r)$, порівнюючи які, можна дійти висновку, що вікно $w_2(r)$ з величиною $f_2 = 18 \text{ кГц}$ є більш перспективним, бо дає можливість зменшити

рівень бокових пелюсток оцінки ІХ підсистеми ГМ на 4 дБ, а вікно $w_1(r)$ з параметром $f_l=19$ кГц — зменшити рівень бокових пелюсток лише на 1 дБ (рис. 9а та рис. 10а). Вікно $w_1(r)$ в інтервалі 0–10 мс є кращим, оскільки забезпечує рівень бокових пелюсток на 6 дБ нижче, ніж вікно $w_2(r)$. Але така різниця не є принциповою, через те, що у випадку вікна $w_2(r)$ рівень цих бокових пелюсток не перевищує -30 дБ, а ширина головного сплеску ІХ на рівні -30 дБ не перевищує 2 мс, а це, у свою чергу, забезпечує достатню точність оцінювання ІХ приміщення для інженерних застосувань.

Зауважимо, що таке застосування вагового вікна $w_2(r)$ можна розглядати як регуляризацію з множителем регуляризації $w_2(r)$ та параметром регуляризації f_2 , а тако як рішення нестійкої оберненої задачі [15].

3.2. Вимірювання бінауральних імпульсних характеристик приміщення та розбірливості мовлення, спотвореного реверберацією.

Метою даного дослідження є оцінювання індексу STI та словесної розбірливості мови W та для різних точок кількох лекційних приміщень, а також зіставлення отриманих результатів між собою та з результатами суб'єктивного оцінювання розбірливості мови.

За умов, що ревербераційна завада переважає над шумовою завадою, оцінювання розбірливості мови модуляційним [104, 105] або формантно-модуляційним [26, 27] методами стає можливим саме після обчислення ІХ приміщення. Початком такого оцінювання є обчислення так званих коефіцієнтів передачі модуляції за формулою Шредера [106]:

$$m_{ki} = \frac{\left| \int_0^{\infty} h_k^2(t) \exp(-j2\pi F_i t) dt \right|}{\int_0^{\infty} h_{ki}^2(t) dt}$$

де $h_k(t)$ — результат фільтрації функції $h_r(t)$ смуговим фільтром (на практиці використовують 7 октавних фільтрів із центральними частотами від 125 Гц до 8 кГц); F_i - частота модуляції (на практиці використовують 14 значень F_i в діапазоні від 0,63 Гц до 12,5 Гц).

Наступними кроками є обчислення ефективних відношень сигнал-шум для кожної смуги та кожної модуляційної частоти:

$$E_{ki} = 10 \lg \frac{m_{ki}}{1 - m_{ki}},$$

середнього індексу передачі модуляції:

$$MT_k = \frac{1}{14} \sum_{i=1}^{14} T_{ki},$$

$$T_{ki} = \begin{cases} 0, & E_{ki} < -15; \\ (E_{ki} + 15)/30, & -15 \leq E_{ki} \leq +15; \\ 1, & E_{ki} > +15. \end{cases}$$

Останнім кроком, згідно модуляційного методу [18], є обчислення індексу передачі мови:

$$STI = \sum_{k=1}^7 \alpha_k \cdot MT_k - \sum_{k=1}^6 \beta_k \cdot \sqrt{MT_k \cdot MT_{k+1}},$$

де α_k – вагові коефіцієнти, β_k – коефіцієнти надмірності, значення яких для октавних смуг із центральними частотами f_0 наведено в табл. 1.

ТАБЛИЦЯ 3.2.1 ВАГОВІ КОЕФІЦІЄНТИ ТА КОЕФІЦІЄНТИ НАДМІРНОСТІ ДЛЯ
ОКТАВНИХ СМУГ

$f_{0k}, \text{Гц}$	125	250	500	1000	2000	4000	8000
α_k	0,085	0,127	0,230	0,233	0,309	0,224	0,173
β_k	0,085	0,078	0,065	0,011	0,047	0,095	-

Також на цьому останньому етапі можна обчислити словесну розбірливість за формантно-модуляційним методом [26]:

$$W = \begin{cases} 1,54 \cdot A^{0,25} [1 - \exp(-11 \cdot A)], & A < 0,15; \\ 1 - \exp\left(\frac{11 \cdot A}{1 + 0,7 \cdot A}\right), & A \geq 0,15; \end{cases}$$

де A – індекс артикуляції.

Організація досліджень

Дослідження виконувалися із застосуванням розробленого комплексу «Штучна голова», описаного у пункті 3.1 даного розділу. Для вимірювань було обрано 2 аудиторії малого та середнього розміру (рис. 3.2.1).

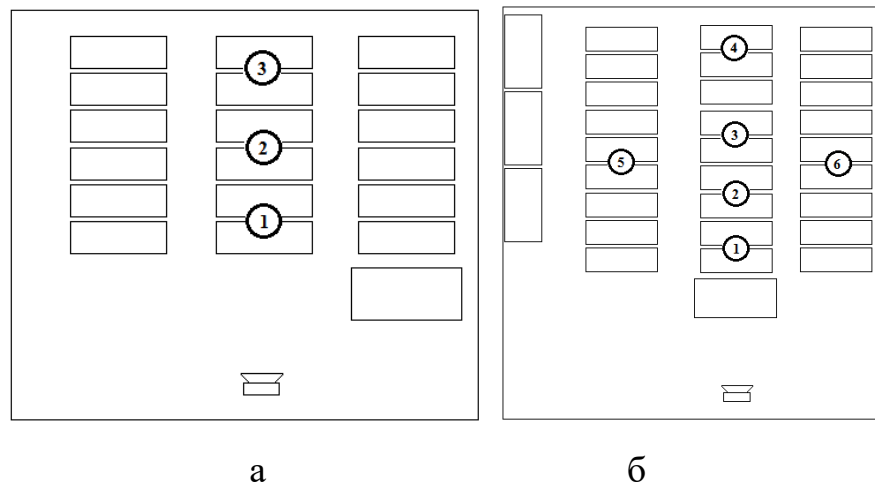


Рис. 3.2.1. Плани малої аудиторії 438 (а) та середньої аудиторії 209 (б)

Мала аудиторія № 438 (рис. 3.2.1а): розміри — $7 \times 7 \times 3,1$ м, два вікна, двері, задня стінка приміщення зашклена, три ряди із 6 парт в кожному ряду та стіл викладача; наповненість приміщення під час запису сигналів становила 11 студентів та викладач. Відстань від гучномовця до номерів 1, 2 та 3 точок, де розміщувалася ШГ, відповідно 2 м, 4 м та 6 м.

Середня аудиторія № 209 (рис. 3.2.1б): розміри — $10 \times 15 \times 3,1$ м, чотири вікна, двері, дві шафи для книг, шафа-гардероб, три ряди із 9 парт в кожному ряду, стіл викладача та 2 додаткових вільних столи, наповненість приміщення під час запису сигналів становила 9 студентів та викладач. Відстань від гучномовця до номерів 1, 2, 3, 4, 5 та 6 точок, де розміщувалася ШГ, відповідно 3 м, 6 м, 9 м, 14 м, 9 м та 9 м.

Результати досліджень

Графіки ІХ розглянутих приміщень та результати вимірювання часу реверберації представлено на рис. 3.2.3, але лише для лівого каналу, бо для правого каналу вони є подібними. Для обчислення часу реверберації (оцінки на рис. 3.2.3) використано обвідну ІХ, побудовану за методом Шредера [26].

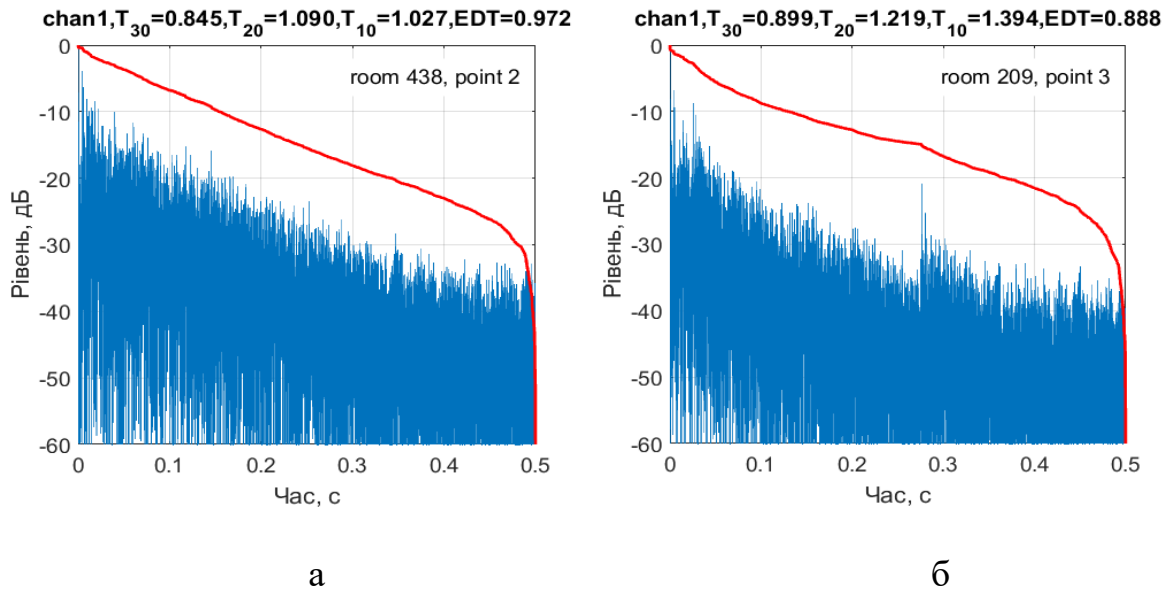


Рис. 3.2.2. ІХ та час реверберації для аудиторій 438 (а) та 209 (б)

При зіставленні ІХ аудиторій 438 (рис. 3.2.2а) та 209 (рис. 3.2.2б) видно, що в аудиторії 209 спостерігається біля 3-х потужних ранніх відлунь на інтервалі часу до 30 мс та ще один виразний сплеск в момент часу 0,27 с, у той час як спад (зниження) енергії в аудиторії 438 є плавнішим. Також бачимо, що в обох приміщеннях час реверберації виявився досить довгим та близьким до 1 с. В аудиторії 209 це пояснюється слабким заповненням порівняно великого приміщення, а в аудиторії 438 — наявністю скляної задньої стінки.

Оцінки характеристик $C_{50} = 10\lg(E_0^{50}/E_0^\infty)$ та $ERB = 10\lg(E_0^{50}/E_0^{10})$ представлено на рис. 3.2.3 та 3.2.4, де $E_a^b = \int_a^b h^2(t)dt$, a та b — час в мілісекундах. Через те, що для $t > 0,5$ с рівень шумової завади переважає над рівнем реверберації (рис. 3.2.2), в якості ∞ при обчисленнях C_{50} використано значення 0,5 с.

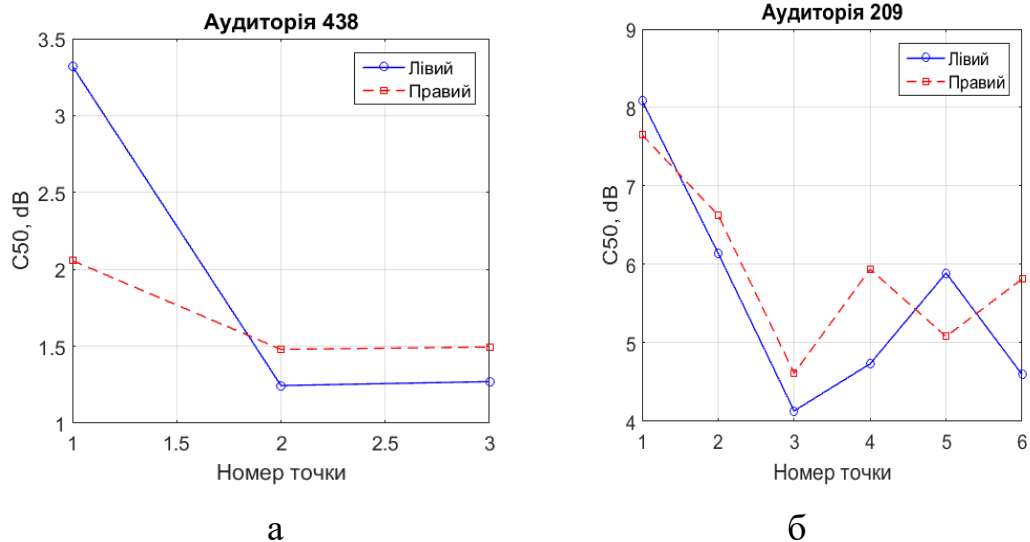


Рис. 3.2.3. Оцінки параметру C_{50} для аудиторій 438 (а) та 209 (б)

Чіткість мовлення, за отриманими оцінками C_{50} , в аудиторії 209 є значно вищою, ніж в аудиторії 438. Ця чіткість швидко спадає із наближенням до центру кімнати, після чого стабілізується на певному рівні. Також цікавим є те, що у цій же аудиторії 209 чіткість мови біля задньої стінки (точка 4) є помітно вищою, ніж у середині кімнати (точка 3), а в аудиторії 438 є подібний ефект, але ледь помітний.

Існування зазначеного ефекту підтверджується поведінкою показника ERB (користь ранніх відбиттів), що характеризує співвідношення енергій ранніх відлунь та прямого звуку. Зауважимо, що якщо брати до уваги, що прямий звук є кращим в сенсі забезпечення високої розбірливості мови порівняно із ранніми відлуннями, то, за рис. 3.2.5а, видно, що в аудиторії 438 розбірливість у лівому каналі є вищою біля задньої стінки (точка 3), ніж у центрі приміщення (точка 2), хоча результати оцінювання ERB у правому каналі свідчать про протилежне. Натомість ситуація в аудиторії 209 (рис. 3.2.5б) не є такою суперечливою: розбірливість мови біля задньої стінки (точка 4) в обох каналах мала би покращитися щодо середини приміщення (точка 3).

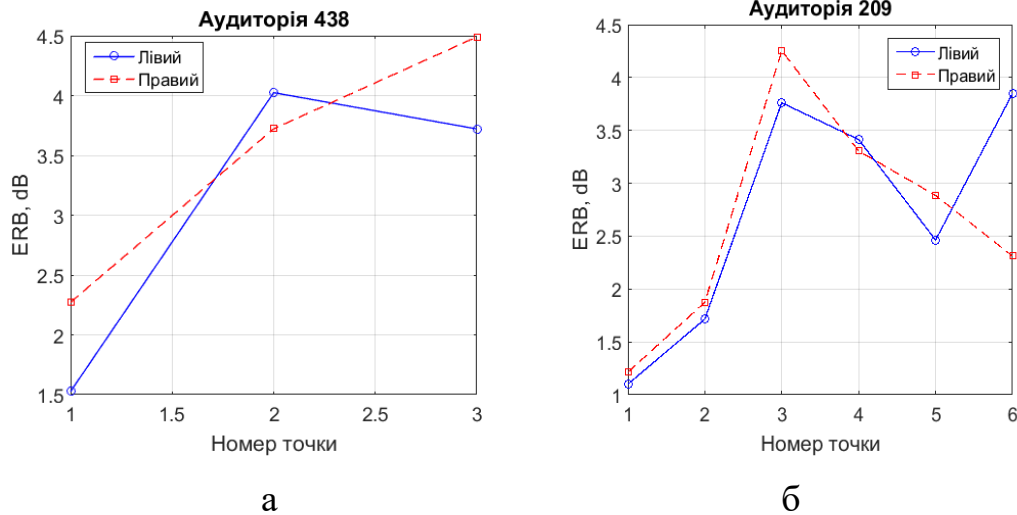


Рис. 3.2.4. Оцінки параметру ERB для аудиторій 438 (а) та 209 (б)

Якщо брати до уваги певну неоднозначність отриманих оцінок C_{50} та ERB, що є непрямыми мірами розбірливості мови, то буде цікаво оцінити розбірливість мови модуляційним та формантно-модуляційним методами у вказаних аудиторіях.

На рис. 3.2.5 представлено результати оцінювання розбірливості мови модуляційним методом, а на рис. 3.2.6 — формантно-модуляційним методом, при якому оцінки були розраховані, використовуючи уточнені коефіцієнти сприйняття, форма яких залежить від смуги частот [24].

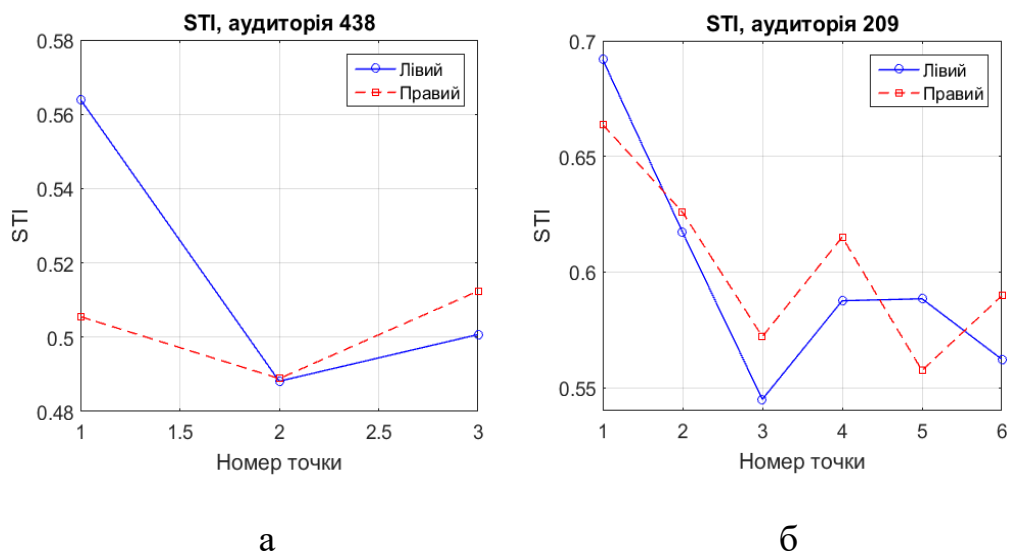


Рис. 3.2.5. Оцінки індексу STI для аудиторій 438 (а) та 209 (б)

Легко також помітити добре узгодження оцінок розбірливості мови, що було отримано модуляційним та формантно-модуляційним методами та їхнє узгодження із оцінками параметру C_{50} (рис. 3.2.3). При цьому слід зауважити, що ефект підвищення розбірливості мови біля задньої стінки аудиторії 438 стосовно середини приміщення, є значно виразнішим для оцінок розбірливості мови (рис. 3.2.5а та 3.2.6а), ніж для коефіцієнту C_{50} (рис. 3.2.3а). Щодо аудиторії 209, де вимірювання виконувалося ще й біля бокових стінок (точки 5 та 6), видно, що саме там (біля бокових стінок аудиторії) розбірливість мови є також вищою, ніж у середині приміщення. У дослідженні [108] було виявлено ефект підвищення розбірливості мови біля стінок приміщення шляхом суб'єктивного оцінювання розбірливості мови.

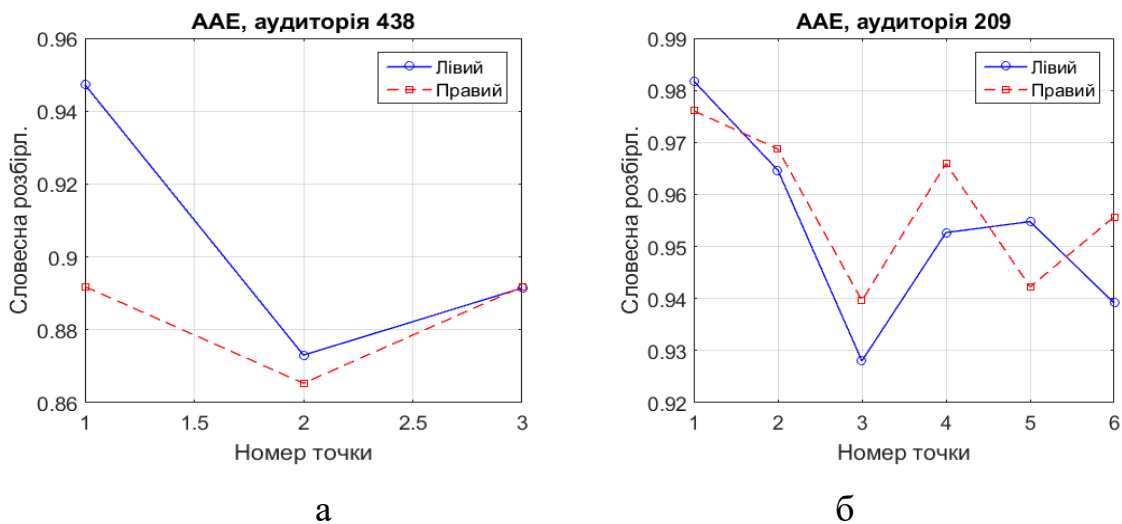


Рис. 3.2.6. Оцінки словесної розбірливості для аудиторій 438 (а) та 209 (б)

Даному ефекту можна дати просте пояснення. До вух людини, що сидить біля стінки, майже одночасно (в наших випадках різниця в часі становила 3–5 мс) надходять прямий та відбитий від стінки звуки. При цьому потужність відбитого від стінки звуку є ненабагато меншою, ніж потужність прямого. Легко зауважити, що на низьких частотах (100–200 Гц) сигнали є практично когерентними, таким чином, при близьких рівнях прямого та відбитого звуку на цих частотах можна очікувати підвищення рівня кінцевого (який є в результаті) звуку до 6 дБ, а на високих частотах, де сигнали спостерігаються практично некогерентними, рівень звуку може підвищитися до 3 дБ. Хоча ці

оцінки є дуже приблизними, вони добре узгоджуються із результатами, що наведені у працях [109, 110]. Щодо середини приміщення, то відбитий від стінки звук більше затримується, ніж прямий звук (в наших випадках 15–20 мс) та послаблюється через розповсюдження в просторі. Зважаючи на те, що у даному випадку про когерентність прямого та відбитого звуків вже не йдеться, то, мабуть, рівень сумарного сигналу має бути меншим, ніж біля стінки. Разом з тим, варто вказати ще й на те, що у зв'язку із помітною затримкою між прямим та відбитим сигналом (відлунням) кімната фактично перетворюється в гребінчастий фільтр (комб-фільтр) [111] із суттєво нерівномірною амплітудно-частотною характеристикою. Таким чином, сумарний звук є ще й спотвореним, а це, в свою чергу, може понизити розбірливість мови в середині кімнати.

Підсумовуючи, можна сказати, що в таких аудиторіях (подібних до розглянутих в даній роботі) користь ранніх відлунь, за розбірливістю мови, є різною, залежно від частин приміщення: біля стін кімнати вона є більшою, ніж в середині.

Висновки до розділу 3

1. Показана принципова можливість створення апаратно-програмного комплексу «Штучна голова» для оцінювання бінауральних імпульсних характеристик приміщення із використанням апаратури непрофесійного рівня.

2. Запропоновано технологію коригування частотної характеристики апаратно-програмного комплексу «Штучна голова», що полягає в еквалізації амплітудно-частотної характеристики підсистеми «гучномовець-мікрофон».

3. Шляхом натурних випробувань комплексу «Штучна голова» перевірено працездатність запропонованої технології коригування частотної характеристики апаратно-програмного комплексу, що містить апаратуру непрофесійного рівня.

4. Результати експериментальних досліджень із застосуванням комплексу «Штучна голова» показали, що розбірливість мови в середині аудиторії може бути суттєво меншою за таку біля стінок, що пояснюється порівняно великим запізненням, відносно прямого звуку, приходу ранніх відбиттів звуку до мікрофонів.

РОЗДІЛ 4. ІНСТРУМЕНТАЛЬНА ОЦІНКА ТА КОРЕКЦІЯ ЯКОСТІ АКУСТИКИ ПРИМІЩЕННЯ

В даному розділі представлені результати інструментальної оцінки якості та корекції якості приміщення. Представлено результати суб'єктивних та об'єктивних оцінок розбірливості мовлення з урахуванням впливу шуму, ранніх та пізніх відбиттів. Розроблено апаратно-програмний комплекс для аналізу напрямків відбиттів звуку у приміщенні, а також продемонстровано працездатність запропонованого методу корекції частотної характеристики приміщення у реальному часі із застосуванням багатоканального компресора.

4.1. Вплив шуму, ранніх та пізніх відбиттів на розбірливість мови

В розділі 1 зазначено, що шум та реверберація негативно впливають на розбірливість мовлення, однак залишається недостатньо дослідженим питання, наскільки сильно ранні відбиття впливають на розбірливість мовлення. Цей недолік було частково усунений в підрозділі 3.2, де експериментальні дослідження виявили цікаве явище, а саме: в центрі приміщення розбірливість виявилася помітно нижчою, ніж біля задньої стінки. Пояснити цей результат можна різним впливом ранніх відбиттів на результуючий сигнал в різних точках приміщення. В даному підрозділі це явище досліджується більш детально, із застосуванням як об'єктивних, так і суб'єктивних методів оцінювання розбірливості мовлення.

У роботі [33] використовувалися складні електроакустичні системи (8-канальна і 29-канальна), в яких слухач розміщувався в центрі безехової кімнати, а гучномовці розташовувалися навколо нього. В обох випадках за допомогою програмного забезпечення ODEON розраховувалися імпульсні реакції приміщення (ІРП). В роботі [35] використана простіша та дешевша експериментальна установка ніж у [33] і [36]. Велика кількість досліджень проводилась у домашніх умовах студентами КПІ ім. Ігоря Сікорського, які

використовували просте обладнання, таке як комп'ютер і навушники, при цьому програмне забезпечення було відносно простим [36].

Запис дикторами односкладних зразків мови типу приголосний-голосний-приголосний проходив у сертифікованій акустичній кімнаті. Мікрофон Superlux ECM 999, для запису чистого мовлення з частотою дискретизації 44,1 кГц і розрядністю 16 біт використовувалися зовнішня звукова карта PRESONUS AudioBox та програмне забезпечення Audacity.

Тестові сигнали синтезувались двома способами: розрахунком згортки бінауральної ІХ $h(t)$ з адитивною сумішшю чистого мовлення та шуму та згортки першої частини бінауральної ІХ $h(t)$ до 50 мс з адитивною сумішшю чистого мовлення та шуму. Мовленнєві сигнали, оброблені спеціальним програмним забезпеченням [40], прослуховувалися через навушники, а склади ідентифікувалися слухачами та фіксувалися за допомогою комп'ютерної клавіатури під управлінням того ж самого програмного забезпечення. Сигнали подавалися слухачам у довільному порядку. Слухачі слухали 3 набори з одного складу; кожен набір містив 50 складів, для 5 різних значень співвідношення сигнал-шум (SNR) та для 6 різних імпульсних характеристик кімнати (RIR), що відповідають 6 відстаням до динаміка. Слухачі виконували кожен етап тестування протягом 10-15 днів, щоб уникнути втоми.

Під час тесту на розбірливість слухачі мали визначити склади та записати їх за допомогою клавіатури. Набір складів є важливою частиною перевірки розбірливості, від якої істотно залежить достовірність результату. В якості прототипу були взяті набори складів стандарту ГОСТ Р 50840-95 [44], на основі яких було розроблено дев'ять українських наборів розбірливості мовлення [40]. Було розроблено три варіанти кожного набору з урахуванням можливості неточного сприйняття та запису складів.

Для досліджень було використано шість бінауральних ІХ з бази даних Aachen Impulse Response (AIR) [45], які належать до аудиторії 10,8 м, 10,9 м, 3,15 м. У кімнаті було 3 скляні вікна, поверхня підлоги - паркет, бетонні стіни,

стілці та дерев'яні столи. Гучномовець був розташований біля місця викладання, а мікрофон був розміщений у різних місцях розміщення студентів на 6 відстанях, що варіювалися від 2,25 м до 10,2 м до динаміка. Час реверберації RT_{60} для ІХ збільшився з 0,70 с до 0,72 с для $d = 2.25 — 4.00$ м і з 0,79 с до 0,83 с для $d = 5.56 — 10.2$ м. Форми хвиль бінауральних ІХ та частотні характеристики на відстанях 2,25 м та 5,56 м наведені на рис. 4.1.1 та 4.1.2 відповідно.

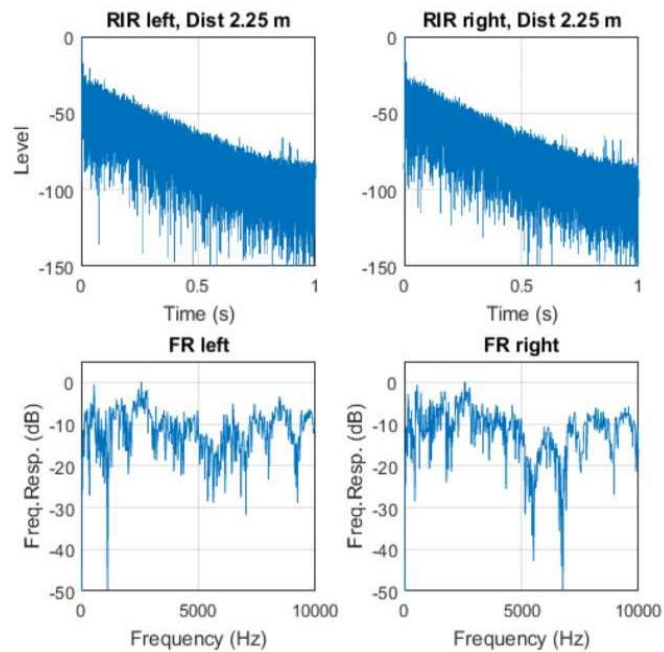


Рис. 4.1.1. Бінауральні ІХ кімнати та АЧХ на відстані 2,25м

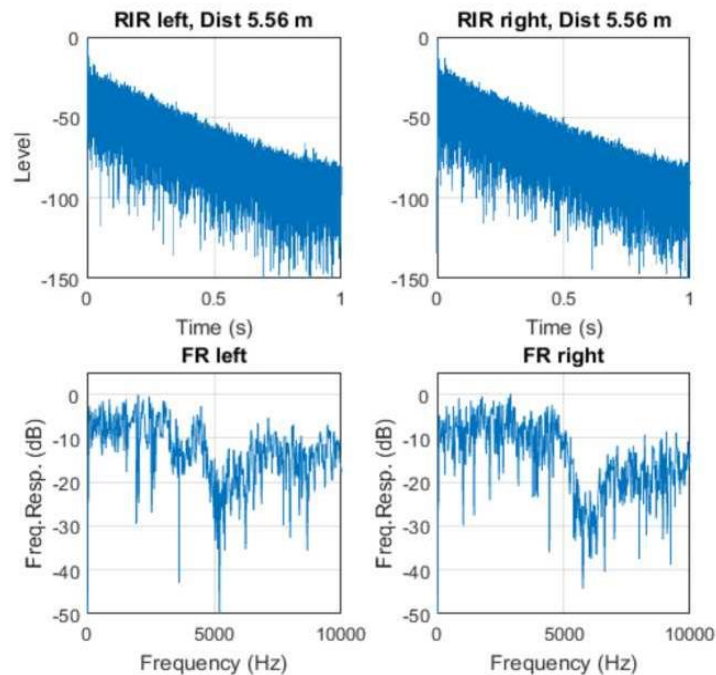


Рис. 4.1.2. Бінауральні ІХ кімнати та АЧХ на відстані 5,56м

Згідно [44] відомо, що розбірливість складів не нижче 80% досягається коли $C_{p0} > -2\text{dB}$. Таким чином, з рис. 4.1.1 видно, що всі ІХ здатні забезпечити хорошу розбірливість мови. Крім того слід зазначити, що значення C_{50} дуже мало змінюються для відстаней 5-10 м, тобто можна зробити припущення, що вплив ранніх відбиттів на розбірливість мови для цих відстаней майже однаковий. Однак результат роботи [45] спростовує ці припущення. З рис 4.1.3 випливає, що відносна (до прямого звуку) інтенсивність ранніх відбиттів монотонно зростає зі збільшенням відстані.

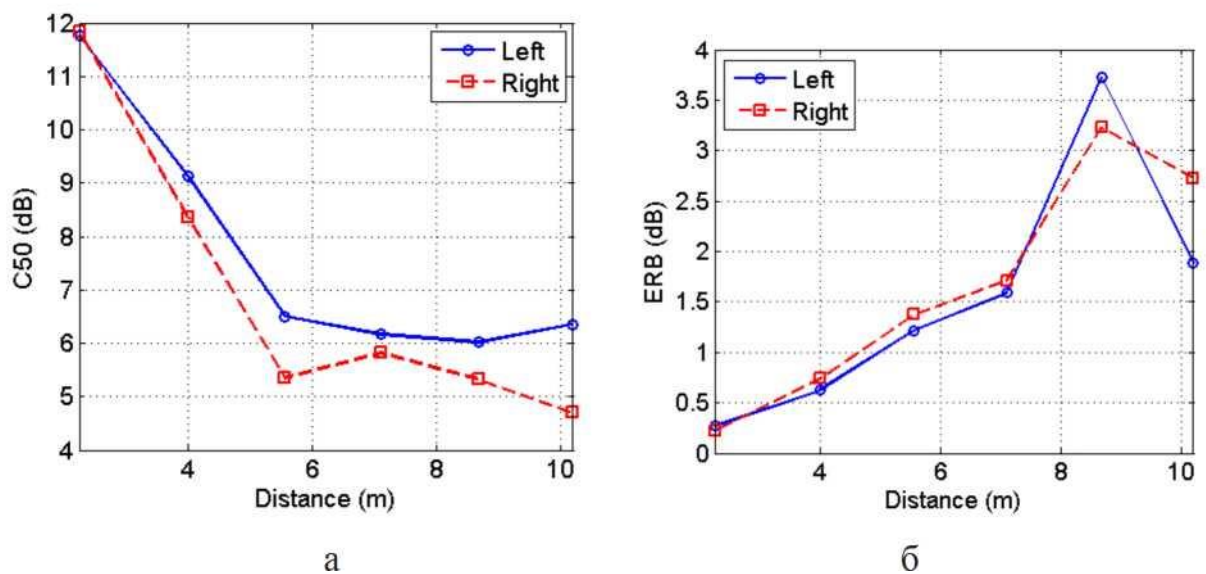


Рис. 4.1.3 - Значення C_{50} та ERB для бінауральних ІХ.

Результати випробувань розбірливості на першому етапі в середньому для 9 слухачів наведені на рис. 4.1.4а, б, а оцінки стандартного відхилення цих результатів наведені на рис. 4.1.4в. Як видно, для всіх розглянутих значень SNR ($\text{SNR} = -15\dots+5\text{дБ}$) розбірливість мови є найвищою для відстані 2,25 м і найнижчою для відстані 5,56 м майже для всіх значень SNR.

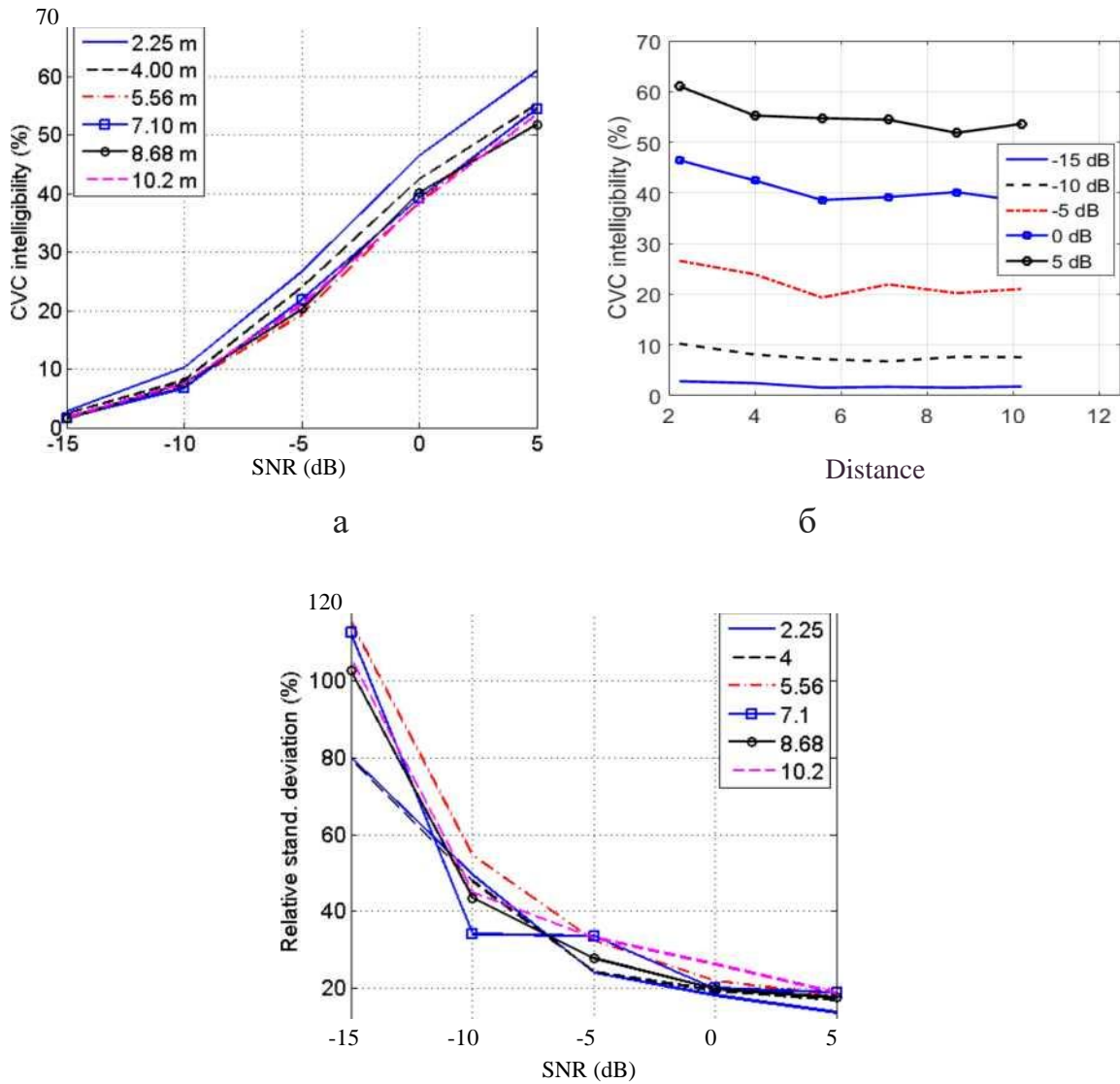


Рис. 4.1.4. Середні оцінки розбірливості мови (а, б) та стандартне відхилення (в) для різних відстаней d для впливу ранніх та пізніх відбиттів

Результати артикуляційних випробувань другого етапу в середньому для 20 слухачів наведені на рис. 4.1.5а, б, а оцінки стандартного відхилення цих результатів наведені на рис. 4.1.5в. Цікаво, що результати не сильно змінилися якісно, але вони помітно змінилися кількісно. Дійсно, розбірливість мови збільшилася на 4-7% у діапазоні значень SNR від -10 дБ до +5 дБ. Водночас розбірливість мови все ще є найвищою для відстані 2,25 м і найнижчою для відстані 5,56 м для всіх значень SNR.

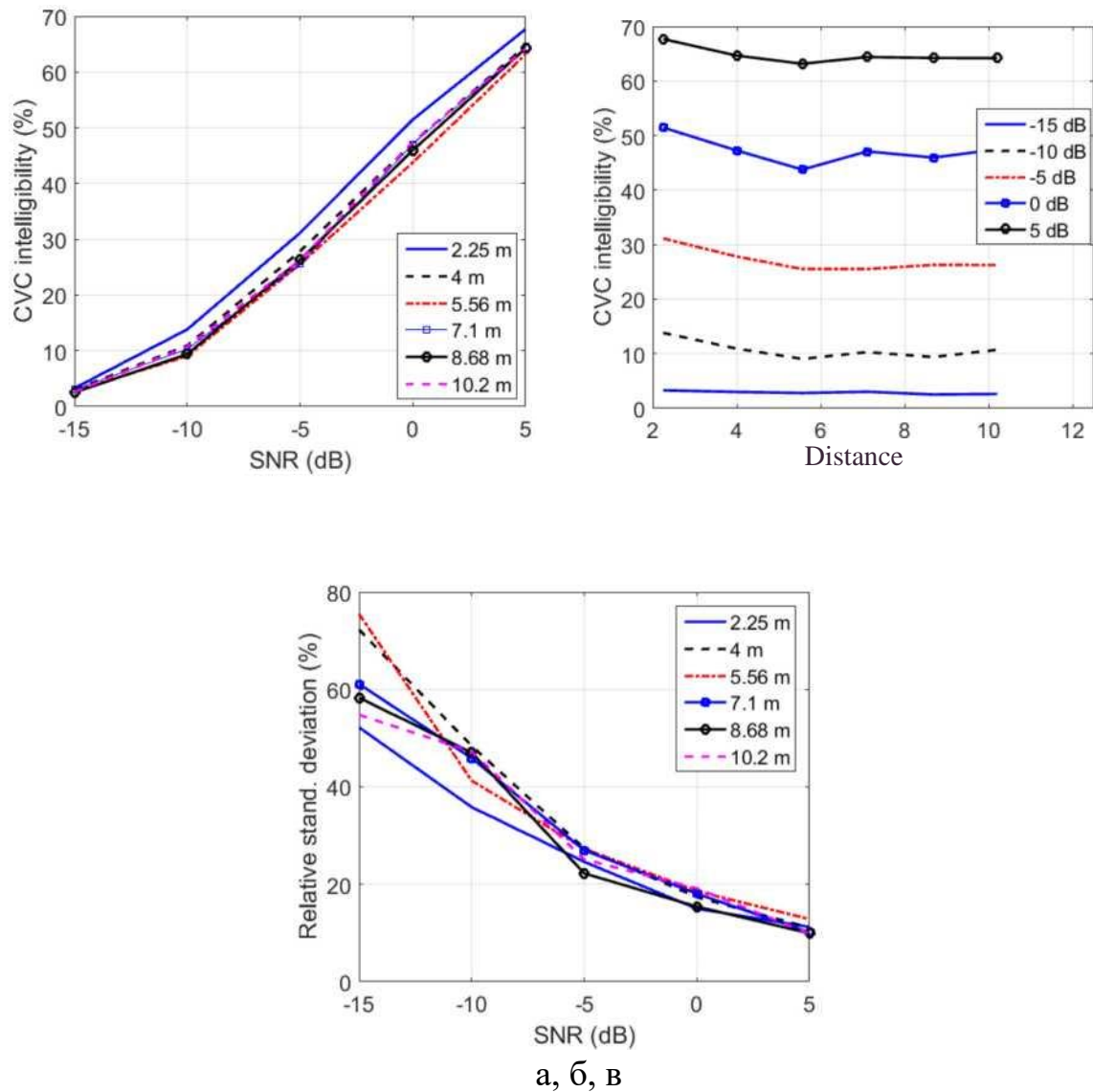


Рис. 4.1.5. Усереднені оцінки розбірливості мови (а, б) та стандартне відхилення (в) для різних відстаней d та для впливу ранніх відбиттів.

Різниця оцінок розбірливості між графіками рис. 4.1.4 та 4.1.5 сягає 8-10 дБ при $\text{SNR}=5$ дБ, тобто при малому рівні фонового шуму. При менших SNR ця різниця не така помітна. Можна зробити висновок, що вплив пізньої реверберації стає найбільш помітним для слухачів при $\text{SNR} \geq 5$ дБ.

Таким чином, слід бути обережним, говорячи про корисність ранніх відбиттів, враховуючи їх руйнівний вплив на спектр мовлення, хоча цей ефект частково компенсується бінауральним сприйняттям.

Результати оцінки розбірливості мови об'єктивними модуляційним та

формантно-модуляційним методами представлені на рис. 4.1.6а, де показані значення розбірливості слів. Подібні результати для об'єктивного модуляційного методу (значення STI індексу передачі мови) показані на рис. 4.1.6б

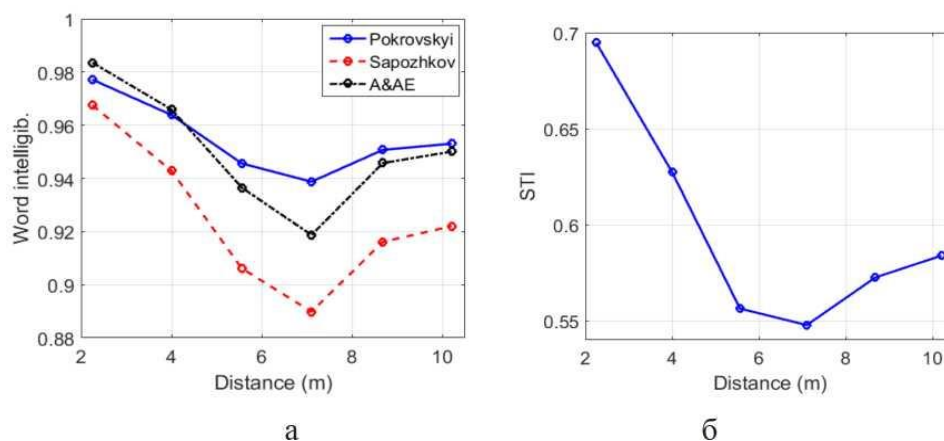


Рис. 4.1.6. Оцінка розбірливості мовлення.

Виявлену різницю в результатах суб'єктивної та об'єктивної оцінки розбірливості мовлення можна пояснити тим, що об'єктивна оцінка розбірливості використовувала спрощений підхід, при якому ефект шумових перешкод не враховувався.

Очевидно, що виявлена відмінність набагато менш важлива порівняно з виявленим фактом, що наявність ранніх відбиттів призводить до зниження розбірливості мови. Як видно, це зниження не може бути компенсоване навіть за допомогою бінаурального прослуховування. Цей результат добре узгоджується з висновками [112] про те, що збільшення інтенсивності ранніх відбиттів не еквівалентно збільшенню інтенсивності прямого звуку. На жаль, в [112] не було зроблено припущень щодо можливої причини цього явища. Виходячи з графіків рис. 4.13, ми можемо спробувати усунути цей недолік і припустити, що причиною цього явища є значна нерівномірність амплітудно-частотної характеристики приміщення, що обумовлено наявністю ранніх відбиттів, що призводить до зниження якості мовленнєвого сигналу. Оскільки це припущення ґрунтується на результатах досліджень лише одного лекційного кабінету, у майбутньому доцільно перевірити його обґрунтованість для кількох кабінетів одного типу.

4.2. Аналіз відбиттів у приміщенні

На сьогоднішній день є досить багато комплексів, що дозволяють оцінити акустичні параметри приміщення різними методами, однак не існує методів та комплексів, що дозволяють локалізувати напрямки, звідки звук приходить до слухача чи мікрофону. На рис. 4.2.1 представлено структурну схему способу оцінки акустичних характеристик приміщення з урахуванням локалізації.



Рис. 4.2.1. Схема способу акустичної оцінки приміщення

Шум, що використовується для тестування, випромінюється акустичною системою — джерелом шуму в приміщенні та приймається групою мікрофонів підключених до АЦП (пристрою реєстрації), який передає форматовану інформацію на ПК з програмним забезпеченням, що реалізовує алгоритми вимірювання рівня акустичних шумів та кутового положення вторинних відбивачів, а також розрахунку часу реверберації,

Для запропонованого способу оцінки акустичних характеристик приміщення та перевірки пристрою для оперативної оцінки акустичних характеристик (ПООАХ) було проведено декілька експериментів у приміщенні (розміри 10х6х4 м, декілька вікон, що виходили на тихе подвір'я та двері, розміщені на протилежній від вікон стіні, що відчинялись у шумний коридор) при температурі повітря 18⁰ С. На рис. 4.2.2 представлено розташування ПООАХ (M1-M4 — мікрофони NTI MiniSPL; звукова карта — M-Audio Fast Track Ultra; ПК — персональний комп'ютер ASUS). Запис звуку з аудіоінтерфейсом в робочу область програмного пакета Matlab забезпечувало програмне забезпечення Cocks Reaper.

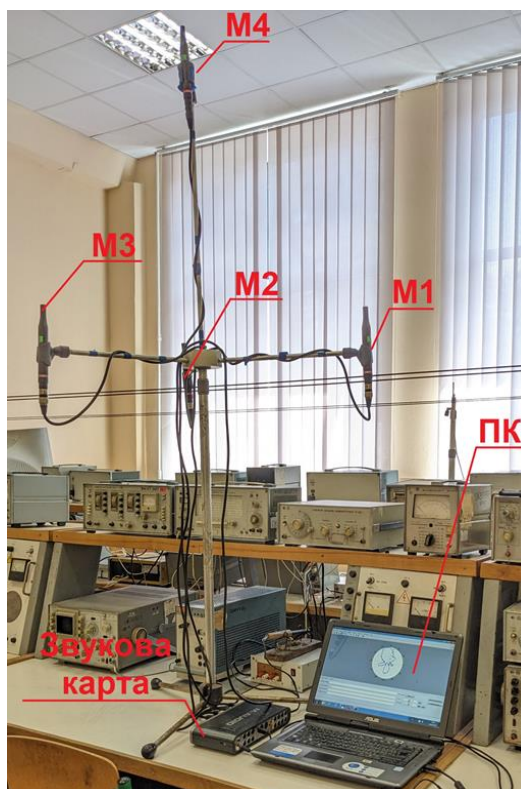


Рис. 4.2.2. Пристрій оперативної оцінки акустичних характеристик

Експериментальні дослідження проходили в три етапи:

- перший етап: досліджувався рівень фоновому шуму та кутовий розподіл у приміщенні за умови закритого приміщення, відчиненого вікна або відчинених дверей;
- другий етап: вимірювався час реверберації за кривими спаду методом перервного шуму в октавній смузі частот білого шуму та встановлювалися характерні часові інтервали формування поля (завершення випромінювання; характерні сегменти кривої спаду, які визначають вплив відбивачів звуку);
- третій етап: за характерними сегментами кривої спаду визначались кутові координати реальних і уявних джерел звуку та досліджувався кутовий розподіл звукового поля в приміщенні.

Перш ніж почати експеримент разом із підсилювачами звукової карти було проведено калібрування мікрофонів, що дало можливість у майбутньому порівнювати інтегральні рівні акустичного поля, отримані ПООАХ та штатним шумоміром Robotron RFT00017.

Активна акустична система (АС) зі смугою частот випромінювання 200—8000 Гц була джерелом шуму. На ПК у програмі Matlab генерувався білий шум різної тривалості.

Фоновий шум у приміщенні досліджувався методом керованої просторової кореляційної характеристики (КПКХ) у діапазоні частот 125—4000 Гц. У декартовій (рис. 4.2.3) та полярній (рис. 4.2.4) системах координат як кутові діаграми (КД) представлено результати першого етапу досліджень, що, у свою чергу, дало змогу оцінити як кутовий спектр значень рівня шумового фону, так і візуалізувати просторове розміщення джерел шуму.

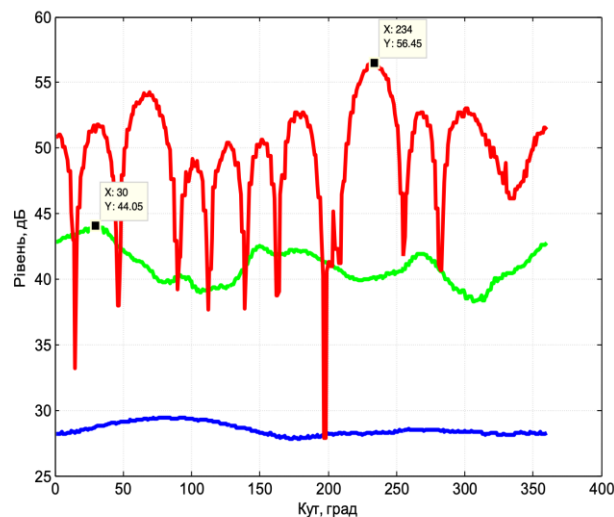


Рис. 4.2.3. Кутовий спектр рівня шумового фону в приміщенні

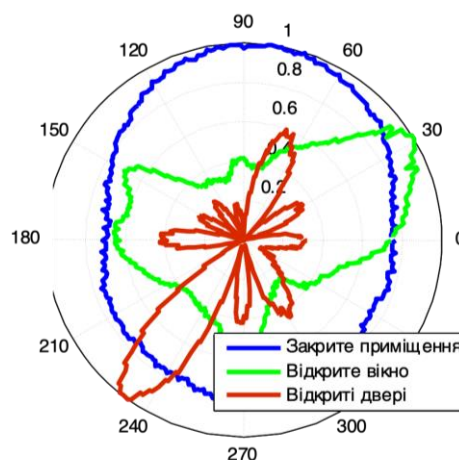


Рис. 4.2.4. Кутова діаграма нормованого шумового фону

Рівномірним поле для закритого приміщення (рис. 4.2.3) є за умови середнього значення рівня кутового спектра $L = 28$ та нерівномірності $\Delta L = 2$. На підвищений рівень шуму зі сторони вікон (сектор кутів 30° — 120°) вказує КД (рис. 4.2.4). Зауважимо, коли вікно відчинене, то максимум КД (рис. 4.2.4) вказує на напрямок джерела випромінювання (вікна) у секторі 30° з кутовим рівнем $L = 44$ (рис. 5), якщо двері відчинені у шумний коридор, то КД (рис. 5; б) має максимум $L = 56$ з кутом приблизно 234° .

За методом переривання шуму [67], будувалася крива спаду та оцінювався час реверберації (наприклад, на рис. 4.2.5 представлено результати, що було отримано для октавної смуги з центральною частотою 1000 Гц).

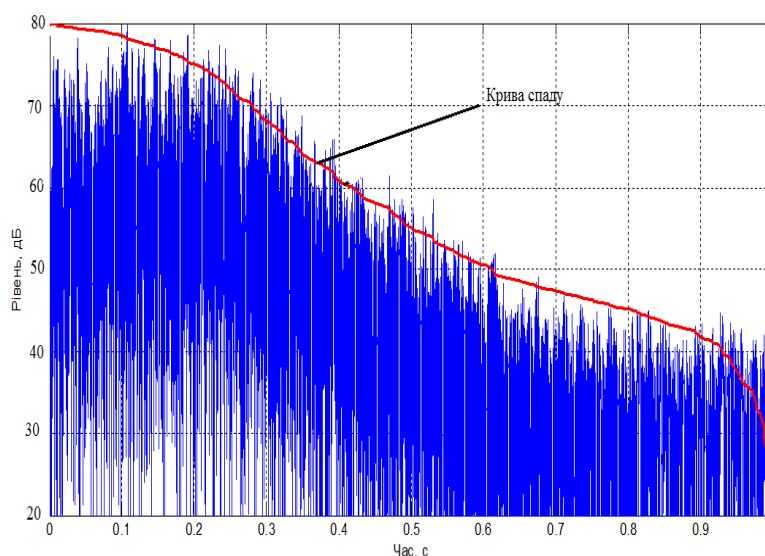


Рис. 4.2.5. Крива спаду ревербераційного процесу

Таким чином, оцінка стандартного часу реверберації, за рівнем 55 дБ (-25 дБ від максимуму), становить $T_{ст} = 1.2с$, час реверберації — $T_{20} = 0.3с$.

Застосовуючи метод КПКХ, проводились дослідження кутового розподілу звукового поля в приміщенні та визначення кутових координат відбивачів звуку, для чого записи звукових файлів трьох мікрофонів сегментувались на відрізки по 100 мс, а, починаючи з моменту вимкнення випромінювача (рис. 4.2.5), фільтрувались октавним фільтром та відповідно

оброблялись. Зазначимо, що результати розрахунків представлено у декартовій (рис. 4.2.6) та полярній (рис. 4.2.7) системах координат.

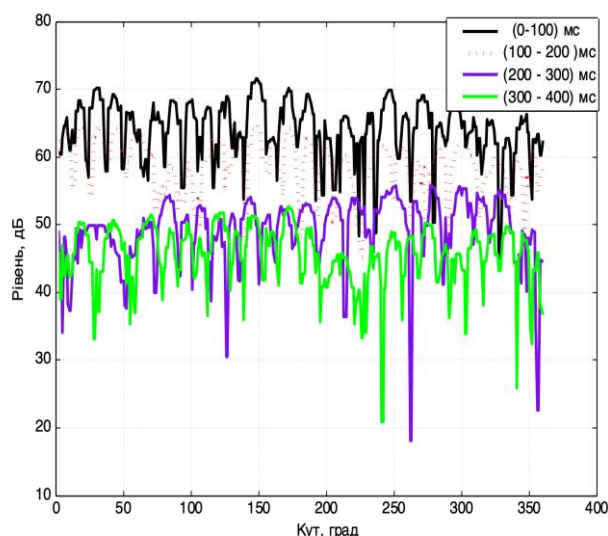


Рис. 4.2.6. Кутові діаграми ревербераційного процесу в окремі проміжки часу

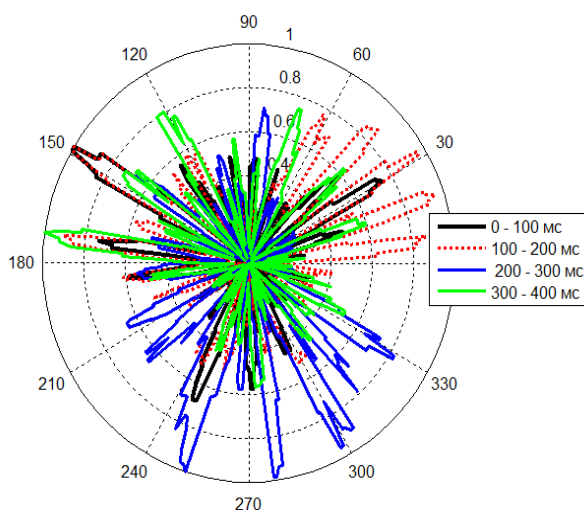


Рис. 4.2.7. Кутове положення відбивачів в окремі проміжки часу

На рис. 4.2.6 подано аналіз кутових діаграм, який дає можливість оцінити кількісний вклад окремих відбивачів в ревербераційний процес на різних часових інтервалах, а їх кутове положення демонструють графіки (нормовані КД) на рис. 4.2.7.

Отже, для оцінки проникнення звуку через огорожувальні конструкції та аналізу просторового розподілу ревербераційного поля в приміщенні застосування методу КПКХ має свої особливості. Зокрема, як було показано вище (рис. 4.2.4), параметри синтезованої КПКХ (ефективний кут розкриття,

бічні пелюстки) залежить від смуги октавних частот, що було вибрано для досліджень (на високих частотах кутові координати відбивачів та каналів проникнення звуку можна виявити більш точно, ніж на низьких). Таким чином, у майбутньому доцільно дослідити кутовий спектр ревербераційного процесу на різних частотах та встановити його відповідність із кутовим розміщенням відбивачів.

4.3. Корекція акустики приміщення шляхом адаптації АЧХ акустичної системи під зміну наповненості приміщення

Зазвичай під час підготовки до публічних заходів, звукоінженери проводять налаштування акустичних систем під особливості приміщення, однак коли захід проходить, то налаштування не змінюють, хоча акустика приміщення змінюється. У першому розділі даної роботи описані методи, які дають можливість коригувати частотну характеристику приміщення. Однак дані методи не дають можливості змінювати параметри системи корекції в реальному часі. В даному розділі представлено результати досліджень можливості усунення даного недоліку.

Структурну схему корекції частотної характеристики приміщення із застосуванням багатоканального компресору зображено на рис. 4.3.1. Згідно цієї схеми, при налаштуванні системи тестовий сигнал у вигляді рожевого шуму подається на вхід 1 багатоканального компресору. Сигнал випромінюється гучномовцем, проходить через приміщення та сприймається парою мікрофонів штучної голови (ШГ), розміщеними у приміщенні. Електричний сигнал з виходу мікрофонів подається на вхід 2 (sidechain) багатоканального компресору. При перевищенні певного порогу рівнем сигналу в певному частотному каналі, сигнал має послабитися компресором, як це зображено на рис. 4.3.2.



Рис. 4.3.1. Структурна схема системи корекції частотної характеристики приміщення із застосуванням багатоканального компресору

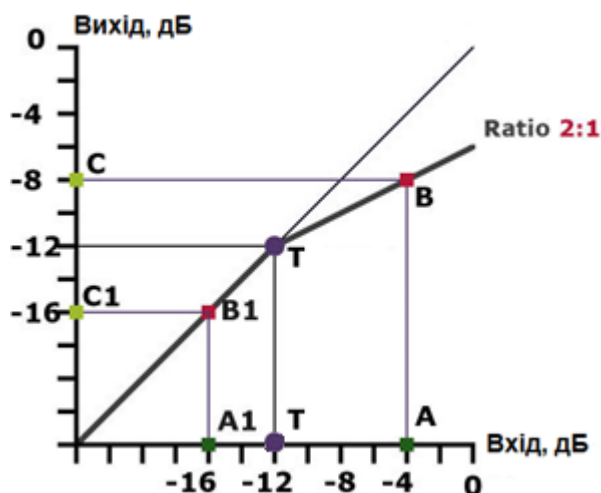


Рис. 4.3.2 Принцип роботи компресору]

На рис. 4.3.2 по осі X відзначений рівень сигналу, що надходить на компресор. По осі Y відзначено рівень вихідного сигналу. Якщо рівень вхідного сигналу перевищить поріг T , зміниться коефіцієнт підсилення сигналу.

Програмним аналогом багатоканального компресору, представленого в патенті [83], є VST-плагін FabFilter Pro-MB [113], який дозволяє виконувати обробку в реальному часі. Він має лінійну фазочастотну характеристику, дозволяє змінювати рівень сигналу у двох режимах, - стиснення (компресія) та розширення (експандія), - а також має режим роботи *sidechain*. В даній роботі обмежуємось лише режимом стиснення як таким, що частіше використовується у реальних задачах. До плюсів даного плагіну також можна віднести його зручність та легкість у налаштуванні.

Алгоритм обробки сигналу за допомогою багатоканального компресору складається із наступних етапів:

1. Сигнал $X(t)$ пропускається через гребінку смугових фільтрів, в результаті отримаємо сигнали $X_1(t) \dots X_N(t)$ на виходах фільтрів.
2. Сигнали $X_1(t) \dots X_N(t)$ подаються на входи компресорів $C_1 \dots C_N$.

3. Керуючими сигналами для компресорів $C_1 \dots C_N$ є сигнали $Y_1 \dots Y_N$, які формуються шляхом пропускання сигналу з виходу мікрофону через іншу гребінку фільтрів.
4. Компресія сигналу $X(n)$ в n -тому каналі компресора відбувається за умови, що рівень сигналу $Y(n)$ перевищує порогове значення T .
5. Параметр Ratio рекомендується обирати від 1.5:1 до 4:1.
6. Сигнали з виходу компресорів підсумовуються.

Результати досліджень

Експериментальні дослідження багатоканального компресору (БК) виконувалися у два кроки. Першим кроком було моделювання спотворюючої дії приміщення на сигнал $X(t)$ та перевірка алгоритму роботи багатоканального компресору. Сигнал $X(t)$ штучно спотворювався шляхом підсилення верхніх частот, в результаті чого формувався сигнал $Y(t)$. Після цього виконувалася корекція сигналу $Y(t)$ з метою відновлення сигналу $X(t)$.

Другим кроком експерименту були натурні випробування у приміщенні. Сигнал $X(t)$ спотворювався приміщенням, після чого виконувалася корекція сигналу $Y(t)$ з урахуванням акустичних властивостей приміщення, з метою відновлення сигналу $X(t)$. Натурні експерименти проводилися у побутовому приміщенні, розміри кімнати 12x5x1.5-3.5м, також було використано наступний перелік обладнання:

1. Цифровий мікшерний пульт Behringer XR18 (МП1);
2. ШГ з двома мікрофонними капсулями типу Sennheiser e614;
3. Персональний комп'ютер з програмним забезпеченням: X-Air-Edit для керування мікшерним пультом, Cockos Reaper, як VST-хост, пакет VST-плагінів FabFilter для обробки сигналу;
4. Аналоговий мікшерний пульт Behringer Xenyx 502 (МП2);
5. Мобільний телефон з програмним генератором рожевого шуму TMSOFT Noise Generator;

6. Джерело звуку у вигляді двохсмугової акустичної системи з бі-амповим включенням Sven Monitor 5;

При модельному тестуванні алгоритму БК тестовий сигнал $X(t)$ у вигляді рожевого шуму подавався на вхід МП2, де піддавався навмисному спотворенню шляхом підсилення високочастотних компонентів (рис. 4.3.3). Далі спотворений сигнал подавався на стерео вхід 1 мікшерного пульта МП1, а на стерео вхід 2 подавався з мікрофонів ШГ. До виходу МП1 було підключено активний гучномовець. Для мінімізації ревербераційного впливу приміщення, мікрофон розміщувався на відстані 5 см від гучномовця.

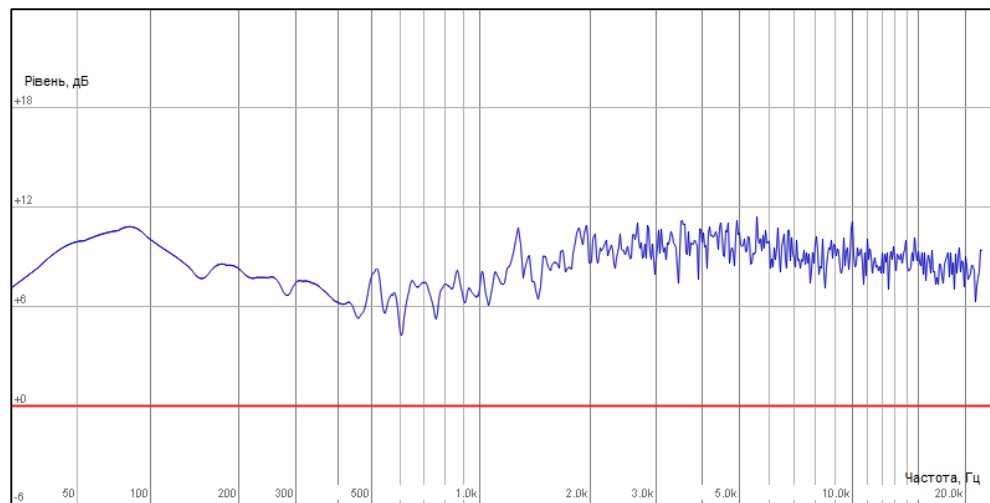


Рис. 4.3.3. Штучно спотворений сигнал

Сигнал $Y(t)$ з виходів мікрофонів був керуючим для БК. Для експериментальних досліджень використовувалось п'ять каналів БК. Обробка сигналів виконувалася згідно вищенаведеному алгоритму. При цьому було обрано наступні значення параметрів БК.

1. Частотні смуги: 20 – 110Гц, 110 – 250; 250 – 600; 600 – 2000; 2000 – 7000, 7000 – 20000.
2. Пороговий параметр T , для кожної смуги, дорівнював -18дБ.
3. Параметр Ratio, для кожної смуги, дорівнював 4:1.
4. Параметр Range, для кожної смуги дорівнював -15дБ

5. Кожна смуга була переведена у режим Sidechain з керуванням від зовнішнього мікрофону.

Інші параметри залишились без змін.

При натурному експерименті тестовий сигнал у вигляді рожевого шуму подавався на стерео вхід 1 мікшерного пульта МП1, до виходу якого була підключена пара активних гучномовців системи Sven Monitor 5. Сигнал $X(t)$ відтворювався гучномовцями, проходив через приміщення та приймався парою мікрофонів ШГ, що були підключені так, як і в попередньому експерименті. Спектр прийнятого сигналу $Y(t)$ зображено на рис. 4.3.4. Обладнання розміщувалось у приміщенні згідно схемі зображених на рис. 4.3.5.

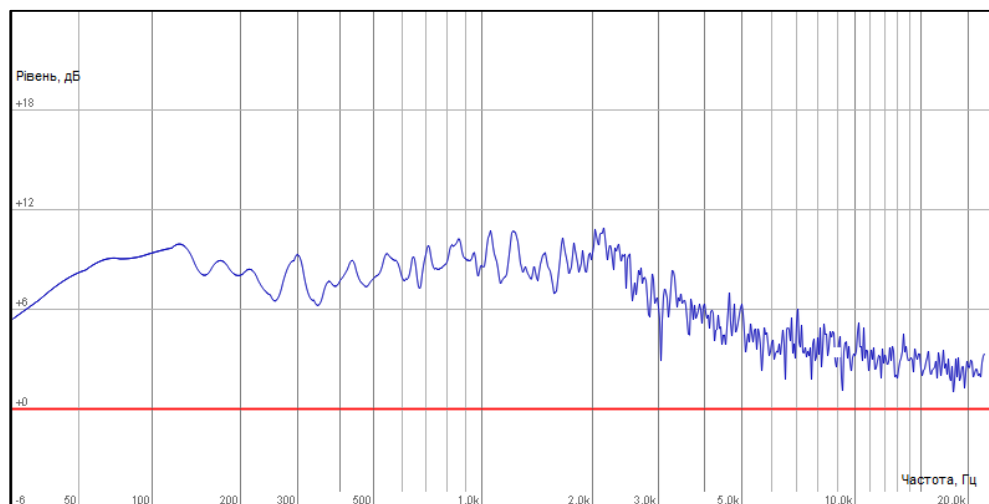


Рис. 4.3.4. Сигнал спотворений приміщенням

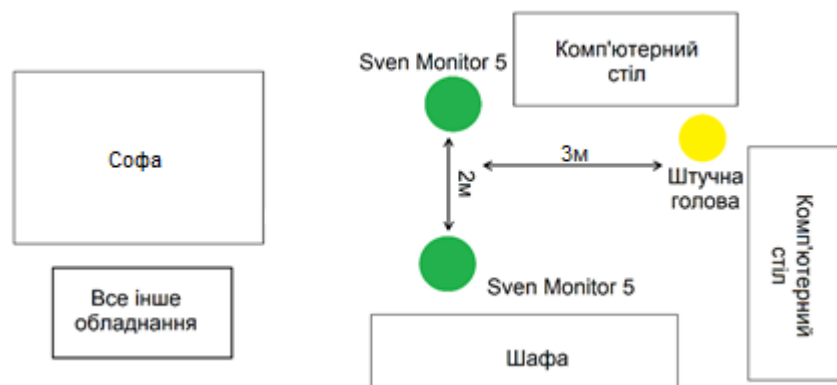
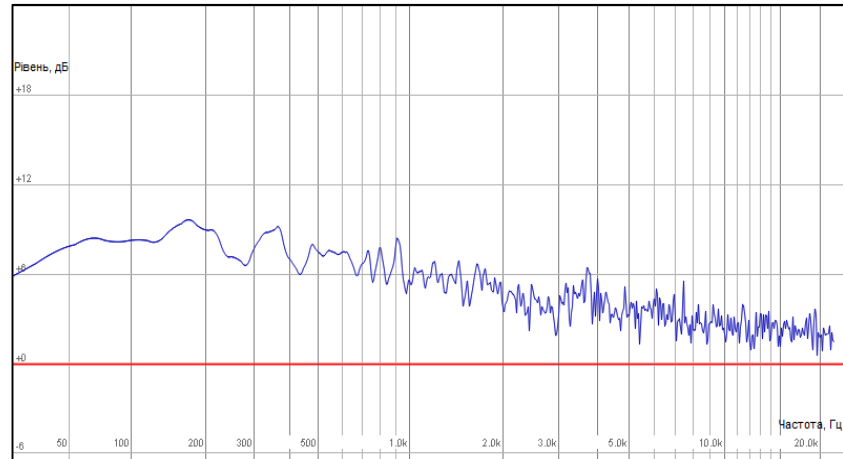
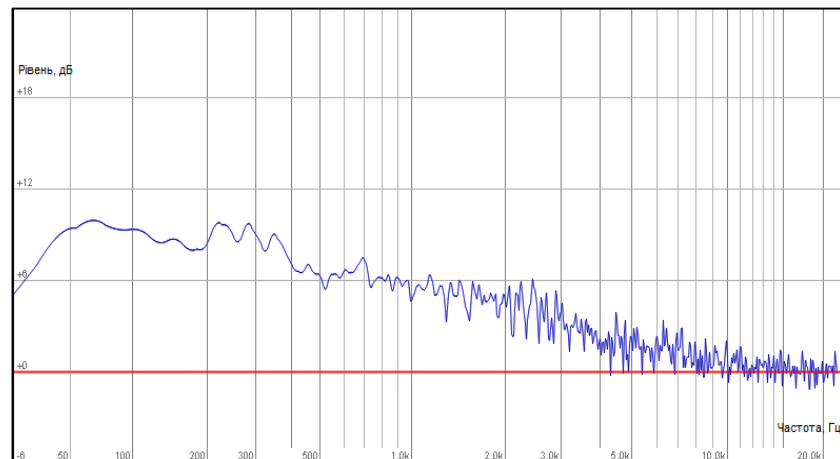


Рис. 4.3.5. Розміщення обладнання у кімнати

Після проведення обробки з вище зазначеними параметрами БК було отримано наступні результати. На рис. 4.3.5 представлений спектр скоригованого сигналу, що спотворювався штучно (а) та приміщенням (б).



(а)



(б)

Рис. 4.3.6. Скориговані спектри рожевого шуму: модельні (а) та натурні (б) дослідження

Аналіз рис. 4.3.6 свідчить, що обрані параметри БК змогли забезпечити відновлення сигналу з похибкою до 3 дБ, яку можна вважати допустимою. Слід зазначити, що для більш точного коригування необхідно використовувати більшу кількість смуг частот та оптимізувати параметри кожної смуги, що і планується реалізувати у наступних роботах.

В ході експериментів виявився суттєвий недолік. Оскільки обробка сигналу відбувається за допомогою комп'ютера, то є місце часу затримки, що обумовлена сумою затримок АЦП + обробка + ЦАП.

У експерименті сумарна затримка дорівнювала 20 мс при частоті дискретизації 48000 Гц. Час затримки обробки становив близько 4 мс, час затримки АЦП+ЦАП дорівнював 16 мс. Зменшення частоти дискретизації призводить до зменшення затримки, однак суб'єктивно відтворений сигнал сприймається як погіршений. Зменшення буферної пам'яті АЦП дозволяє суттєво зменшити затримку, однак у сигнал при цьому вносяться призвуки, схожі на «клацання», що також не припустимо.

Для зменшення затримки обробки існують Live-процесори, по типу Waves Multi Rack SGS1 Combo [114], які дають можливість використовувати VST-хост з нульовою затримкою. Для мінімізації апаратних затримок доцільно використовувати протокол обміну даних Dante [115] у парі з мікшерними консолями Yamaha TF, Yamaha CL, Behringer x32 with Dante Card. Поєднання Dante та Waves дає сумарну затримку біля 0.1 мс при частоті дискретизації у 44100 Гц. Проте недоліком даного поєднання є досить висока вартість.

Висновки до розділу 4

1. Результати суб'єктивної оцінки розбірливості мови залежно від відстані до лектора свідчать, що розбірливість мови є мінімальною в середині аудиторії середнього розміру в широкому діапазоні відношень сигнал/шум. Це пояснюється тим, що до слухачів в середні приміщення прямий звук та перші ранні відбиття приходять із суттєвим запізненням, тоді як до слухачів біля стін прямий звук та перші ранні відбиття приходять майже одночасно.

2. Запропоновано та розглянуто пристрій для дослідження акустичних характеристик приміщень, в основу функціонування якого покладено метод синтезу керованої просторової кореляційної характеристики, що дає можливість оцінювати кутовий розподіл рівня звуку, виявляти огорожувальні конструкції з низьким рівнем звукоізоляції, локалізувати відбивачі звуку та оцінювати їх вклад у ревербераційний процес.

3. Показано можливість локалізації окремих конструктивних елементів, наприклад, вікон або дверей, як каналів проникнення звуку через апробацію макета в лабораторних умовах. Крім того, встановлено кореляцію у часі між процесом реверберації та окремими відбивачами, а також встановлено можливість визначати кутове положення відбивачів.

4. Експериментально перевірено працездатність способу корекції сигналу, спотвореного реверберацією, за допомогою багатоканального компресору. При цьому показано, що використання VST-плагінів для задач акустичної корекції приміщення є дешевим та простим способом вирішення поставленої задачі. Використання VST-хосту з багатоканальним VST компресором з нульовою затримкою призводить до затримки (не менше 4 мс) обробленого сигналу. Використання протоколів Dante та Waves зменшує затримку до мінімуму та дає можливість працювати в реальному масштабі часу.

ВИСНОВКИ

1. Виконано аналіз сучасного стану експертизи та корекції акустичних характеристик приміщень різного призначення, який показав, що на сьогодні залишаються не до кінця розв'язаними або недостатньо обґрунтованими наступні актуальні задачі:
 - вибір найбільш точного та швидкого методу оцінювання розбірливості мовлення (індексу передачі мовлення), що маскується шумом та реверберацією;
 - вибір тривалості тестових сигналів при оцінюванні розбірливості мовлення різними методами;
 - визначення принципової можливості оцінки бінауральних імпульсних характеристик приміщення за допомогою системи "Штучна голова", що містить акустичну апаратуру непрофесійного рівня.
 - локалізація напрямів ранніх відбиттів в приміщенні, що дозволяє уточнити час приходу ранніх відбиттів до слухача;
 - частотна корекція амплітудно-частотної характеристики приміщення в реальному часі, з урахуванням зміни наповненості приміщення.
2. Показано практичне співпадіння, за точністю оцінювання індексу розбірливості мовлення STI в умовах дії шуму та реверберації, повного традиційного методу та повного формантно-модуляційного методу для діапазону відношень сигнал-шум від мінус 28 дБ до плюс 28 дБ та для діапазону тривалості тестових сигналів від 4 с до 64 с.
3. Порівняння швидкого та повного варіантів формантно-модуляційного методу оцінювання індексу розбірливості мовлення в умовах дії шуму та реверберації, в діапазоні значень відношення сигнал-шум від мінус 20 дБ до плюс 20 дБ, показало, що абсолютна похибка оцінювання STI швидким формантно-модуляційним методом не перевищує припустиму

величину 0.03 за умови, що тривалість тестових сигналів є не меншою за 16 секунд.

4. Підтверджено результати попередніх висновків про те, що формантний метод є суттєво точнішим та швидшим, порівняно із модуляційним методом, за умов, що дія шумової завади переважає над дією реверберації. Так, тривалість тестового сигналу формантного методу, що забезпечує прийнятну точність оцінювань індексу розбірливості мовлення STI, може бути близькою до 1 с, тоді як при використанні швидкого формантно-модуляційного методу тривалість тестового сигналу має бути не меншою за 16 с.
5. Запропоновано спеціальний чотирьохкомпонентний тестовий сигнал та відповідні алгоритми його обробки для порівняння двох варіантів непрямого методу оцінювання індексу розбірливості мовлення STI та прямого швидкого формантно-модуляційного методу за точністю вимірювань. Показано, що застосування тестового сигналу у вигляді MLS-сигналу із спектром мовлення дозволяє отримати більшу точність оцінювання індексу розбірливості мовлення STI, ніж у випадку тестового сигналу у вигляді послідовності MLS-сигналу із плоским спектром та шумового сигналу із спектром мовлення.
6. Показано принципову можливість та розроблено технологію використання апаратно-програмного комплексу «Штучна голова», що містить електроакустичну апаратуру непрофесійного рівня, для оцінювання бінауральних імпульсних характеристик приміщення, за умови еквалізації амплітудно-частотної характеристики підсистеми «гучномовець-мікрофон».
7. Виконано дослідження алгоритмів програмно-апаратного комплексу для визначення напрямку прибуття ранніх відбиттів звуку у приміщенні. Запропоновано та перевірено працездатність алгоритмів обробки сигналів у пристрої для дослідження акустичних характеристик

приміщень, в основу функціонування якого покладено метод синтезу керованої просторової кореляційної характеристики.

8. Показано можливість локалізації окремих конструктивних елементів, наприклад, вікон або дверей, як каналів проникнення звуку через апробацію макета в лабораторних умовах. Крім того, встановлено кореляцію у часі між процесом реверберації та окремими відбивачами, а також встановлено можливість визначати кутове положення відбивачів.
9. Виконано дослідження алгоритмів програмно-апаратного комплексу для корекції амплітудно-частотної характеристики приміщення в реальному часі, при цьому експериментально перевірено працездатність способу корекції сигналу, спотвореного реверберацією, за допомогою багатоканального компресору.

Список літератури

1. Dvornyk, Oleksandr Oleksandrovych. 2023. "Real-Time Correction of the Frequency Response of a Public Address System". *Microsystems, Electronics and Acoustics* 28 (2). Kyiv, Ukraine: 280467.1-67. <https://doi.org/10.20535/2523-4455.mea.280467>.
2. Didkovska, Maryna & Dvornyk, Oleksandr & Motorniuk, Daria & Prodeus, Arkadiy. (2020). Artificial Software Complex "Artificial Head". Part 1. Adjusting the Frequency Response of the Path. *Microsystems Electronics and Acoustics*. 22. 56-64.
3. Prodeus, Arkadiy & Dvornyk, Oleksandr & Didkovska, Maryna & Motorniuk, Daria. (2020). Hardware and Software System "Artificial Head". Part 2. Evaluation of Speech Intelligibility in Classrooms. *Microsystems Electronics and Acoustics*. 22. 48-55.
4. Kozeruk, Sergei, & Dvornyk Oleksandr. 2023. "Device for Operational Assessment of Acoustic Characteristics of Room". *Visnyk NTUU KPI Serii A - Radiotekhnika Radioaparotobuduvannia*, no. 94 (December):24-31. <https://doi.org/10.20535/RADAP.2023.94.24-31>.
5. A. Prodeus, O. Dvornyk, A. Naida, M. Didkovska, O. Grebin. (2023). Multicomponent Signal for Comparing Direct and Indirect Methods of Speech Transmission Index Measurement. *Electronics and Control Systems*. 1. 27-35. 10.18372/1990-5548.75.17546.
6. А. Продеус, А. Вітик, О. Дворник, І. Котвицький, О. Чайка, М. Ярошенко, Суб'єктивне оцінювання розбірливості мови на тлі шуму та реверберації. "Мікросистеми, електроніка та акустика", том 23, № 2, 2018, с. 66-73, ISSN 2523-4447, DOI: 10.20535/2523-4455.2018.23.2.128820
7. A. Prodeus, K. Bukhta, P. Morozko, O. Serhiienko, I. Kotvytskyi, O. Dvornyk, "Automated Subjective Assessment of Speech Intelligibility in Various Listening Modes," ISSN 2523-4447. *Microsystems, Electronics and*

- Acoustics, 2018, vol. 23, no. 3, pp.49-57. DOI: <https://doi.org/10.20535/2523-4455.2018.23.3.130367>
8. Продеус, А. М., Дворник, О. О., & Найда, А. С. (2024). Точність оцінки індексу передачі мовлення формантно-модуляційним методом. *Вісник вищих навчальних закладів. Радіоелектроніка* (прийнято до друку) <https://doi.org/10.20535/S002134702306002X>
 9. А. М. Продеус, І. В. Котвицький, М. В. Дідковська, В. С. Дідковський, К. А. Кухарічева, Д. Є. Моторнюк, О. О. Дворник Спосіб виявлення кліпування мовного та музичного сигналів // Патент UA 144291 U, МПК G01R 23/20, опубл. 25.09.2020. Prodeus, Arkadiy & Dvornyk, Oleksandr & Naida, Anton & Didkovska, Maryna.
 10. ISO/TR 4870:1991 Acoustics — The construction and calibration of speech intelligibility tests.
 11. Покровский, Н.Б. Расчет и измерение разборчивости речи [Текст] / Н.Б. Покровский – М.:Связь издат, 1962. – 392 с.
 12. ГОСТ 25902-83 «Зали для глядачів. Метод визначення розбірливості мови».
 13. ГОСТ 51061-97 «Системи низькошвидкісної передачі мови по цифрових каналах. Параметри якості мови і методи вимірювань».
 14. IEC 60268-16:2020 Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index.
 15. Дидковский, В.С. Акустическая экспертиза каналов речевой коммуникации. [Текст]: Монография / В.С. Дидковский, М.В. Дидковская, А.Н. Продеус – К: Имэкс-ЛТД, 2008. – 420 с.
 16. Гавриленко, А.В. Парциальные отношения сигнал-шум в роли критерия, альтернативного разборчивости речи. [Текст] / А.В. Гавриленко, В.С. Дидковский, А.Н. Продеус // Электроника и связь, 2006. – №4. – С.48-53.
 17. Продеус А.Н. О некоторых особенностях развития объективных методов измерений разборчивости речи // Электроника и связь,

- тематический выпуск «Электроника и нано-технологии». – 2010. – №2. – С. 217–223.
- 18.Houtgast T. and Steeneken H. J. M. Evaluation of Speech Transmission Channels by using Artificial Signals. *Acustica*. - 1971. -Vol. 25. - P. 355-367.
 - 19.Steeneken, H.J.M., and Houtgast, T. "A physical method for measuring speech-transmission quality" *J.Acoust. Soc. Am.* 67, 1980. - P. 318-326.
 - 20.Collard J.A. Theoretical Study of the Articulation and Intelligibility of a Telephone Circuit // *Electrical Communication*. – 1929. -- Vol.7. – P. 168.
 - 21.Prodeus A. Formant-Modulation Method of Speech Intelligibility Evaluation: Measuring and Exactness. - *Proceedings of the VII International Conference MEMSTECH 2011*. - Lviv, Polyana, 2011. - Pp.54-60.
 - 22.British Standard BS EN 60268-16. Sound system equipment. Part 16. Objective rating of speech intelligibility by speech transmission index. 2011.
 - 23.A. Prodeus, "Rapid version of a formant-modulation method of speech intelligibility estimation," *Perspective Technologies and Methods in MEMS Design*, Polyana, Ukraine, 2011, pp. 61-63. Available at: <https://ieeexplore.ieee.org/document/5960269>.
 - 24.NTi Audio, Application note. Speech Intelligibility. Measurement with the XL2 analyzer. Dec. 2020. 28 p. Available at: <https://www.nti-audio.com/en>.
 - 25.A. Prodeus, "On possibility of advantages join of formant and modulation methods of speech intelligibility evaluation," 2010 *Proceedings of Vith International Conference on Perspective Technologies and Methods in MEMS Design*, Lviv, Ukraine, 2010, pp. 254-259. Available at: <https://ieeexplore.ieee.org/abstract/document/5499260>.
 - 26.Продеус А.Н., Дронжевская Л.Б., Климков В.А., Шагитова Д.А. Формантный и формантно-модуляционный методы оценки разборчивости речи. Часть 1. Унификация алгоритмов. // *Электроника и связь*, 2010, №6, часть 2. - С.117- 124. Available at: <https://www.twirpx.com/file/479327/>.

- 27.Продеус А.Н., Дронжевская Л.Б., Климков В.А., Шагитова Д.А. Формантный и формантно-модуляционный методы оценки разборчивости речи. Часть 2. Точность и скорость измерений. // Электроника и связь. – 2011. – Т. 64. – №6. – С. 16–24. Available at: <https://www.twirpx.com/file/855519/?mode=full>.
- 28.H. Steeneken and T. Houtgast. Basics of the STI measuring method. Reprint from : Past, present and future of the Speech Transmission Index. ISBN 90-76702-02-0. Available at: https://www.steeneken.nl/wp-content/uploads/2014/04/Basics-STI_2014-V1.pdf.
- 29.Prodeus A. Rapid Version of Formant-Modulation Method of Speech Intelligibility Estimation. - Proceedings of the VII International Conference MEMSTECH 2011. - Lviv, Polyana, 2011. - Pp.61-63. Available at: https://www.researchgate.net/publication/252019968_Rapid_version_of_a_formant-modulation_method_of_speech_intelligibility_estimation.
- 30.Glen M. Ballou, Handbook for Sound Engineers (Fourth Edition), Focal Press, 2008, ISBN 9780240809694, <https://doi.org/10.1016/B978-0-240-80969-4.50002-X>.
- 31.R. Youssef, D. Bard, F. Mahmoud and N. Esa, "Acoustical Quality Assessment of Lecture halls at Lund University, Sweden," in Inter-noise, Melburn, pp. 934-943. 6-19 November, 2014.
- 32.C. D. Marino, D. Fuerth, D. Gignac, A. Lunardi, C. Novak, R. P. Pikul and A.Simone, "Acoustic Enhancement of Proposed Grand Lecture Hall usingComputer Simulation," Canadian Acoustics, vol. 39, no. 1, pp. 43-48, 2011.
- 33.Kunchur, C.R. (2019) Evaluating Room Acoustics for Speech Intelligibility. Open Journal of Applied Sciences, 9, 601-612. <https://doi.org/10.4236/ojapps.2019.97048>.
- 34.W. Yang and J. Bradley, "Effects of Room Acoustics on the Intelligibility ofSpeech in Classrooms for Young Children," J. Acoust. Soc. Am, vol. 125, no.2, p. 922–933, 2009.

- 35.J. Bradley, R. Reich and S. Norcross, "On the Combined Effects of Signal-to-Noise Ratio and Room Acoustics on Speech Intelligibility," J. Acoust. Soc. Am, vol. 106, no. 4, Pt.1, pp. 820-1828, 1999.
- 36.Jeon JY, Fricke FR. A Preliminary Study for a 'Music Intelligibility' Test for Rooms. Building Acoustics. 1994;1(3):195-205. doi:10.1177/1351010X9400100302.
- 37.Jared H. Koshiol, Greg DeBlasio, "Improving lyrical intelligibility in live music and concert settings: Evaluating the application of alternative sound reinforcement techniques", The Journal of the Acoustical Society of America 144, 1762-1762 (2018) <https://doi.org/10.1121/1.5067796/>
- 38.A. Prodeus, M. Didkovska, D. Motorniuk, O. Dvornyk, "The Effects of Noise, Early and Late Reflactions on Speech Intelligibility" Proc. 2020 IEEE 40th Int. Conf. on Electronics and Nanotechnology (ELNANO), pp. 488 - 492, DOI: 10.1109/ELNANO50318.2020.9088854
- 39.A. Продеус, К. Бухта, П. Морозко, О. Сергієнко, І. Котвицький, І. Шербенко, "Автоматизована система суб'єктивного оцінювання розбірливості українського мовлення", Proc. of IEEE 38th Int. Conf. on Electronics and Nanotechnology (ELNANO), с. 533-538, 24-26 квітня, Київ, Україна, 2018
- 40.Jared H. Koshiol, Greg DeBlasio, "Improving lyrical intelligibility in live music and concert settings: Evaluating the application of alternative sound reinforcement techniques", The Journal of the Acoustical Society of America 144, 1762-1762 (2018) <https://doi.org/10.1121/1.5067796>.
- 41.J. Bradley and H. Sato, "Speech Intelligibility Test Results for Grades 1, 3 and 6 Children in Real Classrooms," in ICA, Kyoto, 2004.
- 42.W. Yang and J. Bradley, "Effects of Room Acoustics on the Intelligibility of Speech in Classrooms for Young Children," J. Acoust. Soc. Am, vol. 125, no.2, p. 922–933, 2009.
- 43.A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi and D. Motorniuk, "Automated Subjective Assessment of Speech Intelligibility under Diotic and

- Dichotic Listening," in Lecture Notes on Data Engineering and Communications Technologies, 2021, 48, Cham, Springer, January 2021, pp. 1-16.
44. GOST R 50840-95: Transmission of Speech Through Communication Paths. Methods for Assessing Quality, Intelligibility and Recognizability. Gosstandart of Russia, Moscow, 1996.
 45. "Aachen Impulse Response Database," [Online]. Available: <https://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/aachen-impulse-response-database/>.
 46. M. Jeub, M. Schafer and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in 16th International Conference on Digital Signal Processing, Santorini-Hellas, 5-7 July, 2009.
 47. W. Ahnert and W. Schmidt, Appendix to EASERA Manual: Fundamentals to Perform Acoustical Measurements, Berlin, 2006. - P 1 - 53.
 48. A. Prodeus, M. Didkovska, D. Motorniuk and O. Dvornyk, "The Effects of Noise, Early and Late Reflections on Speech Intelligibility," in Proc. 2020 IEEE 40th Int. Conf. on Electronics and Nanotechnology (ELNANO), Kyiv, pp. 488 – 492. 22-24 April 2020. DOI: 10.1109/ELNANO50318.2020.9088854.
 49. J. Blauert, The Technology of Binaural Listening, Berlin, Heidelberg, New York: Springer-Verlag, 2013.
 50. Odeon Software. [Online]. Available <https://odeon.dk/>
 51. H. Lopacz and R. Marczuk, "A New Approach to the Room Impulse Response Simulation", Archives of Acoustics, vol. 29, no. 1, pp. 61-85, 2004.
 52. P. Malecki, "Spatial Impulse Response Assessment in Room Acoustics Auralization," Acta Physica Polonica A, vol. 128, no. 1-A, pp. 17-21, 2015.
 53. H. Kuttruff, Room Acoustics, 4th ed., London EC4P 4EE: Spon Press, 11 New Fetter Lane, 2000.

- 54.W. G. Gardner, "Reverberation Algorithms," in Application of digital signal processing to audio and acoustics, Norwell, MA, Kluwer Academic Publishers, 1998, pp. 85-132.
- 55.S. V. A. Gari, Investigations on the Influence of Acoustics on Live Music Performance using Virtual Acoustic Methods, Ph.D. dissertation, Germany: Detmold University of Music, 2017.
- 56.H. Haas, "Über den Einfluss eines Einfahechos auf die Hørsamkeit von Sprache," *Acustica*, no. 1, pp. 49-58, 1951.
- 57.M. D. Burkhard and R. M. Sachs , "Anthropometric manikin for acoustic research", *The Journal of the Acoustical Society of America* 58, 214-222 (1975) <https://doi.org/10.1121/1.380648>.
- 58.Willinger, R. & Baumgartner, Daniel & Chinn, B. & Schuller, E.. (2001). New dummy head prototype : Development, validation and injury criteria. *International Journal of Crashworthiness - INT J CRASHWORTHINESS*. 6. 281-294. 10.1533/cras.2001.0178.
- 59.Laws, P. & Platte, H.J.. (1978). METHOD OF PRODUCING A DUMMY HEAD FOR HEAD-RELATED STEREOPHONIC RECORDING TECHNIQUES.. 7-10.
- 60.KU 100 – Dummy Head. Neumann.Berlin. [Online]. Available <https://en-de.neumann.com/ku-100>.
- 61.Zhang, S. & Zheng, X. & Meng, Z.. (2015). The inuence of dummy head on measuring speech intelligibility. 40. 894-901.
- 62.Arend, Johannes & Neidhardt, Annika & Pörschmann, Christoph. (2016). Measurement and Perceptual Evaluation of a Spherical Near-Field HRTF Set.
- 63.Stade, Philipp & Bernschütz, Benjamin & Rühl, Max. (2012). A Spatial Audio Impulse Response Compilation Captured at the WDR Broadcast Studios.
- 64.Pörschmann, Christoph & Arend, Johannes & Gillioz, Raphaël. (2019). How wearing headgear affects measured head-related transfer functions.
- 65.Hořák, Pavel. (2016). Measurement of the reverberation time. ResearchGate. DOI:10.13140/RG.2.2.17246.64321.

- 66.DSTU B V.2.6-86:2009 Konstruktsii budynkiv i sporud. Zvukoizoliatsiia ohorodzhuvalnykh konstruktsii. Metody vymiriuvannia [DSTU B V.2.6-86:2009 Structures of buildings and structures. Sound insulation of enclosing structures. Measurement methods].
- 67.ISO 3382-1:2009. Acoustics. Measurement of room acoustic parameters. Part 1: Performance spaces. *ISO*
- 68.Prodeus A., Didkovska M., Kukharicheva K. and Motorniuk D. (2020). Modeling the Influence of Early Sound Reflections on Speech Intelligibility. *2020 IEEE 6th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC)*, pp. 47-50. doi: 10.1109/MSNMC50359.2020.9255657.
- 69.Wang, Yueyue; et al. (2021). Influence of the acoustic structures placement on the measurements in a reverberation room. *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, pp. 429-433. DOI: 10.3397/IN-2021-1474.
- 70.Aida, Y., Inoue, N., & Sakuma, T. (2019). An experimental study on the influence of incident characteristics on sound insulation between rooms. *Journal of Environmental Engineering (Japan)*, Vol. 84, Iss. 764, pp. 893-902. <https://doi.org/10.3130/aije.84.893>.
- 71.Cecchi, Stefania & Carini, Alberto & Spors, Sascha. (2017). Room Response Equalization—A Review. *Applied Sciences*. 8. 16. 10.3390/app8010016.
- 72.Neely, Stephen & Allen, Jont. (1979). Invertibility of a room impulse response. *The Journal of the Acoustical Society of America*. 66. 165-169. 10.1121/1.383069.
- 73.Karjalainen, M.; Paatero, T. Equalization of loudspeaker and room responses using Kautz filters: Direct least squares design. *EURASIP J. Appl. Signal Process*. 2006, 2007, 060949
- 74.Bank, B. Direct design of parallel second-order filters for instrument body modeling. In *Proceedings of the International Computer Music Conference (ICMC)*, Copenhagen, Denmark, August 2017; pp. 458–465

75. Bank, B. Audio Equalization with Fixed-Pole Parallel Filters: An Efficient Alternative to Complex Smoothing. *J. Audio Eng. Soc.* 2013, 61, 39–49
76. Bank, B. Loudspeaker and room response equalization Using parallel filters: comparison of pole positioning strategies. In *Proceedings of the 51st International Audio Engineering Society Conference: Loudspeakers and Headphones*, Helsinki, Finland, 21–24 August 2013.
77. Haneda, Y.; Makino, S.; Kaneda, Y. Common acoustical pole and zero modeling of room transfer functions. *IEEE Trans. Speech Audio Process.* 1994, 2, 320–32.
78. Haneda, Y.; Makino, S.; Kaneda, Y. Multiple-point equalization of room transfer functions by using common acoustical poles. *IEEE Trans. Speech Audio Process.* 1997, 5, 325–333.
79. Fontana, F.; Gibin, L.; Rocchesso, D.; Ballan, O. Common pole equalization of small rooms using a two-step real-time digital equalizer. In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, 20–20 October 1999; pp. 195–198.
80. V. Bruschi, V. Välimäki, J. Liski and S. Cecchi,. “Linear-Phase Octave Graphic Equalizer”. *J. Audio Eng. Soc.*, vol. 70, no. 6, pp. 435–445, (2022 June).
81. A Low-Latency Quasi-Linear-Phase Octave Graphic Equalizer V Bruschi, V Välimäki, J Liski, S Cecchi - *International Conference on Digital Audio Effects*, 2022.
82. C. Hoffman, “Dynamic EQs vs Multiband Compressors,” Black Ghost Audio, 13-Dec-2018. [Online]. Available: <https://www.blackghostaudio.com/blog/the-difference-between-dynamic-eqs-and-multiband-compressors>.
83. S. Hosokawa, “Multiband compressor,” US 8,924,220 B2, 30-Dec-2014, URL:

- <https://patentimages.storage.googleapis.com/ae/73/10/224c15c9a44c95/US8924220.pdf>
84. Everest, F. Alton, and Ken C. Pohlmann. 2022. Master Handbook of Acoustics. 7th ed. New York: McGraw Hill.
<https://www.accessengineeringlibrary.com/content/book/9781260473599>.
 85. Ballou, G. (Ed.). (2015). Handbook for Sound Engineers (5th ed.). Routledge.
<https://doi.org/10.4324/9780203758281>.
 86. R. A. Orban, "Audio compressor combining feedback and feedforward sidechain processing," J Acoust Soc Am, vol. 100, no. 2, p. 688, Aug. 1996, DOI: <https://doi.org/10.1121/1.416271>
 87. Owsinski, The Mastering Engineer's Handbook, 4th ed. Bobby Owsinski Media Group, 2017.
 88. G. Ballou, Ed., Handbook for Sound Engineers, 5th ed. Routledge, 2015, ISBN: 9781135016661, DOI: <https://doi.org/10.4324/9780203758281>
 89. "DriveRack PA2." [Online]. Available: <https://dbxpro.com/en/products/driverack-pa2>.
 90. M. Jeub, M. Schafer, and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in Proc. Int. Conference on Digital Signal Processing (DSP), Santorini, Greece, 2009, doi: <https://doi.org/10.1109/ICDSP.2009.5201259/>.
 91. J. Bradley, R. Reich, R.S. Norcross, "A just noticeable difference in C50 for speech", Applied Acoustics, (58), pp. 99-108, 1999, doi: [https://doi.org/10.1016/S0003-682X\(98\)00075-9](https://doi.org/10.1016/S0003-682X(98)00075-9).
 92. J. Kotus, B. Kostek, A. Kurowski and P. Szczuko, "A Comparison of STI Measured by Direct and Indirect Methods for Interiors Coupled with Sound Reinforcement Systems," 2018 Joint Conference - Acoustics, Ustka, Poland, 2018, pp. 1-6, doi: <https://doi.org/10.1109/ACOUSTICS.2018.8502277>
 93. IEC 60268-16:2011. Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index.

- 94.Захист акустичної інформації: комп'ютерний практикум [Електронний ресурс] : навч. посіб. для студ. спеціальності 171 «Електроніка» / КПІ ім. Ігоря Сікорського ; уклад. А. М. Продеус. – Електронні текстові дані (1 файл: 2,8 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2021. – 101 с.
Доступ: <https://ela.kpi.ua/handle/123456789/45633>
- 95.Zhu P., Mo F., Kang J. Experimental comparison between direct and indirect measurement methods for the objective rating of speech intelligibility. Proc. The 21st International Congress on Sound and Vibration (ICSV 21), 13-17 July, 2014, Beijing/China. Available a
- 96.D'Orazio D, Rossi E, Garai M. Comparison of different in situ measurements techniques of intelligibility in an open-plan office. Building Acoustics. 2018;25(2):111-122. doi:10.1177/1351010X18776431
- 97.J. Kotus, B. Kostek, A. Kurowski and P. Szczuko, "A Comparison of STI Measured by Direct and Indirect Methods for Interiors Coupled with Sound Reinforcement Systems," 2018 Joint Conference - Acoustics, Ustka, Poland, 2018, pp. 1-6, doi: 10.1109/ACOUSTICS.2018.8502277.
- 98.Zhu P., Tao W., Mo F., Guo F., Lu X., Liu X. Experimental comparison of speech transmission index measurement in natural sound rooms and auditoria. Applied Acoustics, 165 (2020), pp. 1-21. Available at: <https://doi.org/10.1016/j.apacoust.2020.107326>
- 99.Birne D., at al. An international comparison of long-term average speech spectra. J. Acoust.Soc.Am., 96 (4), October 1994, pp.2108-2120.
<http://dx.doi.org/10.1121/1.410152>
100. Morales L., at al. A revised speech spectrum for STI calculations. Applied Acoustics, Volume 132, March 2018, Pages 33-42.
<https://doi.org/10.1016/j.apacoust.2017.11.008>
101. Acoustics Engineering, Technical Note TN008 "DIRAC Stimuli", January 2008. Available at <https://www.acoustics-engineering.com/files/TN008.pdf>

102. IKS: Aachen Impulse Response Database.” [Online]. Available: <https://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/aachen-impulse-response-database/>
103. Дідковський В.С., Дідковська М.В., Продеус А.М. Комп'ютерна обробка акустичних сигналів. Навчальний посібник. – К., «Імекс-ЛТД», 2010. – 420 с.
104. H.J.M. Steeneken, and T. Houtgast, “A physical method for measuring speech-transmission quality,” J.Acoust. Soc. Am. - 1980. - Vol. 67. – P. 318-326.
105. H.J.M. Steeneken, and T. Houtgast, “Validation of the revised STIr method,” Elsevier Speech Communication. – 2002. - Vol. 38. – P. 26-37.
106. M.R. Schroeder, “Modulation Transfer Functions: Definition and Measurement,” Acustica, Vol. 49, 1981, p. 179-182.
107. В.Е. Гантмахер, Н.Е. Быстров, Д.В.Чеботарев. Шумоподобные сигналы. Анализ, синтез, обработка. - СПб.: Наука и Техника, 2005. - 400 с.
108. H. Sato and J. Bradley, “Evaluation of acoustical conditions for speech communication in working elementary school classrooms,” J. Acoust. Soc. Am. 106 (4), Pt. 1, pp. 2064–2077, 2004. DOI: 10.1121/1.2839283.
109. G. Soulodre, N. Popplewell, and J. Bradley, “Combined effects of early reflections and background noise on speech intelligibility,” Journal of Sound and Vibration, vol. 135, No.1, pp. 123-133, 1989.
110. J. Bradley, H. Sato, and M. Picard, “On the importance of early reflections for speech in rooms,” J. Acoust. Soc. Am., vol. 113, no. 6, pp. 3233-3244, June 2003. DOI: 10.1121/1.1570439.
111. Zolzer U. DAFX: Digital Acoustic Effects. - John Wiley & Sons, 2011. – 614 p.
112. E. Habets, N. Gaubitch, and P. Naylor, “Temporal selective dereverberation of noisy speech using one microphone,” Proc. of 2008 IEEE

Int. Conf. on Acoustics, Speech and Signal Processing, pp. 4577-4580, March-April 2008.

113. "FabFilter Pro-MB." [Online]. Available: <https://www.fabfilter.com/products/pro-mb-multiband-compressor-plug-in>.
114. "MultiRack." [Online]. Available: <https://www.waves.com/mixers-racks/multirack>.
115. "Audinate Dante." [Online]. Available: <https://www.audinate.com/>.