

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Сумський державний університет
Факультет електроніки та інформаційних технологій
Кафедра електроніки і комп'ютерної техніки

«До захисту допущено»

Завідувач кафедри

_____ Анатолій ОПАНАСЮК

_____ 2023 р.

КВАЛІФІКАЦІЙНА РОБОТА
на здобуття освітнього ступеня бакалавра

зі спеціальності 172 «Телекомунікації та радіотехніка»,
освітньо-професійної програми «Мережеві та інтернет технології»

На тему:

Методи придушення квазістаціонарних шумів у звукозаписі для
використання в ТКС

групи ТК-91 Радька Даниїла Юрійовича

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

Даниїл РАДЬКО

Керівник, старший викладач,
кандидат фізико-математичних наук, доцент

Олексій Д'ЯЧЕНКО

Суми – 2023

Сумський державний університет

Факультет ЕлІТ

Кафедра електроніки і комп'ютерної техніки

Спеціальність 172 Телекомунікації та радіотехніка

ЗАТВЕРДЖУЮ

Завідувач кафедри

електроніки та комп'ютерної техніки

_____ А. С. Опанасюк.

“ ____ ” _____ 2023 р.

ЗАВДАННЯ

до виконання кваліфікаційної випускної роботи бакалавра

Радька Даниїла Юрійовича

1. Тема роботи :« Методи придушення квазістаціонарних шумів у звукозаписі для використання в ТКС»
затверджена наказом по університету № _____ від “ ____ ” _____ 2023 р.
2. Термін здачі студентом закінченої роботи 10.06.2023.
3. Вихідні дані до роботи:
 - Тональна завада 50 Гц;
 - Тональна завада 100 Гц;
 - Тональна завада 2 кГц;
 - Рожевий шум;
 - Білий шум.
4. Зміст розрахунково-пояснювальної записки (перелік питань, які необхідно вирішити)
 - Вступ
 - Огляд літератури по тематиці проєкту.
 - Шум та його властивості.
 - Дослідження якості подавлення шуму методом експертних оцінок.
 - Висновки
 - Список використаної літератури
5. Перелік обов'язкового графічного матеріалу (з точним зазначенням обов'язкових креслень або плакатів)
 - 1.. Структурна схема лабораторного стенда
 - 2.. Середні показники розбірливості до і після очищення
 - 3.. Результат ефективності шумоподавлення для різних типів завад

Дата видачі завдання: 20.03.2023

Прийняв до виконання студент:

Радько Д.Ю.

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Найменування етапів дипломного проєкту	Термін виконання етапів прокату	Примітка
1.	Огляд літератури по тематиці проєкту	23.03.23	
2.	Розробка схеми алгоритму	01.04.23	
3.	Розробка структурної схеми	05.04.23	
4.	Розробка схеми електричної функціональної	06.05.23	
5.	Розробка принципової електричної схеми	16.05.23	
6.	Оформлення графічної частини	20.05.23	
7.	Оформлення пояснювальної записки	07.06.23	
8.	Рецензування роботи та підготовка до захисту	10.06.23	

Студент-дипломник _____

(підпис)

Керівник проєкту _____

(підпис)

АНОТАЦІЯ

Робота містить : 51 с.; 7 рис.; 3табл.

Алгоритм шумоподавлення, вінерівська фільтрація, апріорне відношення сигнал/шум, апостеріорне відношення сигнал/шум, мовний сигнал

Об'єкт дослідження – Тестування методів шумоподавлення за допомогою метрик сегментної PESQ та відношення сигнал/шум.

Предмет дослідження – Дослідження і оцінка суб'єктивної якості мови. Суб'єктивне сприйняття звуку. Частотне та тимчасове маскування з методами оцінки якості мови.

В кваліфікаційній роботі проведені огляд та дослідження зміни шуму у звукових трактах і їх вплив на якісні показники звукозаписів. Проведено аналіз і дослідження базових методів шумоподавлення

ЗМІСТ

ВСТУП.....	8
1 ЗВУКОЗАПИС.....	10
1.1 Аналоговий та цифровий звукозаписи, характеристики, поняття.....	10
1.2 Звукозапис у цифровому вигляді.....	12
1.3 Цифрова обробка.....	15
1.4 PESQ перетворення.....	18
1.5 Якість зв'язку.....	22
2 ШУМ ТА ЙОГО ВЛАСТИВОСТІ.....	25
2.1 Об'єктивний контроль якості звукозапису.....	26
2.3 Передача інформації разом з фоновим звуком та шумами.....	27
2.3 Розробка методів і алгоритмів шумоочищення звуку.....	28
3 ДОСЛІДЖЕННЯ ЯКОСТІ ПОДАВЛЕННЯ ШУМУ МЕТОДОМ ЕКСПЕРТНИХ ОЦІНОК.....	31
3.1 Методика досліджень.....	31
3.2 Оцінка якості шумоподавлення методом артикуляційного вимірювання розбірливості.....	34
3.3 Загальні показники розбірливості.....	42
ВИСНОВКИ.....	45
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....	49

					ЕЛІТ 6.171.00.10.324 ПЗ			
		<i>№ докум</i>	<i>Підпис</i>	<i>Дата</i>				
<i>Розроб.</i>	<i>Радько</i>				<i>Методи подавлення квазістаціонарних шумів у звукозаписі</i>	<i>Лист.</i>	<i>Аркуш</i>	<i>Листів</i>
<i>Перевір.</i>	<i>Дяченко</i>						4	47
<i>Реценз.</i>								
<i>Н. Контр.</i>								
<i>Затверд.</i>	<i>Опанасюк</i>					СумДУ ТК-91		

ВСТУП

Сучасні методи звукозапису дають змогу отримувати високоякісний звуковий вміст, проте навіть найпередовіші технології не можуть повністю виключити можливість появи квазістаціонарних шумів на записі.

Ці шуми можуть бути спричинені різними факторами, такими як електричні перешкоди, шум від обладнання або механічні вібрації. Квазістаціонарні шуми можна визначити як шуми, що мають постійну частоту та амплітуду протягом тривалого часу. Це означає, що вони не змінюються з часом або не змінюються досить швидко, щоб не розглядатися як випадковий шум.

Цифрова обробка звукових сигналів грає дедалі важливішу роль із поширенням мультимедійних можливостей персональних комп'ютерів. Зародження основних методів цифрової обробки відбувалося у середині 20 століття, коли з'являлися перші комп'ютери. У ті роки великі зусилля під час створення алгоритмів цифрової обробки витрачалися на їх ефективну реалізацію. Зараз, коли потужності навіть персональних комп'ютерів зросли настільки, що дозволяють у реальному часі проводити складну обробку сигналів, на передній план виходить питання якості алгоритмів, а не їх швидкодії.

На сьогоднішній день активно розвиваються такі напрямки цифрової обробки звуку голосу: транскрибування - переведення мови в текст; діаризація - поділ вихідного аудіопотоку на окремі аудіопотоки за дикторами та розпізнавання дикторів; визначення мови мови; визначення емоційного забарвлення мови.

Для спрощення роботи даних високорівневих алгоритмів аналізу і розпізнавання мови, а також для поліпшення якості зв'язку між кінцевими абонентами голосового зв'язку вихідний звук піддається попередній обробці - шумоочищенню. В даний час існує кілька класів методів шумоочищення і шумозаглушення, але усі вони мають недоліки. Продуктивних і добре описаних методів шумоочищення голосу для систем

					ЕлІТ 6.171.00.10.324 ПЗ	Арк
		№ докум.	Підпис	БР		5

реального часу (СРВ), що накладають на використання в них алгоритми низької істотних обмежень, недостатньо. Так, застосування смугових фільтрів і гейтів, що задовольняють вимогам СРЧ щодо часу відгуку алгоритму і передбачуваності його роботи, виявляється марним за умови мінливого в часі вигляду і напрямку шуму, а найефективніші сучасні нейромережеві підходи або вносять у роботу СРВ велику затримку, або вимагають специфічного дорогого обладнання і (або) допоміжних даних.

Компромiсним рішенням, що забезпечує прийнятні швидкість і якість шумоочищення, є розробка алгоритмів фільтрації, які передбачають, що сигнали шуму і голосу некорельовані. Зазвичай такі алгоритми складаються з двох частин: детектора шуму та його безпосереднього пригнічувача.

Мета цієї роботи - розробка набору нових адаптивних методів і алгоритмів шумозаглушення для мовлення на основі найефективніших алгоритмів детектування та придушення шуму.

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	<i>Арк</i>
		<i>№ докум.</i>	<i>Підпис</i>	<i>БР</i>		6

1 ЗВУКОЗАПИС

Звукозапис - це фіксування будь-якого аудіоматеріалу на конкретний носій. З винаходу фонографа, а точніше з 12 серпня 1877 року, і починається точка відліку історії звукозапису. Світ розвивався, а з його вдосконаленням приходили нові ідеї. Поступово задумане втілювалося в життя: став можливим аналоговий запис звуку, а завдяки величезному технічному прогресу, що стався в останні десятиліття, з'явився і цифровий звукозапис. Під час підготовки програм радіомовлення з усіх технологічних операцій найчастіше використовується запис звукових сигналів.

1.1 Аналоговий та цифровий звукозаписи, характеристики, поняття

Звук як такий має аналогову природу. Він поширюється в повітрі і при цьому неминуче спотворюється. На спотворення звуку впливають найрізноманітніші умови: відстань від джерела, швидкість руху відносно нього, особливості відбиття від навколишніх предметів тощо. Людське вухо сприймає звукові коливання в діапазоні від 20 Hz до 20 000 Hz. Однак далеко не кожен може похвалитися такими видатними слуховими можливостями. Основна маса дорослих чує частоти до 16 000-18 000 Hz. Варто уточнити, що навіть частоти, вищі за 6 000-8 000 Hz, зазвичай є лише додатковими гармоніками та призвуками. З іншого боку, якість запису багато в чому визначається саме правильним відтворенням гармонік та інших високочастотних елементів. Під час аналогового запису звукова хвиля, потрапляючи в мікрофон, перетворюється на електричне коливання, яке потім подається або на механічний різець, якщо йдеться про вінілову платівку, або на магнітну голівку, якщо запис здійснюється на магнітну стрічку. Щоб відтворити звук, слід протягнути намагнічену стрічку вздовж магнітної голівки, причому швидкість цього процесу має дорівнювати швидкості запису.

					ЕлІТ 6.171.00.10.324 ПЗ	Арк
						7
		№ докум.	Підпис	БР		

У випадку ж з вінілом для відтворення буде потрібно прогнати голку по канавці, в якій записана інформація. Механічні коливання будуть перетворені в електричні, які передадуться в підсилювач, а з підсилювача відповідно в гучномовці. При уважному ознайомленні з вищевикладеним матеріалом цілком очевидна недосконалість аналогового запису.

При уважному ознайомленні з вищевикладеним матеріалом цілком очевидна недосконалість аналогового запису.

1. Записуючи на магнітну стрічку, слід потурбуватися про якість магнітної голівки і врахувати її калібрування щодо стрічки.
2. Неточності стрічкопротяжного механізму породжують мінливість її швидкості.
3. Не можна не згадати про здатність стрічки розтягуватися, про зміни її характеристик на всьому протязі, про випадкові сторонні частинки на ній тощо.
4. У випадку з вініловою платівкою мають місце детонація, потрапляння пилу в канавки і всілякі механічні пошкодження. Крім того, канавка, так чи інакше, деформується після кожного програвання.
5. Ну і, нарешті, варто згадати, що практично неможливо зробити копію вінілової платівки або магнітного запису без втрати якості. Та й усі аналогові носії з часом старіють і втрачають у якості звучання, навіть якщо їх не використовувати занадто часто.

1.2 Звукозапис у цифровому вигляді

Початкова форма звукового сигналу - безперервна зміна амплітуди в часі - представляється в цифровій формі за допомогою "перехресної дискретизації" - за часом і за рівнем. Строго кажучи, будь-який аналоговий сигнал зрештою теж дискретний як за часом, так і за величиною - наприклад, звуковий сигнал є результатом взаємодії скінченної кількості атомів або молекул газів і твердих тіл,

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	Арк
						8
		<i>№ докум.</i>	<i>Підпис</i>	<i>БР</i>		

електричний сигнал - результатом переміщення скінченної кількості елементарних зарядів (електронів) тощо. Тому, говорячи про дискретизацію аналогового сигналу, мають на увазі значно більший ступінь дискретності, що піддається вимірюванню приладами середньої точності. Згідно з теоремою Котельникова, будь-який безперервний процес з обмеженим спектром може бути повністю описаний дискретною послідовністю його миттєвих значень, які слідують із частотою, що як мінімум удвічі перевищує частоту найвищої гармоніки процесу; частота F_d вибірки миттєвих значень (відліків) називається частотою дискретизації.

З теореми випливає, що сигнал із частотою F_a може бути успішно дискретизовано в часі на частоті $2F_a$ тільки в тому разі, якщо він є чистою синусоїдою, бо будь-яке відхилення від синусоїдальної форми призводить до виходу спектра за межі частоти F_a . Таким чином, для тимчасової дискретизації довільного звукового сигналу (зазвичай має, як відомо, плавно спадаючий спектр), необхідний або вибір частоти дискретизації із запасом, або примусове обмеження спектра вхідного сигналу нижче половини частоти дискретизації. Одночасно з часовою дискретизацією виконується амплітудна - вимір миттєвих значень амплітуди та їхнє представлення у вигляді числових величин із певною точністю (квантування). Точність вимірювання (двійкова розрядність N одержуваного дискретного значення) визначає рівень шуму, що вноситься квантуванням, і динамічний діапазон цифрового сигналу (теоретично, для цифрового сигналу в чистому вигляді, це взаємно-обернені величини, проте будь-який реальний тракт має також і власний рівень шумів та перешкод). Отриманий потік чисел (серій двійкових цифр), що описує звуковий сигнал, називають імпульсно-кодовою модуляцією або ІКМ (Pulse Code Modulation, PCM), бо кожен імпульс дискретизованого за часом сигналу подається власним цифровим кодом. Найчастіше застосовують лінійне

					<i>ЕЛІТ 6.171.00.10.324 ПЗ</i>	Арк
		№ докум.	Підпис	БР		9

квантування, коли числове значення відліку пропорційне амплітуді сигналу. Через логарифмічну природу слуху доцільнішим було б логарифмічне квантування, коли числове значення пропорційне величині сигналу в децибелах, однак це пов'язано з труднощами технічного характеру.

Тимчасова дискретизація й амплітудне квантування сигналу неминуче вносять у сигнал шумові спотворення, рівень яких заведено оцінювати за формулою $6N + 10\lg(F_{\text{дискр}}/2F_{\text{макс}}) + C$ (дБ), де константа C варіюється для різних типів сигналів: для чистої синусоїди це 1.7 дБ, для звукових сигналів - від -15 до 2 дБ. Звідси видно, що до зниження шумів у робочій смузі частот $0 \dots F_{\text{макс}}$ призводить не тільки збільшення розрядності відліку, а й підвищення частоти дискретизації відносно $2F_{\text{макс}}$, оскільки шуми квантування "розмазуються" по всій смузі аж до частоти дискретизації, а звукова інформація займає тільки нижню частину цієї смуги.

У більшості сучасних цифрових звукових систем використовуються стандартні частоти дискретизації 44.1 і 48 кГц, проте частотний діапазон сигналу зазвичай обмежується біля 20 кГц для залишення запасу щодо теоретичної межі. Також найбільш поширене 16-розрядне квантування за рівнем, що дає граничне співвідношення сигнал/шум близько 98 дБ. У студійній апаратурі використовуються вищі роздільні здатності - 18-, 20- і 24-розрядне квантування за частот дискретизації 56, 96 і 192 кГц. Це робиться для того, щоб зберегти вищі гармоніки звукового сигналу, які безпосередньо не сприймаються слухом, але впливають на формування загальної звукової картини. Для оцифрування більш вузькосмугових і менш якісних сигналів частота і розрядність дискретизації можуть знижуватися; наприклад, у телефонних лініях застосовується 7 або 8-розрядне оцифрування з частотами 8...12 кГц.

					<i>ЕлІТ 6.171.00.10.324 ПЗ</i>	Арк
						10
		№ докум.	Підпис	БР		

Подання аналогового сигналу в цифровому вигляді називається також імпульсно-кодовою модуляцією (ІКМ, РСМ - Pulse Code Modulation), тому що сигнал подають як серію імпульсів постійної частоти (тимчасова дискретизація), амплітуда яких передається цифровим кодом (амплітудна дискретизація). РСМ-потік може бути як паралельним, коли всі біти кожного відліку передаються одночасно кількома лініями з частотою дискретизації, так і послідовним, коли біти передаються один за одним з вищою частотою однією лінією.

Сам цифровий звук і речі, що відносяться до нього, заведено позначати загальним терміном Digital Audio; аналогова і цифрова частини звукової системи позначаються термінами Analog Domain і Digital Domain. Не можна не згадати, що цифровий запис не схильний до старіння або якихось інших тимчасових змін. З нього можна створити скільки завгодно копій з однаковою точністю. Як можна помітити з усього вищесказаного, теорія цифрового запису не має на увазі наявності будь-яких у ній вад. На практиці це виглядає так:

1. По-перше, для отримання високої якості потрібне високоякісне оцифрування аналогового звуку, яке головним чином залежить від якості АЦП - аналого-цифрового перетворювача. Високласний мікрофон або дорогі з'єднувальні кабелі не допоможуть у ситуації, коли якість роботи АЦП залишає бажати кращого. Запис відліків з недостатньою точністю, створення вибірок з нерівномірною частотою тощо призведуть до отримання звуку, далекого за якістю від оригіналу, і виправити це вже не вдасться потім нічим.

2. А по-друге, оцифрований звук потрібно ще й якісно відтворити, що можливо тільки за наявності якісного ЦАП - цифро-аналогового перетворювача. Через нерівномірну частоту дискретизації, недостатню точність або відсутність

					<i>ЕлІТ 6.171.00.10.324 ПЗ</i>	<i>Арк</i>
		<i>№ докум.</i>	<i>Підпис</i>	<i>БР</i>		11

інтерполяції звук зіпсується так, що ніяка сучасна акустична система цього не компенсує. Таким чином, можна зрозуміти що, на якість цифрового запису і відтворення головним чином впливає якість перетворювачів.

Перетворювачі, вмонтовані в сучасні (причому, аж ніяк не в найдешевші) аудіоінтерфейси, у своїй основній масі не здатні видавати дійсно високоякісний звук, і з цієї причини багато хто віддає перевагу аналоговому запису. Але, все ж таки, резюмуючи вищесказане, варто зазначити, що цифровий запис має певні і досить виражені переваги, порівняно з аналоговим. Хоча на практиці для отримання дійсно якісного цифрового звуку потрібно витратити чимало коштів на високоякісні перетворювачі.

1.3 Цифрова обробка

Перш за все, необхідно розрізнити "спотворювальні" і "неспотворювальні" види обробки. До перших належать операції, що змінюють форму і структуру сигналу - змішування, посилення, фільтрація, модуляція тощо, до других - операції монтажу (вирізка, вклеювання, накладення) і перенесення (копіювання).

Якість сигналу може страждати тільки під час "спотворювальної" обробки, причому будь-якої - і аналогової, і цифрової. У першому випадку це відбувається внаслідок внесення шумів, гармонійних, інтермодуляційних та інших спотворень у вузлах аналогового тракту, у другому - завдяки кінцевій точності квантування сигналу і математичних обчислень. Усі цифрові обчислення виконуються в деякій розрядній сітці фіксованої довжини - 16, 20, 24, 32, 64, 80 і більше біт; збільшення розрядності сітки підвищує точність обчислень і зменшує помилки заокруглення, проте в загальному випадку не може виключити їх повністю. Кінцева точність квантування первинного аналогового сигналу призводить до того, що навіть у разі абсолютно точного опрацювання отриманого цифрового сигналу квантуване

					ЕЛІТ 6.171.00.10.324 ПЗ	Арк
						12
		№ докум.	Підпис	БР		

значення кожного відліку все одно відрізняється від свого ідеального значення. Для мінімізації спотворень під час обробки в студіях вважають за краще обробляти і зберігати сигналограми на майстер-носіях з підвищеною роздільною здатністю (20, 24 або 32 розряди), навіть якщо результат буде тиражуватися на носії з меншою роздільною здатністю.

Крім власне помилок обчислень і округлення, на точність сильно впливає вибір подання числових відліків сигналу під час обробки. Традиційне подання РСМ з так званою фіксованою точкою (fixed point), коли відліки подаються цілими числами, найзручніше та тягне за собою мінімум накладних витрат, однак точність обчислень залежить від масштабу операцій - наприклад, під час множення утворюються числа вдвічі більшої розрядності, які потім доводиться приводити назад до розрядності початкових відліків, а це може призвести до переповнення розрядної сітки. Компромісним варіантом слугує проміжне збільшення розрядності відліків (наприклад, 16->32), що знижує ймовірність переповнення, однак вимагає більшої обчислювальної потужності, обсягу пам'яті та вносить додаткові спотворення при зворотному зниженні розрядності. Крім того, зниженню похибки сприяє правильний вибір послідовності комутативних (таких, що допускають перестановку) операцій, групування дистрибутивних операцій, врахування особливостей роботи конкретного процесора тощо.

Іншим способом збільшення точності є перетворення відліків у форму з плаваючою крапкою (floating point) з поділом на значущу частину - мантису і показник величини - порядок. У цій формі всі операції зберігають розрядність значущої частини, і множення не призводить до переповнення розрядної сітки. Однак, як саме перетворення між формами з фіксованою і плаваючою крапкою, так і обчислення в цій формі вимагають на порядки більшої швидкодії процесора, що сильно ускладнює їх використання в реальному часі.

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	Арк
						13
		№ докум.	Підпис	БР		

Незважаючи на те, що якість сигналу (відносно вихідного аналогового) неминуче, хоч і несуттєво, погіршується під час будь-якого "спотворювального" цифрового опрацювання, деякі операції за певних умов є повністю й однозначно оборотними. Наприклад, посилення сигналу за амплітудою втричі полягає в множенні кожного відліку на три; якщо цю операцію виконували з фіксованою крапкою і водночас не виникло переповнення, за допомогою ділення на три потім можна буде повернути всі відліки до початкового стану, тим самим повністю відновивши початковий стан сигналу. Сказане вище демонструє, що погіршення якості під час цифрової обробки, що "спотворює", зовсім не обов'язково накопичується з часом, хоча в більшості реальних застосувань відбувається саме так. Крім того, це не означає, що будь-яка операція цифрового посилення завжди буде однозначно оборотною - це залежить від багатьох особливостей застосування операції.

Під час порівняння актуальності аналогової та цифрової обробки потрібно мати на увазі, що обидва види вносять у сигнал деяку кількість похибок, які можна розглядати як шум, проте кожен вид обробки має свої особливості. Будь-який елемент аналогового тракту вносить шум завжди, проте його величина коливається залежно від якості тракту та виду обробки. При цифровій обробці "первинний" шум завжди вноситься самим фактом квантування, а всередині цифрового тракту його можуть вносити тільки операції з обмеженою точністю (наприклад, описане множення на три шуму не додає - воно лише посилює раніше внесені шуми, і шум квантування в тому числі).

З цього випливає, що точні операції не змінюють співвідношення сигнал/шум цифрового сигналу, однак можуть збільшувати помилки первинного квантування. Таким чином, збільшення точності цифрового опрацювання саме по собі не гарантує гарної якості сигналу - необхідне ще й досить точне квантування.

					<i>ЕЛІТ 6.171.00.10.324 ПЗ</i>	<i>Арк</i>
		<i>№ докум.</i>	<i>Підпис</i>	<i>БР</i>		14

Наприклад, 20-розрядне квантування в поєднанні з 24-розрядним опрацюванням уже може успішно змагатися з багатьма аналоговими пристроями, тоді як таке саме опрацювання за 16-розрядного квантування істотно їм поступається. Якісно реалізоване цифрове опрацювання може давати істотно менший рівень спотворень, ніж таке саме аналогове, хіба що це будуть спотворення різних видів.

1.4 PESQ перетворення

Перцептивна оцінка якості мовлення (Perceptual Evaluation of Speech Quality, PESQ) - це сімейство стандартів, що містить методологію тестування для автоматизованої оцінки якості мовлення, яку відчуває користувач системи телефонії. PESQ використовується для об'єктивного тестування якості голосу виробниками телефонів, постачальниками мережевого обладнання та телекомунікаційними операторами.

Цей алгоритм являє собою об'єктивну методику визначення якості мовленнєвого зв'язку в телефонних системах, яка прогнозує результати суб'єктивної оцінки якості цього виду зв'язку слухачами-експертами. Для визначення якості передавання мови в PESQ передбачено порівняння вхідного, або еталонного, сигналу з його спотвореною версією на виході системи зв'язку. Цей процес схематично показано на Рисунок 1.1:

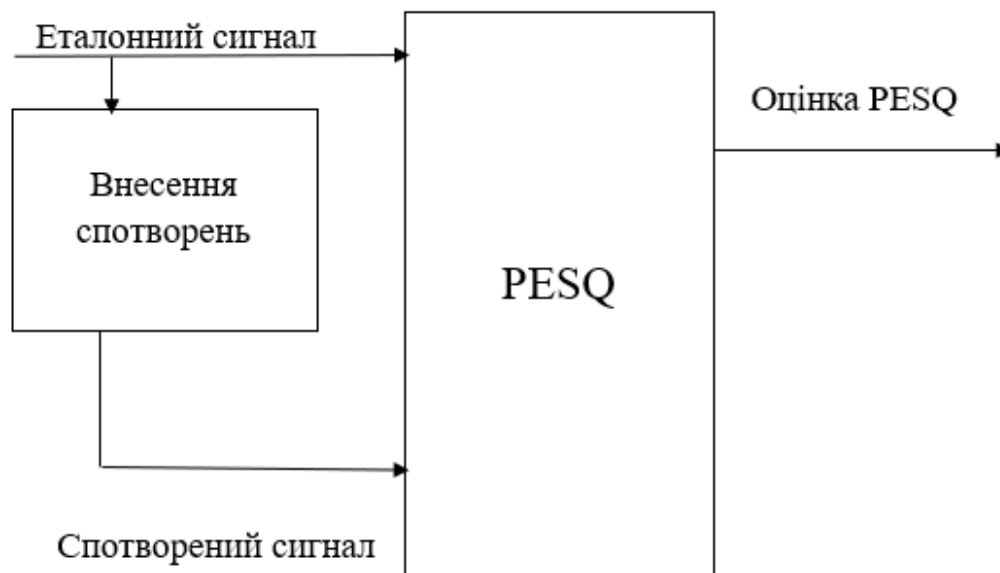


Рисунок 1.1 - Порівняння вхідного, або еталонного, сигналу з його спотвореною версією на виході системи зв'язку

У PESQ закладено безліч нових розробок, що вигідно вирізняє його з-поміж попередніх алгоритмів оцінювання якості роботи мовних кодеків, наприклад, PSQM і MNB [MCE-T P.861]. Ці інновації дають змогу впевнено використовувати PESQ як для визначення якості наскрізного (end-to-end) передавання мовлення, так і для оцінювання впливу на якість зв'язку окремих елементів мережевого обладнання, включно з кодеками.

У процесі розроблення стандарту PESQ фахівці MCE-T відбирали найкращі методи визначення якості мовного зв'язку з погляду кореляції результатів, які вони дають, з оцінками MOS за різних умов зв'язку, що є гарантією хорошої роботи стандартизованого алгоритму при тестуванні звичайних (фіксованих і мобільних) мереж та систем передавання пакетних даних.

В алгоритмі PESQ враховуються такі причини погіршення якості сигналу: спотворення його під час кодування, помилки під час передавання, втрата пакетів, час затримки передавання пакетів і флуктуація цього часу, фільтрація сигналу в аналогових мережевих компонентах.

					<i>ЕлІТ 6.171.00.10.324 ПЗ</i>	Арк
						16
		<i>№ докум.</i>	<i>Підпис</i>	<i>БР</i>		

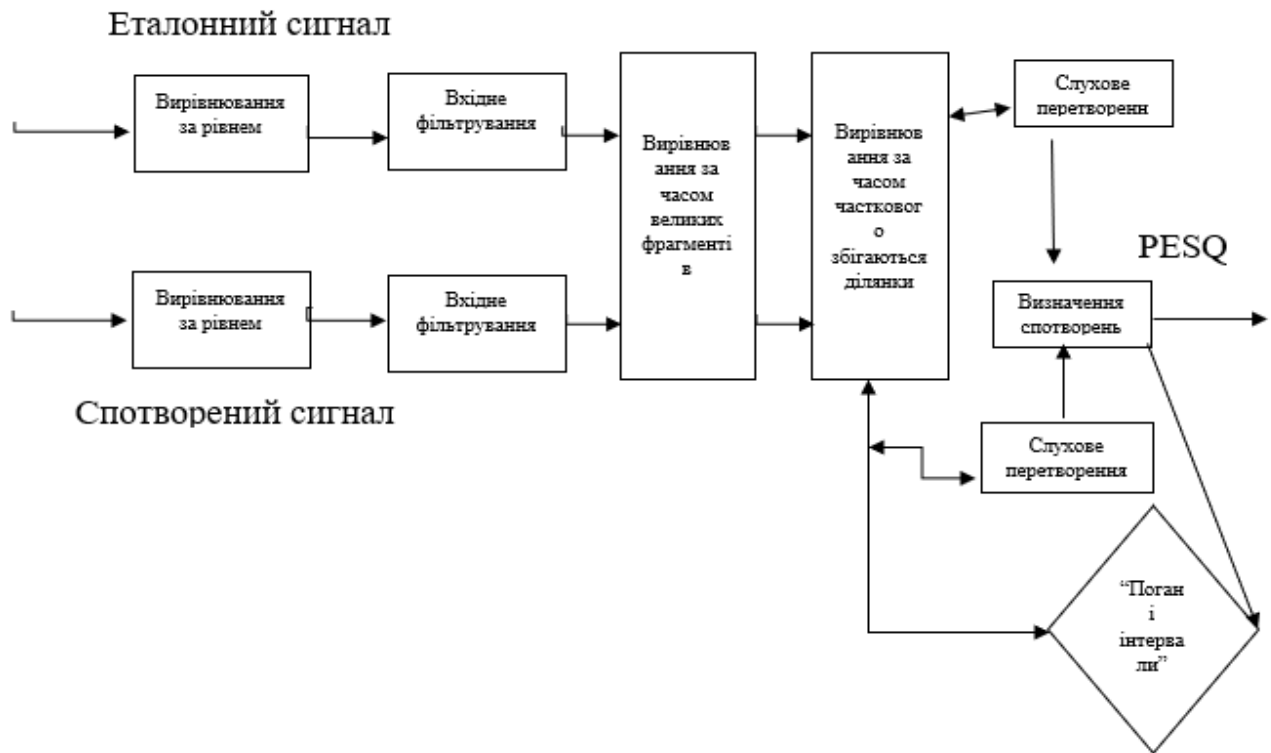


Рисунок 1.2 - Процес оброблення сигналів у PESQ

1. Вирівнювання за рівнем.

Для коректного порівняння вхідного і вихідного мовних сигналів їхній рівень потужності потрібно вирівнювати. Це необхідно, оскільки вхідний сигнал не може бути будь-якого певного рівня, і коефіцієнт посилення тестованої системи невідомий до проведення випробувань. У PESQ прийнято, що рівень прослуховуваного сигналу постійний і дорівнює 79 дБ звукового тиску в точці ERP (Ear Reference Point) [MCE-T P.830, розділ 8.1.2]. Для доведення до зазначеного рівня посилюються обидва сигнали - вхідний і вихідний.

1. Вхідне фільтрування

Аналогові з'єднання часто тією чи іншою мірою фільтрують передані ними сигнали. Наприклад, передавальна частина телефонної трубки зазвичай фільтрує мовний сигнал, маючи амплітудно-частотну характеристику (АЧХ), яка схожа на стандартну АЧХ Modified IRS (Intermediate Reference System) send [МСЕ-Т Р.830]. Як правило, це допустимо, оскільки такого роду обробка сигналу чинить менший вплив на якість зв'язку, ніж спотворення сигналу, що виникають під час його кодування.

2. Вирівнювання за часом

У системі зв'язку може мати місце змінна затримка передачі сигналів. Щоб коректно порівнювати вхідний і вихідний сигнали, вони мають бути вирівняні відносно один одного за часом. У PESQ моделюється прослуховування сигналу, а інформації про час затримки останнього в мережі в ньому немає. Для ідентифікації мовних частин сигналу та відкидання шуму в PESQ виявляється голос. Вирівнювання за часом проводиться в три етапи:

На першому етапі PESQ вирівнює великі фрагменти активної мови, ідентифіковані детектором голосу. Ці фрагменти можуть містити паузи, тривалість яких не перевищує заздалегідь визначеного порогового значення (200 мс). У цьому процесі виявляється затримка передавання великих фрагментів вихідного сигналу, порівнюваного із вхідним.

На другому етапі PESQ вирівнює невеликі ділянки мови (кадри), що частково збігаються за часом. Цей процес виявляє затримку, яка непостійна протягом передавання великого фрагмента активної мови, у пакетних мережах така затримка може бути досить значною.

Третій етап проводиться після операції слухового перетворення. На цьому етапі повторно вирівнюються так звані "погані інтервали" (фрагменти мови з дуже великими спотвореннями). Цей крок підвищує точність роботи алгоритму під час

					ЕліТ 6.171.00.10.324 ПЗ	Арк
						18
		№ докум.	Підпис	БР		

використання невеликої кількості файлів, під час передачі яких неправильно визначається варіація затримки під час початкового процесу вирівнювання за часом.

1.5 Якість зв'язку

Це до певної міри штучні методи обробки цифрового звукового сигналу, спрямовані на поліпшення суб'єктивної якості звучання ціною очевидного погіршення його об'єктивних характеристик (передусім - коефіцієнта нелінійних спотворень і співвідношення сигнал/шум).

Dithering (згладжування) полягає в додаванні до сигналу невеликої кількості шуму (псевдовипадкового цифрового сигналу) різного спектра (білий, рожевий тощо). При цьому помітно послаблюється кореляція помилок квантування з корисним сигналом ("розсіюються" помилки округлення) і, незважаючи на деяке збільшення шуму, суб'єктивна якість звучання помітно підвищується. Рівень шуму, що додається, обирається залежно від завдання і коливається від половини молодшого розряду відліку до кількох розрядів. Noise Shaping (формування шуму) полягає в перетворенні сильно зашумленого корисного сигналу з метою витіснення суто шумових компонент у надтональну область із виділенням у нижній частині спектра основної енергії корисного сигналу. По суті, Noise Shaping є одним із видів PWM (Pulse Width Modulation - широтно-імпульсна модуляція, ШІМ) з дискретною шириною імпульсу. Сигнал, оброблений цим методом, вимагає обов'язкової фільтрації з придушенням високих частот - це виконується або цифровим, або аналоговим способом.

Основне застосування Noise Shaping знаходить у сфері подання цифрових сигналів відліками меншої розрядності з підвищеною частотою проходження. У delta-sigma ЦАП для підвищення частоти проходження відліків збільшується в десятки разів частота дискретизації, на якій з вихідних багаторозрядних відліків формуються серії відліків з розрядністю 1...3. Низькочастотна частина спектра потоку цих відліків з високою точністю повторює спектр вихідного сигналу, а високочастотна містить в

					ЕЛІТ 6.171.00.10.324 ПЗ	Арк
						19
		№ докум.	Підпис	БР		

основному чистий шум. У разі перетворення цифрового сигналу до відліків нижчої розрядності на тій самій частоті дискретизації Noise Shaping виконується разом з операцією Dithering'a. Оскільки в цьому разі підвищення частоти дискретизації неможливе, натомість спектр шуму, що додається, формують таким чином, щоб його низька середньочастотна частина максимально точно повторювала слабку частину сигналу, що міститься у молодших розрядах відліків, які відсікаються. Завдяки цьому основна енергія шуму витісняється у верхню частину робочого діапазону частот, а в найбільш чутній ділянці залишаються цілком розбірливі сліди слабого сигналу, який інакше виявився б повністю знищеним. Незважаючи на те, що об'єктивні спотворення збереженого таким чином слабого сигналу дуже великі, його суб'єктивне сприйняття залишається цілком прийнятним, даючи змогу сприймати на слух компоненти, рівень яких менший за молодший розряд відліку.

По суті, Dithering і Noise Shaping є окремими випадками однієї технології - з тією різницею, що в першому випадку використовується білий шум з рівномірним спектром, а в другому - шум зі спектром, спеціально сформованим під конкретний сигнал. Ця технологія призводить до "нестандартного" використання цифрового формату, заснованого на особливостях людського слуху.

2 ШУМ ТА ЙОГО ВЛАСТИВОСТІ

Шуми, крім ілюстрації звукової предметності кадру, виконують сюжетно-драматургічну функцію, беручи активну участь у формуванні атмосфери фільму.

Шуми діляться на види:

- внутрішньокадрові (синхронні) - шуми, які супроводжують видимий у кадрі рух;
- закадрові (несинхронні) - шуми, джерело яких знаходиться поза кадром;
- фонові - другопланові шуми, які не потребують синхронності.

					ЕлІТ 6.171.00.10.324 ПЗ	Арк
		№ докум.	Підпис	БР		20

Закадрові та фонові шуми характеризуються протяжністю звучання, залежно від драматургії є сильним виразним, художнім засобом.

Характеристики всіх шумів складаються з кількох понять:

- Тембр - характерне забарвлення звуку, що дає змогу відрізнити один від одного два однотипні шуми (кінський тупіт по бруківці або по дерев'яному мосту).
- Ритм - періодичне чергування звуків за тривалістю звучання і пауз (стукіт коліс на стиках рейок).
- Темп - швидкість, з якою чергуються звуки (потяг, що швидко і повільно йде).

Функції шумів різноманітні. Шуми можуть вказати час і місце дії, у деяких випадках конкретніші й образніші, ніж музика (виття вітру, шум дощу, моря, дзвін тощо). З їхньою допомогою режисер може створити звуковий підтекст кадру, більш тонко змалювати характер дійової особи. Шуми можуть сприйматися "вухами" героїв на екрані, тобто ми чуємо, так як чують дійові особи.

Простір сприймається як реальність у тому разі, якщо він має звучання. Відсутність деяких видів шумів так само здатна виконувати образну функцію, є найсильнішим драматургічним елементом.

Шуми відповідають за створення атмосфери всередині кадру. Атмосфера, поряд з іншими виразними засобами, сприяє створенню цілісного образу фільму. Так само як і музичний лейтмотив, у кіно може існувати шумовий лейтмотив, лейтобраз, лейттема. Шуми, піддавшись частотній, просторовій, тимчасовій (відлуння) та ін. обробці, можуть створити атмосферу фантастики, міражу, сну, асоціацій тощо. Так само в образний звукоряд можна внести поняття "музичний шум", "шумо-музика". Синтезування шумів сильно розкриває, розкріпає можливості звукорежисера, дає змогу створити унікальні шумові фактури.

					ЕліТ 6.171.00.10.324 ПЗ	Арк
		№ докум.	Підпис	БР		21

Рисунок 1.3 – Принцип роботи корелометра

Безладний рух стрілки навколо і близько нуля "0" характерний для ідеального, з погляду співвідношення фаз, сигналу. Відхилення в бік (-) свідчить про наявність у стереосигналі протифазних компонентів. Відхилення в бік (+) свідчить про наявність моносигналу.

Спектроаналізатор - прилад для вимірювання відносного розподілу енергії звукових коливань у смузі чутних частот. Відображає обертоновий, гармонійний склад:

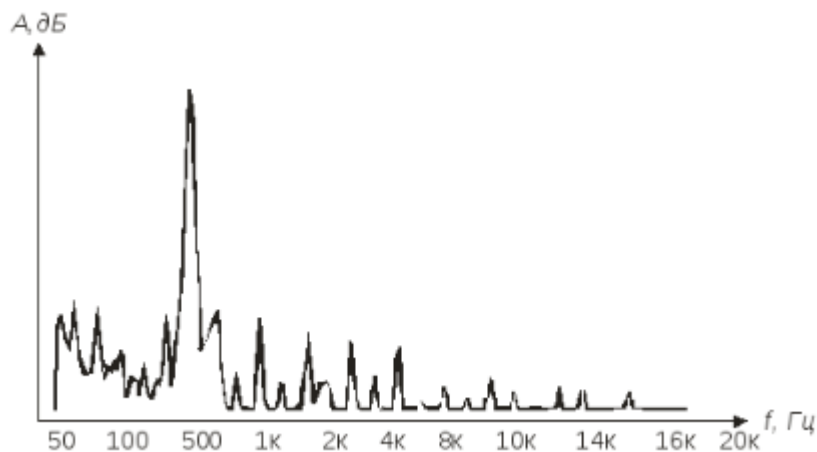


Рисунок 1.4 – Спектроаналізатор

2.2 Передача інформації разом з фоновим звуком та шумами

Передача мови в аудіовізуальній програмі - завдання, що потребує особливої уваги, під час якого необхідно зберегти атмосферу, інтонації, тембр, індивідуальні особливості.

У документальному, неігровому кіно збереження особливостей мовлення є важливим завданням. Частотна, темброва, динамічна корекція має бути дуже акуратною, щоб кадр не втратив документальності.

					ЕліТ 6.171.00.10.324 ПЗ	Арк
						23
		№ докум.	Підпис	БР		

В ігровому, постановочному кіно мова - один із найважливіших виразних засобів. Чіткість, розбірливість, музикальність, інтонаційні та темброві особливості реплік, монологів, діалогів у кадрі спрямовані на максимальне розкриття художнього образу. Голос за кадром - передає думки, почуття, настрій, тому для запису голосу за кадром слід залучати диктора або актора з поставленим голосом, чіткою дикцією, багатим тембром, вимовою та інтонаційною переконливістю.

Розбірливість мови можна визначити як - відношення кількості почутого тексту до кількості переданого тексту (вимірюється у відсотках).

Недостатня чіткість може бути з низки причин:

- фонетична неопрацьованість мови, перенасиченість звукової партитури шумами, музикою з маскувальним ефектом;
- голосові особливості диктора, актора (дикція, скоромовка, акцент);
- невдалий запис і/або монтаж (робота звукорежисера)
- низькі якісні параметри тракту звукопередачі (приміщення, зайвий шум на природі, нелінійні спотворення...)

2.3 Розробка методів і алгоритмів шумоочищення звуку

На сьогоднішній день найактивніше використовуються дві реалізації алгоритмів шумоочищення: Speex (SpeexDSP) і WebRTC. Для оцінки спектра шуму в Speex застосовується імовірнісний алгоритм Кохена - Бердуго MCRA (Minima-Controlled Recursive Averaging), а для безпосереднього придушення отриманого спектра шуму - алгоритм Ефраїма - Малаха. У WebRTC для детектування шуму здійснюється квантильне оцінювання, а для шумопридушення - класична вінерівська фільтрація.

Дослідження методів безпосереднього шумоочищення і методів детектування шуму показали, що існує низка алгоритмів, які працюють не гірше, а нерідко й ефективніше за ті, які використовуються в Speex і WebRTC. Так, для детектування шуму може застосовуватися подальший розвиток MCRA - розробка Рангачарі-Лойзю MCRA2, а

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	Арк
						24
		№ докум.	Підпис	БР		

для безпосереднього шумопоглинання - одна із сучасних модифікацій вінерівської фільтрації.

Об'єднаємо детектор MCRA2 і методи шумозаглушення, описані в формулах 1.1 і 1.2, і опишемо принципи їх роботи.

Нехай вхідний сигнал обробляється алгоритмом БПФ із 50 %-вим перекриттям і, наприклад, віконною функцією Ханна, що задається такою формулою:

$$win(n) = 0,5(1 - \cos(\frac{2\pi n}{N-1})), \quad n \in [0...N-1], \quad (2.1)$$

де N - розмір вікна.

На початку роботи MCRA2 (як і MCRA) виконується рекурсивне згладжування вхідного спектра потужності. Воно необхідне для придушення малих локальних мінімумів і максимумів спектра потужності:

$$P(\lambda, k) = \alpha_{sm} P(\lambda - 1, k) + (1 - \alpha_{sm}) P(\lambda, k), \quad (2.2)$$

Де $P(\lambda, k)$ і $P(\lambda - 1, k)$ - спектр потужності поточної та попередньої згладженої смуги k ; $[0,7, 0,9]$ α_{sm} - коефіцієнт згладжування

Далі здійснюється пошук мінімумів енергії в смугах спектра. На відміну від MCRA, що здійснює пошук у коротких вікнах (0,4-1 с), MCRA2 виконує безперервний пошук мінімумів. Цей процес може бути описаний такою формулою:

$$P_{min}(\lambda, k) = \begin{cases} P(\lambda, k), & P_{min}(\lambda - 1, k) \geq P(\lambda, k), \\ \gamma P_{min}(\lambda - 1, k) + \frac{1 - \gamma}{1 - \beta} (P(\lambda, k) - \beta P(\lambda - 1, k)), & \text{else,} \end{cases} \quad (2.3)$$

Де $P(\lambda, k)$ і $\min P(\lambda - 1, k)$ - спектри мінімумів потужностей поточної та попередньої смуги з номером k ;

$\gamma \in (0, 0, 1, 0)$ - ваговий параметр нових смуг;

$\beta \in (0, 0, 1, 0)$ - коефіцієнт, що відповідає за швидкість адаптації алгоритму до мінливого шуму;

На підставі поточних значень потужності та мінімуму потужності може бути проведено приблизну оцінку апостеріорного відношення зашумленого сигналу до шуму (ОСШ). На її основі здійснюється бінарна оцінка наявності мови в даній смузі:

$$S_{pp}(\lambda, k) = \begin{cases} 1,0 & \text{при } \frac{P(\lambda, k)}{P_{\min}(\lambda - 1, k)} > \delta, \\ 0,0 & \text{при } else, \end{cases} \quad (2.4)$$

Де $S_{pp}(\lambda, k)$ - імовірність наявності мови в поточній смузі з номером k ;

δ - порогове значення апостеріорного ОСШ для розділення наявності мови та її відсутності. МСРА використовує фіксоване значення δ , тоді як МСРА2 пропонує використовувати різні значення δ для різних смуг: малі

δ - для смуг з високою ймовірністю наявності мови (0-3000 Гц) і великі для решти.

3 Дослідження якості подавлення шуму методом експертних оцінок

В даному розділі проводиться оцінка якості шумоподавлення мовних сигналів від стаціонарних і квазістаціонарних шумів досліджуваним алгоритмом методом експертних оцінок. Проводяться вимірювання розбірливості мови методом артикуляційних вимірювань.

					ЕЛІТ 6.171.00.10.324 ПЗ	Арк
						26
		№ докум.	Підпис	БР		

3.1 Методика досліджень

Для визначення ефективності шумоподавлення в мовних сигналах, необхідно провести ряд експертних досліджень, за підсумками яких виноситься оцінка якості та ефективності розробленого алгоритму і його програмної реалізації.

Для дослідження використані методики, рекомендовані стандартом «Передача мови по трактах зв'язку. Методи оцінки якості, розбірливості та впізнаваності». Стандарти регламентують отримання комплексної оцінки якості передачі мови, заснованої на методах вимірювання показників розбірливості та якості мови. Чіткість голосу можна визначити через відносна кількість (у відсотках) правильно прийнятих елементів (складів, слів, фраз) спеціальних артикуляційних таблиць.

Для проведення дослідження використовувався лабораторний стенд, структурна схема якого показана на рис. 3.1.

У ПЕОМ проводиться запис вихідного мовного сигналу, генерація завади, зашумлення вихідного сигналу, шумозаглушення. Звукова карта відтворює всі записані звукові файли.

Для експериментального дослідження параметра розбірливості мови в магістерській роботі були проведені порівняльні артикуляційні випробування на артикуляційних таблицях складів російської мови [5] для наступних маскуючих адитивних завад:

					ЕліТ 6.171.00.10.324 ПЗ	Арк
						27
		№ докум.	Підпис	БР		

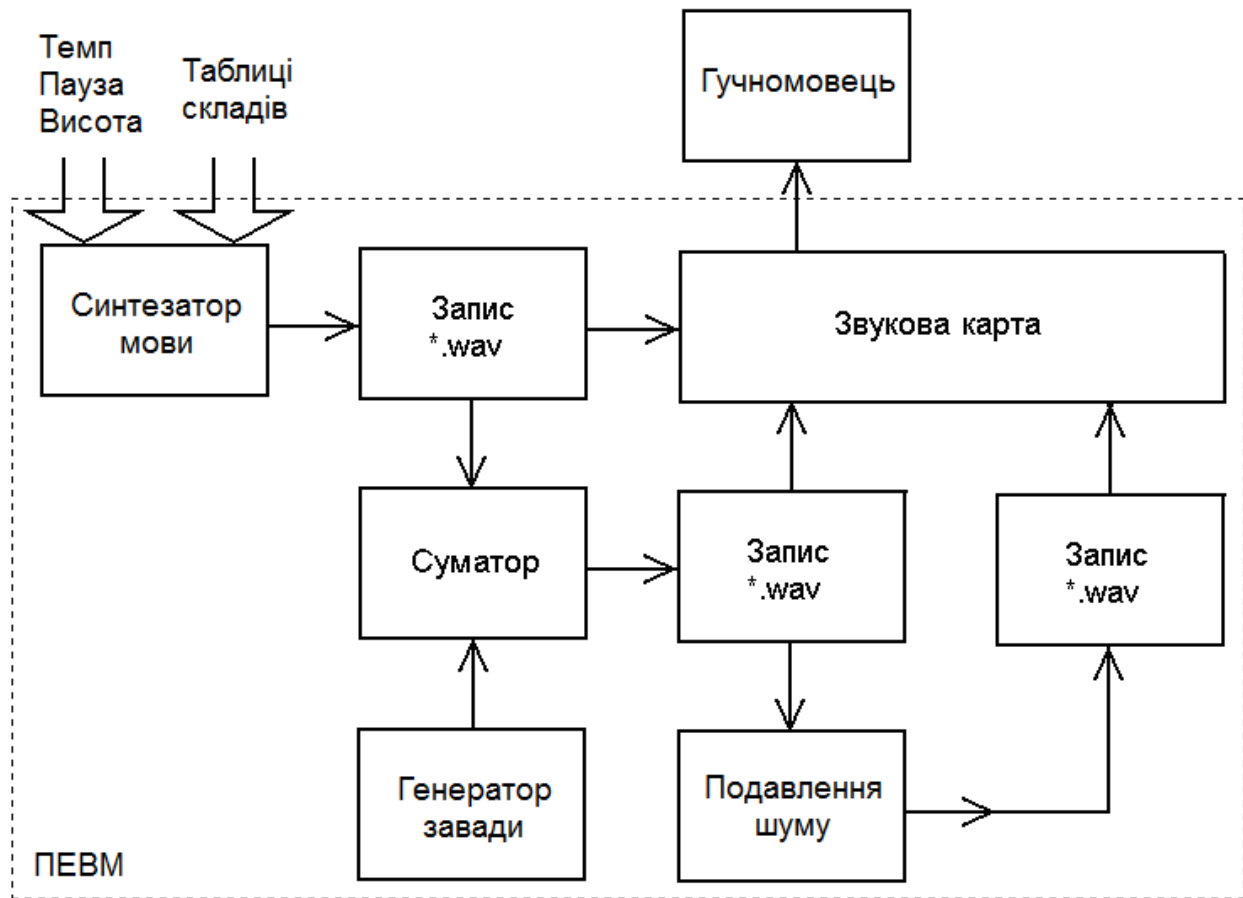


Рисунок 3.1 – Структурна схема лабораторного стенда

- тональна завада 50 Гц,
- тональна завада 100 Гц,
- тональна завада 2 кГц,
- рожевий шум,
- білий шум.

Вибрані шуми є квазістаціонарними (шум натовпу, моря, верстатів, двигунів, ефіру – при передачі по каналах зв'язку), а тому представляють найбільший практичний інтерес.

Розглянемо методику досліджень.

1. Артикуляційні випробування проводилися на оцифрованої базі стандартних артикуляційних таблиць складів зі стандартом. На відміну від стандартної методики,

згідно з якою склади повинні диктуватися професійним диктором, в даному експерименті використано програмний синтезатор мови IVONA Voices 1.6.70.

Застосований підхід має ряд переваг. По-перше, знайти у відкритому доступі звукозаписи артикуляційних таблиць не виявилось можливим.

По-друге, залучення професійних дикторів матеріально затратно, особливо для досліджень в рамках магістерської атестаційної роботи.

По-третє, дикторами спочатку відпрацьовується техніка читання для досягнення зачитування складів рівним голосом, чітко, але без підкреслення з постійним рівнем мови, з регламентованими паузами між словами. При використанні синтезатора мови всі ці умови легко досяжні.

Крім того, компанія IVONA Software є абсолютним лідером в області мовних технологій. Застосований синтезатор IVONA за якістю синтезованої мови значно перевершує всі аналоги, оскільки над створенням голосів IVONA працювали професійні диктори.

2. Мовні дані і сигнали завад перетворювалися в цифрову форму з параметрами: частота дискретизації 22050 Гц, глибина квантування 16 біт. У смузі частот оцифрованого сигналу присутні перші п'ять формант і антиформанта мови. Смуга частот сигналів завади відповідала п'яти октавам (не менше 250-8000 Гц).

Використано синтезовані звукозапису дикторів (Д1, Д2 і Д3), серед яких один жіночий і два чоловічих голоси. Кожним диктором зачитується по 5 таблиць складів. Склади начитуються в ритмі: 1 склад в $(3 \pm 0,3)$ с, протягом всієї таблиці. Синтезовані таблиці складів зберігаються в пам'яті ПЕОМ в форматі * .wav. В результаті отримуємо 15 записаних файлів незашумлених мовних сигналів (МС).

3. Для кожного з сигналів розраховувалися нормовані значення спектральної щільності потужності, які, для сигналів завади і мовних сигналів підсумовувалися програмно, з урахуванням відношення сигнал-завада, розрахованого за загальноприйнятою методикою. При розрахунку виключалося вплив

					<i>ЕлІТ 6.171.00.10.324 ПЗ</i>	Арк
						29
		№ докум.	Підпис	БР		

псофометричного фільтра, електроакустичних перетворювачів і акустики приміщення.

4. При артикуляційних дослідженнях для сигналів таблиць складів було вибрано відношення сигнал-завада +10 дБ.

При наявності 5 таблиць, зачитаних 3 дикторами, і 5-ти видів завад в результаті отримуємо 75 зашумлених і 75 оброблених мовних сигналів.

5. В якості аудиторів працювали студенти Харківського національного університету радіоелектроніки (3 людини – А1, А2 і А3), один хлопець і дві дівчини у віці 20-23 роки, які не мають дефектів слуху. Бригада аудиторів не проходила тренування на прослуховування зашумлених таблиць.

Час роботи бригади становить не більше 4 год за один день. Після прийому 5 таблиць (одного диктора) робиться перерва 5-10 хв.

З огляду на роботу в акустичних шумах, аудитори приступають до вимірювань через 5-10 хв. після перебування в умовах шуму

6. При аналізі результатів аудиту маскованих мовних даних відсоток розбірливості розраховувався за методикою, рекомендованою в стандарті. При розрахунку відсотка розбірливості маскованих фонем враховувалися фонем, які були правильно розпізнані в правильно і неправильно розпізнаних словах.

Дослідження проводилися в обсязі п'яти таблиць з 31 складу, які збалансовані для розмовної слов'янської мови. Загальна кількість фонем, проаналізованих в експерименті $3 \cdot 31 \cdot 5 = 465$.

За результатами вимірювань, в рамках досліджень, виноситься оцінка якості виробленої очищення мовного сигналу реалізованим алгоритмом.

3.2 Оцінка якості шумоподавлення методом артикуляційного вимірювання розбірливості

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	Арк
						30
		№ докум.	Підпис	БР		

3.3 Загальні показники розбірливості

На першому етапі аудитори А1, А2 і А3 прослуховують зашумлені мовні сигнали МС і записують їх у бланк. В разі, коли аудитор не зрозумів переданого складу, він ставить прочерк у відповідній пронумерованій рядку в бланку прийнятих складів.

На другому етапі вимірювань розбірливості артикуляційних методом проводиться цикл вимірювань.

Цикл вимірювань включає в себе результат прийому всіх аудиторів від всіх дикторів по 5*k* таблиць (*k* = 1,2,3) мовних сигналів МС2 с завадами і МС3 – оброблених розробленим алгоритмом шумоподавлення.

1. Для кожного вимірювання обчислюється середнє значення розбірливості (*S*) за формулою:

$$S = \frac{1}{N} \sum_{i=1}^N Si, \quad (3.1.)$$

де *Si* – результат одиничного вимірювання, % (диктор-таблиця-аудитор),

N – число одиничних вимірювань.

В рамках проведених досліджень: *Si* – це відсоток вірно розпізнаних складів в одній таблиці від одного аудитора, що прослухав одного диктора, а *N* – одне одиничне вимірювання тобто одна таблиця, тому отримуємо п'ять одиничних вимірювань (для МС3).

3.2.1 Обчислення середньої розбірливості для оброблених мовних сигналів (МС3)

Вимірювання розбірливості А1-Д1 в умовах п'яти типів завод:

Тональна завада 50 Гц А1-Д1

1) Обчислимо середнє значення розбірливості (3.1.) За формулою (3.2.)

					ЕЛІТ 6.171.00.10.324 ПЗ	Арк
						31
		№ докум.	Підпис	БР		

$$S = \frac{1}{5} \sum_{i=1}^5 (74 + 78 + 68 + 74 + 70) = \frac{1}{5} \cdot 364 = 72,8 \%. \quad (3.2)$$

2) Обчислимо середньоквадратичне відхилення (СКВ) за формулою:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (S_i - S)^2}{N - 1}}. \quad (3.3)$$

З виразу (3.3)

$$\sigma = \sqrt{\frac{(74 - 72,8)^2 + (78 - 72,8)^2 + (68 - 72,8)^2 + (74 - 72,8)^2 + (70 - 72,8)^2}{5 - 1}} = 3,9.$$

3) Знаходимо одиночні вимірювання, для яких $|S_i - S| > 2$, з метою виключення сумнівних значень і обчислення нового середнього значення:

а) $|74 - 72,8| \dots 2 \cdot 4,89$, т.ч. $1,2 < 7,8$;

б) $|78 - 72,8| \dots 2 \cdot 4,89$, т.ч. $5,2 < 7,8$;

в) $|68 - 72,8| \dots 2 \cdot 4,89$, т.ч. $4,8 < 7,8$;

г) $|74 - 72,8| \dots 2 \cdot 4,89$, т.ч. $1,2 < 7,8$;

д) $|70 - 72,8| \dots 2 \cdot 4,89$, т.ч. $2,8 < 7,8$.

Сумнівних значень S_i не виявлено.

Тональна завада 100 Гц А1-Д1

1) Обчислимо середнє значення розбірливості за формулою (3.1):

$$S = \frac{1}{5} \sum_{i=1}^5 (70 + 68 + 68 + 74 + 72) = \frac{1}{5} \cdot 352 = 70,4 \%. \quad (3.4)$$

2) Обчислимо середньоквадратичне відхилення (СКО) по (3.3)

$$\sigma = \sqrt{\frac{(70 - 70,4)^2 + (68 - 70,4)^2 + (68 - 70,4)^2 + (74 - 70,4)^2 + (72 - 70,4)^2}{5 - 1}} = 3,03. \quad (3.5)$$

3) Знаходимо одиничні вимірювання, для яких $|S_i - S| > 2\sigma$:

					ЕЛІТ 6.171.00.10.324 ПЗ	Арк
		№ докум.	Підпис	БР		32

- а) $0,4 < 3,03$;
- б) $2,4 < 3,03$;
- в) $2,4 < 3,03$;
- г) $4,4 < 3,03$;
- д) $2,4 < 3,03$.

Сумнівних значень S_i не виявлено.

Тональна завада 2000 Гц А1-Д1

1) Обчислимо середнє значення розбірливості (S) за формулою (3.1)

$$S = \frac{1}{5} \sum_{i=1}^5 (66 + 72 + 64 + 76 + 72) = \frac{1}{5} \cdot 350 = 70 \%. \quad (3.6)$$

2) Обчислимо середньоквадратичне відхилення (СКВ) по (3.3)

$$\sigma = \sqrt{\frac{(66-70)^2 + (72-70)^2 + (64-70)^2 + (76-70)^2 + (72-70)^2}{5-1}} = 4,89. \quad (3.7)$$

3) Знаходимо одиничні вимірювання, для яких $|S_i - S| > 2$, з метою виключення сумнівних значень і обчислення нового середнього значення:

- а) $|66-70| \dots 2 \cdot 4,89$, т.ч. $4 < 9,78$;
- б) $|72-70| \dots 2 \cdot 4,89$, т.ч. $2 < 9,78$;
- в) $|64-70| \dots 2 \cdot 4,89$, т.ч. $6 < 9,78$;
- г) $|76-70| \dots 2 \cdot 4,89$, т.ч. $6 < 9,78$;
- д) $|72-70| \dots 2 \cdot 4,89$, т.ч. $2 < 9,78$.

Сумнівних значень S_i не виявлено.

Рожевий шум А1-Д1

1) Обчислимо середнє значення розбірливості (S) по (3.1):

$$S = \frac{1}{5} \sum_{i=1}^5 (70 + 66 + 64 + 70 + 72) = \frac{1}{5} \cdot 342 = 68,4 \%. \quad (3.8)$$

2) Обчислимо СКВ за формулою (3.3):

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	Арк
		№ докум.	Підпис	БР		33

$$\sigma = \sqrt{\frac{(70 - 68,4)^2 + (66 - 68,4)^2 + (64 - 68,4)^2 + (70 - 68,4)^2 + (72 - 68,4)^2}{5 - 1}} = 3,3. \quad (3.9)$$

3) Знаходимо одиничні вимірювання, для яких $|Si - S| > 2\sigma$:

а) $1,6 < 6,6$;

б) $2,4 < 6,6$;

в) $4,4 < 6,6$;

г) $1,6 < 6,6$;

д) $3,6 < 6,6$.

Сумнівних значень Si не виявлено.

Білий шум А1-Д1

1) Обчислимо середнє значення розбірливості (S) по (3.1):

$$S = \frac{1}{5} \sum_{i=1}^5 (64 + 66 + 60 + 68 + 62) = \frac{1}{5} \cdot 320 = 64 \% . \quad (3.10)$$

2) Обчислимо СКВ по формулі (3.3)

$$\sigma = \sqrt{\frac{(64 - 64)^2 + (66 - 64)^2 + (60 - 64)^2 + (68 - 64)^2 + (62 - 64)^2}{5 - 1}} = 3,16. \quad (3.11)$$

3) Знаходимо поодинокі вимірювання, для яких $|Si - S| > 2$, для виключення сумнівних значень і обчислення нового середнього значення:

а) $0 < 6,3$;

б) $2 < 6,3$;

в) $4 < 6,3$;

г) $4 < 6,3$;

д) $2 < 6,3$.

Сумнівних значень Si не виявлено.

Після себе прийому усіма аудиторами від всіх дикторів по 5к таблиць МСЗ, і обчислень середньої розбірливості був отриманий результат, показаний в табл. 3.1.

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	Арк
		№ докум.	Підпис	БР		34

Таблиця 3.1 – Результат прослуховування артикуляційних таблиць (відсоток вірно розпізнаних складів в очищених записах мови МС3)

Тональна завада 50 Гц				Тональна завада 100 Гц			
	A1	A2	A3		A1	A2	A3
Д1	72,8	76,8	74,0	Д1	70,0	72,4	68,0
Д2	74,4	77,2	75,2	Д2	70,8	73,6	71,6
Д3	73,2	80,8	77,6	Д3	72,4	73,2	71,6
Тональна завада 2000 Гц				Рожевий шум			
	A1	A2	A3		A1	A2	A3
Д1	70,0	73,6	68,4	Д1	68,4	72,0	67,2
Д2	72,0	73,2	71,36	Д2	72,0	72,8	70,8
Д3	72,4	73,2	67,6	Д3	69,2	73,6	64,4
Білий шум							
	A1	A2	A3				
Д1	64,0	66,0	64,4				
Д2	67,2	69,6	66,0				
Д3	67,6	70,8	66,4				

3.2.2 Обчислення середньої розбірливості для зашумлених мовних сигналів (МС2)

Вимірювання розбірливості А1 - Д1 в умовах п'яти типів завад.

Тональна завада 50 Гц А1-Д1

1) Обчислимо середнє значення розбірливості (S) по (3.1):

$$S = \frac{1}{5} \sum_{i=1}^5 (60 + 56 + 57 + 60 + 58) = \frac{1}{5} \cdot 290 = 58 \%. \quad (3.12)$$

2) Обчислимо середньоквадратичне відхилення (СКВ) по (3.3):

$$\sigma = \sqrt{\frac{(60-58)^2 + (56-58)^2 + (56-58)^2 + (60-58)^2 + (58-58)^2}{5-1}} = 2. \quad (3.13)$$

3) Знаходимо одиничні вимірювання, для яких $|S_i - S| > 2$, з метою виключення сумнівних значень і обчислення нового середнього значення:

а) $|60-58| \dots 2 \cdot 2$, т.ч. $2 < 4$;

б) $|56-58| \dots 2 \cdot 2$, т.ч. $2 < 4$;

в) $|56-58| \dots 2 \cdot 2$, т.ч. $2 < 4$;

г) $|60-58| \dots 2 \cdot 2$, т.ч. $2 < 4$;

д) $|58-58| \dots 2 \cdot 2$, т.ч. $2 < 4$.

Сумнівних значень S_i не виявлено.

Тональна завада 100 Гц А1-Д1

1) Обчислимо середнє значення розбірливості (S) по (3.1):

$$S = \frac{1}{5} \sum_{i=1}^5 (54 + 56 + 54 + 56 + 54) = \frac{1}{5} \cdot 274 = 54,8 \%. \quad (3.14)$$

2) Обчислимо СКВ за формулою:

$$\sigma = \sqrt{\frac{(54-54,8)^2 + (56-54,8)^2 + (54-54,8)^2 + (56-54,8)^2 + (54-54,8)^2}{5-1}} = 1,1. \quad (3.15)$$

3) Знаходимо одиничні вимірювання, для яких $|S_i - S| > 2$:

а) $|54-54,8| \dots 2 \cdot 1,1$, т.ч. $0,8 < 2,2$;

б) $|56-54,8| \dots 2 \cdot 1,1$, т.ч. $1,2 < 2,2$;

в) $|54-54,8| \dots 2 \cdot 1,1$, т.ч. $0,8 < 2,2$;

г) $|56-54,8| \dots 2 \cdot 1,1$, т.ч. $1,2 < 2,2$;

д) $|54-54,8| \dots 2 \cdot 1,1$, т.ч. $0,8 < 2,2$.

Сумнівних значень S_i не виявлено.

Тональна завада 2000 Гц А1-Д1

1) Обчислимо середнє значення розбірливості (S) по (3.1):

					ЕЛІТ 6.171.00.10.324 ПЗ	Арк
		№ докум.	Підпис	БР		36

$$S = \frac{1}{5} \sum_{i=1}^5 (56 + 54 + 58 + 56 + 58) = \frac{1}{5} \cdot 282 = 56,4 \%. \quad (3.16)$$

2) Обчислимо СКВ по формулі:

$$\sigma = \sqrt{\frac{(56-56,4)^2 + (54-56,4)^2 + (58-56,4)^2 + (56-56,4)^2 + (58-56,4)^2}{5-1}} = 1,67. \quad (3.17)$$

3) Знаходимо одиничні вимірювання, для яких $|S_i - S| > 2$:

- а) $|56-56,4| \dots 2 \cdot 1,67$, т.ч. $0,4 < 1,67$;
- б) $|54-56,4| \dots 2 \cdot 1,67$, т.ч. $2,4 < 1,67$;
- в) $|58-56,4| \dots 2 \cdot 1,67$, т.ч. $1,6 < 1,67$;
- г) $|56-56,4| \dots 2 \cdot 1,67$, т.ч. $0,4 < 1,67$;
- д) $|58-56,4| \dots 2 \cdot 1,67$, т.ч. $1,6 < 1,67$.

Сумнівних значень S_i не виявлено.

Рожевий шум А1-Д1

1) Обчислимо середнє значення розбірливості (S) по (3.1):

$$S = \frac{1}{5} \sum_{i=1}^5 (50 + 54 + 48 + 52 + 56) = \frac{1}{5} \cdot 260 = 52 \%. \quad (3.18)$$

2) Обчислимо СКВ за формулою (3.3):

$$\sigma = \sqrt{\frac{(50-52)^2 + (54-52)^2 + (58-52)^2 + (52-52)^2 + (56-52)^2}{5-1}} = 3,16. \quad (3.19)$$

3) Знаходимо одиничні вимірювання, для яких $|S_i - S| > 2$, з метою виключення сумнівних значень і обчислення нового середнього значення:

- а) $|50-52| \dots 2 \cdot 3,16$, т.ч. $2 < 6,32$;
- б) $|54-52| \dots 2 \cdot 3,16$, т.ч. $2 < 6,32$;
- в) $|58-52| \dots 2 \cdot 3,16$, т.ч. $4 < 6,32$;
- г) $|52-52| \dots 2 \cdot 3,16$, т.ч. $0 < 6,32$;
- д) $|56-52| \dots 2 \cdot 3,16$, т.ч. $4 < 6,32$.

Сумнівних значень S_i не виявлено.

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	Арк
		№ докум.	Підпис	БР		37

Білий шум А1-Д1

1) Обчислимо середнє значення розбірливості S по (3.1):

$$S = \frac{1}{5} \sum_{i=1}^5 (52 + 54 + 56 + 54 + 56) = \frac{1}{5} \cdot 272 = 54,4 \% \quad (3.20)$$

2) Обчислимо СКВ за формулою (3.3):

$$\sigma = \sqrt{\frac{(52 - 54,4)^2 + (54 - 54,4)^2 + (56 - 54,4)^2 + (54 - 54,4)^2 + (56 - 54,4)^2}{5 - 1}} = 1,67. \quad (3.21)$$

3) Знаходимо одиничні вимірювання, для яких $|S_i - S| > 2$:

а) $|52 - 54,4| \dots 2 \cdot 1,67$, т.ч. $2,4 < 3,34$;

б) $|54 - 54,4| \dots 2 \cdot 1,67$, т.ч. $0,4 < 3,34$;

в) $|56 - 54,4| \dots 2 \cdot 1,67$, т.ч. $1,6 < 3,34$;

г) $|54 - 54,4| \dots 2 \cdot 1,67$, т.ч. $0,4 < 3,34$;

д) $|56 - 54,4| \dots 2 \cdot 1,67$, т.ч. $1,6 < 3,34$.

Сумнівних значень S_i не виявлено.

Після прийому усіма аудиторами від всіх дикторів по 5к таблиць МС2, і обчислень середньої розбірливості був отриманий результат, показаний в табл. 3.2.

Таблиця 3.2 – Результат прослуховування артикуляційних таблиць (відсоток вірно розпізнаних складів в зашумлених записах мови МС2)

Тональна завада 50 Гц				Тональна завада 100 Гц			
	A1	A2	A2		A1	A2	A3
Д1	58,0	62,8	58,8	Д1	54,8	52,8	55,2
Д2	59,2	54,4	56,4	Д2	51,2	50,0	52,4
Д3	56,8	57,2	54,8	Д3	53,2	55,2	54,8

Продовження табл. 3.2

Тональна завада 2000 Гц			Рожевий шум				
	A1	A2	A3		A1	A2	A3
Д1	56,4	54,4	55,2	Д1	52,0	54,0	54,4
Д2	54,8	56,0	55,2	Д2	52,8	54,4	53,6
Д3	52,4	54,4	56,2	Д3	54,8	52,4	54,0
Білий шум							
	A1	A2	A3				
Д1	54,4	54,0	53,6				
Д2	55,6	53,2	51,2				
Д3	52,8	51,6	50,4				

3.2.3 Загальні показники розбірливості

Усереднюючи отримані результати, наведені в табл. 3.1 і табл. 3.2, отримаємо середній показник розбірливості за двома видами – МС3 і МС2. Результати усереднення дані в табл. 3.3 і на рис. 3.2.

Таблиця 3.3 – Середній показник розбірливості, %

Завади	МС3	МС2	МС3 - МС2
50 Гц	75,77	57,6	18,17
100 Гц	71,5	53,3	18,2
2000 Гц	71,3	55,0	16,3
Рожевий шум	70,0	53,6	16,4
Білий шум	68,9	52,9	16,0
Усього			17,014

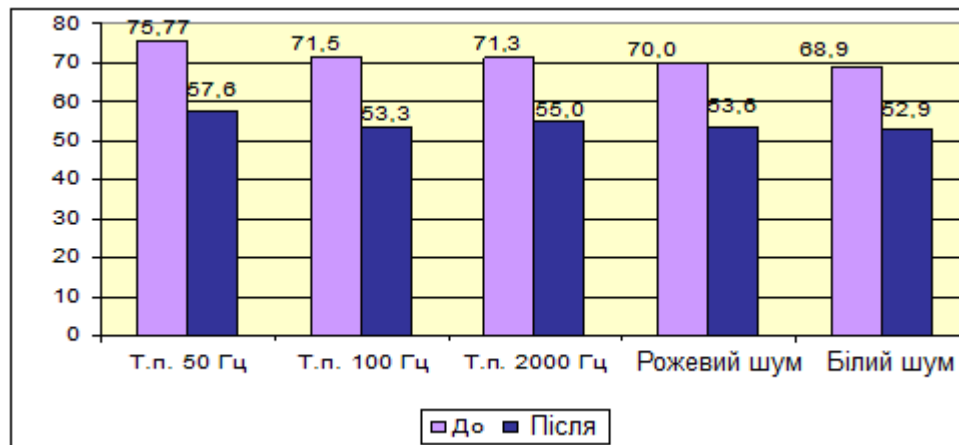


Рисунок 3.2 – Середні показники розбірливості до і після очищення

З рис.3.2 і табл. 3.3 випливає, що результат очищення мовних сигналів дослідженим алгоритмом шумоподавлення підвищив середнє значення розбірливості мови на 17%, з них:

- для тональної завади 50 Гц – розбірливість підвищилася на 18,2%;
- для тональної завади 100 Гц – розбірливість підвищилася на 18,2%;
- для тональної завади 200 Гц – розбірливість підвищилася на 16,3%;
- для рожевого шуму – розбірливість підвищилася на 16,4%;
- для білого шуму – розбірливість підвищилася на 16,0%.

Результати досліджень свідчать про ефективність розробленого алгоритму.

За результатами в табл. 3.3, також можна судити і про ефективність шумоподавлення по відношенню до різних типів завад.

Так, наприклад, результат розбірливості при тональній заваді 100 Гц, підвищився на 18,2% після очищення, що є найвищим результатом, ніж результат розбірливості МС після очищення в умовах білого шуму який склав 16,8% (рис.3.3).

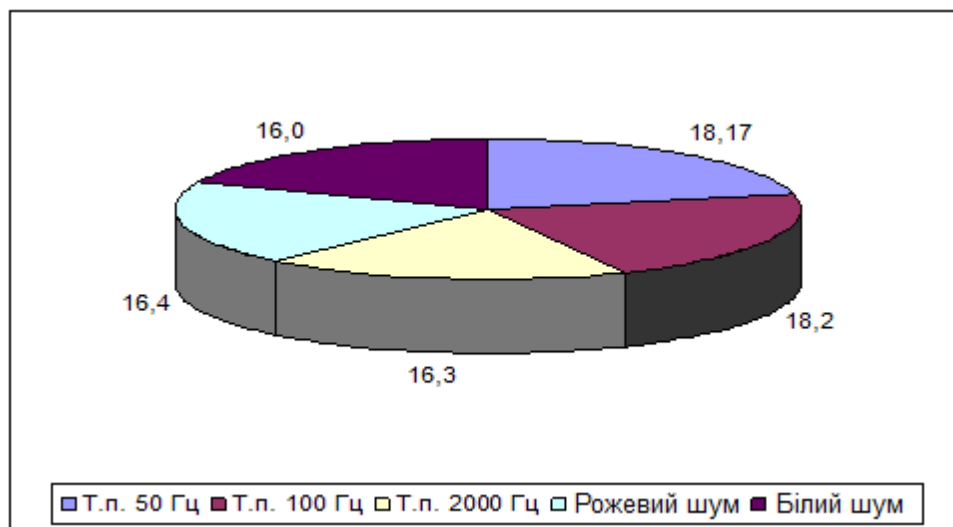


Рисунок 3.3 – Результат ефективності шумоподавлення для різних типів завад

Таким чином, розроблений алгоритм шумозаглушення за результатами проведених досліджень більш ефективний в порядку зменшення:

- тональна завада 100 Гц;
- тональна завада 50Гц;
- рожевий шум;
- тональна завада 2000 Гц;
- білий шум.

На підставі наведених результатів, можна судити про досить високу величину очищених мовних сигналів, перед зашумленими, що свідчить про хорошу якість і ефективності системи очищення за допомогою досліджуваного алгоритму очищення мовних сигналів.

ВИСНОВКИ

Зменшення шуму в звукових сигналах є дуже актуальним завданням. В роботі розглядається мовний сигнал, спотворений адитивною стаціонарною завадою. Завада може бути викликана акустичним шумом під час запису, шумом в каналі зв'язку або згенерована цілеспрямовано активною системою захисту мовної інформації.

В роботі вирішені такі завдання:

1. У більшості сучасних цифрових звукових систем використовуються стандартні частоти дискретизації 44.1 і 48 кГц, проте частотний діапазон сигналу зазвичай обмежується біля 20 кГц для залишення запасу щодо теоретичної межі, по-перше, для отримання високої якості потрібне високоякісне оцифрування аналогового звуку, яке головним чином залежить від якості АЦП - аналого-цифрового перетворювача. Висококласний мікрофон або дорогі з'єднувальні кабелі не допоможуть у ситуації, коли якість роботи АЦП залишає бажати кращого. Запис відліків з недостатньою точністю, створення вибірок з нерівномірною частотою тощо призведуть до отримання звуку, далекого за якістю від оригіналу, і виправити це вже не вдасться потім нічим. А по-друге, оцифрований звук потрібно ще й якісно відтворити, що можливо тільки за наявності якісного ЦАП - цифро-аналогового перетворювача. Через нерівномірну частоту дискретизації, недостатню точність або відсутність інтерполяції звук зіпсується так, що ніяка сучасна акустична система цього не компенсує. Таким чином, можна зрозуміти що, на якість цифрового запису і відтворення головним чином впливає якість перетворювачів.

2. Проведено аналіз існуючих алгоритмів шумоочищення для покращення розбірливості мовних сигналів та встановлено, що на сьогоднішній день найактивніше використовуються дві реалізації алгоритмів шумоочищення: Speex (SpeexDSP) і WebRTC. Для оцінки спектра шуму в Speex застосовується імовірнісний алгоритм Кохена - Бердуго MCRA (Minima-Controlled Recursive Averaging), а для

					ЕЛІТ 6.171.00.10.324 ПЗ	Арк
						42
		№ докум.	Підпис	БР		

безпосереднього придушення отриманого спектра шуму - алгоритм Ефраїма - Малаха. У WebRTC для детектування шуму здійснюється квантильне оцінювання, а для шумопридушення - класична вінерівська фільтрація.

Дослідження методів безпосереднього шумоочищення і методів детектування шуму показали, що існує низка алгоритмів, які працюють не гірше, а нерідко й ефективніше за ті, які використовуються в Speex і WebRTC. Так, для детектування шуму може застосовуватися подальший розвиток MCRA - розробка Рангачарі-Лойзю MCRA2, а для безпосереднього шумопоглинання - одна із сучасних модифікацій вінерівської фільтрації.

3. Встановлено причини передачі інформації разом з фоновим звуком та шумами, та виявлено, що недостатня чіткість може бути з низки причин:

- фонетична неопрацьованість мови, перенасиченість звукової партитури шумами, музикою з маскувальним ефектом;
- голосові особливості диктора, актора (дикція, скоромовка, акцент);
- невдалий запис і/або монтаж (робота звукорежисера)
- низькі якісні параметри тракту звукопередачі (приміщення, зайвий шум на природі, нелінійні спотворення).

4. Зроблено огляд і класифікація найбільш поширених методів подавлення шуму в мовних сигналах. Для реалізації і подальших досліджень обрані методи, засновані на відніманні амплітудних спектрів, які вважаються одними з кращих - вони дають хороші результати і не вимагають великих обчислювальних ресурсів.

5. Розроблено покроковий алгоритм спектрального віднімання, що стосується шумоподавлення в мовних сигналах. Він складається з наступних етапів: розкладання сигналу за допомогою короткочасного перетворення Фур'є (STFT); оцінка спектра шуму; «віднімання» амплітудного спектра шуму з амплітудного спектра сигналу; зворотне перетворення STFT і синтез результуючого сигналу.

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	Арк
						43
		№ докум.	Підпис	БР		

Сигнал розбивається на ряд інтервалів за допомогою ковзного вікна. Оскільки динамічний діапазон мовного сигналу не перевищує 40 дБ вирішено використовувати вікно Хеммінга, бічні пелюстки спектру якого не перевищують -42 дБ. Розмір вікна для забезпечення частотного розрізнення 12 Гц при частоті дискретизації 22,050 кГц повинен складати 8192 відліки. Оцінка спектра шуму здійснюється вручну на часовому сегменті, де присутній тільки шум.

6. Для кількісної оцінки ефективності розробленого алгоритму використаний метод експертних оцінок. Використана методика прослуховування артикуляційних таблиць, рекомендована стандартом. Для дослідження використовувався лабораторний стенд на основі персонального комп'ютера. У ПЕОМ проводився запис мовного сигналу, генерація завади, зашумлення вихідного сигналу, шумоподавлення. Звукова карта відтворює всі записані звукові файли. У середовищі MATLAB згенеровані такі види маскуючих адитивних завад: тональна завада 50 Гц, 100 Гц і 2 кГц, рожевий і білий шум.

На підставі наведених результатів, можна судити про досить високу розбірливість очищених мовних сигналів, у порівнянні із зашумленими, що свідчить про хорошу якість і ефективність виробленого шумоподавлення за допомогою дослідженого алгоритму.

7. Також в роботі запропоновані універсальні адаптивні методи шумоочищення звуку базуються на найефективніших сучасних алгоритмах детектування та придушення шуму. Розроблені алгоритми можуть працювати без використання спеціалізованих апаратних засобів та допоміжної інформації у режимі реального часу. Усі алгоритми реалізовані мовою С без застосування зовнішніх залежностей.

Тестування за допомогою метрик сегментної ОСШ та PESQ показало їх високу ефективність та перевага над поширеними реалізаціями шумоочищення Speex і WebRTC за якістю шумоочищення та порівнянню з ними швидкість роботи.

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	Арк
						44
		№ докум.	Підпис	БР		

У висновку можна додати, що багато серйозних студій звукозапису, як і раніше, використовують аналогову апаратуру для запису звукових доріжок на магнітну стрічку. Справа в тому, що цифровий звукозапис, за своєю природою, неминуче втрачає якусь частину оригінального звучання. Оскільки цифровий файл є послідовністю зафіксованих в окремі моменти часу показників, він завжди пропускає інформацію про звук, який мав місце між двома зафіксованими станами. Багато гурманів якісного відео та музики не можуть змиритися з таким станом речей. Досягнута сьогодні частота дискретизації дозволяє настільки часто фіксувати окремі кадри та зміни звуку, що більшість людей вже не можуть відрізнити цифровий запис від аналогового — за своєю якістю вони вже практично зрівнялися. Однак цифрова технологія має одну незаперечну перевагу, яка дозволяє з упевненістю говорити, що майбутнє належить саме їй. Після того, як ви зробили початковий цифровий запис, ви можете копіювати його знову і знову, і кожна нова копія нічим не відрізнятиметься від оригіналу. Цифровий запис ніколи сам собою не змінюється.

					ЕліТ 6.171.00.10.324 ПЗ	Арк
		№ докум.	Підпис	БР		45

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Нікамін В.А. Цифровая звукозапись. Технологія і стандарти. – СПб: Наука і техніка, 2002. – 256 с.
2. Лукін А . Введення в цифрову обробку сигналів (Математичні основи). Методическое пособие, 2002, 44 с.
3. Квінт І. Sound Forge 9. – СПб.: Пітер, 2009. – 167 с.
4. Степаненко О.С. Adobe Audition 3.0 створення фонограми та обробка звуку. – М.: Діалектика, 2010. – 414 с.
5. Ковалгин Ю.А., Вологдин Э.И. Цифрове кодування звукових сигналів. – СПб,: Корона Принт., 2004. – 240 с.
6. Рабінер Л., Гоулд Б. Теорія і застосування цифрової обробки сигналів. – М.: Мир, 1978. – 848 с.
7. Лядов М . Лукин А. RightMark Audio Analyzer 5.5. Руководство пользователя. – IXBT.com, 2005. – 17 с.
8. Горюнов В.А., Стась А.Н. Розробка та монтаж аудіозаписів за допомогою Audacity (ПО для обробки та монтажу аудіозаписів): Навчальне посібіє. — Москва: 2008. — 40 с.
9. Цифровая звукозапись. Под редакцией Дж. Мааса і М. Веркамена. пер. з англ. – М.: Мир, 2004. – 352 с.
10. Самовчитель по музичним редакторам: WaveLab, Sound Forge, Cubase, Acid. – М.: ІДДК, 2007. – 212 с.
11. Лукин А . Подавление широкополосного шума: історія та нові розробки [Електронний ресурс] // Звукорежиссер. №10. – 2008. – <http://audioproducer.625-net.ru/files/601/523> (дата звернення 01.12.2021).
12. Кен Гендрі. Системы шумоподавления [Електронний ресурс] // Звукорежиссёр, №6, 2004. – <http://rus.625-net.ru/archive/z0604/noise.htm> (дата звернення 01.12.2021).
13. Ашихмін В.О. Дослідження системи корекції звуку з урахуванням характеристик приміщення // 23-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у ХХІ столітті». – Харків, 16–18 квітня 2019 р. – с. 96-97
14. Буйницький Д.В. Апаратно-програмний комплекс для дослідження акустичних пристроїв // 23-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у ХХІ столітті». – Харків, 16–18 квітня 2019 р. – с. 92-93.

					ЕЛІТ 6.171.00.10.324 ПЗ	Арк
						46
		№ докум.	Підпис	БР		

15. Тушев В.О. Дослідження впливу фазових проявів аудіоапаратури на якість звучання // 23-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у XXI столітті». – Харків, 16–18 квітня 2019 р. – с. 84-85.
16. Шевченко П. Д. Система вимірювання та аналізу акустичних характеристик приміщень: атестаційна робота на здобуття освітнього ступеня магістра. – Х.: ХНУРЕ. – 2019 р. – 68 с.
17. Методичні вказівки з виконання атестаційної магістерської роботи за спеціальністю 8.05090102 «Апаратура радіозв'язку, радіомовлення і телебачення». Освітньо-кваліфікаційний рівень – магістр / Упоряд. В.М. Карташов, В.А. Тихонов, І.В. Савченко – Харків: ХНУРЕ, 2012. – 68 с.
18. V. Kartashov, V. Oleynikov, I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin, I. Selieznov. Use of Acoustic Signature for Detection, Recognition and Direction Finding of Small Unmanned Aerial Vehicles. 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). 2020. 4 p.
19. В. Н. Олейников, О.В. Зубков, В.М. Карташов, И.В. Корытцев, С.И. Бабкин, С.А. Шейко, И.С. Селезнев. Экспериментальная оценка эффективности алгоритмов пеленгования беспилотных летательных аппаратов по акустическому излучению. Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2019. – Вып. 199. – С. 29 – 37.
20. V. Kartashov, V. Oleynikov, I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin. Processing of Wide Band Acoustic Signals During Detection of Unmanned Aerial Vehicles // 2020 IEEE Ukrainian Microwave Week (UkrMW). Kharkiv, Ukraine, September 21 - 25, 2020. Volume 1 on 2020 IEEE 12th International Conference on Antenna Theory and Techniques (ICATT). pp. 35-39.
21. Сідоров Г.І., Шейко С.О., Шаповалов С.В., Полонська А.С., Дмитренко А.І. Акустичний метод вимірювання турбулентного стану атмосферного прикордонного шару // Радиотехніка: Всеукр. міжвід. наук.- техн. зб. 2018. – Вип. 192. – С. 46–50.
22. Valerii V. Semenets, V. M. Kartashov, V. I. Leonidov. Registration of refraction phenomenon in the problem of acoustic sounding of atmosphere in airports zone. Telecommunications and Radio Engineering. Volume 77, Issue 5, 2018. – P. 461-468.

					<i>ЕліТ 6.171.00.10.324 ПЗ</i>	Арк
		№ докум.	Підпис	БР		47