

## АНОТАЦІЯ

Небожук Р.Б. Метод виділення огинаючої складової мовних сигналів для ефективного їхнього кодування. – Рукопис. Кваліфікаційна робота магістра, Тернопільський національний технічний університет імені Івана Пулюя, Тернопіль, 2019.

Роботу присвячено питанням розроблення методу виділення огинаючої складової мовних сигналів для ефективного їхнього кодування. Показано, що для ефективної роботи вкодерів при кодуванні мовних сигналів важливим є виділення огинаючої складової часової реалізації мовного сигналу або його спектру. Проаналізовано методи виділення огинаючої складової та запропоновано використання з цією метою перетворення Гільберта. Проведено моделювання процесу виділення огинаючої складової в середовищі Matlab.

Ключові слова: голосовий сигнал, огинаюча складова, кодування.

## ABSTRACT

Nebozhuk RB A method of allocating the envelope component of speech signals to effectively encode them. – Manuscript. Master's qualifying work, Ternopil Ivan Puluj National Technical University, Ternopil, 2019.

The work is devoted to the development of the method of selection of the envelope component of speech signals for their effective coding. It is shown that for the efficient operation of encoders when encoding speech signals, it is important to select the envelope component of the temporal realization of the speech signal or its spectrum. The methods of selection of the envelope component are analyzed and the use of the Hilbert transform for this purpose is proposed. The process of selection of the envelope component in the Matlab medium is simulated.

Keywords: voice signal, envelope component, coding.

## ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ ТА ПОЗНАЧЕНЬ

АЧХ – амплітудно-частотна характеристика;

ГС – голосовий сигнал;

КЛП – коефіцієнти лінійного передбачення

ОТ – основний тон;

ЦСП – цифровий сигнальний процесор;

ФНЧ – фільтр низьких частот;

ЧОТ – частота основного тону.

## ЗМІСТ

ВСТУП.....	9
РОЗДІЛ 1. ЗАДАЧА КОДУВАННЯ ГОЛОСОВИХ СИГНАЛІВ.....	12
1.1 Кодування голосових сигналів.....	12
1.2 Типи вокодерів та особливості їхнього функціонування.....	13
1.3 Порівняння методів кодування голосових сигналів.....	18
1.4 Висновки до розділу 1.....	21
РОЗДІЛ 2. ПРИНЦИПИ РОБОТИ ВОКОДЕРІВ.....	22
2.1 Принцип роботи параметричних вокодерів.....	22
2.2 Смугові (канальні) вокодери.....	23
2.3 Ортогональні вокодери.....	27
2.4 Ліпредери (вокодери з лінійним передбаченням).....	36
2.5 Формантний вокодер.....	52
2.6 Задача виділення огинаючої голосових сигналів.....	64
2.7 Висновки до розділу 2.....	64
РОЗДІЛ 3. МАТЕМАТИЧНІ ОСНОВИ ВИДІЛЕННЯ ОГИНАЮЧОЇ СИГНАЛУ.....	66
3.1 Метод виділення огинаючої складової сигналів.....	66
3.2 Визначення перетворення Гільберта.....	68
3.3 Висновки до розділу 3.....	73
РОЗДІЛ 4. ВИДІЛЕННЯ ОГИНАЮЧОЇ ГОЛОСОВОГО СИГНАЛУ З ДОПОМОГОЮ ПЕРЕТВОРЕННЯ ГІЛЬБЕРТА.....	74
4.1 Програмна реалізація методу виділення огинаючої.....	74
4.2 Результати виділення огинаючої голосового сигналу.....	75
4.3 Висновки до розділу 4.....	79
РОЗДІЛ 5. СПЕЦІАЛЬНА ЧАСТИНА.....	80

5.1	Методика проведення медико-біологічних досліджень.....	80
5.2	Обґрунтування вибору УДК напряму наукового дослідження.....	84
РОЗДІЛ 6. ОБҐРУНТУВАННЯ ЕКОНОМІЧНОЇ ЕФЕКТИВНОСТІ.....		86
6.1	Науково-технічна актуальність науково-дослідної роботи.....	86
6.2	Розрахунок витрат на проведення науково-дослідної роботи.....	87
6.3	Науково-технічна ефективність науково-дослідної роботи.....	93
6.4	Висновки до розділу 6.....	97
РОЗДІЛ 7. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ.....		98
7.1	Охорона праці.....	98
7.2	Безпека в надзвичайних ситуаціях.....	100
РОЗДІЛ 8. ЕКОЛОГІЯ.....		110
8.1	Актуальність екологічних проблем.....	110
8.2	Вплив на довкілля при здійсненні технологічного процесу монтажу та складання приладу.....	111
8.3	Заходи по охороні навколишнього середовища.....	112
ЗАГАЛЬНІ ВИСНОВКИ.....		114
Бібліографія.....		116
ДОДАТКИ		

## ВСТУП

**Актуальність роботи.** Через значний розвиток технологій цифрового зв'язку, зокрема цифрових телефонних мереж, комп'ютерної телефонії, супутникових засобів зв'язку, актуальним стає питання ефективного стиснення мовних сигналів для збільшення швидкості обміну даними, оскільки застосування цифрових методів подання, обробки і передачі призводить до багаторазового збільшення займаної смуги частот і, як наслідок цього, до багаторазового збільшення швидкості передачі інформаційних повідомлень.

Для цієї задачі використовуються різного роду пристрої кодування мовних сигналів, які називаються вокодерами. Ці пристрої окрім кодування передаваного сигналу проводять і синтез мовного сигналу та приймаючій стороні каналу передачі даних. Початково вокодери розроблялись з метою економії частотних ресурсів радіолінії системи зв'язку при передачі мовних повідомлень. Економія досягалася за рахунок того, що замість власне мовного сигналу передавалися тільки значення його певних параметрів, які на приймальній стороні керували так званим синтезатором мови.

Розглянуто принципи роботи параметричних вокодерів, зокрема смугових (канальних), ортогональних, вокодерів з лінійним передбаченням, формантних вокодерів тощо та встановлено, що важливим для зниження обсягів переданих даних є виділення тих інформативних параметрів мовних сигналів, яких буде достатньо для відтворення вихідного мовного сигналу на приймаючій стороні. При цьому такими параметрами відповідно до акустичної теорії мовотворення є частота основного тону та параметри огинаючої складової мовного сигналу в часі або спектра такого сигналу (у випадку застосування формантних та фазових вокодерів). При цьому сам мовний сигнал розглядається як складне амплітудно модульоване коливання. Для наближення до природного звучання в структуру

таких сигналів можна додатково вводити шумові складові із заданими законами розподілу.

При цьому важливим є обґрунтування способу виділення огинаючої складової мовного сигналу.

**Мета і задачі дослідження.** *Метою дослідження є обґрунтування методу виділення огинаючої складової мовних сигналів для ефективного їхнього кодування. Досягнення цієї мети вимагає розв'язання таких задач:*

1. Провести аналітичний огляд літературних джерел за тематикою дослідження;
2. Провести аналіз принципів функціонування вокодерів із різними методами кодування мовних сигналів;
3. Проаналізувати відомі методи виділення огинаючої складової мовного сигналу;
4. Запропонувати метод виділення огинаючої складової мовного сигналу;
5. Провести експериментальну верифікацію запропонованого методу.

*Об'єкт дослідження:* процес виділення огинаючої складової мовних сигналів для ефективного їхнього кодування.

*Предмет дослідження:* метод виділення огинаючої складової мовних сигналів.

**Наукова новизна одержаних результатів.** Запропонований метод виділення огинаючої складової мовних сигналів шляхом використання перетворення Гільберта та наступної низькочастотної фільтрації дає можливість проводити ефективно виділення огинаючої складової для кодування голосових сигналів у сучасних вокодерах.

**Практичне значення одержаних результатів.** Одержані результати можуть бути використані при розробленні методів ефективного кодування мовних сигналів у вокодерах.

**Публікації.** Викладені в роботі результати доповідей і обговорювалися на VII науково-технічній конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології».



## РОЗДІЛ 1

### ЗАДАЧА КОДУВАННЯ ГОЛОСОВИХ СИГНАЛІВ

#### 1.1 Кодування голосових сигналів

Ефективно кодування голосових сигналів виконується в області стиснення та передачі голосових сигналів. В обох випадках функцію кодування виконують такі пристрої, як вокодери.

Ядром вокодера, - цифрової моделі мовного тракту, - є цифровий фільтр зі змінними параметрами, АЧХ якого змінюється в часі, відображаючи зміну резонансних властивостей мовного тракту. В аналізаторі вокодера вимірюються характеристики цього фільтра, а в синтезаторі фільтр відтворюється за результатами цих вимірів (рис.1.1). Крім того, в аналізаторі визначається, який звук має місце в даний момент часу - голосний або приголосний, а також вимірюється частота основного тону голосного звуку. У синтезаторі на основі цієї інформації формується збудливий сигнал: для приголосних звуків це шум, для голосних - періодична послідовність імпульсів з періодом, рівним періоду основного тону.

Вирішальну роль в розбірливості синтезованого мовного сигналу грає інформація про резонансні властивості мовного тракту. Так, збуджуючи фільтр синтезатора тільки шумом або тільки періодичною послідовністю імпульсів, можна отримати цілком розбірливий мова. Відмінність лише в тому, що в першому випадку ми почуємо шепіт, у другому - «голос робота».

## 1.2 Типи вокодерів та особливості їхнього функціонування

Вокодер (англ. voice coder – кодувальник голосу) - пристрій синтезу голосу на основі довільного сигналу з багатим спектром. Спочатку вокодери були розроблені з метою економії частотних ресурсів радіолінії системи зв'язку при передачі мовних повідомлень. Економія досягається за рахунок того, що замість власне голосового сигналу передаються тільки значення його певних параметрів, які на приймальній стороні керують синтезатором мови. Основу синтезатора мови складають три елементи:

- генератор тонального сигналу для формування голосних звуків;
- генератор шуму для формування приголосних;
- система формантних фільтрів для відтворення голосу абонента.

Вокодер являє собою пристрій, що здійснює параметричне компандування голосових сигналів. Компресія мовних сигналів на передавальному кінці каналу зв'язку проводиться в аналізаторі, що виділяє з голосового сигналу повільно мінливі складові, які передаються по каналу зв'язку у вигляді кодових посилок. На приймальному кінці за допомогою місцевих джерел сигналів, керованих прийнятими параметрами, синтезується голосовий сигнал.

Робота вокодерів заснована на моделюванні людської мови з урахуванням її характерних особливостей. Замість безпосереднього вимірювання амплітуди вокодер перетворює вхідний сигнал в якийсь інший, схожий на вихідний. Причому вимірювані характеристики голосового сигналу використовуються для підгонки параметрів в прийнятій моделі голосового сигналу. Саме ці параметри і передаються приймачу, який по ним відновлює вихідний голосовий сигнал. По суті, мова йде про синтез голосу. Природно, що вимірювання спотворень відношення сигнал/шум не є коректне для вокодерів, і, отже, необхідні інші суб'єктивні оцінки, такі, як середня експертна оцінка, діагностичний римований тест, діагностична оцінка прийнятності та ін.

Вокодери можна розділити на два класи: голосоелементні і параметричні.

У голосоелементних вокодерах при передачі розпізнаються виголошені елементи мови (наприклад, фонемні) і передаються тільки їх номери. На прийомі ці елементи створюються за правилами голосотворення або беруться з пам'яті пристрою. Область застосування фонемних вокодерів - лінії командного зв'язку, голосове управління і розмовляючі автомати інформаційно-довідкової служби. Практично в таких вокодерах відбувається автоматичне розпізнавання слухових образів, а не визначення параметрів мови.

У параметричних вокодерах з голосового сигналу виділяють два типи параметрів:

- параметри, що характеризують огинаючу спектра голосового сигналу, (фільтрову функцію);
- параметри, що характеризують джерело голосових коливань (генераторну функцію), - частоту основного тону, її зміну в часі, моменти появи і зникнення основного тону, шумового сигналу.

За цими параметрами на прийомі синтезують мова.

За принципом визначення параметрів фільтрової функції голосу розрізняють вокодер:

- смугові каналні (channel);
- формантні;
- ортогональні;
- ліпредери (з лінійним передбаченням голосу);
- гомоморфні.

У смугових вокодерах спектр голосу ділиться на 7-20 смуг (каналів) аналоговими або цифровими смуговими фільтрами. Більше число каналів в вокодері дає велику натуральність і розбірливість. З кожного смугового фільтра сигнал надходить на детектор і фільтр низьких частот з частотою зрізу  $F_{зр}$ . Таким

чином, сигнали на виході кожного каналу змінюються з частотою менше  $F_{зр}$ . Їх передача можлива в аналоговому або цифровому вигляді.

У формантних вокодерах обвідна спектра голосу описується комбінацією формант (резонансних частот голосового тракту). Основні параметри формант - центральна частота, амплітуда і ширина смуги частот.

У ортогональних вокодерах обвідна миттєвого спектру розкладається в ряд по вибраній системі ортогональних базисних функцій. Обчислені коефіцієнти цього розкладання передаються на приймальну сторону. Поширення отримали гармонійні вокодери, що використовують розкладання в ряд Фур'є.

Вокодери з лінійним передбаченням (LPC - Linear Prediction Coding) або ліпредери, засновані на оригінальному математичному апараті. Вони набули найбільшого поширення.

Гомоморфна обробка дозволяє розділити генераторну і фільтрову функції, що утворюють голосовий сигнал.

Через складність визначення параметрів генераторної функції з'явилися напіввокодери (VE - Voice Excited Vocoder), в яких замість сигналів основного тону і тон-шуму використовується смуга голосового сигналу. Смуга частот до 800 ... 1000 Гц кодується АДІКМ, АДМ (адаптивна дельта модуляція) або за допомогою лінійного передбачення малого порядку, а в деяких моделях передається в аналоговому вигляді. Відомі різні типи напіввокодерів-ліпредерів: VELP - Voice Excited Linear Prediction; RELP - Residual Excited Linear Prediction.

Вокодери VELP використовують голосове збудження і коефіцієнти лінійного передбачення (КЛП). У вокодерах RELP по вихідному сигналу також обчислюються КЛП. Так як КЛП описує фільтрову функцію, то сигнал помилки (залишку) передбачення містить інформацію про генераторну функцію мови і передається на прийомну сторону (можливим є його стиснення методами АДІКМ, АДМ або за допомогою лінійного передбачення малого порядку).

Якість голосу вокодерів є функцією швидкості передачі, продуктивності і затримки обробки. Якщо вокодери призначені для телефонії через Інтернет, розробники продукції повинні враховувати ці характеристики, між якими існує суворозалежність. Наприклад, низькошвидкісні вокодери зазвичай мають більшу затримку і більш низьку якість мовлення, ніж високошвидкісні.

Швидкість. Так як вокодер спільно використовує канал зв'язку або часто перевантажену мережу підприємства або Інтернет з іншими інформаційними потоками, максимальна швидкість повинна була б бути якомога нижче, особливо для додатків малих офісів. В даний час більшість вокодерів працюють на фіксованій швидкості незалежно від характеристик вхідного сигналу, проте метою сучасних розробок є вокодер зі змінною швидкістю. Для додатків при одночасній передачі мови і даних компромісом є створення алгоритмів стиснення пауз в якості частини стандарту кодування. Спільним рішенням є використання фіксованої швидкості для мови і низької швидкості для фонових шумів. Спосіб виконання механізму стиснення пауз важливий для підвищення якості передачі мови, проте часто вигаш від компресії пауз не реалізується. Проблемою є те, що при великих фонових шумах складно провести різницю між голосом і шумом. Інша проблема полягає в тому, що якщо механізм стиснення пауз неправильно виявив стан голосу, початок мовлення може бути «відрізаний», що значно погіршує розбірливість кодованої мови.

Спосіб генерації комфортного шуму повинен бути таким, щоб кодер і декодер залишалися синхронізованими, навіть якщо протягом деякого інтервалу часу передача даних не здійснюється. Це дозволяє згладити переходи між сегментами активної і неактивної мови.

Продуктивність алгоритму. Вокодери частот виконуються на основі цифрових сигнальних процесорів (ЦСП). Відповідно до комп'ютерної термінології їх продуктивність може бути виміряна в млн. операцій за секунду, обсязі пам'яті з довільним доступом ОЗП і обсязі ПЗП. Продуктивність визначає

вартість вокодера, тому при визначенні типу вокодера для тих чи інших додатків розробник повинен зробити відповідний вибір. У випадках, коли вокодер спільно використовує процесор з іншими додатками, розробник повинен вирішити, скільки ресурсів можна виділити для вокодера. Вокодери, що використовують менше 15 млн. операцій/с, вважаються низькопродуктивні. А ті, що використовують 30 або більше млн. операцій/с - високопродуктивними.

Збільшення продуктивності призводить до збільшення вартості і великих витрат енергії. Енергетичні витрати важливі для додатків в портативній апаратурі, так як при великих витратах енергії скорочується час між підзарядками батарей або виникає необхідність використовувати батареї більшої ємності, що, в свою чергу, призводить до більшої ціни і ваги.

### 1.3 Порівняння методів кодування голосових сигналів

Всі методи цифрового кодування голосу можна розділити на дві категорії: кодери форми сигналу і кодери джерела. У свою чергу схеми кодування голосу можуть бути розділені на три основні класи. Хоча існують і інші, три виділених класи є найбільш дослідженими. Основне завдання і функція цих схем кодування - проаналізувати вхідний сигнал, видалити надмірність і відповідним чином закодувати інформативні частини сигналу. Для зниження швидкості цифрового потоку доводиться розробляти все більш складні методи усунення надмірності.

Найбільш поширені стратегії кодування ГС наведені на рис. 1.1.

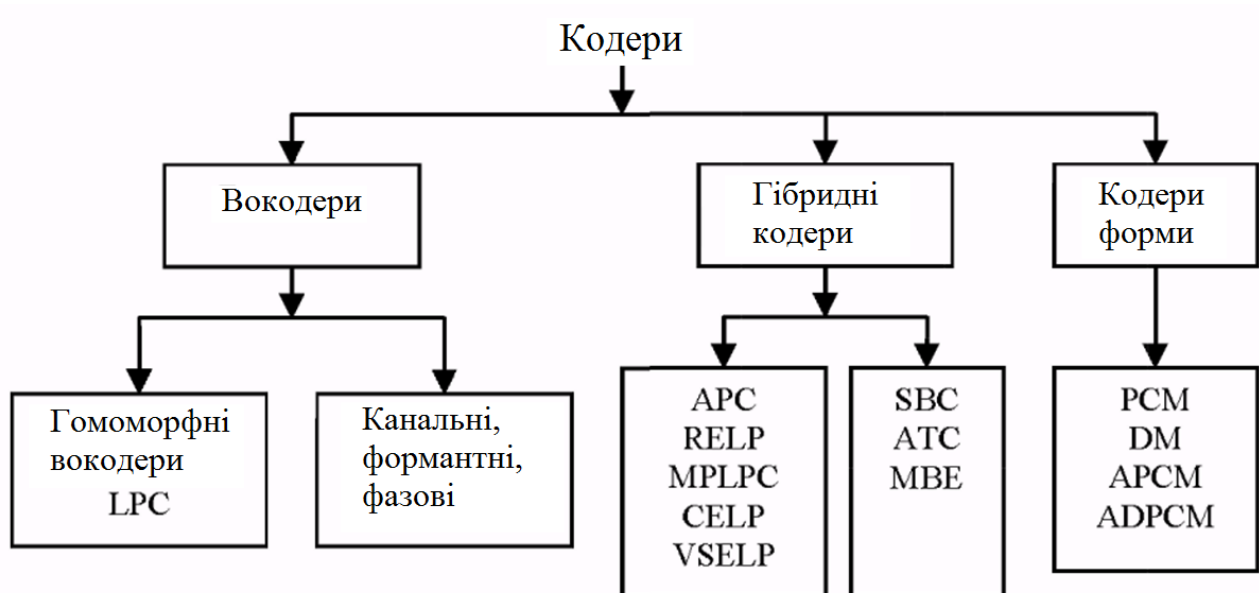


Рис. 1.1. Ієрархія систем кодування ГС

Розглянемо параметри порівняння методів кодування мови. До найбільш важливих відносяться наступні параметри.

**Швидкість** - це діапазон необхідних швидкостей передачі. Для систем з більш низькою бітовою швидкістю потрібна менша смуга частот, з цієї причини вони забезпечують більш високу ефективність використання спектра і потужності і в кінцевому підсумку призводять до стільникових систем радіозв'язку з підвищеною місткістю.

**Якість.** Критерій, що використовується при порівнянні - наскільки добре звучить мова в ідеальних умовах - так звана чиста мова, без помилок передачі, тільки при кодуванні. Слід зазначити, що якість є суб'єктивним результатом вимірювання і оцінювання.

3. Порогова ймовірність помилки на біт. Більш високе значення порогової помилки веде до більш робастної структури системи, а, отже, до більш низьких вимог до відношення сигнал/шум і збільшення ємності мережі.

4. Затримки кодування системи передачі мови - це фактор, тісно пов'язаний з характеристикою якості системи. Затримка кодування складається з

алгоритмічної затримки (буферизації мови для аналізу), обчислювальної затримки (часу, необхідного для обробки збережених мовних вибірок) і затримки в процесі передачі. Тільки перші дві відносяться до підсистеми кодування мови, хоча дуже часто схема кодування побудована так, що передача може початися навіть до того, як буде повністю оброблено інформацію в уже згаданому фреймі.

У додатках типу телеконференцій може виникнути необхідність з'єднання декількох учасників через багатоточковий пристрій управління. Це вимагає декодування кожного потоку даних, підсумовування декодованих сигналів, а потім перекодування результуючого сигналу. Цей процес подвоює затримку і, одночасно, зменшує якість мовлення через неоднократне кодування. Така система телемосту може допустити максимальну затримку в одну сторону 100 мс, так як міст призведе до подвоєння затримки системи в одну сторону до 200 мс.

5. Складність і споживання енергії. У міру розроблення «витончених» алгоритмів кодування неминуче зростає складність обчислень.

6. Обробка інших сигналів звукового діапазону. Канали зв'язку, спочатку призначені для передачі ГС, нерідко використовуються для передачі іншого роду сигналів, таких як сигнали від модему, факсу та ін. Статистичні параметри форми і частотний спектр таких сигналів зовсім відмінні від відповідних параметрів мови, тому алгоритм обробки повинен оперувати з обома видами сигналів. При розробці алгоритму можливість перетворення немовних сигналів звукового діапазону часто відходить на другий план і розглядається на заключному етапі, що, безумовно, є помилкою, особливо при розробці кодерів для мереж загального користування.

Якість мови і швидкість передачі - два конфліктуючих фактора. Чим нижча швидкість мовного кодера, тобто більш високий ступінь компресії сигналу, тим більше страждає якість - в даному випадку ступінь достовірності. Для систем, які



працюють в телефонних мережах загального користування і подібних до них, основною вимогою є якість кодування (кодери форми). Однак для закритих мереж, таких, як приватні комерційні або військові системи зв'язку, вимоги до якості можуть бути знижені з метою підвищення швидкості. Хоча зазвичай потрібна абсолютна якість, цим нехтують для іншої групи стандартів, мета яких - домогтися більш високого ступеня стиснення (гібридні кодери). Наприклад, в системах рухомого радіозв'язку високий ступінь стиснення при середній якості - це вирішальний фактор при виборі стандарту кодування. Ця якість зазвичай приймається в розрахунок як при хороших, так і при поганих умовах передачі. Виявляється, що перший фактор оцінити просто, по зменшенню швидкості цифрового потоку. Визначити і оцінити спотворення важче. Причина полягає в тому, що ГС сприймає не пристрій, а людина. На жаль, відсутність моделей сприйняття не дозволяє розробити об'єктивні методи оцінки спотворень мови.

У цьому випадку використовують суб'єктивні оцінки групи експертів, яким пропонують для прослуховування фрагменти декодувати мови. Усереднена експертна оцінка (УЕО) виводиться з оцінок окремих експертів за п'ятибальною шкалою: 5 - відмінно, 4 - добре, 3 - задовільно, 2 - незадовільно, 1 - погано.

У таблиці 1.1 наведені основні характеристики і усереднені експертні оцінки (УЕО) для різних кодерів.

## Порівняльний аналіз алгоритмів кодування

Кодер	Швидкість, кбіт/с	Затримка, мс	Продуктив- ність процесора, MIPS	УЕО
ПКМ з компандуванням	64	0,5	0,2	4,3
ДПКМ	40	1	3	4,1
АДПКМ	32	1	5,6	4,1
Вокодер LD-CELP	16	5	16	3,6
Вокодер GSM RPE	13	15	10	3,8
Вокодер CELP	8	20	15	3,7
Вокодер LPC	2,4	20	10	2,5

## 1.4 Висновки до розділу 1

Проаналізовано мовний сигнал, як засіб обміну інформацією. Розглянуто питання функціонування та організації систем цифрової телефонії, в якій саме мовний сигнал є переносником інформації.

Проаналізовано принципи функціонування та кодування мовних сигналів в вокодерах, зокрема цифрових.

Встановлено, що важливим є обґрунтування способу кодування мовних сигналів з метою збереження їх інформативності та підвищення швидкості обміну даними.

## РОЗДІЛ 2

### ПРИНЦИПИ РОБОТИ ВОКОДЕРІВ

#### 2.1 Принцип роботи параметричних вокодерів

Розглянемо узагальнену блок-схему параметричного вокодера, зображену на рисунку 2.1:

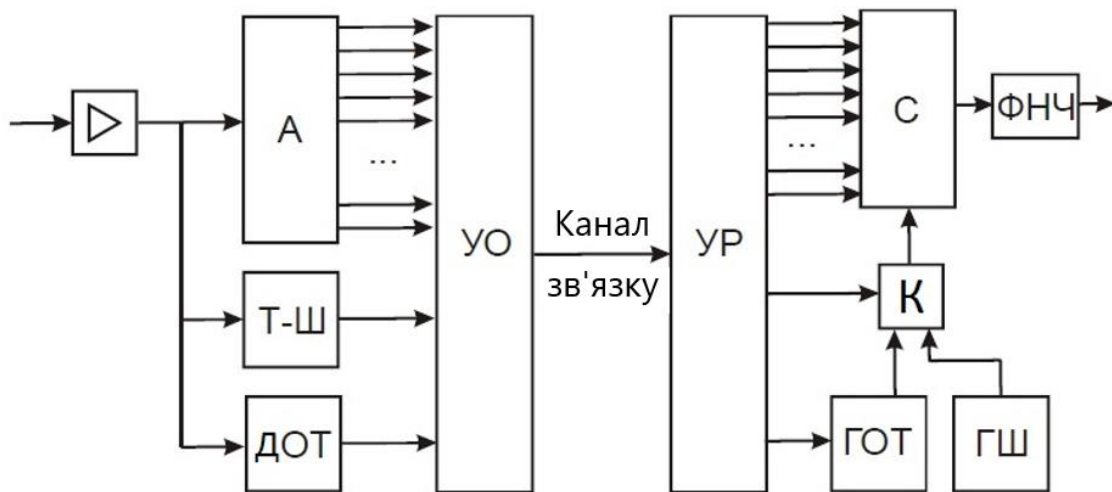


Рис. 2.1. Блок-схема параметричного вокодера

На схемі: А - аналізатор вхідного сигналу, який на основі сегмента вхідний реалізації ГС знаходить параметри, що підлягають передачі. Пристрій Т-Ш (виявляч «Тон-шум») здійснює розрізнення типу сегмента сигналу - вокалізованих він, або фрикативний, а ДОТ - детектор основного тону (блок оцінки параметрів основного тону) в разі, якщо сегмент відноситься до вокалізованих типу. УО - пристрій об'єднання сигналів для передачі через канал зв'язку, УР - пристрій поділу сигналів, переданих через канал. На приймальній стороні на основі прийнятих параметрів відбувається синтез ГС. Для цього

використовуються генератор основного тону, К - ключ. З - синтезатор, в якому відтворюється мовний сигнал.

## 2.2 Смугові (каналні) вокодери

У смугових вокодера мова ділиться на до 20 каналів за допомогою смугових фільтрів. Аналізатор визначає величину середньої інтенсивності ГС в кожній смузі. Ці величини передаються через канал зв'язку. Чим більше число каналів використовується у вокодері, тим більше натуральність і розбірливість мови. Блок-схема смугового вокодера зображена на малюнку 8:

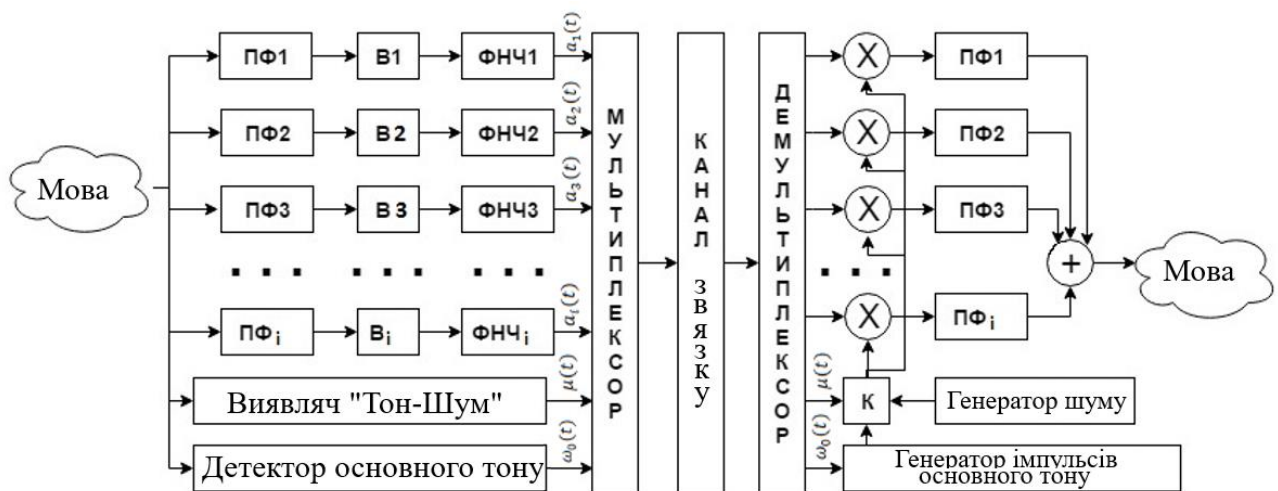


Рис. 2.2. Блок-схема смугового вокодера

ГС надходить на гребінку із смугових фільтрів (ПФ1 - ПФi). До виходів смугових фільтрів підключені випрямлячі (детектори) і згладжуючі фільтри НЧ. Тим самим отримуємо сигнали  $a_i(t)$ ,  $k = \overline{1, l}$ , характеризують амплітуду ГС в заданій смузі частот і є повільно мінливими в порівнянні з вихідним ГС. Крім того, на передавальному кінці ГС надходить на детектор основного тону. На виході отримуємо сигнал частоти основного тону  $\omega_0(t)$ . Виявляч «Тон-шум»

визначає характер ГС: вокалізований або фрикативний. Для вокалізованих звуків склад спектра звуків мови дискретний, і характер сигналу визначається як «тон». Для невокалізованих звуків (наприклад, шиплячих) склад звуків неперервний: характер сигналу «шумовий». Зформований на виході детектора сигнал позначимо:

$$\mu(t) = \begin{cases} 0, & \text{якщо звук шумовий} \\ 1, & \text{якщо вокалізований} \end{cases}$$

Слід зауважити, що визначник «Тон-шум» і детектор основного тону повинні взаємодіяти один з одним. Якщо сегмент ГС ідентифікований як шумовий, то немає сенсу здійснювати оцінку частоти основного тону, і  $\omega_0(t) = 0$ .

В результаті через канал зв'язку передаються: частота основного тону  $\omega_0(t)$ , тип сигналу  $\mu(t)$ , середні інтенсивності сигналу в обраних смугах частот  $a_i(t)$ ,  $k = \overline{1, l}$ . В мультиплексорі перераховані сигнали об'єднуються і перетворюються в форму, прийнятну для передачі по каналу зв'язку. На приймальній стороні виконується поділ сигналів за допомогою демультимплексора і перетворення їх в форму, необхідну для роботи синтезатора ГС.

Синтезатор ГС на приймальній стороні має в своєму складі генератор ГС, що складається з генератора імпульсів основного тону і генератора шуму. Генератор імпульсів використовується як джерело тональних сигналів для вокалізованих звуків і управляється сигналом  $\omega_0(t)$  таким чином, що частота проходження імпульсів дорівнює частоті основного тону ГС на передавальному кінці. Генератор шуму використовується як джерело шумових сигналів для невокалізованих ділянок мови. Перемикач тон/шум управляється сигналом  $\mu(t)$ .

Сигнал з ключа надходить на гребінку смугових фільтрів, таких же, які використовувалися на передавальному кінці, а з їх виходів - на помножувачі. Модульованими коливаннями тут є виходи смугових фільтрів, а модулюючими - сигнали  $a_i(t)$ ,  $k = \overline{1, l}$ . З виходів помножувачів сигнали надходять на ще одну гребінку смугових фільтрів, що служать для зменшення впливів побічних продуктів модуляції, що виникають в помножувачах. Сукупність схем і вузлів перетворення сигналу в межах кожної з частотних смуг називається спектральним каналом смугового вокодера. Вихідні сигнали спектральних каналів підсумовуються в вихідному підсилювачі і формують вихідний ГС.

При виборі характеристик смугових фільтрів виходять з того, що зі збільшенням крутизни затухання поза межами смуги пропускання підвищується точність вимірювання спектрів. Однак при цьому підвищується час перехідних процесів в смуговому фільтрі, що спотворює швидкі спектральні зміни і призводить до реверберації синтезованої мови. Тому має сенс застосування смугових фільтрів, що мають з одного боку хорошу роздільну здатність, а з іншого - монотонну імпульсну характеристику без сплесків.

Смуга пропускання ФНЧ повинна забезпечити передачу часових змін сигналів в спектральних каналах і при цьому досить сильно придушити коливання гармонік основного тону.

Кількість каналів смугового вокодера можна вибирати зі співвідношення розбірливості і швидкості передачі мови. Так, якщо використовувати від 16 до 20 спектральних каналів, то їх ширина буде порівнянна з шириною частотних груп слуху, отже вокодер буде мати достатню розбірливість.

Зазвичай схеми смугових вокодерів доповнюють пристроями лінійного передбачення. Тим самим отримують ЛПК-вокодери.

Суть методу лінійного передбачення полягає в тому, що прогнозована величина ГС  $\tilde{\lambda}(h)$  в момент опитування  $h$  лінійна сума попередніх відліків:

$$\tilde{\lambda}(h) = \sum_{m=1}^p a_m \tilde{\lambda}(h-m) \quad (2.1)$$

де  $\tilde{\lambda}(h-m)$  – ГС в попередні моменти опробування,  $a_m$  – коефіцієнти передбачення,  $m = 1, 2, \dots, p$ .

Інтервали часу між моментами опробування визначаються частотою дискретизації  $F_d = 1/(t_h - t_{h-1})$ . В момент  $h$ , коли відомі не лише  $\tilde{\lambda}(h-m)$ , але і істинні значення ГС  $\lambda(h)$ , можна визначити похибку передбачення:

$$e_h = \lambda_h - \tilde{\lambda}_h \quad (2.2)$$

Задовільної якості синтезованих ГС можна досягти при використанні не менше десяти коефіцієнтів передбачення  $a_m$ , що в (2.1) відповідає  $p = 10 \dots 12$ .

Каналом зв'язку передаються коефіцієнти передбачення. Їх значення використовуються в якості параметрів в рекурсивному цифровому фільтрі, на вхід якого подаються сигнали збудження.

Таким чином, якість (розбірливість, натуральність) синтезованої мови залежить від числа спектральних каналів, на які поділяється в аналізаторі спектр ГС. Так, 16-канальний смуговий вокодер дозволяє відтворити спектральну огинаючу сигналу з точністю до ширини частотних груп слуху, в межах яких вухо не помічає переміщення максимуму спектра. При цьому досягається швидкість передачі, що становить  $3 \text{ (біта)} \times 50 \text{ (Гц)} \times 16 \text{ (каналів)} = 2,4 \text{ кбіт/с}$ . Число каналів (фільтрів) і ширина смуги пропускання частот можуть варіюватися, в відповідність з цим буде змінюватися і якість відтворення звуку. Смугові вокодери забезпечують високу розбірливість мови (до 85% розбірливості складів), але натуральність її, як і в інших системах вокодерів, залишає бажати кращого.

До основних недоліків смугових вокодерів, що погіршує розбірливість синтезованої мови, відносять появу великих інтерференційних спотворень

обвідної спектра, що виникають через неузгодженість ФЧХ смугових фільтрів, а також спотворення спектральної картини в формантній області, пов'язані з попаданням форманти в міжфільтрову зону.

### 2.3 Ортогональні вокодери

Принцип дії ортогональних вокодерів базується на можливості представлення спектральної обвідної ГС у вигляді суми ортогональних функцій. У цих випадках, на відміну від інших вокодерів, спектральна огинаюча на прийомі відтворюється не по окремим ординатам, а у вигляді суми тих же ортогональних функцій, тобто .:

$$\Phi(\omega, t) = \sum_{i=0}^m a_i(t) \xi_i(\omega) \quad (2.3)$$

де  $\xi_i(\omega)$  – і-тий член ортогонального ряду;  $a_i(t)$  – його коефіцієнти. При цьому відповідно до умов ортогональності:

$$a_i(t) = \int_{\omega_1}^{\omega_2} \Phi(\omega, t) \xi_i(\omega) d\omega \quad (2.4)$$

де  $[\omega_1, \omega_2]$  – інтервал розкладання спектральної обвідної, тобто переданий діапазон частот.

Застосування теорії наближення функцій дозволяє підібрати найбільш підходящі ортогональні функції для подання спектральних огинаючих еталонного ГС, створюваного на виході мікрофонів хорошої якості, хоча б в телефонній смузі частот.

Орієнтуючись на такий ГС можна легко встановити, що кількість функцій, «потрібних» для наближення, обмежена. Відповідними, очевидно, будуть функції, які при найменшому числі членів суми будуть давати найбільшу точність наближення. Крім того, при виборі функцій наближення важливо мати



на увазі і можливість простого апаратурного виділення їх з ГС і відтворення на приймальному кінці.

До теперішнього часу найбільш повно розроблений метод подання спектральної обвідної у вигляді суми гармонійних функцій. На основі гармонійного представлення розроблений гармонічний вокодер. За допомогою з'єднання передаються коефіцієнти розкладання в тригонометричний ряд усередненої за 25-20 мс спектральної обвідної ГС. За цими коефіцієнтами на приймальній стороні здійснюється відновлення спектральної обвідної. Число коефіцієнтів залежить від необхідної точності відтворення спектральної обвідної.

При реалізації вокодера в цифровому варіанті коефіцієнти ряду Фур'є можна обчислити прямим застосуванням формули (2.4). У звичайному варіанті можна використовувати матричну схему перерахунку ординат спектра в коефіцієнти ряду Фур'є відповідно до виразів:

$$a_0 = \frac{1}{2n} \sum_{k=0}^{2n-1} y_k; \quad a_m = \frac{1}{n} \sum_{k=0}^{2n-1} y_k \cos \frac{km\pi}{n}; \quad b_m = \frac{1}{n} \sum_{k=0}^{2n-1} y_k \sin \frac{km\pi}{n}$$

де  $y_k$  – ординати спектральної густини,  $m$  – індекс коефіцієнта Фур'є,  $a_m, b_m$  – коефіцієнти Фур'є. З формул випливає, що якщо відомі ординати спектральної обвідної, то досить кожен з них помножити на відповідні множники, визначені для заданого числа коефіцієнтів розкладання, і потім алгебраїчно підсумувати добутки. Отримані суми і будуть коефіцієнтами  $a_m, b_m$ . Все це можна реалізувати за матричною схемою за допомогою резисторних дільників і суматорів. Спосіб отримання ординат спектральної обвідної можна запозичити з реалізації смугового вокодера. Блок-схема гармонічного кодера на основі аналізатора смугового кодера зображена на рисунку 2.3:

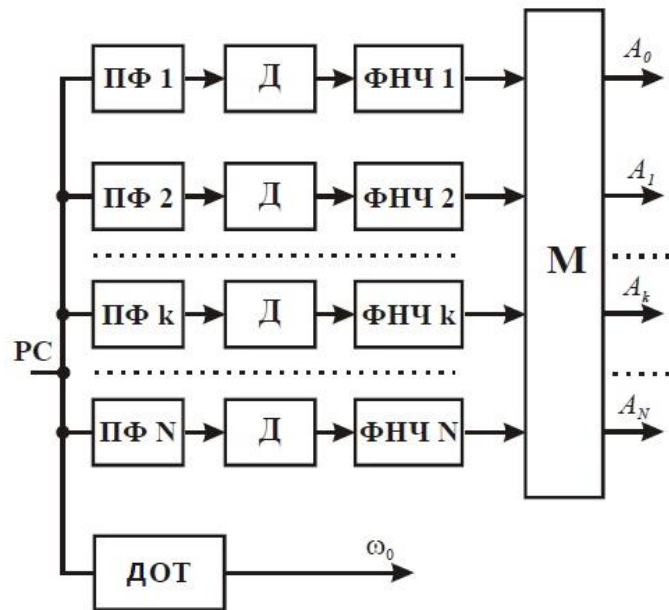


Рис.2.3. Блок-схема гармонічного кодера

На схемі: ПФ - смугові фільтри, Д - детектори, ДОТ - детектор (виявляч) основного тону, М - матрична схема перерахунку ординат спектра в коефіцієнти ряду Фур'є.

Синтез ГС на приймальній стороні тракту зводиться до відтворення форми спектральної обвідної вихідного ГС для сигналів, що надходять від місцевих джерел мовних коливань, що мають монотонну форму спектральної обвідної. В якості таких джерел коливань використовується (в залежності від сигналу тон-шум) або генератор дискретного спектра (генератор ОТ), або генератор шуму.

У гармонічному вокодері для синтезу ГС використовуються періодичні фільтри, на які подається сигнал збудження від джерела коливань (ОТ або шум). Фільтри управляються сигнал-параметрами, які приходять з лінії. Для цього періодичні фільтри складаються з лінії затримки з відводами через рівні часові затримки модуляторів і суматора. За допомогою лінії затримки створюються гармонічні функції різних порядків, модулятори служать для управління амплітудами цих функцій за допомогою сигнал-параметрів. Підсумовування

сигналів, що надходять від усіх модуляторів, дозволяє отримати сигнал, який має спектральну огинаючу вихідного ГС (якщо сигнал-параметри в лінії не зазнали спотворень). На малюнку 10 наведена блок-схема синтезатора косинусного типу:

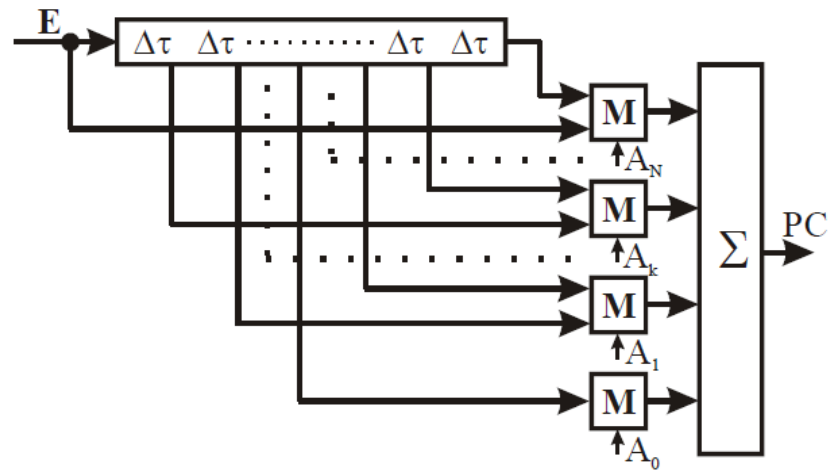


Рис.2.4. Блок-схема синтезатора косинусного типу

Синтезатор косинусного типу відтворює спектральну огинаючу, що виражається сумою косинусів:

$$\begin{aligned}
 E_{\text{ВЫХ}} = E \exp(-in\omega\Delta\tau) \{ & A_0 + A_1[\exp(i\omega\Delta\tau) + \exp(-i\omega\Delta\tau)] \\
 & + A_2[\exp(i2\omega\Delta\tau) + \exp(-i2\omega\Delta\tau)] + \dots + A_n[\exp(in\omega\Delta\tau) \\
 & + \exp(-in\omega\Delta\tau)] \} = E \exp(-in\omega\Delta\tau) \left[ A_0 + 2 \sum_{m=1}^n A_m \cos(2\pi m f \Delta\tau) \right]
 \end{aligned}$$

де  $E$  - сигнал, що подається на вхід лінії затримки від місцевого джерела мовних коливань.

На рисунку 2.5 наведена схема синтезатора з періодичними фільтрами, описаними повним тригонометричним поліномом:

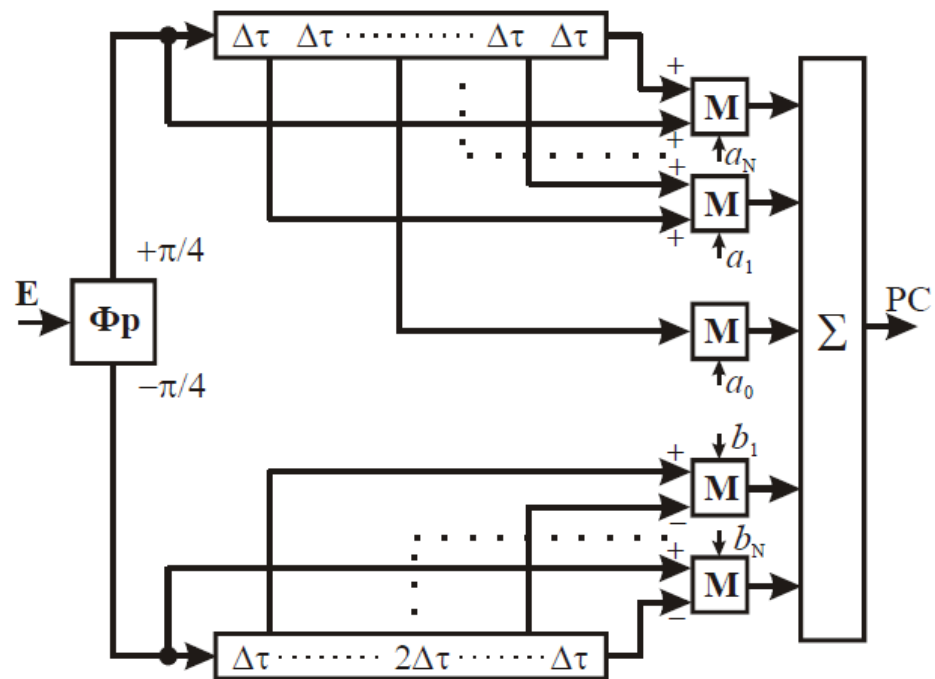


Рис. 2.5. Схема синтезатора з періодичними фільтрами

На відміну від наведеної вище схеми синтезатора в неї входять фільтри, які синтезують і синусоїдальні функції, а також широкосмуговий фазорозщеплювач сигналу (Фр), що приходить від джерела мовних коливань з поворотом фаз всіх складових коливань на  $\pm \pi/4$ . Загальний час затримки  $\tau$  визначається добутком числа виводів (рівним подвоєному числу сигнал-параметрів) і значення часу затримки між виводами, тобто  $\tau = 2n\Delta\tau$ , а значення  $\Delta\tau$  визначається верхньою частотою переданого частотного діапазону голосу, тобто  $\Delta\tau = 1/f_{max}$ .

У гармонічному вокодері існують спотворення, схожі за звучанням з спотвореннями в смуговому вокодері, що пов'язані з тим, що при підсумовуванні сигналів, які приходять від віддалених виводів лінії затримки, спостерігаються інтерференційні спотворення. З метою зменшення спотворень можна пробувати подавати на матричну схему визначення коефіцієнтів ряду Фур'є не відліки спектральної щільності, а деякі їх функції, наприклад, логарифм або корінь. В

результаті виходять модифікації гармонічного, наприклад, косинусно-логарифмічний вокодер, блок-схема якого приведена на рисунку 2.6:

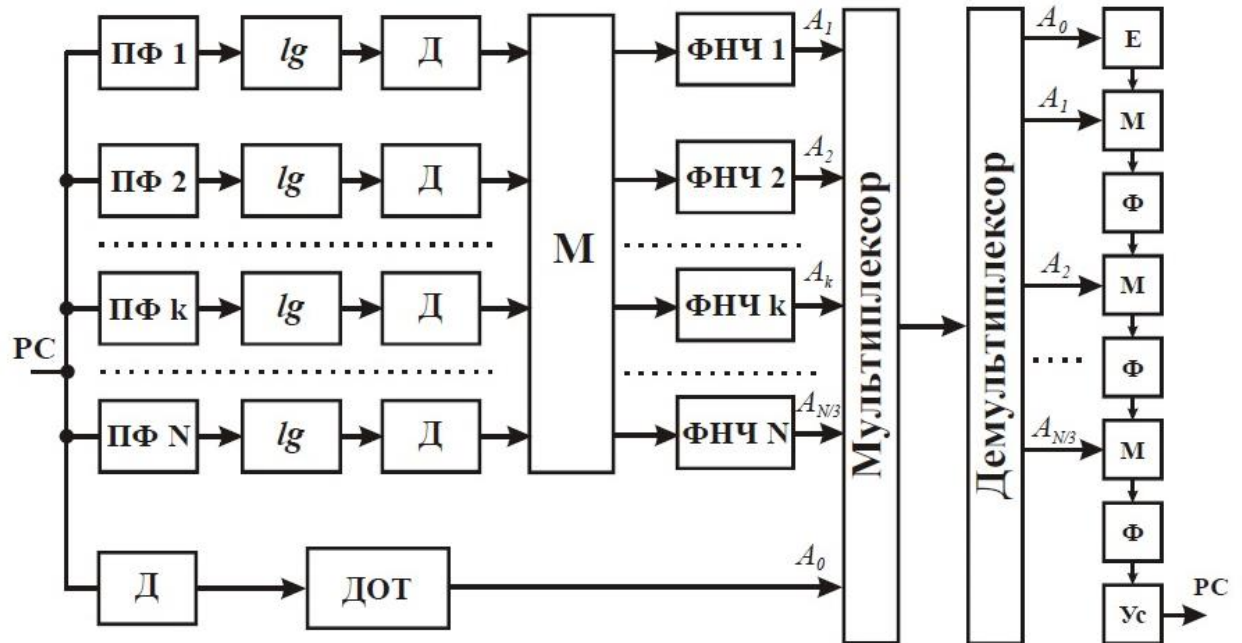


Рис. 2.6. Блок-схема косинусно-логарифмічного вокодера, де в аналізатор додані блоки логарифмування  $lg$  і позначено: М - модулятори, Ф - фазовирівнювач, Ус – підсилювач

Для отримання високої точності розкладу в суму Фур'є доцільно брати 18 смуг рівної ширини в діапазоні до 3500 Гц, а число членів тригонометричного ряду - не більше п'яти. При цьому шостий член - постійна складова - визначався в лінійному масштабі і передавався по каналу зв'язку без обробки за матричною схемою.

Перевагою цього вокодера є послідовний спосіб побудови синтезатора, що дозволяє уникнути інтерференційних спотворень між окремими складовими спектра.

Ще одним з недоліків, описаних вище схем гармонічного вокодера, є застосування рівномірного частотного масштабу для передачі спектральної обвідної, внаслідок чого точність передачі її не відповідає суб'єктивному масштабу сприйняття по частотному діапазону. Цей же недолік є і у смугового вокодера. Він може бути усунутий вибором ширини смуг фільтрів аналізатора і синтезатора в суб'єктивному масштабі. Для аналізатора гармонічного вокодера технічно не важко перейти на суб'єктивний масштаб по частоті. Для цього треба взяти смугові фільтри з середніми частотами, розподіленими по частотному діапазону рівномірно в суб'єктивному масштабі по частоті. Складніше йде справа з приймальною частиною гармонічного вокодера - синтезатором, якщо використовується аналогова обробка сигналу. Звичайні лінії затримки дають рівномірний масштаб по частотному діапазону, а тут необхідний нерівномірний масштаб. Однак, при цифровій реалізації синтезатора зазначені труднощі легко подолати.

У ортогональних вокодерах, на відміну від смугових, сигнал-параметри нерівнозначні і некорельовані один з одним. Тому спотворення одного з них іноді призводить до значних спотворень передачі. Особливо помітно ця властивість ортогональних вокодерів позначається на якості - передачі мови при невеликому числі переданих параметрів. Іншим недоліком гармонічних вокодерів, що також виявляється при невеликому числі переданих параметрів, є спотворення форми формантних кривих: замість резонансних виходять криві синусоїдальної форми і з'являються додаткові максимуми - помилкові форманти, як показано на рисунку 2.7 а):

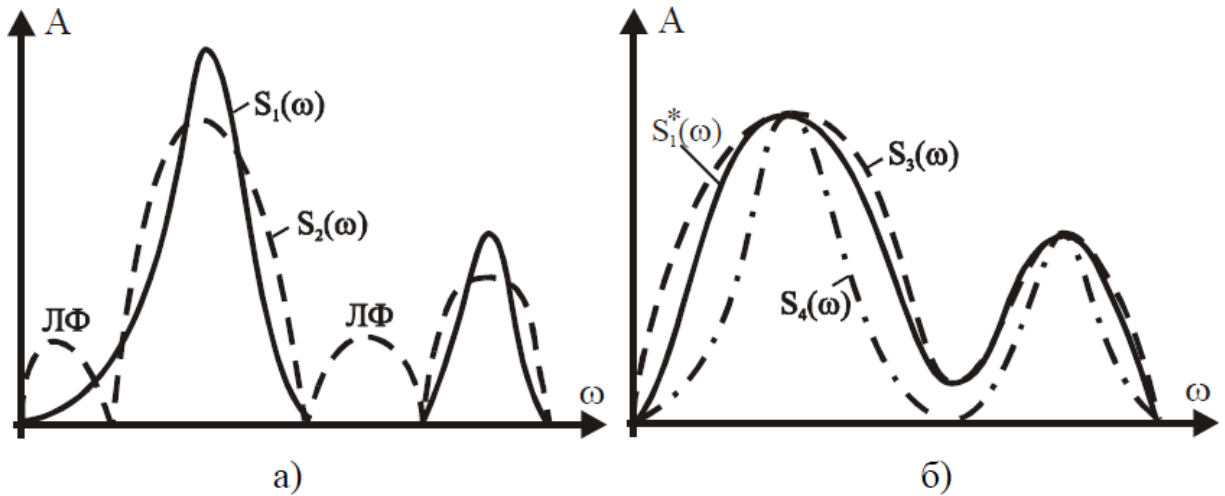


Рис. 2.7. Спотворення форми формантних кривих

Поява таких спотворень надає звучанню синтезованої мови назалізації. Щоб наблизити форму спектральної обвідної до типової формантної структури можна використовувати квадратичний ортогональний (гармонійний) вокодер (рис. 2.8):

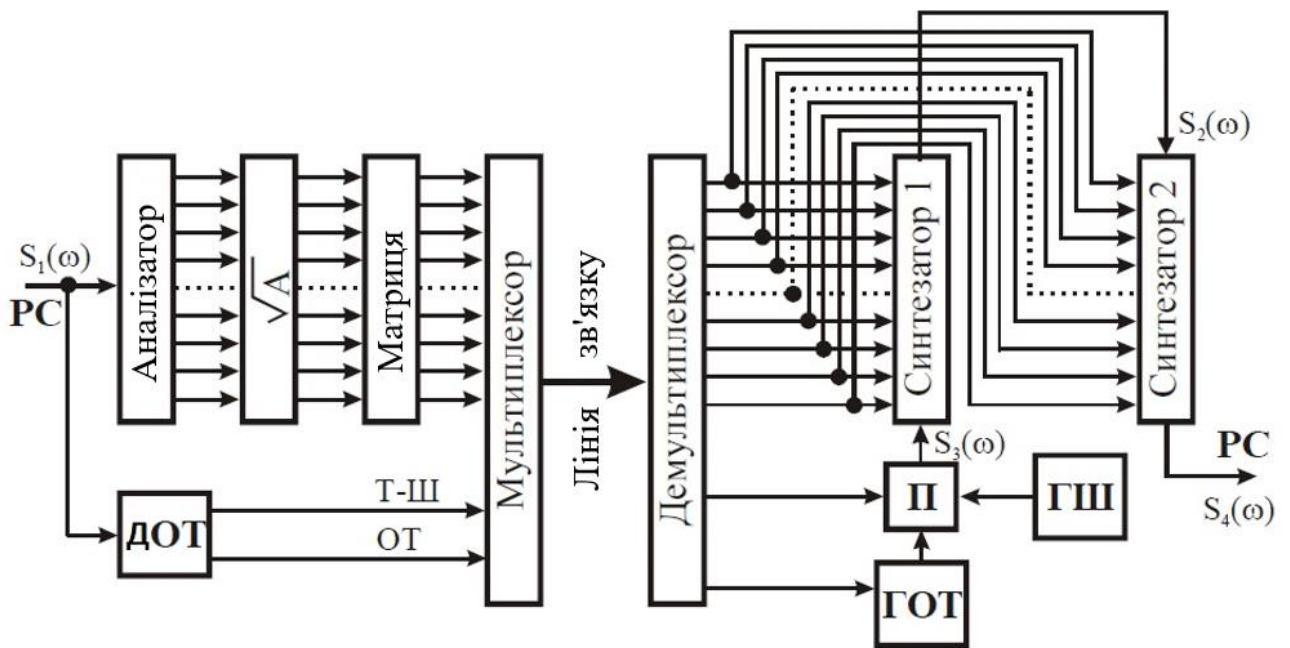


Рис. 2.8 Блок-схема квадратичного ортогонального вокодера

В аналізаторі цього вокодера, на відміну від описаного вище лінійного гармонічного вокодера, кожна з  $n$  ординат спектра  $y_k$  піддається нелінійному перетворенню - вилучення квадратного кореня з її величини. Тільки після цієї операції за отриманими значеннями змінених ординат  $\sqrt{y_k}$  визначаються коефіцієнти ортогонального ряду  $\xi_i$ , в такій же матриці, як і в лінійному гармонічному вокодері. Таким чином, спектральна огинаюча  $S_1^*(\omega) = \sqrt{S_1(\omega)}$  розкладається в суму із  $n$  членів:

$$\sqrt{S(\omega)} \approx S_3(\omega) = \sum_{i=1}^m a_i(t) \xi_i(\omega)$$

де  $\xi_i(\omega)$  – члени суми;  $a_i = \sum_{k=1}^n \sqrt{y_k} \xi_i(\omega_k)$ ,  $\omega_k$  – середня частота смуги, для якої визначена  $k$ -та ордината спектра. Крива  $S_1^*(\omega)$  на рис. 13 б) має більш згладжену форму, ніж вихідний спектр  $S_1(\omega)$ , і тому краще апроксимується гармонічними функціями, ніж при звичайному перетворенні (рис. 13 а)). Якщо звести в квадрат спектр  $S_3(\omega)$ , то отримаємо спектр  $S_4(\omega)$ , що збігається з  $S_1(\omega)$ . В результаті такого перетворення в спектральній обвідній з'являються максимуми спектра, подібні за формою з формантними максимумами.

На приймальній стороні сигнал-параметри керують двома синтезаторами лінійного типу (див. Рис. 2.7). На вхід лінії затримки першого синтезатора подається сигнал від джерела мовних коливань, тобто від генератора ОТ або генератора шуму, а на вхід лінії затримки другого синтезатора подається сигнал з виходу першого синтезатора. Кожен з синтезаторів має керовану частотну характеристику:

$$K_1(\omega) = K_2(\omega) = \sum_{i=0}^m a_i \xi_i(\omega)$$

Якщо на вхід першого синтезатора подається спектр  $S_1(\omega) = 1$ , то спектральна огинаюча вихідного сигналу буде:



$$S_4(\omega) = S_1(\omega)K_1(\omega)K_2(\omega) = [S_3(\omega)]^2 \approx S_1^*(\omega)$$

В результаті цієї операції відбувається «квадрування» обвідної спектра, отримана спектральна огинаюча стає близькою до вихідної як за формантними кривими, так і внаслідок усунення помилкових формант. Експериментально доведено, що в даному випадку розбірливість є кращою в порівнянні з лінійним вокодером.

#### 2.4 Ліпредери (вокодери з лінійним передбаченням)

Вокодери з лінійним передбаченням є найпоширенішими в сучасних системах передачі мови. Метод лінійного передбачення запропонований в 1960-х роках і отримав потужний розвиток в 1980-х, в тому числі в прямому зв'язку з розробкою мовних кодеків для цифрових систем стільникового зв'язку. Лінійне передбачення є одним з найбільш ефективних методів аналізу ГС. Цей метод стає домінуючим при оцінці основних параметрів ГС, таких, як період основного тону, форманти, спектр, а також при скороченому поданні мови з метою її низькошвидкісної передачі і економного зберігання. Важливість методу обумовлена високою точністю одержуваних оцінок і простотою обчислень.

Саме методи кодування з лінійним передбаченням і застосовуються в стільниковому зв'язку. Розглянемо спочатку модель мовного тракту, яка дозволяє використовувати методи лінійного передбачення для аналізу і синтезу ГС. Блок-схема моделі мовного тракту зображена на рис. 2.9:

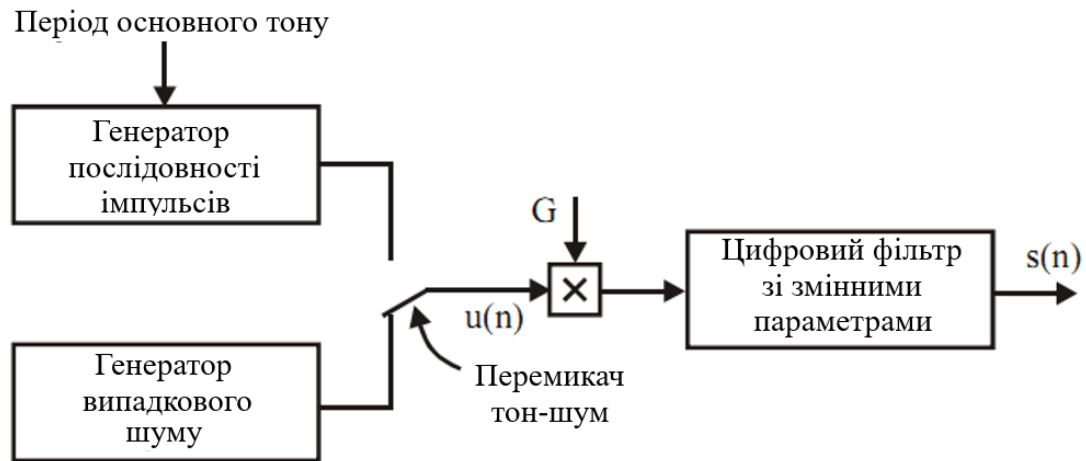


Рис. 2.9. Блок-схема моделі голосотворення

Загальний спектр, обумовлений випромінюванням, мовним трактом і збудженням, описується за допомогою лінійної системи зі змінними параметрами з передавальною функцією

$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \sum_{k=1}^p a_k z^{-k}}. \quad (2.5)$$

Ця система збуджується імпульсною послідовністю для вокалізованих звуків мови і шумом для невокалізованих. Таким чином, модель має наступні параметри: класифікатор вокалізованих і невокалізованих звуків, період основного тону для вокалізованих сегментів, коефіцієнт підсилення  $G$  і коефіцієнти  $a_k$  цифрового фільтра. Всі ці параметри повільно змінюються в часі.

Вираз (2.5) являє собою передавальну функцію так званої полюсної моделі без урахування нулів. Відомо, що якщо порядок  $p$  моделі досить великий, то полюсна модель дозволяє отримати досить точний опис майже для всіх звуків мови. Головна перевага цієї моделі полягає в тому, що як параметр  $G$ , так і коефіцієнти можна оцінити безпосередньо з використанням дуже ефективних з обчислювальної точки зору алгоритмів. Для системи рис. 1 відлік ГС  $s(n)$  пов'язаний з сигналом збудження  $u(n)$  простим різницеvim рівнянням

$$s(n) = \sum_{k=1}^p \alpha_k s(n-k) + Gu(n). \quad (2.6)$$

Лінійний передбачувач з коефіцієнтами  $\alpha_k$  визначається як система, на виході якої маємо

$$\tilde{s}(n) = \sum_{k=1}^p \alpha_k s(n-k), \quad (2.7)$$

де  $\tilde{s}$  – передбачене значення ГС;  $n$  - номер часового відліку;  $p$  - порядок передбачення (число коефіцієнтів лінійного передбачення);  $k$  - коефіцієнти лінійного передбачення - вагові коефіцієнти, що використовуються в лінійній комбінації. Системна функція передбачувача  $p$ -го порядку представляє собою поліном виду

$$P(z) = \sum_{k=1}^p \alpha_k z^{-k}. \quad (2.8)$$

Різниця між істинним (відомим точно)  $s(n)$  і передбаченим  $\tilde{s}(n)$  значеннями відліку визначає похибку (залишок) передбачення:

$$e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{k=1}^p \alpha_k s(n-k). \quad (2.9)$$

З виразу (2.9) видно, що похибка передбачення являє собою сигнал на виході системи з передавальною функцією.

$$A(z) = 1 - \sum_{k=1}^p \alpha_k z^{-k} = 1 - P(z). \quad (2.10)$$

Фільтр похибки, передбачення  $A(z)$  є зворотним фільтром для системи  $H(z)$ , що відповідає рівнянню (2.5), тобто

$$H(z) = \frac{G}{A(z)}. \quad (2.11)$$

Вираз (2.11) можна трактувати наступним чином: лінійна система зі змінними параметрами (фактично - модель фільтра голосового тракту)  $H(z)$  збуджується імпульсною послідовністю для вокалізованих звуків і шумом для невокалізованих. При цьому сигнал збудження формується адаптивним фільтром-аналізатором  $A(z)$  (21), адаптація якого здійснюється шляхом зміни коефіцієнтів лінійного передбачення.

Таким чином, в процесі параметричного представлення ГС в системі кодер/декодер використовуються два цифрових фільтра: в кодері джерела в фільтрі-аналізаторі  $A(z)$  визначаються коефіцієнти передбачення, а в декодері на основі цих коефіцієнтів за допомогою цифрового фільтра  $H(z)$  синтезується еквівалент голосового тракту. Шляхом збудження цього «еквівалента тракту» формується синтезоване мовлення. Зауважимо, що фільтр-аналізатор і фільтр-синтезатор є рекурсивними, оскільки значення сигналу на їх виходах визначається лише попередніми відліками ГС.

Пряме  $z$ -перетворення послідовності  $s(n)$  визначається рівнянням:

$$S(z) = \sum_{k=-\infty}^{\infty} s(k)z^{-k} \quad (2.12)$$

А в результаті  $z$ -перетворення різницевого рівняння (20) маємо:

$$R(z) = S(z)A(z), \quad (2.13)$$

де  $R(z)$  -  $z$ -перетворення вибірки (відліку) помилки передбачення  $e(n)$  на виході фільтра-аналізатора, що має передавальну функцію (2.13).

Оскільки амплітудно-частотна характеристика фільтра-аналізатора  $A(z)$  зворотна до частотної характеристики голосового тракту (отже, обернена і до обвідної спектра вхідного сигналу), такий фільтр називається інверсним. Його блок-схема показана на рис. 2.10. Значення коефіцієнтів передбачення  $a_k$ , є параметрами цього фільтра. Вони залишаються постійними на інтервалі аналізованого сегмента мови (як правило 20 мс), оскільки лінійний передбачувач переналаштовується (тобто адаптується) не під кожен мовний відлік, а під їх послідовність. Тому помилка мінімізується протягом усього сегмента.

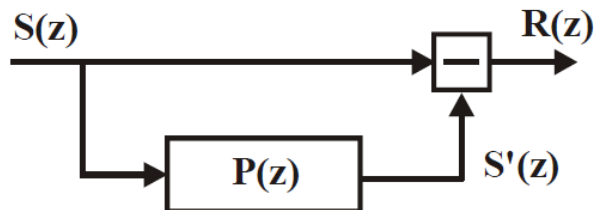


Рис.2.10 Блок-схема фільтра-аналізатора

Інверсний фільтр застосовується в кодері для усунення надмірності ГС. Пропускаючи через цей фільтр вихідний сигнал на виході отримуємо сигнал залишку передбачення  $e(n)$  (інакше - перший залишковий сигнал). Цей «залишок» містить періодичну складову ГС, що відповідає основного тону, оскільки в ньому усунені внутрішні кореляційні зв'язки. У той же час відповідно до моделі голосотворення вихідний сигнал інверсного фільтра з оптимально підібраними параметрами буде являти собою сигнал збудження  $e(n) = Gu(n)$ ,  $R(z)$ , подібний до сигналу збудження  $u(n)$  на вході фільтра голосового тракту на рис. 2.10.

Синтезуючий фільтр знаходиться в декодері і здійснює формування ГС. «Потрібна» настройка цього фільтра забезпечується передачею на приймальну сторону коефіцієнтів передбачувача  $P(z)$ , які використовуються в цей момент в

кодері, а на вхід синтезуючого фільтра подається сигнал збудження  $R(z)$ . Очевидно, що він повинен бути максимально «схожий» на сигнал залишку передбачення, отриманий в кодері.

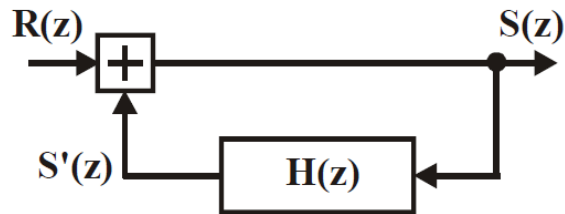


Рис. 2.11 Модель фільтра синтезатора

Модель фільтра-синтезатора (рис. 2.11) визначається рівнянням (2.14), що описує модель фільтра голосового тракту. Передавальна функція фільтра-синтезатора  $H(z)$  є зворотною до передавальної характеристики фільтра-аналізатора з точністю до скалярного коефіцієнта підсилення  $G$ . Це впливає з виразу (2.14):

$$S(z) = R(z) / A(z) = R(z)H(z) \quad (2.14)$$

Такий запис рівняння моделі голосотворення пояснює механізм отримання синтезованого сигналу  $S(z)$  на виході фільтра при впливі на його вхід сигналу збудження  $R(z)$ .

При виконанні аналізу до відповідальних рішень відноситься вибір порядку фільтра лінійного передбачення (предиктор)  $p$ . При кодуванні вибірки ГС зазвичай беруться з частотою 8кГц, що забезпечує значення верхньої граничної частоти спектра ГС 4кГц. У смузі 4 кГц максимальне число виділених формант зазвичай дорівнює чотирьом, що вимагає застосування фільтра як мінімум 8-го порядку. Зазвичай використовується 10-полісний фільтр. Тому в таких системах резонанси формант і форма основного спектру моделюються точно. В особливо точних схемах використовуються фільтри більш високого порядку. На рис. 2.12

показані огинаючі спектри сигналів, синтезовані за допомогою фільтрів для різних  $\rho$ :

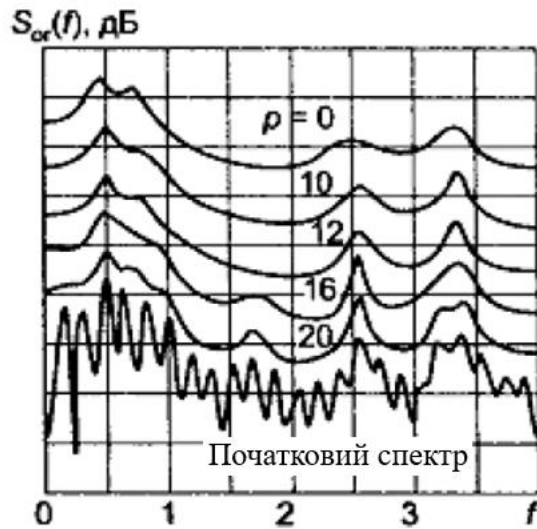


Рис. 2.12 Графіки огинаючих спектру сигналу

Це свідчить про те, що лінійне передбачення можна розглядати як метод короткочасної спектральної оцінки ГС.

Основне завдання аналізу на основі лінійного передбачення полягає в безпосередньому визначенні параметрів  $a_k$  за ГС з метою отримання хороших оцінок його спектральних властивостей шляхом використання рівняння (22). Оскільки властивості ГС змінюються в часі, коефіцієнти передбачення повинні оцінюватися на коротких сегментах мови.

Основним підходом є визначення параметрів передбачення таким чином, щоб мінімізувати дисперсію похибки передбачення на короткому сегменті сигналу. При цьому передбачається, що отримані параметри є параметрами системної функції  $H(z)$  в моделі голосотворення.

Подібний підхід призводить до корисних результатів. По-перше, якщо  $a_k = \alpha_k$ , тоді  $e(n) = Gu(n)$ . Для вокалізованих звуків це означає, що  $e(n)$  буде складатися з послідовності імпульсів з частотою основного тону, і  $e(n)$  буде

досить малим майже весь час. Тому в даному випадку мінімізація похибки передбачення дозволить отримати необхідні коефіцієнти.

По-друге, подібна мінімізація призводить до лінійної системи рівнянь, вирішення яких порівняно легко призводить до отримання параметрів передбачення. Крім того, отримані параметри складають ґрунтовну основу для точного опису сигналу.

Для визначення коефіцієнтів передбачення  $a_k$  необхідно вирішити систему  $p$  лінійних рівнянь з  $p$  невідомими коефіцієнтами. Для цього власні похідні від  $[e^2(n)]_{cp}$  по  $a_k$  прирівнюються до нуля, в результаті чого виходить система лінійних рівнянь:

$$B_t = \sum_{k=1}^p a_k C_{k,t}, \quad 0 \leq t \leq p \quad (2.15)$$

де коефіцієнти  $B_t, C_{k,t}$  будуть розглянуті нижче.

Система з  $p$  лінійних рівнянь з  $p$  невідомими коефіцієнтами  $a_k$  і методи її вирішення виявляються дещо різними в залежності від того, на якому сегменті ГС виконується пошук коефіцієнтів передбачення. Якщо припустити, що сигнал дорівнює нулю поза деяким інтервалом (сегмента)  $0 \leq n \leq N-1$ , то його зручно записати як  $s_m(n) = s(m+n)w(n)$  де  $w(m)$  — вікно кінцевої тривалості, наприклад Хеммінга, рівне нулю поза інтервалом,  $m$  - номер розглянутого сегмента сигналу. Тривалість вікна вибирається рівною 10 ... 30 мс, тобто декілька періодів основного тону. Тоді для визначення коефіцієнтів передбачення необхідно вирішити систему рівнянь (2.16), де короткочасна автокореляційна функція ГС:

$$B_t = D_t, \quad C_{kt} = D_{k-t} = D_{t-k}, \quad D_t = \sum_{n=0}^{N-1-t} s(n)s(n+t) \quad (2.16)$$



Матриця  $C$  розміром  $p \times p$  є симетричною і такою, що елементи на будь-якій діагоналі рівні між собою, що може бути використано для отримання ефективного алгоритму вирішення системи рівнянь (2.16). Це перший з методів лінійного передбачення, названий автокореляційним.

Сегмент ГС можна визначити без множення на прямокутне вікно. Можна вважати, що ГС не дорівнює нулю поза межами даного сегменту, а сегмент визначається інтервалом, на якому обчислюється середній квадрат похибки. Тоді коефіцієнти рівняння (2.16) обчислюються як:

$$B_t = D_{t0}, \quad C_{kt} = D_{kt}, \quad D_{tk} = \sum_{n=0}^{N-1} s(n-t)s(n-k) = \sum_{n=-k}^{N-1-k} s(n)s(n+k-t) \quad (2.17)$$

Для обчислення  $C_{kt}$  для всіх необхідних значень:  $t$  і  $k$  необхідно використовувати значення  $s(n)$  на інтервалі, тобто без обмеження послідовності відліків вікном кінцевої тривалості. Таким чином, обчислення  $C_{kt}$  призводить не до автокореляційної, а до взаємокореляційної функції між двома дуже схожими, але не однаковими сегментами ГС кінцевої тривалості. Матриця  $C$  розміром  $p \times p$  є квазікореляційно симетричною. Метод аналізу, заснований на викладеному вище способі обчислення  $C_{kt}$ , називається коваріаційним, оскільки матриця має властивості коваріаційної матриці.

Отримані тим або іншим методом значення коефіцієнтів фільтра передбачення  $a_k$  передаються по лінії зв'язку. При цьому для забезпечення прийнятної точності відновлення значень ГС в приймачі виявилось необхідним використовувати від 8 до 10 бітів для цифрового представлення кожного коефіцієнта. Тому замість цих коефіцієнтів зазвичай передають так звані коефіцієнти відображення, або коефіцієнти часткової кореляції (partial correlation, PARCOR) - коефіцієнти  $J_k, k = 1 \dots p$ , число яких дорівнює числу коефіцієнтів передбачення  $a_k$ , і які пов'язані з ними взаємно однозначними нелінійними функціональними співвідношеннями. Коефіцієнти відбиття  $J_k$  в

силу своїх статистичних характеристик (вони мають менший динамічний діапазон) вимагають для кодування всього 6 бітів на кожен коефіцієнт. Це призводить до більш економного використання лінії зв'язку. Наприклад, якщо порядок фільтра передбачення  $p = 10$ , то загальне число бітів, що виділяються для одного сегмента, так само 72, включаючи по 6 бітів для кодування значень коефіцієнта підсилення і періоду основного тону. Можна підрахувати при цьому, що якщо значення параметрів фільтра будуть перераховуватися кожні 15 або 30 мс, то результуюча швидкість дорівнюватиме 4800 або 2400 біт / с.

Кодування коефіцієнтів відбиття може бути покращене додатково шляхом застосування нерівномірного квантування, коли до кодування значення цих коефіцієнтів піддаються нелінійному перетворенню, яке зменшує чутливість коефіцієнтів відбиття до помилок квантування. В якості такого перетворення зазвичай вибирають зворотній гіперболічний тангенс або так зване логарифмічне відношення площ:

$$r_k = \log \frac{1 - J_k}{1 + J_k}$$

Синтезуючий фільтр  $H(z)$  має ту ж структуру, що і фільтр-аналізатор (інверсний) і визначається тим же набором параметрів (коефіцієнтів передбачення  $a_k$ , або коефіцієнтів відбиття  $J_k$ , або логарифмічних відношень площ  $r_k$ ), але входи і виходи в аналізуючому і синтезуючому фільтрах міняються місцями. Якщо на вхід синтезуючого фільтра подати сигнал збудження, то на його виході буде отриманий ГС.

На основі вищевикладеного функціонують вокодери з лінійним передбаченням (ліпредери, LPC-вокодери). Перші реалізації кодеків з лінійним передбаченням були призначені для передачі даних на низьких швидкостях - 2,4 і 4,8 кбіт / с. На швидкості 2,4 кбіт / с забезпечувався прийнятний рівень розбірливості мови, проте якість, природність і впізнаваність мови були незадовільні. Оскільки цей метод сильно залежить від точного відтворення

людської мови, його реалізації, як правило, не підходять для сигналів немовного походження.

Різні варіанти алгоритмів кодування відрізняються набором переданих параметрів фільтра, методом формування сигналу збудження тощо. Узагальнена блок-схема кодера і декодера зображена на рис. 2.13.

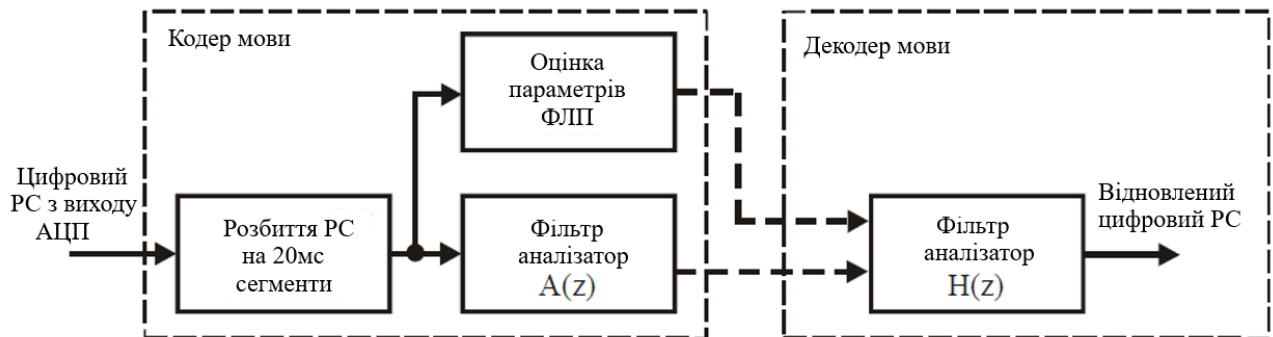


Рис. 2.13 Узагальнена блок-схема кодера / декодера

Узагальнена процедура кодування мови зводиться до наступного алгоритму:

- 1) виконується аналого-цифрове перетворення ГС;
- 2) оцифрований ГС «нарізається» на сегменти тривалістю 20 (рідше 30) мс;
- 3) для кожного сегмента обчислюються параметри фільтра лінійного передбачення і параметри сигналу збудження;

4) обчислюється залишок передбачення як вихідний сигнал фільтра з передавальною характеристикою  $A(z)$  з параметрами, отриманими з оцінки для даного сегмента; цей залишок може виступати як сигнал збудження на приймальній стороні;

5) параметри фільтра і параметри сигналу збудження кодуються за певним законом і передаються в канал зв'язку.

При декодуванні мови прийнятий або згенерований сигнал збудження пропускається через синтезуючий фільтр відомої структури, параметри якого

передані одночасно з сигналом збудження. Сигнал на вхід аналізуючого фільтра надходить безпосередньо з виходу АЦП, а вихідний сигнал, який синтезує фільтр, потрапляє на вхід ЦАП. Такий опис пояснює принцип дії вокодера, але не є технічно докладним. Практичні схеми значно складніші, і це пов'язано в основному з наступними двома моментами.

По-перше, ГС володіє двома видами внутрішніх кореляційних зв'язків - короткочасною і довготривалою надмірністю, тому в переважній більшості сучасних мовних кодеків використовується два передбачувачі: короткочасний (STP) і довготривалий (LTP) (рис. 2.14).

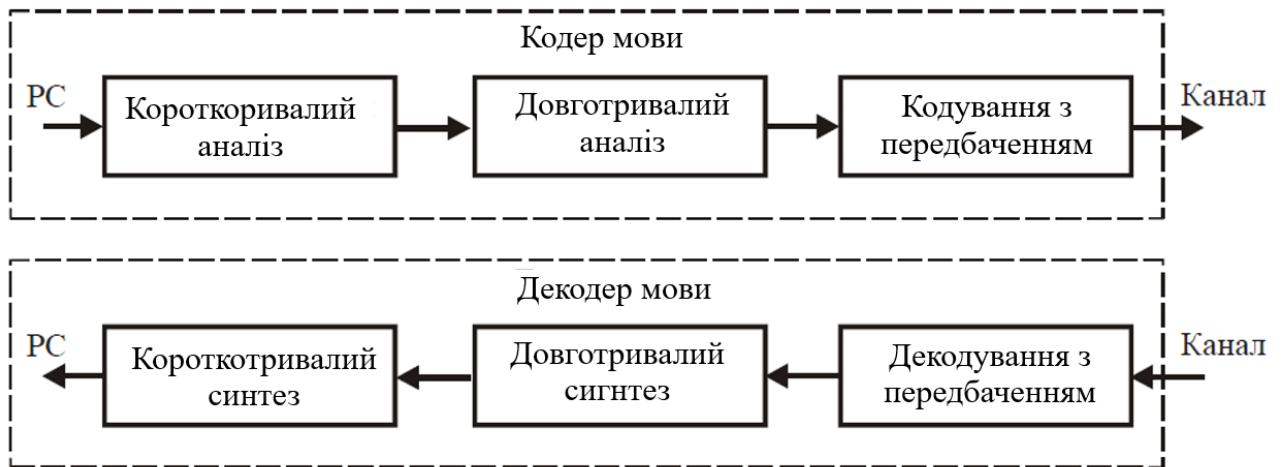


Рис.2.15 Передбачувачі мовних кодеків

Перший передбачувач (STP), що враховує короткочасну надмірність ГС, пов'язаний з кореляціями між близько розташованими відліками сигналу. Здійснюючи короткочасний аналіз, він визначає огинаючу спектра і має передавальну функцію  $1/A(z)$ , де вираз  $A(z)$  визначається співвідношенням (26). Короткотривалий передбачувач STP є основним фільтром аналізу лінійного передбачення. Адаптивна зміна параметрів цього фільтра є необхідна для відображення змін спектра ГС. Період адаптації при цьому зазвичай складає 20 ... 30 мс, а порядок передбачення  $p$  зазвичай вибирається рівним 8-12.

Другий, довготривалий передбачувач (LTP) моделює довготривалу кореляцію ГС. Здійснюючи довготривалий аналіз, він визначає точну структуру ГС і пов'язаний з кореляцією двох відрізків сигналу між собою, реально - двох сусідніх періодів основного тону.

Цей фільтр, який ще називають тоновим, також є адаптивним і зазвичай має малий (в порівнянні з STP) період адаптації - кожні 5 ... 10мс. Період основного тону мови змінюється в широких межах. На практиці забезпечується формування частоти основного тону в межах від приблизно 55 до 500 Гц. Поєднання двох передбачувачів з різними характеристиками дозволяє значною мірою усунути залишкову надмірність і наблизити залишок передбачення за своїми статистичними характеристиками до дискретного білого шуму. При цьому на приймальний бік передаються параметри сигналу збудження (залишок передбачення) і параметри обох (STP і LTP) передбачувачів.

По-друге, використання залишку передбачення в якості сигналу збудження виявляється недостатньо ефективним, так як вимагає для кодування занадто великого числа бітів. Тому практичне застосування знаходять більш економічні (по завантаженню каналу зв'язку, але аж ніяк не з обчислювальних витрат) методи формування сигналу збудження.

Таким чином, передбачувач розпадається на короткотривалий та довготривалий передбачувачі. Вони розміщуються один за іншим, причому кращі результати виходять при розміщенні довготривалого передбачувача після короткотривалого. Робиться це для усунення періодичності, яка ще зберігається в сигналі залишку короткотривалого передбачення. Таким шляхом відбувається згладжування динамічних змін в сигналі залишку.

Розглянемо структурну схему вокодера з лінійним передбаченням (ліпредера) більш детально (рис. 2.16):

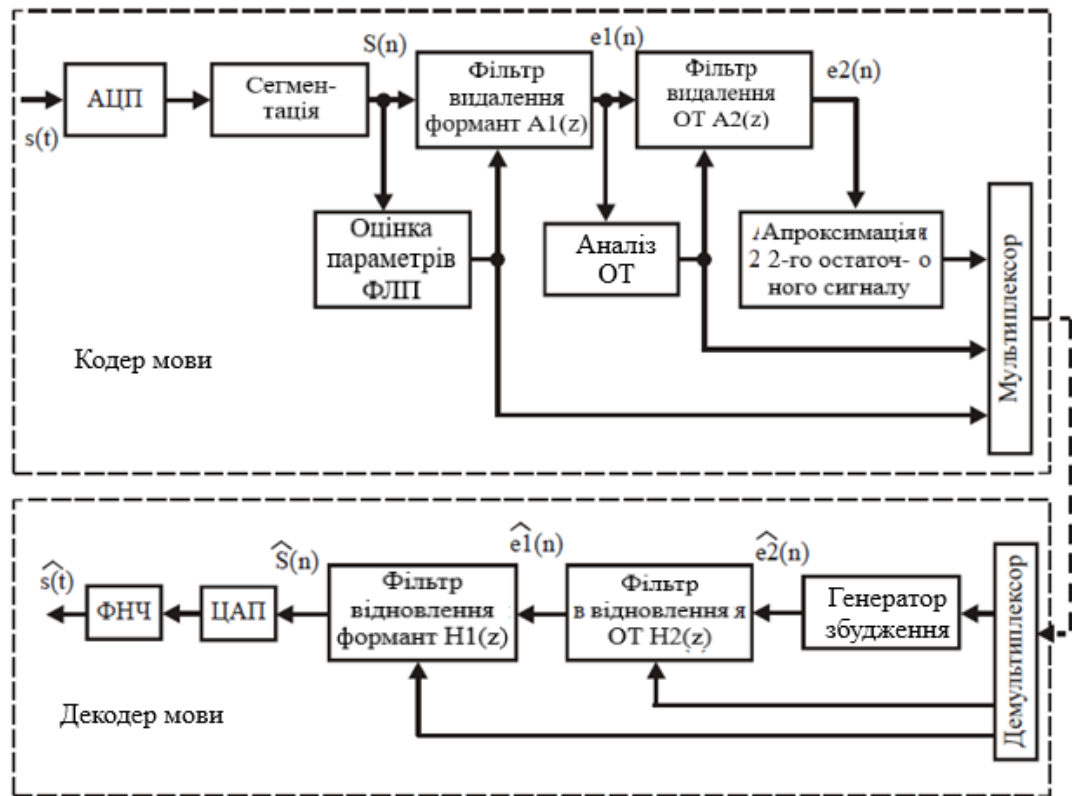


Рис. 2.16 Структурна схема ліпредера

На підготовчому етапі виконують аналого-цифрове перетворення ГС і сегментацію цифрового потоку: для подальшої обробки вибирають відліки сигналу на інтервалі тривалістю 20 мс, що при частоті дискретизації 8 кГц забезпечує в кожному сегменті 160 оброблюваних відліків. Після сегментації відліків ГС в кодері послідовно виконуються наступні три процедури:

1) короткотривалий (формантний) аналіз з використанням процедури лінійного передбачення, в результаті чого отримують перший залишковий сигнал  $e_1(n)$ ;

2) довготривалий аналіз з використанням лінійного передбачення для визначення параметрів основного тону, в результаті чого отримують другий залишковий сигнал  $e_1(n)$ , близький за своїми характеристиками до шумового, оскільки між відліками цього сигналу кореляція є малою;

3) апроксимація другого залишкового сигналу з метою формування сигналу збудження.

У першій процедурі оцінку поточного відліку  $s(n)$  визначають як суму  $p$  попередніх відліків. Порядок передбачення  $p$  вибирають рівним 8-12. Визначення коефіцієнтів передбачення  $a_k$  фільтра-аналізатора проводять в блоці формантного аналізу. Попередньо обчислені значення коефіцієнтів передбачення використовують в фільтрі видалення формант кодера, на виході якого отримують сигнал  $e_1(n)$ , вільний від квазіперіодичних складових - формант. Інформацію про форманти несуть передані на приймальний кінець параметри фільтру  $a_k$ , або пов'язані з ними коефіцієнти часткової кореляції (коефіцієнти відображення).

У другій процедурі з урахуванням того, що основний тон характеризується всього двома параметрами - амплітудою і періодом, передавальна функція фільтра видалення основного тону  $A_2(z)$  описується більш простим у порівнянні з виразом

$$A_2(z) = 1 - gz^{-d} \quad (2.18)$$

де  $g$  - єдиний коефіцієнт передбачення, що характеризує амплітуду основного тону. Затримка  $d$  визначає період основного тону. Відомо, що значення основного тону для різних голосів може змінюватися майже в 10 разів - від 2 до 18 мс. Ця обставина створює чимало труднощів при оцінці основного тону, так як слух дуже чутливий до його спотворень. Методів вимірювання основного тону відомо дуже багато, і в той же час метод, який не потребує надмірної затримки, поки не знайдений.

Незважаючи на відносну простоту виразу (2.18), аналіз і видалення основного тону є більш складною процедурою в порівнянні з формантним аналізом. Це обумовлено істотно великим періодом ОТ і складністю виявлення кореляції між відліками на великому часовому інтервалі. Крім того, період і

амплітуда основного тону дуже важливі для точного відновлення мови. Саме тому на етапі довготривалого аналізу сегмент мови поділяють на 4 підсегменти по 5 мс, що містять по 40 відліків. Параметри  $g$  і  $d$  визначають для кожного підсегмента окремо і використовують в фільтрі видалення основного тону. Їх також передають на приймальний кінець в декодер, де використовують при синтезі ГС.

Вирішувана задача третьої процедури - при мінімальному обсязі інформації про сигнал збудження забезпечити прийнятну якість відновленого сигналу. Для досягнення цього обробку другого залишкового сигналу проводять окремо для кожного підсегмента з 40 відліків.

Передані по каналу зв'язку параметри апроксимації другого залишкового сигналу, параметри основного тону  $g$  і  $d$ , коефіцієнти формантного фільтра  $a_k$  надходять на відповідні блоки декодера (рис. 2.17). У будь-якому випадку по каналу зв'язку замість самої мови передають так чи інакше виділені і квантовані параметри короткотривалого і довготривалого передбачення, амплітуду і період основного тону, параметри збудження. У декодері ліпредера по прийнятим параметрам відновлюють сигнал збудження, пропускають його через синтезуючий фільтр і відновлюють мову.

Синтез сигналу починають з відновлення другого залишкового сигналу  $e_{2\text{восст}}(n)$ , що виконується генератором збудження. Відновлений сигнал дещо відрізняється від другого залишкового сигналу  $e_2(n)$  в кодері через похибки апроксимації. Відновлений сигнал  $e_{2\text{восст}}(n)$  пропускають через фільтр відновлення основного тону, передавальну характеристику якого  $H_2(z)$  встановлюють по зворотній характеристиці (29) фільтра видалення основного тону кодера:

$$H_2(z) = \frac{1}{A_2(z)}. \quad (2.19)$$



На виході цього фільтра отримують відновлений перший залишковий сигнал  $e_{1\text{восст}}(n)$ , який включає основний тон. Нарешті, фільтр відновлення формант з передавальною функцією  $H(z)$  відновлює формантні складові сигналу.

Відновлений сигнал досить близький до вихідного сигналу на вході кодера  $s(n)$ . Виконавши цифро-аналогове перетворення і пропустивши сигнал через ФНЧ, отримують відновлений аналоговий сигнал. Всі процедури обробки сигналу в кодері і декодері виконуються цифровими методами. Показані на рис. 27 модулі ліпредера фактично є блоками програмного забезпечення.

## 2.5 Формантний вокодер

У формантних вокодерах спектральна огинаюча ГС апроксимується комбінацією декількох простих резонансних кривих. Принципи побудови форматного вокодера багато в чому аналогічні принципам природного голосотворення і прийому мови. Оскільки голосовий тракт являє собою комплекс резонаторів, резонансні частоти і добротності яких змінюються в процесі мовлення відповідно до сигналів, що йдуть з центральної нервової системи, то і в форматного вокодера відбувається виділення з ГС керуючих сигналів (сигнал-параметрів), які на прийомі впливають на резонансні контури і відтворюють необхідну огинаючу спектра. Основними параметрами форманти є частота, рівень і ширина її спектра на рівні -3 дБ. Голосні звуки повністю характеризуються формантами. Ряд приголосних звуків, особливо вибухових, характеризуються не самими формантами, а формантними переходами, тобто тенденцією і швидкістю зміни формант, а для глухих приголосних замість формант краще користуватися різного порядку моментами частотного спектра.

В існуючих формантних вокодерах з ГС виділяються не більше трьох-чотирьох формант, якщо не брати до уваги введення в деякі вокодер спеціальних пристроїв для визначення моментів спектра в області частот вище 4000 Гц,

призначених головним чином для аналізу щільних звуків. При управлінні трьома параметрами резонансних контурів (резонансною частотою, амплітудою коливань і добротністю) можна найбільш точно в порівнянні з іншими параметричними методами наближення відтворити на виході вокодера спектральну огинаючу. З урахуванням двох параметрів основного тону (частоти і рівня) і навіть чотирьох формант (кожна з трьох параметрами) необхідно передати 14 сигналів (сигнал-параметрів), не рахуючи сигналу тон-шум. Отже, за кількістю сигналів формантний вокодер тоді буде відповідати 12-канальному смуговому вокодеру. А так як для передачі формантних параметрів необхідні смуги є вузькими ніж для смугових вокодерів, щодо займаної ширини каналу перевагу матимуть формантні вокодери в порівнянні з смуговими.

Основними характеристиками формантних параметрів є їх діапазони і швидкість зміни в часі. Так, для кожної з формантних частот важливо знати розподіл її по частотному діапазону або принаймні діапазон, в якому вона майже завжди знаходиться протягом передачі звуків мови. Для рівнів кожної з формант важливо знати їх розподіл по динамічному діапазону. Це - статичні характеристики, але важлива передача і динамічних характеристик формант.

Велике значення для синтезу мови мають формантні переходи, і в першу чергу швидкість переходів, час переходу не так важливий для голосних, а для приголосних важливі обидва фактори. Для вибухових звуків сам вибух і аспірація менш важливі, ніж перехід. Важливе значення мають і такі характеристики параметрів формант, як кореляція між частотами формант, а також між їх амплітудами і часовими змінами.

Рівні формант, особливо вищих, корельовані в досить значній мірі. Так, в середньому коефіцієнти кореляції рівнів третьої і четвертої формант становлять 0,96, а першої та другої 0,77. Це означає, що практично одночасна передача рівнів двох верхніх формант необов'язкова.

З усіх опублікованих даних про ширину смуги, що зайнята формантами відомо, що ширина смуги в середньому для першої форманти голосних на рівні - 3 дБ складає 50 Гц, для другої 72 і для третьої 125 Гц. Менш визначені дані про ширину смуги формант для приголосних звуків, особливо глухих. Їх можна визначити з перехідних характеристик мовного тракту, а також по моментам спектра.

Згідно з рядом досліджень спостерігається неточність оцінки більшості приголосних звуків по формантним частотам, яка визначається як частоти максимумів спектра. З огляду на нерівномірність спектральної обвідної для цих звуків і швидкої їх зміни в часі частота максимуму спектра не завжди характеризує місце розташування формантної частоти в частотному діапазоні. А в ряді випадків при наявності декількох максимумів визначення формантної частоти взагалі є довільним. Внаслідок цього використовується метод оцінки спектрів приголосних звуків по їх моментам різних порядків ( $M_0$ ,  $M_1$  і  $M_2$ ). У той же час для деяких приголосних (зокрема, сонорних) було доцільно зберегти і оцінку по формантним частотам як частотам максимумів спектра.

Перераховані моменти могли б використовуватися і для голосних звуків, при чому, в цьому випадку їх слід було б визначати для кожної формантної області. Відповідно інтенсивність форманти замінюється моментом нульового порядку, замість формантної частоти слід брати центроїд форманти (практично це і буде зважена частота форманти), а замість ширини форманти - дисперсію спектру в цій області. У ряді практичних схем виділення параметрів формант голосних звуків застосовується цей спосіб, але в загальному випадку використання тільки моментів спектра в формантних областях неминуче обмежувало б можливості. Моменти  $M_0$ ,  $M_1$  і  $M_2$  глухих турбулентних звуків можуть розглядатися як параметри однієї з їх формант. Потужність цих звуків зосереджена головним чином в області частот понад 1200 Гц. Дзвінки приголосні

звуки характеризуються комплексом параметрів, тобто як моментами, так і формантами.

Основні труднощі при виділенні формантних параметрів пояснюються двома причинами. По-перше, частотні діапазони формантних максимумів перекриваються. По-друге, формантні частоти іноді так сильно зближуються, що сусідні форманти практично зливаються в один формантний сплеск, і тому їх неможливо розділити. Розглянемо принципи формантного аналізу мови, тобто виділення часових огинаючих її формантних параметрів (включаючи і моменти спектра) і принципи синтезу мови за цими параметрами, тобто відновлення вихідного ГС. Існують дві основні групи методів цього аналізу. За першими методами ГС після його поділу на формантні області надходить в аналізатор, який безпосередньо вимірює формантну частоту і рівень форманти в кожній формантній області. За другими методами ГС надходить на комплект паралельно або послідовно включених вузькосмугових фільтрів, що аналізують спектр мови за принципом його аналізу в вусі людини. В цьому випадку визначається місце розташування найбільших максимумів спектра і загальний рівень ГС. Всі інші методи являють собою видозміну цих двох.

З параметрів формант найбільш важливими є формантні частоти, в той час як інші параметри - рівень і ширина смуги - грають у впізнаваності звуків значно меншу роль. Тому в подальшому найбільша увага приділятиметься методам точного виділення формантних частот. Але перш за все треба уточнити поняття формантної частоти. За одними визначеннями за формантну приймається частота максимуму спектра. Але в формантній області можуть бути і інші максимуми через неточне вимовляння звуку мови і особливостей мовного тракту. За іншими визначеннями формантою вважають середню частоту спектра в заданій формантній області, тобто центроїд спектра. За третім визначенням частота форманти відповідає полюсу передавальної функції голосового тракту. Даним визначенням відповідають різні методи вимірювання формантної частоти. Одним

з найбільш простих і найпоширеніших методів вимірювання формантної частоти є метод -метра (ро-метра). Він заснований на вимірюванні частоти переходів кривої миттєвого значення звукового тиску через нуль, тобто на тому ж принципі, на якому засновано більшість частотомірів. Сигнал піддається граничному обмеженню, потім диференціюється, в результаті виходять короткочасні імпульси. Ці імпульси затягуються на 15 мс і випрямляються, отримана ступінчаста крива згладжується. Напряга на виході такого пристрою для синусоїдального сигналу пропорційна його частоті, тому при повільній зміні частоти сигналу ця напряга буде слідувати за зміною частоти. Складніше йде справа при складному сигналі. У цьому випадку напряга на виході ро-метра залежить не тільки від частот складових, але і від форми спектра.

Найбільш часті випадки неточності роботи систем виділення формантних частот за допомогою ро-метрів спостерігаються при аналізі гармонійних спектрів з двома складовими, дещо віддаленими по частоті, але близькими за амплітудою.

Для дискретного спектра у вигляді двох складових ро-метр показуватиме частоту, близьку до частоти з максимальною амплітудою. Точніше, за однакової кількості  $\omega_1 A_1 = \omega_2 A_2$  число переходів через нуль визначається середньою арифметичною частотою. Якщо  $\omega_1 A_1 < \omega_2 A_2$ , то ро-метр показує частоту, близьку до вищої, а при  $\omega_1 A_1 > \omega_2 A_2$  — до нижчої. Для більш складних спектрів покази ро-метра вже ближчі до середньоквадратичної частоти. Відомо, що похибка вимірювання частоти тим менша, чим меншим є інтервал по частоті між аналізованими складовими. На цій підставі була запропонована схема виділення формантних частот після зсуву їх спектру в область високих частот (до 20-100 кГц). Якщо спектр змістити вгору по частоті, то відносне значення інтервалів між складовими спектра відповідно зменшується. У цих умовах вимірювання формантних частот за допомогою ро-метра виходить значно точніше. Цей метод носить назву методу ВЧ ро-метра. Точність визначення частот формант даним

методом в найбільш не вигідному випадку (дві складові перебувають на краях формантного діапазону) все ж вище, ніж для НЧ ро-метра в найкращому разі.

Для телекомунікацій метод вимірювання формантних частот має багато спільного з методом ВЧ ро-метра. Використовуються основні співвідношення, одержувані при частотному детектуванні сигналів ЧМ. У схему пристрою входить амплітудний обмежувач і дискримінатор з Z-подібною частотною характеристикою. Напруга на виході дискримінатора пропорційна частоті сигналу.

Теоретичний аналіз показує, що обидва методи в своїй основі ідентичні, хоча і оперують різними величинами. Дійсно, сигнал після зсуву по частоті вгору і граничного амплітудного обмеження набуває вигляду сигналу ЧМ. Якщо дискримінатор має ідеальну характеристику, то сигнал після нього представляє похідну фази сигналу ЧМ. А ро-метр вимірює щільність переходів через нуль, тобто середню частоту сигналу ЧМ, рівну середній частоті вихідного сигналу.

Одночасно з методом ро-метра був розроблений метод визначення середньозваженої частоти спектра для приголосних звуків шляхом вимірювання коефіцієнта кореляції ГС, спектр якого зміщений в область високих частот. Основою цього способу є наступна обставина.

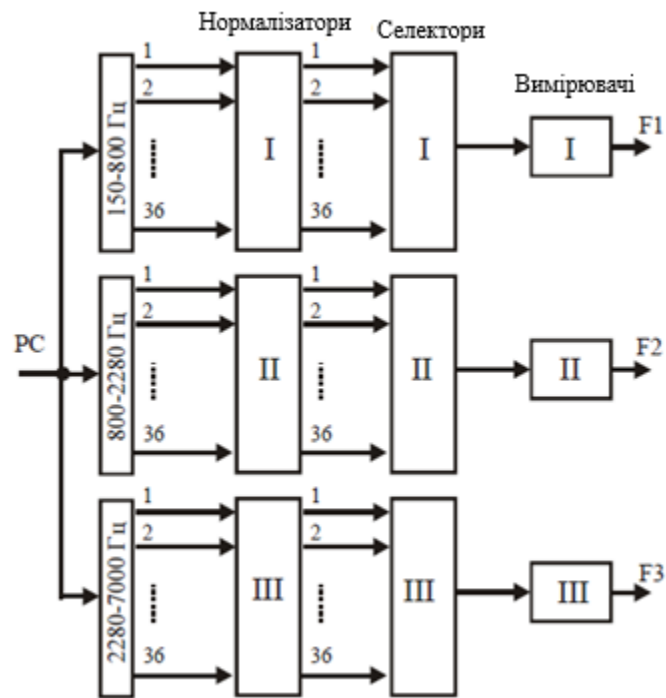
Якщо вибрати часовий зсув  $\tau$  при вимірюванні коефіцієнта кореляції таким чином, щоб будь-яка зі складових спектра, наприклад  $\omega_1$ , задовольняла умові  $\cos(\omega_1\tau) \approx \omega_1\tau - 3\pi/2$ , то коефіцієнт кореляції буде пов'язаний лінійною функціональною залежністю із середньозваженою частотою спектра. Отже, вимірюючи коефіцієнт кореляції сигналу при зсуві в часі на  $\tau$ , можна визначити середньозважену частоту спектра.

До другої групи методів вимірювання формантних частот відноситься метод фільтрації за допомогою комплексу вузькосмугових фільтрів. Можна, наприклад, застосувати комплект з 36 фільтрів з шириною смуги 100 Гц в діапазоні 150-1000 Гц і з розширенням її до 450 Гц (за шкалою Кеніга) в діапазоні

1000-7000 Гц. Один спосіб вимірювання формантних частот (рис. 2.20 а) полягає в послідовному перемиканні виходів фільтрів зі швидкістю перемикання 100 раз в секунду і фіксації місця розташування в діапазоні частот, відповідних спектральним максимумів.



а)



б)

Рис. 2.20. Способи вимірювання формантних частот

Другий спосіб (рис. 2.20 б) полягає в групуванні вузькосмугових фільтрів за трьома формантними областями I, II, III (150- 800, 800-2280 і 2280-7000 Гц). У кожній області шляхом перемикання фільтрів знаходять максимум спектра. При

цьому для дзвінких звуків третя форманта перебувала в діапазоні 2280- 3100 Гц, а для глухих - в діапазоні 3000-7000 Гц. Цей спосіб виявився точнішим порівняно з першим, хоча й трохи більш складним. Для збільшення точності вимірювання при перемиканні фільтрів застосовується електронна схема шукача максимальної напруги, що буде на виходах смугових фільтрів. Для стійкості роботи селекторів перед кожним з них включається нормалізуючий пристрій, в завдання якого входить усереднення напруг всіх фільтрів даної групи і видача в селектор тільки різниці між напругою в кожному каналі і середнім. На один вихід селектора надходить імпульс напруги на початку кожного інтервалу опитування, а на іншій - в момент появи максимальної напруги на його вході за даний інтервал часу опитування. У вимірювачі створюється імпульс напруги з тривалістю, пропорційною часу між цими двома імпульсами. Після деякого згладжування в фільтрі НЧ з частотою зрізу 10 Гц напруга на виході аналізатора буде пропорційна поточній частоті форманти.

Для подальшого підвищення точності виділення межі діапазону формант були зроблені змінними в межах ширини смуги суміжних фільтрів. У цьому випадку визначення формантних частот відбувається послідовно: спочатку визначається місце розташування самого низькочастотного максимуму, після чого визначається друга формантна частота, а потім вже третя.

Порівнюємо методи паралельного і послідовного аналізу. Перші не дають високої точності визначення частоти формант, але зате їх помилка майже не залежить від різниці в голосах і частоти основного тону. Таким чином, їх показання мають менший розкид. Друга група методів для голосів з гарною дикцією і низьким основним тоном дає високу точність, проте у них великий розкид в показаннях для різних голосів.

Одним з точних методів вимірювання формант в реальному часі є кепстральний, аналогічний методу виділення частоти основного тону, але без другого перетворення Фур'є. Для отримання частот формант сегменти ГС



послідовно подаються на аналізатор спектру зі скануванням по частоті. Тому вихідний сигнал аналізатора для кожного сегмента розгортається в часі. Цей сигнал пропускають через логарифматор.

Для дзвінкого звуку залежність має вигляд сплесків, що мають максимуми на частотах кожної з гармонік основного тону, а для глухого звуку ця залежність буде мати максимуми і мінімуми на частотах формант і антиформант відповідно (звичайно, в цьому випадку можуть бути і помилкові максимуми і мінімуми, які не відповідають їм). Якщо з'єднати між собою вершини сплесків, що відповідають частотам гармонік спектра, то максимуми цієї обвідної будуть відповідати формантним частотам з вищесказаним застереженням. Поєднуючи між собою максимуми обвідної спектра наступних один за одним сегментів ГС, отримують траєкторію для кожної формантної частоти. При цьому легко виключити помилкові форманти і врахувати близько розташовані формантні частоти.

Якщо аналізується аналоговий сигнал, то граничний пристрій, включений після логарифматора, відзначає кожен сплеск, і компаратор порівнює амплітуду сплеску з амплітудою попереднього сплеску в цьому сегменті.

Якщо амплітуда сплеску починає зменшуватися, то запам'ятовується час, що минув з моменту початку аналізу даного сегмента. Він повинен бути пропорційним до формантної частоти. Так отримують відліки можливих формантних частот. Порівнюючи їх з відповідними даними для попередніх сегментів, уточнюють значення формантних частот і виключають з них помилкові, отримуючи в результаті їх траєкторію. Звичайно, в разі близькості формант різних порядків по частоті можливі помилки, але все ж ця траєкторія, як правило, дає можливість виключення помилкових формант з точністю 3-5%. При цифровому методі сигнал з аналізатора спектра піддається дискретизації в часі та квантуванню за рівнем, і потім піддається обробці за тією ж програмою, що й для аналогового сигналу.

Аналіз мови на основі лінійного передбачення при використанні його частки оцінювання формантних частот вокалізованого сигналу має як переваги, так і недоліки. Форманти можна оцінити за коефіцієнтами передбачення двома способами. Перший полягає в факторизації полінома передбачення на основі отриманих коренів і винесенні рішення про те, які з коренів описують форманти, а які - форму спектра. Інший спосіб полягає в оцінюванні спектра і використанні методу виділення максимумів.

Особлива перевага, властиве методу лінійного передбачення в формантному аналізі полягає в тому, що як центральні частоти формант, так і їх смуги можна оцінювати досить точно за допомогою факторизації полінома передбачувача. Оскільки порядок полінома  $p$  вибирається заздалегідь, кількість комплексно-сполучених полюсів становить  $p / 2$ . Таким чином, згадана вище проблема класифікації коренів полінома з метою визначення того, які з коренів описують форманти, в даному випадку виявляється значно менш складною, ніж при використанні подібних методів, наприклад кепстрального згладжування. Крім того, побічні полюси легко усуваються внаслідок того, що смуга відповідних їм формант виявляється у багато разів більше, ніж можна було очікувати для звичайного ГС.

Недоліком методу лінійного передбачення є використання для опису спектра сигналу полюсної моделі. Так, хоча для носових звуків і виходить непоганий опис спектра полюсною моделлю, збіг коренів полінома і дійсних формантних частот неочевидний. Інша складність полягає в тому, що хоча оцінки ширини формант можна визначити з використанням отриманих коренів, проте незрозуміло, як вони співвідносяться з істинними формантами. Це пояснюється тим, що отримана оцінка залежить від розташування і тривалості інтервалу аналізу і методу аналізу.

З урахуванням цих переваг і недоліків запропонований ряд методів оцінювання формантних частот з використанням лінійного передбачення як на

основі методу виділення максимумів в спектрі, так і на основі факторизації полінома передбачувача. Після вибору сукупності формантних параметрів встановлюється відповідність між формантними параметрами і номерами формант, як і у всіх інших методах аналізу. Сюди входять і вимоги нерервності формант, необхідності передспотворень для виключення взаємного поглинання однієї форманти іншою і використання методів загострення піків в спектрі за допомогою переміщення старшого параметра лінійного передбачення до межі одиничного кола.

Більш вдалим виявився метод аналізу через синтез, запропонований на початку 1960-х років. Цей метод досить широко застосовується в сучасних системах низького кодування мови. Його суть полягає в порівнянні короткочасного спектра ГС з генерованим спектром, що змінюється по закладеній в алгоритм стратегії, яка призводить до найкращого погодження із спектром вихідного сигналу. Успішність багато в чому залежить від обраної стратегії наближення.

Як раніше згадувалося, в існуючих формантних вокодерах з ГС виділяються не більше трьох-чотирьох формант; в деяких вокодерах визначається наявність спектра в області частот вище 4000 Гц, що необхідно для аналізу щілинних і фрикативних звуків. При невеликому збитку для індивідуальності звучання обмежуються передачею параметрів тільки трьох формант, а також інтонаційних параметрів «тон-шум» і основного тону. У цьому випадку число сигналів-параметрів дорівнюватиме 8.

Структурна схема формантного вокодера показана на рис. 2.21:

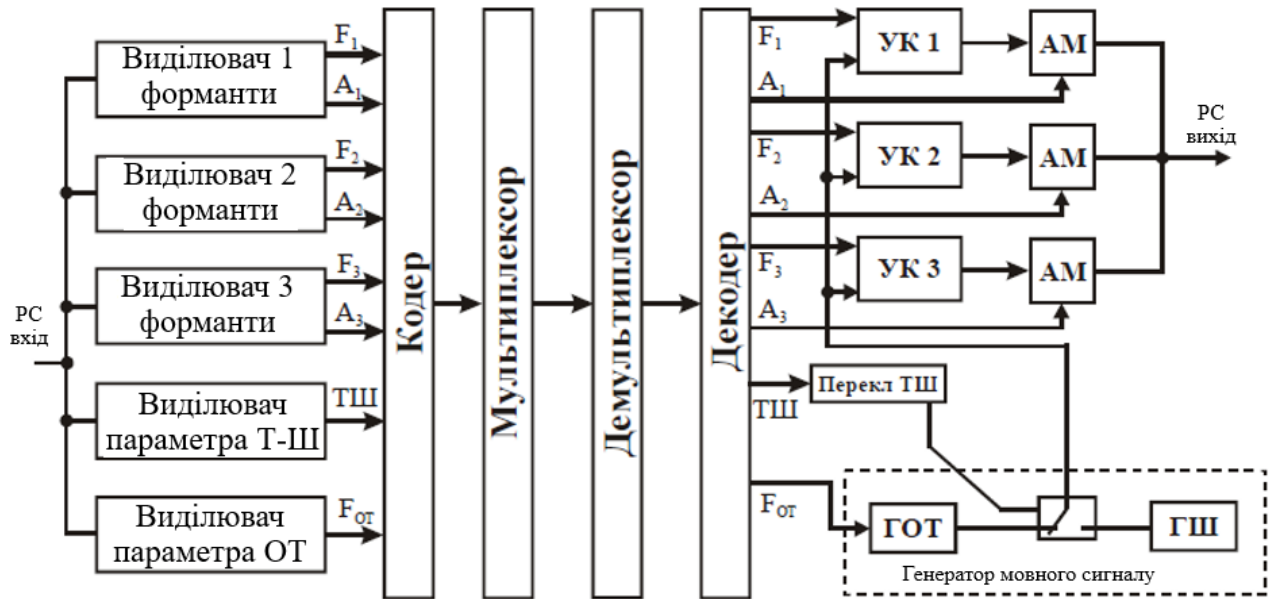


Рис. 2.21. Структурна схема формантного вокодера

На передавальній стороні в аналізаторі вокодера виділяються структурні сигнали-параметри перших трьох формант ( $A_1, F_1, A_2, F_2, A_3, F_3$ ), а також інтонаційні параметри (Т-Ш і ОТ). В якості оцінки амплітуди форманти використовується її усереднений рівень, що виділяється за допомогою смугових фільтрів, випрямлячів і ФНЧ (рис. 2.22).

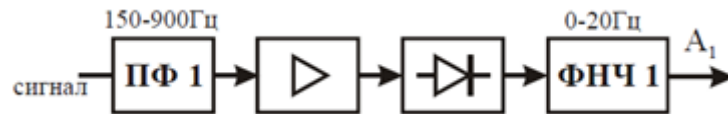


Рис. 2.22. Блок-схема виділення оцінки амплітуди форманти

Для виділення формантних частот застосовуються метод ро-метра, а також методи для телекомунікацій та фільтраційний. Синтезатор формантного вокодера (рис. 2.22) складається з трьох керованих резонансних контурів (КК) з плавною перебудовою їх частоти під надійшовших параметрів. На модуляторах відбувається взаємодія резонансного і коливального процесів в залежності від рівня відповідного сигналу параметра  $A_1, A_2$  або  $A_3$ . Вважається, що за допомогою формантного вокодера можна отримувати досить високоякісний

синтез мовлення, як по розбірливості, так і по натуральності - навіть при швидкості цифрового потоку близько 1200 біт / с. При цьому коефіцієнт компресії формантного вокодера не менше ніж удвічі вище, ніж смугового.

Сучасні вокодери забезпечують досить хорошу якість мовлення при швидкості передачі 4800 ... 2400 біт / с і якість мови, придатну для ведення службових переговорів, при швидкості передачі 1200 ... 2400 біт / с. Формантні і смугові вокодери знаходять застосування також при цифровій передачі телефонних сигналів по КХ-каналах радіозв'язку.

## 2.6 Задача виділення огинаючої голосових сигналів

Як показав огляд типів методів кодування голосових сигналів у вокодерних системах, важливим для зниження обсягів переданих даних є виділення тих інформативних параметрів ГС, яких буде достатньо для відтворення вихідного ГС на приймаючій стороні. При цьому такими параметрами відповідно до акустичної теорії мовотворення є частота основного тону та параметри огинаючої складової голосового сигналу в часі. При цьому сам голосовий сигнал розглядається як складне амплітудно модульоване коливання. Для наближення до природного звучання в структуру таких сигналів можна додатково вводити шумові складові із заданими законами розподілу. При цьому важливим є обґрунтування способу виділення огинаючої складової сигналу.

## 2.7 Висновки до розділу 2

Розглянуто принципи роботи параметричних вокодерів, зокрема смугових (канальних), ортогональних, вокодерів з лінійним передбаченням, формантних вокодерів тощо та встановлено, що важливим для зниження обсягів переданих даних є виділення тих інформативних параметрів голосових сигналів, яких буде

достатньо для відтворення вихідного голосового сигналу на приймаючій стороні. При цьому такими параметрами відповідно до акустичної теорії мовотворення є частота основного тону та параметри огинаючої складової голосового сигналу в часі. При цьому сам голосовий сигнал розглядається як складне амплітудно модульоване коливання. Для наближення до природного звучання в структуру таких сигналів можна додатково вводити шумові складові із заданими законами розподілу. При цьому важливим є обґрунтування способу виділення огинаючої складової сигналу.

## РОЗДІЛ 3

## МАТЕМАТИЧНІ ОСНОВИ ВИДІЛЕННЯ ОГИНАЮЧОЇ СИГНАЛУ

## 3.1 Метод виділення огинаючої складової сигналів

У теорії і практиці радіозв'язку широко застосовуються амплітудно-модульовані сигнали - сигнали виду

$$x(t) = A(t) \cos(\omega_0 t + \varphi),$$

де передана інформація міститься в функції  $A(t)$ , що називається обвідною сигналу  $x(t)$ . Частота  $\omega_0$  є несучою гармонійного коливання, та зазвичай набагато вище верхньої граничної частоти низькочастотного коливання  $A(t)$ , що дає привід називати такі сигнали вузькосмуговими. Наприклад, верхня гранична частота мовного сигналу близька значенням 3,5 кГц, а музичного сигналу - 20 кГц, тоді як несучі частоти сучасних радіопередавачів і радіоприймачів складають, як мінімум, сотні кГц.

Оскільки корисна інформація міститься в структурі обвідної  $A(t)$ , мета роботи АМ-радіопередачі полягає в «обчисленні» цієї обвідної за результатами спостереження сигналу  $x(t)$ .

Розглянемо найпростіший сигнал у вигляді гармоніки:

$$x(t) = A \cos \omega t,$$

огинача якого, як очевидно, дорівнює  $A$ . Оскільки  $\cos^2 \omega t + \sin^2 \omega t = 1$ , виникає думка про використання якогось допоміжного сигналу

$$\tilde{x}(t) = A \sin \omega t,$$

за допомогою якого огинача тепер може бути визначена так:

$$A = \sqrt{x^2(t) + \tilde{x}^2(t)}.$$

Оскільки

$$\tilde{x}(t) = A \sin \omega t = A \cos\left(\omega t - \frac{\pi}{2}\right),$$

приходимо до висновку, що допоміжний сигнал в даному найпростішому випадку може бути отриманий з основного шляхом його пропускання через фазозсуваючу (на  $\frac{\pi}{2}$ ) систему. Стосовно до комплексного подання гармонійних сигналів:

$$\sin \omega t = \cos\left(\omega t - \frac{\pi}{2}\right) = \frac{e^{j(\omega t - \frac{\pi}{2})} + e^{-j(\omega t - \frac{\pi}{2})}}{2} = \frac{e^{j\omega t} + e^{-j\omega t}}{2} H(\omega) = \cos \omega t H(\omega)$$

це означає, що частотна характеристика фазозсуваючої системи має вигляд:

$$H(\omega) = \begin{cases} \exp[-j(\pi/2)], & \omega > 0; \\ 0, & \omega = 0; \\ \exp[j(\pi/2)], & \omega < 0; \end{cases} = \begin{cases} -j, & \omega > 0; \\ 0, & \omega = 0; \\ j, & \omega < 0; \end{cases} = -j \operatorname{sgn}(\omega)$$



### 3.2 Визначення перетворення Гільберта

Узагальнимо розглянутий вище простий приклад на випадок довільного сигналу  $x(t)$ . З цією метою згадаємо, що за допомогою перетворення Фур'є сигнал  $x(t)$  можна представити у вигляді нескінченної суми комплексних гармонік нескінченно малої комплексної амплітуди  $X(f)df$  :

$$x(t) = \int_{-\infty}^{\infty} X(f) \exp(j2\pi ft) df , \quad (3.1)$$

Де

$$X(f) = F[x(t)] = \int_{-\infty}^{\infty} x(t) \exp(-j2\pi ft) dt . \quad (3.2)$$

Позначимо  $\tilde{X}(f)$  - перетворення Фур'є допоміжного сигналу  $\tilde{x}(t)$  :

$$\tilde{X}(f) = F[\tilde{x}(t)] = \int_{-\infty}^{\infty} \tilde{x}(t) \exp(-j2\pi ft) dt , \quad (3.3)$$

отриманого з сигналу  $x(t)$  шляхом пропускання через фазозсуваючу систему:

$$\tilde{X}(f) = (-j \operatorname{sgn} f) X(f) = H(f) X(f) \quad (3.4)$$

Тоді

$$\tilde{x}(t) = \int_{-\infty}^{\infty} \tilde{X}(f) \exp(j2\pi ft) df = \int_{-\infty}^{\infty} (-j \operatorname{sgn} f) X(f) \exp(j2\pi ft) df = \int_{-\infty}^{\infty} H(f) X(f) \exp(j2\pi ft) df \quad (3.5)$$

Комплексна функція  $H(f)$  може бути подана у вигляді:

$$H(f) = |H(f)| \exp(j\phi), \quad (3.6)$$

де

$$|H(f)| = 1, \quad f \neq 0, \quad (3.7)$$

$$\phi(f) = \begin{cases} -\pi/2, & f > 0; \\ \pi/2, & f < 0. \end{cases} \quad (3.8)$$

Таким чином, перетворення Гільберта можна трактувати як операцію над спектром сигналу, при якій амплітуда спектру практично не змінюється (за винятком придушення постійної складової), але зате фаза зазнає істотних змін - стрибком зменшується на  $\pi/2$  для позитивних частот і збільшується на  $\pi/2$  для негативних частот.

В іншому випадку, оскільки

$$F[1/\pi] = -j \operatorname{sgn} f = \begin{cases} -j, & f > 0; \\ 0, & f = 0; \\ j, & f < 0, \end{cases} \quad (3.9)$$

перетворення Гільберта функції  $x(t)$  може бути представлено у вигляді згортки  $x(t)$  і  $1/(\pi)$ :

$$\tilde{x}(t) = H[x(t)] = \int_{-\infty}^{\infty} \frac{x(u)}{\pi(t-u)} du = x(t) \otimes (1/\pi t). \quad (3.10)$$

Помноживши обидві частини співвідношення (4) на  $j \operatorname{sgn}(f)$ , отримаємо зворотне перетворення Гільберта:

$$j \operatorname{sgn}(f) \tilde{X}(f) = X(f).$$

Перетворення Гільберта дуже зручне при аналізі електричних кіл і сигналів. Нехай маємо комплексний сигнал (такий сигнал називають аналітичним):

$$z(t) = x(t) + j\tilde{x}(t). \quad (3.11)$$

представляючи сигнал  $z(t)$  у вигляді

$$z(t) = A(t) \exp[j\theta(t)], \quad (3.12)$$

приходимо до висновку, що огинаючу  $A(t)$  сигналу  $x(t)$  можна визначити співвідношенням

$$A(t) = [x^2(t) + \tilde{x}^2(t)]^{1/2}, \quad (3.13)$$

а миттєву фазу сигналу  $x(t)$  - відношенням

$$\theta(t) = \operatorname{arctg} \left[ \frac{\tilde{x}(t)}{x(t)} \right]. \quad (3.14)$$

Тоді миттєва частота сигналу  $x(t)$  визначиться співвідношенням

$$f(t) = \left( \frac{1}{2\pi} \right) \frac{d\theta(t)}{dt}. \quad (3.15)$$

Таким чином, завдяки застосуванню перетворення Гільберта дуже просто визначаються такі поняття як огибаюча, миттєва фаза і миттєва частота сигналу  $x(t)$  навіть в тих випадках, коли форма сигналу  $x(t)$  істотно відрізняється від форми гармонійного сигналу.

Перерахуємо кілька властивостей перетворення Гільберта [1].

1) Лінійна властивість:

$$H[ax(t) + by(t)] = a\tilde{x}(t) + b\tilde{y}(t) \quad (3.16)$$

для довільних функцій і постійних.

2) Властивість зсуву:

$$H[x(t - a)] = \tilde{x}(t - a). \quad (3.17)$$

3) Перетворення Гільберта:

$$H[\tilde{x}(t)] = -x(t). \quad (3.18)$$

4) Властивість парності і непарності: якщо  $x(t)$  - парна функція, тоді  $\tilde{x}(t)$  - непарна функція. І навпаки, якщо  $x(t)$  - непарна функція, тоді  $\tilde{x}(t)$  - парна функція.

5) Властивість подібності:

$$H[x(at)] = \tilde{x}(at) . \quad (3.19)$$

6) Енергетична властивість:

$$\int_{-\infty}^{\infty} x^2(t) dt = \int_{-\infty}^{\infty} \tilde{x}^2(t) dt \quad (3.20)$$

6) Властивість ортогональності:

$$\int_{-\infty}^{\infty} x(t)\tilde{x}(t) dt = 0 . \quad (3.21)$$

7) Властивість згортки:

$$H[x(t) \otimes y(t)] = \tilde{x}(t) \otimes y(t) = x(t) \otimes \tilde{y}(t) . \quad (3.22)$$

8) Відсутність коммутативності:

$$F\{H[x(t)]\} \neq H\{F[x(t)]\} , \quad (3.23)$$

тобто операції перетворення Фур'є і перетворення Гільберта можна міняти місцями.

9) Властивість модуляції:

$$\begin{aligned} H[x(t) \cos 2\pi f_0 t] &= x(t) \sin 2\pi f_0 t, \\ H[x(t) \sin 2\pi f_0 t] &= -x(t) \cos 2\pi f_0 t. \end{aligned} \quad (3.24)$$

Остання властивість справедлива для сигналів з обмеженим спектром:

$$X(f) = X(f) \left[ \text{rect} \left( \frac{f - f_0}{2f_s} \right) + \text{rect} \left( \frac{f + f_0}{2f_s} \right) \right],$$

при виконанні умови

$$f_0 > f_s.$$

### 3.3 Висновки до розділу 3

Розглянуто метод виділення огибаючої складного сигналу з використанням перетворення Гільберта.

Проаналізовано математичний апарат перетворення Гільберта та встановлено, що з допомогою такого перетворення визначати такі поняття як огибаюча, миттєва фаза і миттєва частота сигналу навіть в тих випадках, коли форма сигналу істотно відрізняється від форми гармонійного сигналу.

Проаналізовано властивості перетворення Гільберта та показано, що це перетворення є оптимальним для виділення огибаючої складової голосових сигналів.

## РОЗДІЛ 4

### ВИДІЛЕННЯ ОГІНАЮЧОЇ ГОЛОСОВОГО СИГНАЛУ З ДОПОМОГОЮ ПЕРЕТВОРЕННЯ ГІЛЬБЕРТА

#### 4.1 Програмна реалізація методу виділення огинаючої

Класичний спосіб обчислення обвідної сигналу полягає в застосуванні перетворення Гільберта:

$$C_n = (X_n^2 + XG_n^2)^{\frac{1}{2}},$$

де  $XG_n$  - перетворення Гільберта сигналу  $X_n$ .

Після стаціонаризації знову знаходимо частоту гармонійного сигналу.

Програма обчислення обвідної по Гільберту:

```
L=length(x1);
x2=zeros(1,L+100);
for n=1:L,
    x2(n)=x1(n);
end;
cn=sqrt((real(hilbert(x2))).^2+(imag(hilbert(x2))).^2);
cn(L+1:L+100)=[];
m=1:L;
plot(m,cn)
```

Графік усіченого сигналу і його обвідної по Гільберту показаний на рис.4.1.

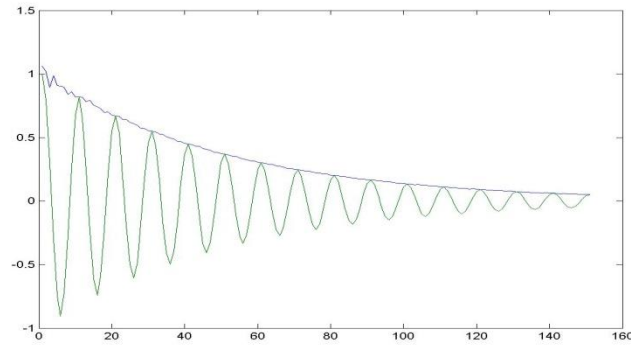


Рис. 4.1. Приклад застосування перетворення Гільберта

#### 4.2 Результати виділення огинаючої голосового сигналу

Текст програми виділення огинаючої голосового сигналу наведено в Додатку А.

Вигляд голосового сигналу, для якого проводились виділення огинаючої, наведено на рис. 4.2.

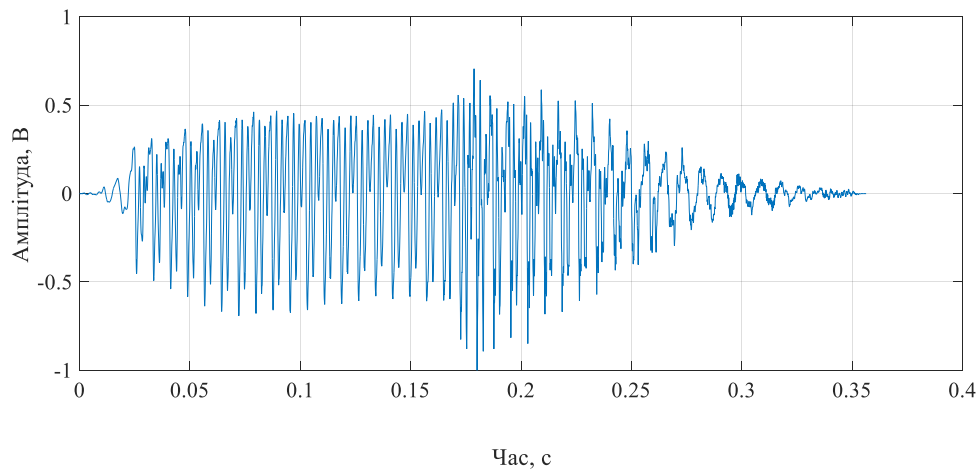


Рис. 4.2 Реєстрограма голосового сигналу

Вигляд вибірки із реєстрограми голосового сигналу наведено на рис. 4.3.



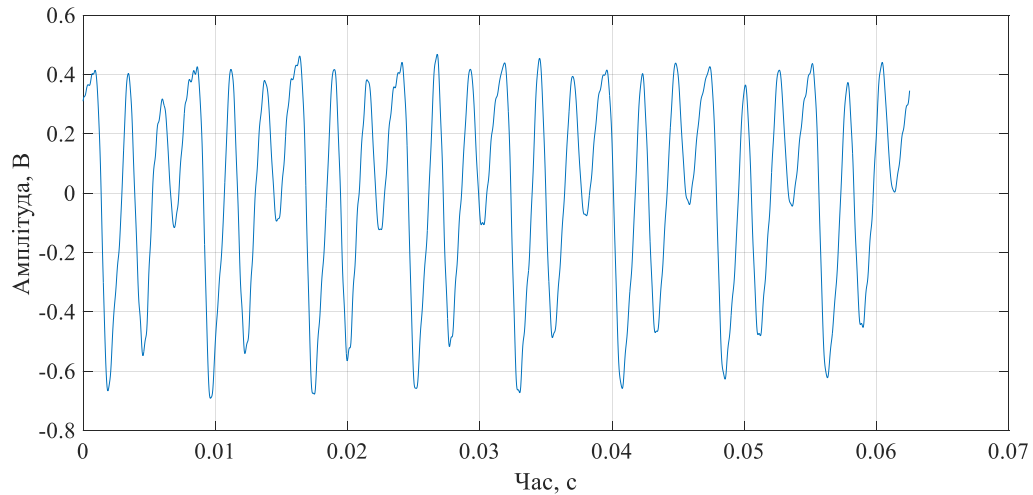


Рис. 4.3. Вигляд вибірки із реєстрограми голосового сигналу

Відповідно до рис. 4.3, несуча складова голосового сигналу є близькою до гармонічної, відповідно в спектрі вона буде проявлятися у вигляді форманти. Обвідна буде проявлятися на низьких частотах спектру сигналу. Спектр сигналу наведено на рис. 4.4.

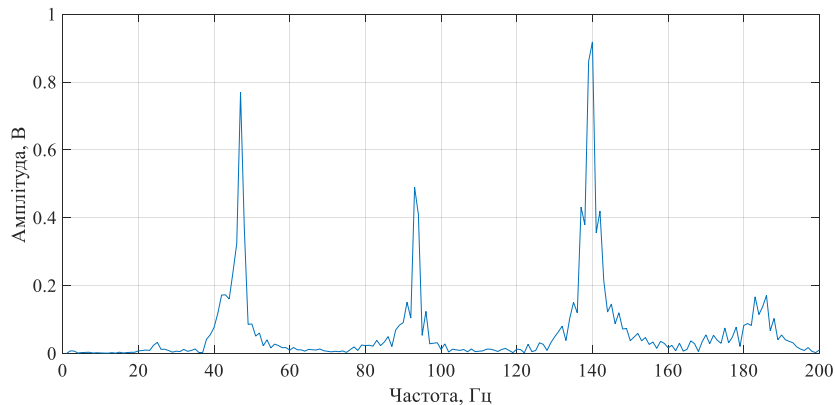


Рис. 4.4. Спектр голосового сигналу

Перший максимум характеризує параметри обвідної складової голосового сигналу.

Застосуємо перетворення Гільберта до вихідного голосового сигналу.

Результат застосування перетворення Гільберта наведено на рис. 4.5.

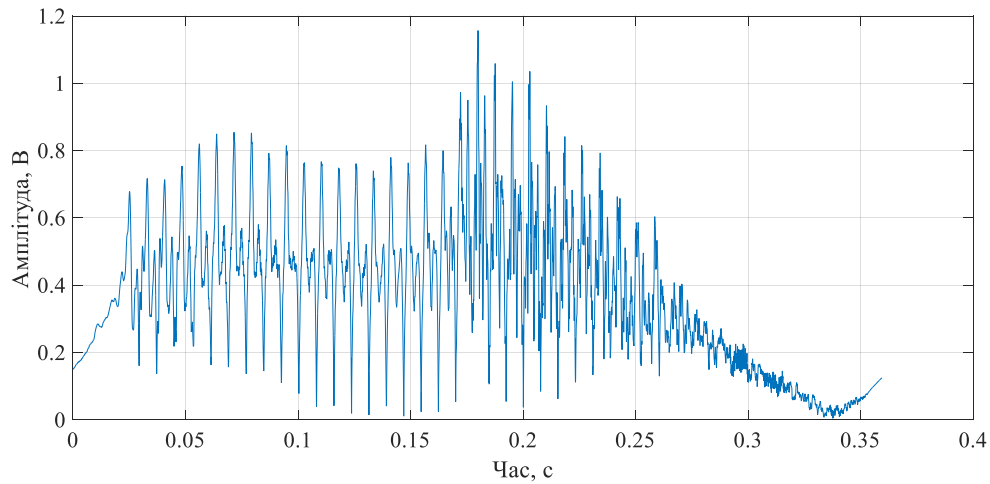


Рис. 4.5. Результат застосування перетворення Гільберта

Спектр отриманого сигналу наведено на рис. 4.6.

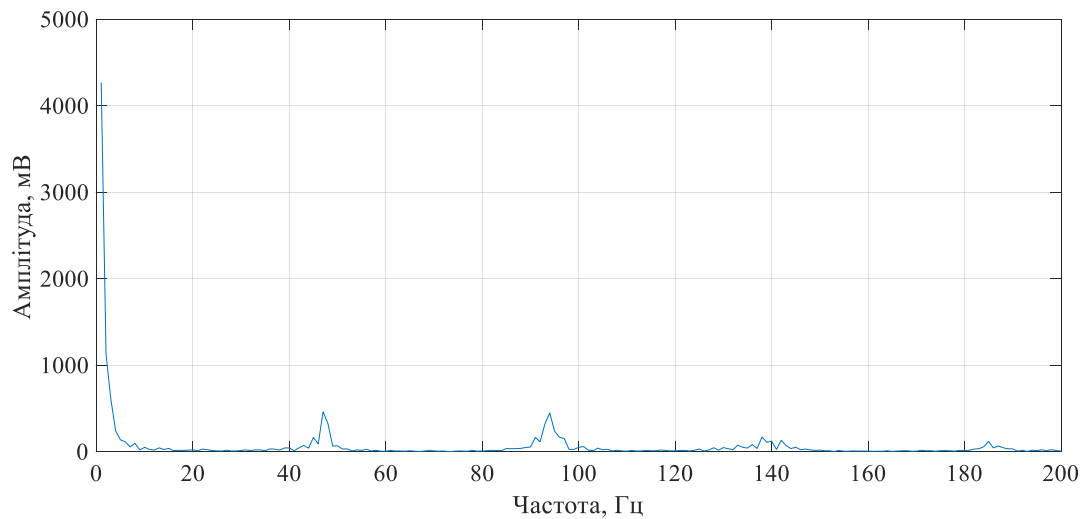


Рис. 4.6. Спектр результуючої огибаючої

Відповідно до рис. 4.6, на частотах більше 40 Гц розміщуються залишкові складові формант сигналу. Для згладжування форми огибаючої необхідно її пропустити через фільтр низьких частот з частотою зрізу 40 Гц.

Результат фільтрації перетворення Гільберта низькочастотним фільтром з частотою зрізу 40 Гц наведено на рис. 4.7.

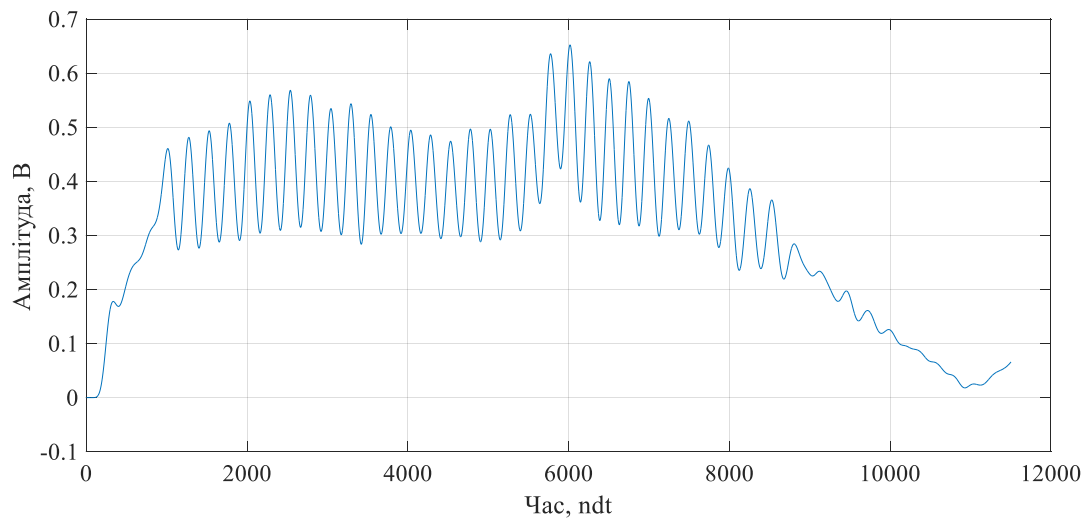


Рис. 4.7. Результат фільтрації перетворення Гільберта низькочастотним фільтром з частотою зрізу 40 Гц

Результат фільтрації перетворення Гільберта низькочастотним фільтром з частотою зрізу 20 Гц наведено на рис. 4.8.

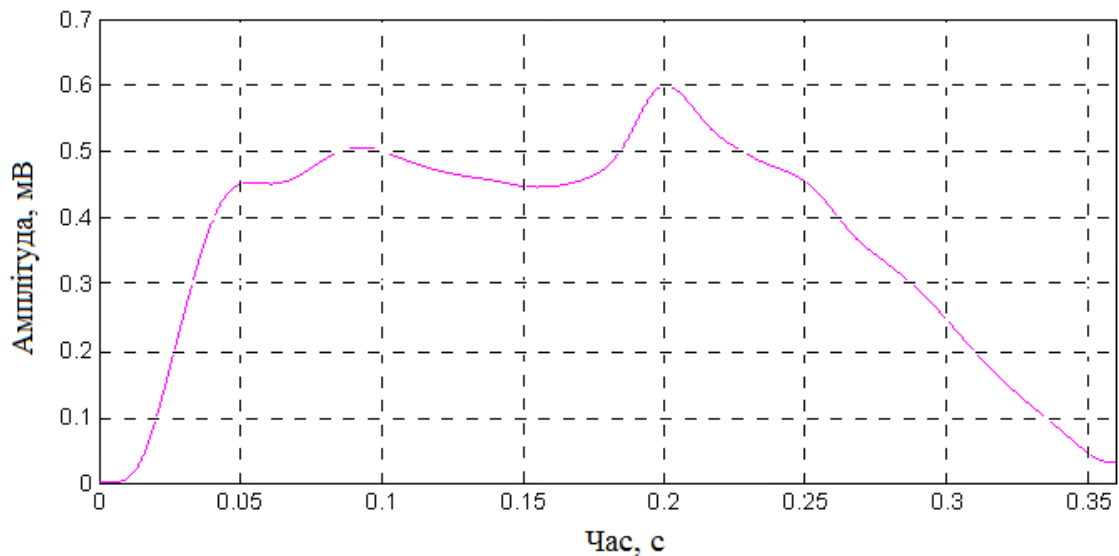


Рис. 4.8. Результат фільтрації перетворення Гільберта низькочастотним фільтром з частотою зрізу 20 Гц

Результат фільтрації є задовільним і дозволяє проводити перетворення Гільберта для виділення огинаючої складової голосового сигналу.

### 4.3 Висновки до розділу 4

Завантажено в середовище Matlab голосовий сигнал. Встановлено, що несуча складова голосового сигналу є близькою до гармонічної, відповідно в спектрі вона буде проявлятися у вигляді форманти. Обвідна буде проявлятися на низьких частотах спектру сигналу.

Застосовано перетворення Гільберта до вихідного голосового сигналу. Встановлено, що перший максимум характеризує параметри обвідної складової голосового сигналу. На частотах більше 40 Гц розміщуються залишкові складові форманти сигналу. Для згладжування форми огинаючої пропущено результат перетворення через фільтр низьких частот з частотою зрізу 40 Гц. Потім додатково застосовано низькочастотну фільтрацію сигналу фільтром з частотою зрізу 20 Гц.

Результат фільтрації є добрим і дозволяє застосувати перетворення Гільберта для виділення огинаючої складової голосового сигналу.

## РОЗДІЛ 5

### СПЕЦІАЛЬНА ЧАСТИНА

#### 5.1 Методика проведення медико-біологічних досліджень

У розділі наведено обґрунтовано структуру системи для відбору голосових звуків для задачі корекції звуковимови людини, а також правила підготовки до дослідження і методикою його проведення.

Перед початком проведення експерименту для відбору голосових звуків необхідно провести його планування [82, 83]. Воно передбачає таку організацію експериментальних досліджень, яка дасть можливість зібрати необхідні дані, використати для їх аналізу статистичні методи і зробити об'єктивні висновки. Проведення експерименту, відповідно, повине включати наступні етапи [82]:

- формулювання задачі. На цьому етапі необхідно зкоректувати всі уявлення про мету експерименту;
- вибір факторів і рівнів. На цьому етапі необхідно визначити незалежні змінні і фактори, які будуть досліджуватися в експерименті, а також значення та рівні цих факторів;
- проведення експерименту;
- аналіз даних.

Перші два етапи проведення експерименту були розглянуті і проаналізовані в першому та другому розділах. На третьому етапі необхідно провести вибір схеми активного експерименту для проведення відбору сигналів та обґрунтувати характеристики технічних засобів, що в неї входять.

Акустичний сигнал від суб'єкта перетворюється мікрофоном в електричний та подається на вхід звукової картки комп'ютера. Після попереднього опрацювання сигнал відображається на екрані монітора.

Важливим є питання обґрунтування вибору конкретного типу мікрофона та його технічних характеристик з метою врахування суттєвих з точки зору діагностичної цінності характеристик голосових звуків та забезпечення коректності результатів їх наступного опрацювання.

Відбір виконується мікрофонами електродинамічного типу. Необхідно розглянути вимоги, яким повинен відповідати мікрофон для реєстрації цих звуків та обґрунтувати вибір конкретного його типу.

5.1.1. Обґрунтування вибору типу мікрофона. Наведено класифікацію типів мікрофонів відповідно до способу перетворення акустичних коливань.

До основних характеристик і параметрів мікрофонів, що визначають їх якість та область застосування, відносяться наступні [87]:

- чутливість - відношення напруги на виході мікрофона до звукового тиску, що діє на нього;
- динамічний діапазон - різниця між рівнями граничного звукового тиску і власних шумів;
- робочий частотний діапазон;
- частотна характеристика;
- характеристика направленості - залежність чутливості мікрофона від кута між його акустичною віссю і напрямом на джерело звуку.

Одним з найважливіших показників при виборі мікрофона є характеристика його просторової направленості. Графічно її зображають в полярних координатах у вигляді діаграми направленості в горизонтальній площині.

За виглядом характеристики направленості мікрофони поділяють на три основні типи: ненаправлені, двосторонньо і односторонньо направлені. Для прикладу, в Додатку Б наведено характеристики направленості поширених мікрофонів МКЭ-2 та МД-78. Перший є мікрофоном конденсаторного типу,

другий – електродинамічного.

Для відбору звуків необхідно використати ненаправлений або односторонньо направлений мікрофон, що зменшить негативний вплив зовнішніх акустичних завад (шумів).

Поряд з діаграмою направленості, іншою не менш важливою характеристикою мікрофона є його частотна характеристика. Принциповою вимогою до частотної характеристики є її рівномірність.

Аналізуючи частотний діапазон ряду електродинамічних мікрофонів, встановлено, що їх частотний діапазон становить в середньому 60-12 кГц, тоді як у випадку приголосних звуків важливим є діапазон низьких частот 20-200 Гц, де розміщуються перші форманти, а у випадку свистячих звуків – частоти до 14 кГц. Крім того, виявлено, що в області низьких частот мікрофони мають значний спад частотної характеристики, що призводить до зниження амплітуди першої форманти, яка для чоловічого голосу може знаходитись на частотах в діапазоні 70-200 Гц.

На основі проведеного аналізу приймаємо рішення про використання для відбору голосових звуків мікрофона конденсаторного МКЭ-2.

Сигнал з мікрофона подається на вхід звукової карти, де проходить попереднє опрацювання та оцифрування сигналу. Проведемо обґрунтування основних параметрів АЦП звукової картки.

5.1.2. Обґрунтування вибору параметрів АЦП звукової картки. Основними параметрами АЦП звукових карток є розрядність та частота дискретизації вхідного сигналу.

Так, для восьмирозрядного АЦП звукової карти  $N=256$ ,  $D=48$  дБ, для шістнадцятирозрядного АЦП  $N=65536$ ,  $D=96$  дБ. АЦП з більшою розрядністю (18 та 20) використовуються в професійних звукових картках для комплексної обробки звуку, монтажу, конвертування тощо.

Від числа рівнів квантування, а відповідно і розрядності залежить похибка квантування (шум квантування).

Для нормованого сигналу відносна величина максимальної похибки квантування рівна  $1/N$ . Цією ж величиною, представленою в логарифмічних одиницях (децибелах), оцінюється рівень шумів квантування АЦП звукової картки.

Для трьохрозрядного АЦП  $N=8$ , і  $D=-18$  дб; для восьмирозрядного -  $N=256$ ,  $D=-48$  дб; для шістнадцятирозрядного -  $N=65536$ ,  $D=-96$  дб; для вісімнадцятирозрядного АЦП  $N=262144$ ,  $D=-108$  дб; і для двадцятирозрядного АЦП  $N=1648576$ ,  $D=-120$  дб. Ці цифри наочно демонструють, що із зростанням розрядності АЦП шум квантування зменшується. Прийнятним вважається 16-розрядне представлення сигналу, що є на сьогодні стандартним для відтворення звуку, записаного в цифровій формі. З погляду зниження рівня шумів квантування подальше збільшення розрядності АЦП не є доцільним, оскільки рівень шумів, що виникають з інших причин (теплові шуми, а також імпульсні перешкоди, що генеруються елементами схем комп'ютера і розповсюджуються або по колах живлення, або у вигляді електромагнітних хвиль), все одно виявляється значно вищим, ніж -96 дб.

Вибір частоти дискретизації сигналу проведемо на основі теореми про відліки та технічних характеристик мікрофона.

Людський голос здатний відтворити звук, частотою до 20 кГц. Враховуючи те, що вибраний для запису звуків мікрофон, згідно з паспортними даними, працює в діапазоні частот до 16 кГц, частота дискретизації повинна становити не менше 32 кГц.



## 5.2 Обґрунтування вибору УДК напряму наукового дослідження

Універсальна десяткова класифікація (УДК) є міжнародною системою класифікації документів. Вона відповідає найістотнішим вимогам до класифікації (міжнародність, універсальність) та надає можливість відобразити новітні досягнення науки й техніки без будь-яких суттєвих змін в її структурі. Такої гнучкості не має жодна з існуючих систем класифікації.

Наявність детально розробленої системи допоміжних таблиць визначників, здатність відобразити нові поняття за допомогою розподілу рубрик від загального до конкретного також роблять систему УДК гнучкою. Це дає змогу багатоаспектно розкривати зміст матеріалів за допомогою комбінування індексів. Застосування визначників безмежно розширює можливості класифікації та відкриває нові для детальної класифікації матеріалу.

В основі структури УДК – принцип десяткових дробів. Для позначення рубрик застосовують арабські цифри, зрозумілі в усіх країнах, що робить УДК загальнодоступною міжнародною системою. Десятковий принцип структури дає змогу безмежно розширювати її за допомогою приєднання нових цифрових позначень до існуючих, не змінюючи системи загалом.

Індекси УДК побудовані так, що кожна наступна цифра, що приєднується до індексу, не змінює попереднє значення, а лише уточнює, позначаючи конкретніше поняття.

Отже, тема наукового дослідження включає у своїй структурі дві сторони: Тема включає у своїй структурі дві сторони:

- 1) Медична (голосові сигнали).
- 2) Метод статистичного опрацювання

Згідно з класифікатором УДК (сайт - <http://teacode.com/online/udc/>), **медична сторона** класифікується наступною послідовністю дій, при виборі номера:

- 1) УДК 61 - Медицина. Охорона здоров'я. Пожежна справа;
- 2) УДК 612 - Фізіологія. Порівняльна фізіологія;
- 3) УДК 612.78 - Голосовий апарат. Голос. Мова

**А технічна сторона** класифікується наступною послідовністю:

- 1) УДК 51 - Математика
- 2) УДК 519.2 - Теорія ймовірностей і математична статистика
- 3) УДК 519.21 - Теорія ймовірностей і випадкові процеси
- 4) УДК 519.218 - Випадкові процеси спеціального виду

Отже, загальний номер УДК буде мати наступний номер:

**УДК 612.7:519.218**

Отже, у розділі описано методіку проведення медико-біологічного дослідження та обґрунтовано вибір УДК тематики за напрямом наукового дослідження

## РОЗДІЛ 6

### ОБГРУНТУВАННЯ ЕКОНОМІЧНОЇ ЕФЕКТИВНОСТІ

#### 6.1 Науково-технічна актуальність науково-дослідної роботи

Наукові дослідження, які є основою наступних стадій інноваційних процесів, класифікують по трьом видам: фундаментальні, пошукові та прикладні.

Фундаментальні дослідження проводять з метою отримання систематизованих даних щодо певної науково-технічної проблеми, виявлення нових закономірностей і принципів розвитку світу, обґрунтування нових понять, створення нових теорій.

Пошукові дослідження розвивають фундаментальні розробки з метою їх практичної використання, тобто вони спрямовані на конкретний науково-технічний результат.

Прикладні наукові дослідження, в свою чергу, базуються на пошукових і проводяться для розробки нових чи удосконалення існуючих технологічних процесів; створення матеріалів з особливими властивостями; принципово нових зразків машин, обладнання, приладів, оснащення, високотехнологічних наукомістких виробництв.

І, нарешті, розробки – технологічні, дослідно-конструкторські, проектні, організаційні роботи, які включають створення техніко-економічної документації для освоєння нововведень (нових технологій, нової продукції та виробництв, споруд, прогресивних методів організації та управління виробництвом) та їх дослідно-експериментального випробування.

Основне завдання економічного обґрунтування – довести, що тема досліджень, яку опрацьовує магістрант, має, перш за все, наукову, технічну, а також економічну, соціальну або екологічну значущість і сприяє тим самим зростанню темпів науково-технічного прогресу в цілому. З цією метою

акцентується увага на масштабах виробництва і використання продукції, на підвищення якості або удосконалення виробництва якої направлена тема магістерської роботи.

У разі, коли дослідження має фундаментальний або фундаментально-пошуковий характер необхідно висвітлити науково-технічне значення даної сфери знань та перспективи, які розкривають дослідження по темі магістерської роботи.

Ця частина економічного розділу повністю формується на основі критичного опрацювання фахових публікацій останніх років, які присвячені питанням, що стосуються теми дослідження. Всі викладки цієї частини повинні спиратись на конкретні кількісні оцінки експлуатаційних та технологічних властивостей матеріалів та виробів, обсягів їх виробництва та використання, режимів технологічних процесів, ринкової вартості виробів та технологічних матеріалів, сировини, енергоресурсів тощо з відповідним посилками в тексті на першоджерела.

Результатом цього розділу має стати чітко сформульована науково-технічна проблема, на вирішення якої повинна бути направлена дана дослідницька робота. Таким чином, сформульована проблема і тема науково-дослідницької роботи повинні знаходитись у логічній єдності між собою.

## 6.2 Розрахунок витрат на проведення науково-дослідної роботи

Розрахунок усіх витрат організації-виконавця НДР, пов'язаних з виконанням теми, дає можливість встановити її собівартість або кошторисну вартість. Кошторис розробляє виконавець робіт на основі календарного плану проведення досліджень і затверджує замовник або орган, що забезпечує фінансування робіт. Як правило, кошторис складається до початку виконання робіт і тому називається плановим.

6.2.1 Витрати на оплату праці. Витрати за цією статтею включають заробітну плату безпосередніх виконавців теми, а заробітна плата адміністративно-управлінського персоналу, працівників дослідних виробництв включаються в кошторисну вартість теми через статтю «Накладні витрати». Крім цього, слід враховувати, що для тем, які фінансуються за рахунок держбюджету прибуток не планується і тому в дану статтю витрат включається тільки основна заробітна плата (без премій та інших виплат, що здійснюються із прибутку). Витрати на оплату праці розраховують на основі даних про трудомісткість окремих робіт по темі (табл. 1.1) та посадових окладів безпосередніх їх виконавців.

Загальна трудомісткість робіт, що виконуються безпосередньо студентом (інженером - дослідником), визначається навчальним планом відповідного напрямку підготовки.

Таблиця 6.1

Трудомісткість робіт по темі НДР

Найменування робіт по темі дослідження	Трудомісткість за виконавцями, людино-днів					
	Провідний науковий співробітник	Старший науковий співробітник	Молодший науковий співробітник	Інженер	Лаборант	Студент
1. Уточнення та конкретизація завдань по темі дослідження	2	2	1	–	–	2
2. Аналіз науково-технічних публікацій з теми	1	2	3	–	–	7
3. Розроблення математичної моделі сигналу	3	3	4	–	–	5
4. Розроблення Компонентного методу опрацювання сигналу	3	3	4	–	–	5
5. Експериментальні дослідження сигналу	2	2	2	2	2	2
6. Формування звіту по НДР	5	7	7	7	7	7
Разом за виконавцями теми	16	20	21	9	9	28

Подальші розрахунки витрат на оплату праці проводиться за алгоритмом, зрозумілим із табл. 6.2.

Середньоденна заробітна плата за категоріями виконавців розраховується шляхом ділення їх посадового місячного окладу на 21,2 (де 21,2 – усереднене число робочих днів за місяць).

Таблиця 6.2

## Розрахунок витрат на оплату праці

Посада виконавців теми	Планова трудомісткість, люд-днів	Заробітна плата, грн		
		Посадовий місячний оклад	Середньоденна зарплата	Усього за виконавцями
1.Провідний науковий співробітник	15	4289,70	202,34	3035,10
2.Старший науковий співробітник	18	3334,80	157,30	2831,40
3. Молодший науковий співробітник	19	1802	85	1615
4. Інженер	8	1683	79,39	635,12
5. Лаборант	8	1302	61,42	491,36
6. Студент	28	1302	61,42	1719,76
Разом оплата праці з теми				10327,74

6.2.2 Відрахування на соціальні заходи. До цієї статті витрат належать виплати у вигляді єдиного соціального внеску, які здійснює організація – виконавець теми в пенсійний фонд в розмірі 37,26%, що становить 3848,12 грн. від загальних витрат на оплату праці.

Базою вказаного нарахування слугують загальні витрати на оплату праці по темі (табл.6.2).

6.2.3 Обладнання, необхідне для проведення досліджень. В даній статті враховують вартість усіх видів матеріалів, необхідних для проведення НДР, з вирахуванням вартості зворотних відходів.

Тематика дослідницьких робіт, які виконуються на факультеті контрольно-вимірювальних та радіокомп'ютерних систем, передбачає використання, перш за все, комп'ютерної діагностичної системи, комп'ютерів для опрацювання кардіосигналів сигналів та формування матеріалів звітності, оргтехніки та інші.

Розрахунки зведено за формою у табл.6.3

Таблиця 6.3

Розрахунки витрат на обладнання

Найменування обладнання	Одиниця виміру	Кількість	Ринкова ціна за одиницю, грн	Сума,грн.
1. ПК (системний блок, монітор, клавіатура, мишка, кабель живлення)	шт	1	8200	8200
2. Принтер лазерний	шт	1	1600	2200
3. Кабель для підключення до ПК	шт	1	120	120
Загальні витрати на матеріали				10520

6.2.4 Енергоносії для проведення досліджень. На підприємстві електроенергія використовується для освітлення, живлення медобладнання, комп'ютерної техніки та оргтехніки.

$$Z_{cm} = \sum_{i=1}^n P_i \cdot k_i \cdot t_i \cdot C_i,$$

(6.1)

де  $P_i$  – витрата  $i$ -го виду матеріального ресурсу, натуральні одиниці;

$C_i$  - ціна за одиницю  $i$ -го виду матеріального ресурсу, грн;

$k_i$  – коефіцієнт використання потужності  $i$ -го виду матеріального ресурсу;

$t_i$  – час роботи  $i$ -го виду матеріального ресурсу;

$i$  - вид матеріального ресурсу;

$n$  - кількість видів матеріальних ресурсів.

Якщо для проведення НДР використовується електрообладнання, то необхідно розрахувати витрати на електроенергію за формою (6.1), наведеною в таблиці 6.4.

Таблиця 6.4

Витрати на електроенергію

Найменування обладнання	Паспортна потужність, Вт	Коефіцієнт використання потужності	Час роботи обладнання для розробку АІС, год	Ціна електроенергії, Грн/ (кВт/год)	Сума, грн.
ПК (системний блок, монітор, клавіатура, мишка, кабель живлення)	265	0,5	102	2,5	4700,63
Принтер лазерний	700	0,25	3	1,25	656,25
Лампи розжарювання (освітлення)	150	0,85	10	1,25	1593,75
РАЗОМ витрати на електроенергію					6950,63

6.2.5 Витрати на службові відрядження. Дані витрати складаються із фактичних витрат на службові відрядження штатних працівників, зайнятих виконанням НДР: витрат на проїзд до місця відрядження і назад; витрат на проживання у готелі; добових витрат, які розраховуються на кожний день перебування у відрядженні, враховуючи час перебування в дорозі, та деякі інші.



Під час виконання НДР здійснюються ряд відряджень, які пов'язанні із доповідями на конференціях, які наведено у таблиці 6.5.

Таблиця 6.5

## Приблизні витрати на службові відрядження

Тип відрядження	Кількість	Приблизна вартість відрядження
Конференція	5	1000
Здача звітів НДР	1	200
Впровадження результатів НДР	3	300
Всього	—	1500

6.2.6. Розроблення планової калькуляції кошторисної вартості теми. Планова калькуляція вартості проведення досліджень по темі складається на підставі виконаних розрахунків та нормативних даних (табл.6.6).

Таблиця 6.6

## Планова калькуляція кошторисної вартості НДР

Найменування статей витрат	Сума, грн	Обґрунтування
1	2	3
1.Витрати на оплату праці	10327,74	Відповідно до розрахунків
2.Відрахування на соціальні заходи	3848,12	Відповідно до діючих загальнодержавних нормативів
3.Обладнання для проведення досліджень	30890	Відповідно до розрахунків
4.Енергоносії для проведення досліджень	4615,22	Відповідно до розрахунків
5.Витрати на службові відрядження	1500	Відповідно до розрахунків
6.Інші невраховані прямі витрати по темі	5118,1	10% від суми прямих розрахованих витрат по темі
7.Кошторисна вартість теми	56299,19	Сума попередніх статей

Кінцевим результатом науково-дослідницьких робіт є досягнення наукового, науково-технічного, економічного, соціального, екологічного та інших видів ефектів.

### 6.3 Науково-технічна ефективність науково-дослідної роботи

Економічна оцінка фундаментальних і пошукових НДР у вартісному вимірі, як правило, неможливо, бо ймовірність доведення результатів таких досліджень до конкретного практичного застосування невелике. Для таких досліджень рекомендується визначати науковий та науково-технічний ефект, який враховує результати наукових досліджень та їх значущість для прискорення науково-технічного прогресу та розвитку національної економіки.

Науковий та науково-технічний ефект рекомендується оцінювати коефіцієнтом науково-технічної ефективності ( $E_{нт}$ ) за допомогою формули:

$$E_{нт} = \frac{\sum B_i \cdot B_{ij}}{\sum B_i \cdot B_{ij}^{\max}}, \quad (6.2)$$

де  $B_i$  – нормативні значення коефіцієнтів вагомості факторів науково-технічної ефективності (табл. 6.7);

$B_{ij}$  – середнє значення балу, який виставляється експертами  $i$ -му фактору;

$B_{ij}^{\max}$  – максимально можливе значення балу (табл. 6.8);

$i$  – порядковий номер фактору;

$j$  – відповідна характеристика  $i$ -го фактора.

Нормативні значення коефіцієнтів вагомості факторів науково-технічної ефективності наведені в табл. 6.7.

Нормативні значення коефіцієнтів вагомості факторів  
науково-технічної ефективності

Фактори ( $i$ )	Коефіцієнти вагомості ( $B_i$ )
1.Новизна очікуваних або одержаних результатів	0,25
2.Глибина наукового опрацювання	0,16
3.Ступінь ймовірності успіху	0,09
4.Перспективність використання результатів	0,25
5.Масштаб можливої реалізації результатів	0,15
6.Завершеність одержаних результатів	0,10
Разом	1,00

Характеристика факторів науково-технічної ефективності НДР наведена в табл. 6.8.

Таблиця 6.8

Характеристика факторів науково-технічної ефективності НДР

Фактор наукової та науково-технічної ефективності	Характеристика фактора	Оцінка фактора	
		Якісна	Бальна $A_{ij}^{\max}$
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
1.Новизна одержаних або передбачуваних результатів	Одержані принципово нові результати, раніше невідомі в науці, розроблена нова теорія, відкрита нова закономірність	Висока	10
	Встановлені деякі часткові закономірності, методи, способи, які дозволяють створити принципово нові види техніки	Середня	7
	Позитивне вирішення поставлених задач на підставі простих узагальнень, аналіз зв'язків між факторами, розповсюдження відомих наукових принципів на об'єкти	Недостатня	3
	Опис окремих елементарних фактів, передача та поширення отриманих раніше результатів, реферативні огляди	Тривіальна	1

## Продовження таблиці 6.8

1	2	3	4
2.Глибина наукового опрацювання	Проведена значна кількість експериментів по нетрадиційним методикам, виконані складні теоретичні розрахунки, підтверджені експериментальними даними	Істотна	10
	Проведена обмежена кількість розрахунків по відомим методикам, виконані теоретичні розрахунки невисокої складності, частково перевірені експериментальними даними	Середня	6
	Проведена недостатня кількість експериментів, виконані прості теоретичні розрахунки без експериментальної перевірки	Несуттєва	1
3.Стіпень ймовірності успіху	Висока ймовірність повного вирішення поставлених задач НДР	Значна	10
	Середня ймовірність вирішення більшості експериментальних або теоретичних задач	Помірна	6
	Низька ймовірність вирішення поставлених задач, отримання позитивних результатів сумнівне	Незначна	1
4.Масштаб використання результатів	Результати можуть бути використані в багатьох наукових напрямках, мають значення для розвитку суміжних наук	Широкий	10
	Результати можуть бути використані в конкретному науковому напрямку при розробці нових технічних рішень, спрямованих на суттєве підвищення продуктивності суспільної праці	Достатньо широкий	8
	Результати будуть використані при проведенні наступних НДР, при розробці нових технічних рішень в конкретній галузі	Достатній	5
5.Ступінь реалізації результатів	Строк впровадження, роки: До 2	Висока	10
	До 4	Середня	7
	До 6	Достатня	4
	Більше 6	Недостатня	2
6.Завершення одержаних результатів	Авторське свідоцтво, стаття в фаховому виданні, методика, інструкція, класифікатор, стандарти, нормативи.	Висока	10
	Технічне завдання на прикладну НДР	Середня	8
	Рекомендації, розгорнутий аналіз, пропозиції	Достатня	6
	Огляд, інформаційне повідомлення	Недостатня	3

Кількісна оцінка факторів науково-технічної ефективності НДР здійснюються експертним шляхом за десятибальною шкалою і визначається як середньоарифметичне. Отримані результати зводять за формою табл. 6.9.

Таблиця 6.9

### Результати розрахунків науково-технічної ефективності НДР

Фактори науково-технічної ефективності	Характеристика фактора	Розрахунок $B_{ij}$			$B_{ij}^{\max}$
		Експертні оцінки		$B_{ij}$	
		1	2		
1.Новизна очікуваних або одержаних результатів	Встановлені деякі часткові закономірності, методи, способи, які дозволяють створити принципово нові види техніки	3	3	3	10
2.Глибина наукового опрацювання	Проведена обмежена кількість розрахунків по відомим методикам, виконані теоретичні розрахунки невисокої складності, частково перевірені експериментальними даними	6	6	6	10
3.Ступінь ймовірності успіху	Середня ймовірність вирішення більшості експериментальних або теоретичних задач	6	6	6	10
4.Перспективність використання результатів	Результати можуть бути використані в багатьох наукових напрямках, мають значення для розвитку суміжних наук	10	10	10	10
5.Масштаб можливої реалізації результатів	До 2 років	10	10	10	10
6.Завершеність одержаних результатів	Рекомендації, розгорнутий аналіз, пропозиції	6	6	6	10

Розраховане за формулою 6.2 значення  $E_{nm}$  буде відображати рівень наукової та науково-технічної ефективності конкретної теми фундаментального чи пошукового дослідження:

$$E_{nm} = \frac{0.25 \cdot 3 + 0.16 \cdot 6 + 0.09 \cdot 6 + 10 \cdot 0.25 + 10 \cdot 0.15 + 6 \cdot 0.1}{1 \cdot 10} = 0,685 .$$

Загальну оцінку магістерської НДР можна здійснити, користуючись даними табл. 6.10.

Таблиця 6.10

Загальна оцінка наукової та науково-технічної ефективності  
фундаментальних та пошукових НДР

Загальна оцінка наукової та науково-технічної ефективності		Можливі рекомендації по результатам виконання НДР
Розраховане значення $E_{нт}$	Загальна якісна оцінка ефективності	
0,91-1,00	Відмінно	Оформлення авторського свідоцтва, публікація у фаховому виданні, продовження досліджень по даній тематиці
0,76-0,90	Дуже добре	
<b>0,61-0,75</b>	<b>Добре</b>	<b>Рекомендації можуть бути сформульовані після ретельного аналізу отриманих результатів</b>
0,36-0,60	Достатня	Переглянути технічне завдання у разі продовження досліджень по даній темі
Менш 0,35	Незадовільна	Здійснити всебічний аналіз отриманих результатів по темі

#### 6.4 Висновки до розділу 6

У розділі на підставі виконаних розрахунків та нормативних даних встановлено, що планова калькуляція вартості проведення досліджень по темі становить 56299,19 грн., а кількісна оцінка науково-технічна ефективність науково-дослідної роботи, яка здійснюється експертним шляхом за десятибальною шкалою і визначається як середньоарифметичне, що складає 0,685 від максимального числа 1, а рекомендації по результатам виконання НДР можуть бути сформульовані після ретельного аналізу отриманих результатів.

## РОЗДІЛ 7

### ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

#### 7.1 Охорона праці

Оскільки під час проведення дослідження на тему: «Метод виділення огиноючої складової мовних сигналів для ефективного їхнього кодування» використовується персональний комп'ютер, як основна складова для виділення огиноючої складової мовних сигналів, тому розглянуто питання аналізу шкідливих і небезпечних факторів під час роботи з персональним комп'ютером.

Оскільки екран монітора є джерелом світла, тому необхідно організувати освітлення та розташувати його так, щоб у полі зору користувача не було інших яскравих джерел світла, а також освітленість екрану не збільшувалась за рахунок будь-якого стороннього джерела світла.

Спектр випромінювання комп'ютера містить у собі рентгенівську, ультрафіолетову й інфрачервону області спектра, а також широкий діапазон електромагнітних хвиль інших частот. Небезпека рентгенівських променів вважається зараз спеціалістами досить малою, оскільки цей вид променів поглинається склом екрана. Проте необхідно пам'ятати, що рентгенівське випромінювання, навіть у мізерно малих інтенсивностях, сприяє іонізації повітря.

Інфрачервоне випромінювання шкідливо діє на зір, втомлюючи очі, при тривалій дії порушує нормальне сприйняття кольору оком. Рентгенівське випромінювання шкідливо впливає на кісткові тканини і кровотворні функції кісткового мозку. Електромагнітне випромінювання, яким супроводжується показ зображення на моніторі комп'ютера, негативно впливає на органи зорового сприйняття людини.

Під час роботи комп'ютерів суттєво зменшується в робочій зоні кількість негативно заряджених іонів. Це обумовлено дією електричних полів, насамперед електростатичного. За санітарними нормами рівень концентрації аероіонів, відхилення від якого складає небезпеку здоров'ю людини знаходиться в діапазоні від 600 АІ (аероіонів) до 50000 АІ в куб.см. Рівні позитивних і негативних іонів у повітрі приміщень з ВДТ мають відповідати санітарно-гігієнічним нормам N 2152-80 (дод.3).

Забезпечення нормальних умов праці на робочому місці. При плануванні та організації робочого місця дослідника, який виконує науково-дослідну роботу, згідно санітарних норм, існує ряд правил, які необхідно виконувати для забезпечення нормальної працездатності упродовж усього робочого дня. Монітори комп'ютерів орієнтовані так, щоб кількість бліків була мінімальна, а відстань до очей оператора оптимальна згідно «правил охорони праці під час експлуатації електронних машин» з наказу державного комітету України з промислової безпеки, охорони праці та гірничого нагляду від 19.04.2010р.

Площу приміщень, в яких розташовують відеотермінали, визначають згідно з чинними нормативними документами з розрахунку на одне робоче місце, обладнане відеотерміналом: площа - не менше 6,0 м<sup>2</sup>, обсяг - не менше 20,0 м<sup>3</sup>, з урахуванням максимальної кількості осіб, які одночасно працюють у зміні.

Приміщення, що оснащені дисплеями (ділянки підготовки даних, дисплейні зали), мають північну або північно-східну орієнтацію. Коли у діючих обчислювальних центрах такі приміщення орієнтовані переважно на південні, південно-західні або південно-східні румби, передбачають обов'язкове розташування на вікнах сонцезахисних пристроїв (жалюзі, штори). Розраховується КПО за методикою, викладеною в СНиП II-4-79.

В приміщеннях, які оснащені відеотерміналами, стіни пофарбовані у кольори пастельних тонів з коефіцієнтом відбиття, що знаходиться у межах 40-



60%. Забарвленню поверхонь надано матову фактуру. Значення напруженості електромагнітних полів на робочих місцях з ВДТ мають відповідати нормативним значенням (ГДР N 3206-85, ГДР N 4131-86, СН N 5802-91, ГОСТ 12.1.006-84), (дод.6). Інтенсивність потоків інфрачервоного випромінювання має не перевищувати допустимих значень відповідно до СН 4088-86, ГОСТ 12.1.005-88

Робочі місця, які обладнані ПК, розташовують таким чином, щоб у поле зору користувача не потрапляли вікна та освітлювальні прилади. Тому відеотермінали розміщено під кутом 90-105° до площини зовнішньої стіни. В поле зору користувача ПК не повинні потрапляти і інші поверхні робочого приміщення, які мають властивості віддзеркалювання. Співвідношення яскравості екрана та поверхонь, які оточують, повинно складати 5:1, причому, безпосереднє покриття робочого стола повинно мати коефіцієнт відбиття 0,25-0,4.

Отже, враховуючи вище сформульовані рекомендації по охороні праці потрібно забезпечити безпечні умови праці при роботі за комп'ютером.

## 7.2 Безпека в надзвичайних ситуаціях

### 7.1.1 Дії персоналу підприємства при виникненні надзвичайних ситуацій

#### 1. Загальні положення

Типова Інструкція розроблена згідно з Кодексом цивільного захисту та інших нормативно-правових документів.

Усі працівники підприємства, не залежно від займаних посад, повинні знати і суворо виконувати вимоги Типової Інструкції щодо дій персоналу при загрозі або виникненні надзвичайних ситуацій природного або техногенного характеру. За

невиконання вимог Інструкції персонал підприємства може бути притягнутий до адміністративної відповідальності.

2. Характеристика можливої обстановки в районі підприємства при виникненні надзвичайної ситуації

У розділі перелічуються можливі джерела потенційної небезпеки на території самого підприємства або поблизу від нього, додається характеристика можливої обстановки при виникненні надзвичайних ситуацій, яка пов'язана з руйнуванням або іншим негативним впливом.

Відомості про джерела і характер їх дії на підприємство надають районні державні адміністрації, міські ради.

3. Порядок оповіщення адміністрації та персоналу про загрозу виникнення надзвичайних ситуацій

Оповіщення адміністрації, робітників та службовців підприємства щодо надзвичайних ситуацій проводиться по завчасно розробленій схемі.

Адміністрація у неробочий час оповіщається по телефону (вказується відповідальний виконавець). У залежності від обстановки, оповіщається і решта персоналу.

У робочий час персонал підприємства оповіщається про надзвичайну ситуацію (вказується яким способом).

При отриманні відповідної інформації вмикають сирени, інші звукові та світлові засоби, що буде означати подання попереджувального сигналу "Увага всім", після чого негайно приводяться у готовність гучномовці, радіотрансляційні та телевізійні приймачі для прослуховування початкового повідомлення.

Кожний працівник підприємства повинен знати сигнали оповіщення цивільного захисту та вміти правильно діяти в умовах загрози та виникнення надзвичайних ситуацій.

4. Порядок укриття персоналу у захисних спорудах цивільного захисту

На випадок виникнення надзвичайної ситуації, яка пов'язана із загрозою або початком забруднення повітря хімічно небезпечною, радіоактивною речовиною всі працівники підприємства підлягають укриттю у захисних спорудах цивільного захисту (вказується адреса, кому належить).

Для термінового укриття працівників у разі забруднення хімічно небезпечною речовиною використовуються власні загерметизовані приміщення (вказується адреса), забезпечується перебування у ньому осіб, без подачі повітря протягом години.

При отриманні інформації щодо радіоактивної небезпеки працівники укриваються (вказується приміщення, адреса), яке забезпечує захист осіб, що переходять від ураження іонізуючим випромінюванням при радіоактивному зараженні.

#### 5. Порядок видачі персоналу засобів індивідуального захисту

Засоби індивідуального захисту (вказуються які) видаються після отримання відповідного розпорядження або за рішенням керівника підприємства (вказується місце видачі).

Працівники, які отримали такі засоби, повинні перевірити їх стан, провести підбір та мати постійно при собі або на робочому місці.

Протигази переводяться у бойовий стан за командою або самостійно, при наявності небезпеки забруднення повітря.

#### 6. Порядок виділення автомобільного транспорту для проведення евакуаційних заходів

При проведенні негайної евакуації персоналу з небезпечних зон залучається увесь наявний службовий, а також особистий транспорт працівників підприємства, які повинні надавати його у розпорядження адміністрації для негайної евакуації із небезпечної зони працівників та відвідувачів підприємства (вказується вид та тип транспортних засобів, місця посадки на транспорт та маршрут руху).

7. Вимоги до персоналу щодо додержання протиепідемічних заходів при загрозі розповсюдження особливо небезпечних інфекційних захворювань

Якщо на території підприємства або поблизу його виникла небезпека розповсюдження особливо небезпечних інфекційних захворювань, усі працівники повинні суворо дотримуватись вимог санітарно-епідеміологічної служби щодо проведення термінової профілактики та імунізації, ізоляції і лікування виявлених хворих, дотримуватися режиму, який запобігає розповсюдженню інфекції. При необхідності працівники, які прибули на роботу, повинні проходити санітарну обробку (проводити дезінфекцію або зміну одягу) (вказується місце її проведення), а водії транспортних засобів - здійснювати спеціальну обробку автотранспорту (вказується місце її проведення), а також виконувати інші вимоги та заходи, які перешкоджають розповсюдженню особливо небезпечних інфекційних захворювань.

8. Заходи щодо зберігання матеріальних цінностей у період загрози та виникнення надзвичайних ситуацій

Усі працівники підприємства повинні вжити необхідних заходів щодо зберігання матеріальних цінностей при загрозі або виникненні надзвичайних ситуацій.

У період виконання заходів захисту від надзвичайних ситуацій або при ліквідації їх наслідків необхідно вживати заходи, які направлені на попередження або зменшення можливих збитків підприємству від надзвичайних ситуацій, на забезпечення охорони майна та обладнання.

Відповідальність за організацію охорони покладається на (вказується посада, прізвище).

9. Особливості дій працівників при деяких надзвичайних ситуаціях

При загрозі ураження хімічно небезпечною речовиною оповіщаються усі працівники та відвідувачі, які знаходяться на території підприємства.

Вентиляційні установки та кондиціонери терміново вимикаються, зачиняються вікна, двері, квартирки, приміщення герметизується. Вихід із будівлі й вхід до неї припиняється до особливого розпорядження адміністрації.

Працівникам видаються засоби індивідуального захисту, одночасно приймаються заходи щодо забезпечення відвідувачів ватно-марлевими пов'язками.

Відповідальними призначаються за:

- забезпечення герметизації приміщень (посада, прізвище);
- забезпечення працівників та відвідувачів засобами індивідуального захисту (посада, прізвище).

При виявленні у приміщенні, де укриваються працівники, хімічно небезпечної речовини, працівники повинні вийти (вказати куди) або з дозволу адміністрації залишити зону забруднення. Виходити з неї необхідно тільки у засобах індивідуального захисту і рухатися в напрямку, перпендикулярному напрямку вітру.

При виникненні пожежі на підприємстві всі працівники зобов'язані суворо виконувати вимоги Інструкції з пожежної безпеки, евакуацію проводити згідно з Планом евакуації (вказується покажчик напрямку руху).

Відповідальність за дотриманням заходів пожежної безпеки та організацію дій персоналу при загрозі або виникненні пожежі покладається на (посада, прізвище).

Працівник (посада, прізвище) організовує на території підприємства контроль за радіаційною обстановкою за допомогою побутового дозиметру (називається тип приладу) і постійно інформує про результати вимірювань адміністрацію підприємства, управління з питань надзвичайних ситуацій.

При перевищенні гранично допустимих норм опромінення організується облік доз опромінювання. Відповідальний за виконання цього заходу (посада, прізвище).

Скорочується до мінімуму вхід у будівлю та вихід із неї. Контроль за дотриманням режиму поведінки й роботи працівників, який дозволяє

максимально понизити наслідки радіоактивного опромінення покладається на (посада, прізвище).

При загрозі або виникненні катастрофічних стихійних лих призначений працівник підприємства за розпорядженням адміністрації повинен зупинити виробництво, виконати необхідні протипожежні заходи, відімкнути від електромережі електрообладнання, підготуватися до евакуації або вивезення у безпечні місця найбільш цінних матеріальних засобів.

Контроль за обстановкою на території підприємства при стихійних лихах і за прийняті заходи захисту персоналу покладається на (посада, прізвище).

Якщо з'явилися постраждалі - надається перша медична допомога.

Залучаються санітарні дружини підприємства або санітарні пости, які створені. Приймаються заходи щодо госпіталізації постраждалих до спеціалізованих медичних закладів.

При надходженні анонімної інформації про загрозу на території підприємства або поблизу нього терористичного акту, працівник, який прийняв її, повинен терміново доповісти керівнику підприємства й у правоохоронні органи, і діяти згідно з розпорядженнями і рекомендаціями.

#### 7.1.2 Надзвичайні ситуації природного характеру. Геологічні ситуації

Землетруси, повені, зсуви, селеві потоки, бурі, урагани, снігові заноси, лісові пожежі лише протягом останніх 20 років забрали життя більше трьох мільйонів чоловік. За даними ООН, за цей період майже один мільярд жителів нашої планети зазнав збитків від стихійних лих.

Серед надзвичайних ситуацій природного походження на Україні найчастіше трапляються:

– геологічні небезпечні явища (зсуви, обвали та осипи, просадки земної поверхні);

– метеорологічні небезпечні явища (зливи, урагани, сильні снігопади, сильний град, ожеледь);

- гідрологічні небезпечні явища (повені, паводки, підвищення рівня ґрунтових вод та ін.);
- природні пожежі лісових та хлібних масивів;
- масові інфекції та хвороби людей, тварин і рослин.

Особливості географічного положення України, атмосферні процеси, наявність гірських масивів, підвищень, близькість теплих морів зумовлюють різноманітність кліматичних умов: від надлишкового зволоження в західному Поліссі - до посушливого - в південній Степовій зоні. Виняткові кліматичні умови на Південному березі Криму, в горах Українських Карпат та Криму. Внаслідок взаємодії всіх цих факторів виникають небезпечні стихійні явища. В окремих випадках вони мають катастрофічний характер.

Надзвичайні ситуації природного походження в Україні поділяються на: геологічні, географічні, метеорологічні, агрометеорологічні, морські гідрологічні, гідрологічні небезпечні явища, природні пожежі, епідемії, епізоотії, епіфітотії.

Стихійні явища часто виникають в комплексі, що значно посилює їх негативний вплив. Небезпечні природні явища, переважно, визначаються трьома основними групами процесів - ендогенними, екзогенними та гідрометеорологічними.

Стихійні лиха, що мають місце на території України, можна поділити на прості, що включають один елемент (наприклад, сильний вітер, зсув або землетрус) та складні, що включають декілька процесів однієї групи або кількох груп, наприклад, негативних атмосферних та геодинамічних екзогенних процесів, ендогенних, екзогенних та гідрометеорологічних процесів у поєднанні з техногенними.

Геологічно небезпечні явища

*Землетруси (earthquakes)* – коливання земної кори, що виникають внаслідок вибухів у глибині землі, розламів шарів земної кори, активної

вулканічної діяльності. Область підземного удару викликає пружні коливання (сейсмічні хвилі), що поширюються по землі у всіх напрямках. Область землі, з якої виходять хвилі землетрусу, називають центром, а розташовану на поверхні землі ділянку - епіцентром землетрусу. Інтенсивність землетрусу вимірюється в балах за шкалою Ріхтера, а у останні роки у нашій країні та у ряді європейських держав використовують 12-бальну міжнародну шкалу MSK-64. Інтенсивність землетрусу зменшується до периферії зони катастрофи. Осередки землетрусів знаходяться на глибині 30–60 км, а інколи на глибині до 700 км. В залежності від причин і місця виникнення землетрусу поділяються на тектонічні, вулканічні, обвальні і моретруси.

Землетруси захоплюють великі території і характеризуються:

- руйнуванням будівель і споруд, під уламки яких потрапляють люди, виникненням масових пожеж і виробничих аварій;
- затопленням населених пунктів і цілих районів;
- отруєнням газами при вулканічних виверженнях;
- ураженням людей і руйнуванням будівель уламками вулканічних гірських порід;
- ураженням людей і виникненням осередків пожеж у населених пунктах від вулканічної лави;
- провалом населених пунктів при обвальних землетрусах;
- руйнуванням і змиванням населених пунктів хвилями цунамі;
- негативною психологічною дією.

Сейсмоактивні зони оточують Україну на південному заході і півдні: Закарпатська, Вранча, Кримсько-Чорноморська та Південно-Азовська.

У сейсмічному відношенні найбільш небезпечними областями в Україні є Закарпатська, Івано-Франківська, Чернівецька, Одеська та Автономна Республіка Крим. На теренах Закарпаття відзначаються осередки землетрусів з інтенсивністю 6-7 балів (за шкалою Ріхтера) у зонах Тячів – Сигет, Мукачево –



Свалява. Закарпатська сейсмоактивна зона характеризується проявом землетрусів, що відбуваються у верхній частині земної кори на глибинах 6-12 км з інтенсивністю в епіцентрі 7 балів. Шестибальні землетруси зафіксовані також у Прикарпатті (Буковина). Прикарпаття відчуває вплив від району Вранча (Румунія).

У платформній частині України виділено ряд потенційно сеймотектонічних зон з інтенсивністю 4-5,5 бала. За інженерно-сейсмічними оцінками приріст сейсмічності на півдні України перевищує 1,5 бала, і у зв'язку з чим було визначено, що в окремих районах 30-50% забудови не відповідає сучасному рівню сейсмічного та інженерного ризику.

Попередити землетруси точно поки що неможливо. Прогноз справджується лише у 80 випадках і має орієнтовний характер. Серед всіх стихійних лих за даними ЮНЕСКО землетруси займають перше місце в світі за заподіяною економічною шкодою і кількістю загиблих.

### 7.1.3 Долікарська допомога при задусі, утопленні, заваленні землею

Асфіксія (задуха) — припинення надходження кисню в легені протягом 2-3 хв. Зупиняється газообмін в легенях, настає кисневе голодування, людина непритомніє. За цим — зупинення серця і настає смерть.

Асфіксія може виникнути внаслідок стискання (рукою, шнурком) гортані і трахеї (задушення), заповнення гортані і трахеї водою (утоплення), слизовими масами, блювотинням, землею; закривання входу в гортань чужорідним тілом чи запалим язиком (при наркозі або без свідомості); параліч дихального центру від отрути, вуглекислого газу, снодійних засобів; внаслідок прямої травми головного мозку (електрошок, блискавка, рана); внаслідок дифтерії, грипу, ангіни.

При рятуванні втопленника його беруть за волосся, перевертають лицем догори і пливуть, не даючи зачепити себе. Потерпілого кладуть животом на зігнуте коліно так, щоб голова була нижче грудної клітки, видаляють з ротової порожнини і гортані воду, блювотні маси, водорості.

Енергійно стискають грудну клітку, видаляють воду з трахей і бронхів. В утопленика параліч легенів настає через 4-5 хв., а серце працює 15 хв. Тоді потерпілого кладуть на рівну поверхню, роблять штучне дихання і непрямий масаж серця.

При набряку гортані спостерігається шумне затруднення дихання, посиніння шкіри. Накладають холодний компрес на карк, ноги занурюють в гарячу воду. Підшкірно вводять 1 мл. 2% розчину димедролу. При потребі лікарі роблять трахіостомію — вводять трубку в розсічену трахею.

При заваленні землею стискається грудна клітка й ускладнюється відтік крові, тиск крові у венозній системі зростає, що призводить до розриву дрібних вен обличчя і шиї. Різко погіршується дихання. Розвивається синдром травматичного роздавлення. В тканинах накопичуються токсичні речовини, які викликають важку інтоксикацію, аніроз, порушують функції серця, печінки, нирок. При потребі очищують рот, гортань, накладають джгути, терміново транспортують в стаціонар.

## РОЗДІЛ 8

### ЕКОЛОГІЯ

#### 8.1 Актуальність екологічних проблем

Розвиток людської цивілізації неможливий без раціональної взаємодії з природою. Людина отримує все необхідне для життя: енергію, продукти харчування, матеріали, бере від природи емоційну та естетичну наснагу.

Цілеспрямованість дії людини на природу зумовлює не тільки позитивний вплив, а й призводить і до негативних наслідків. Тому сьогодні вкрай необхідна стратегія охорони довкілля, широке екологічне виховання населення.

Екологія – це наука про відносини між рослинними і тваринними організмами та навколишнім середовищем.

Важливим об'єктом вивчення екології є біосфера. Біосфера – оболонка Землі, у межах якої існує життя. Основним елементом біосфери є людина. Людина – вищий ступінь розвитку живих організмів на Землі, суб'єкт суспільно-історичної діяльності та культури. Визначальною особливістю людини є розробка і створення знарядь виробництва і використання їх для впливу на навколишнє середовище. Біологічна єдність людини і біосфери полягає в тому, що людина є однією з ланок кругообігу у природі. Речовини, які бере людський організм з оточуючого середовища (повітря, води, їжі), використовуються у процесі життєдіяльності, а потім (звичайно у зміненому вигляді) вони виділяються і повертаються у зовнішнє середовище, де внаслідок біогеохімічних перетворень розпадаються на складові елементи, які знову надходять у біологічний кругообіг.

Поки людство було нечисленним, а його діяльність примітивною, люди могли впливати на навколишнє середовище тільки локально, тобто істотно не змінювати біологічного кругообігу речовин у масштабах планети і динамічної

рівноваги в цілому. При цьому середовище самоочищалося, що не впливало негативно на організм людини та її здоров'я.

Із збільшенням чисельності людства і розселенням його по всій планеті потреби у природних ресурсах постійно зростають. Людина, щоб забезпечити свої потреби і поліпшити умови життя, за допомогою запасу знань починає відбирати у природи все більше її ресурсів, іноді на шкоду собі. На відміну від інших істот людина бере не тільки ті речовини, які потрібні їй для фізіологічних потреб (кисень, воду, продукти харчування), а й використовує величезну кількість копалин, лісових та інших матеріалів, потрібних для функціонування сотень тисяч електростанцій, фабрик, заводів. Їхні відходи потрапляють до навколишнього середовища у кількості, що перевищує можливості його самоочищення.

Проблема “Людина - біосфера” має два основних аспекти:

1. Техніко-економічний, пов'язаний із зростаючим виснаженням природних ресурсів планети, що ставить перед ученими чимало проблем щодо пошуку нових джерел енергії тощо.

2. Екологічний, пов'язаний із забрудненням навколишнього середовища і порушенням біологічної рівноваги у системі людина – біосфера.

8.2 Вплив на довкілля при здійсненні технологічного процесу монтажу та складання приладу

При монтажі та складанні даного приладу в процесі багатьох технологічних операцій має місце забруднення довкілля різними відходами. Ці відходи потрапляють в повітря і воду.

В даному технологічному процесі використовуються :

- Лудіння
- Пайка

- Лакування
- Промивка
- Сушка

При проведенні операції лудіння і пайки в повітря виділяються пари олова і свинцю, при використанні припою ПОС-61. Небезпека свинцю для людини визначається його значною токсичністю та здатністю накопичуватися.

Лакування друкованого вузла приладу і його сушка приводить до викидів в атмосферу парів органічних речовин.

В навколишнє середовище небезпечні речовини можуть потрапляти не лише у вигляді випаровувань, а й через каналізацію з водою, яка використовується для промивки плат. Отже операції промивки забруднюють водні стани іонам хімічних речовинами.

### 8.3 Заходи по охороні навколишнього середовища

Для зменшення шкідливого впливу на навколишнє середовище при монтажі і складанні приладу необхідно провести заходи щодо знищення цього впливу.

Використання сучасної елементної бази при виробництві даного електронного апарату дозволило скоротити кількість дискретних елементів, для виробництва яких необхідно задіяти різні типи виробництв, які можуть виявитись шкідливими для навколишнього середовища. Зменшення кількості матеріалів, які необхідні при технологічному процесі, дозволило не тільки зменшити собівартість, але й величину відходів при виробництві. При монтажі і складанні даного апарату не використовується обладнання, яке виділяє в навколишнє середовище токсичні речовини, оксиду вуглецю. Кількість місць пайки зменшена, що дозволяє підвищити економію електроенергії.

Для монтажу і складання приладу використовують такі види енергоресурсів: природний газ для обігріву виробничих приміщень та електроенергія для технологічних цілей та освітлення.

Зменшення забруднення повітря під час проведення операції пайки досягається за допомогою застосування витяжки, а також вентиляції з фільтруючими, адсорбуючими елементами.

Під час проведення операцій лакування також застосовують вентиляцію з фільтруючими елементами.

Зменшення теплового забруднення під час операції сушки досягається застосуванням припливно-витяжної вентиляції.

Щоб зменшити забруднення водних станів іонами міді, домішками кислот і іншими хімічними речовинами необхідно проводити очищення води.

При очистці води під час промивки використовують фільтрацію електрохімічним способом. Даний спосіб очистки стічної води має ряд переваг перед іншими: малий об'єм очистки споруд, простота експлуатації внаслідок високої ступені автоматизації електрохімічних процесів.

Отже, при проведенні виробничих процесів монтажу і складання приладу необхідно використовувати технологічні методи і засоби, які створюють мінімальний вплив на навколишнє середовище.

## ЗАГАЛЬНІ ВИСНОВКИ

В результаті виконання роботи отримано наступні результати.

Розглянуто принципи роботи параметричних вокодерів, зокрема смугових (канальних), ортогональних, вокодерів з лінійним передбаченням, формантних вокодерів тощо та встановлено, що важливим для зниження обсягів переданих даних є виділення тих інформативних параметрів голосових сигналів, яких буде достатньо для відтворення вихідного голосового сигналу на приймаючій стороні. При цьому такими параметрами відповідно до акустичної теорії мовотворення є частота основного тону та параметри огинаючої складової голосового сигналу в часі. При цьому сам голосовий сигнал розглядається як складне амплітудно модульоване коливання. Для наближення до природного звучання в структуру таких сигналів можна додатково вводити шумові складові із заданими законами розподілу. При цьому важливим є обґрунтування способу виділення огинаючої складової сигналу.

Розглянуто метод виділення огинаючої складного сигналу з використанням перетворення Гільберта. Проаналізовано математичний апарат перетворення Гільберта та встановлено, що з допомогою такого перетворення визначати такі поняття як огинаюча, миттєва фаза і миттєва частота сигналу навіть в тих випадках, коли форма сигналу істотно відрізняється від форми гармонійного сигналу.

Проаналізовано властивості перетворення Гільберта та показано, що це перетворення є оптимальним для виділення огинаючої складової голосових сигналів.

В ході експериментальних досліджень проведено завантаження в середовище Matlab голосовий сигнал. Встановлено, що несуча складова голосового сигналу є близькою до гармонічної, відповідно в спектрі вона буде

проявляться у вигляді форманти. Обвідна буде проявляться на низьких частотах спектру сигналу.

Застосовано перетворення Гільберта до вихідного голосового сигналу. Встановлено, що перший максимум характеризує параметри обвідної складової голосового сигналу. На частотах більше 40 Гц розміщуються залишкові складові форманти сигналу. Для згладжування форми огинаючої пропущено результат перетворення через фільтр низьких частот з частотою зрізу 40 Гц. Потім додатково застосовано низькочастотну фільтрацію сигналу фільтром з частотою зрізу 20 Гц.

Результат фільтрації є добрим і дозволяє застосувати перетворення Гільберта для виділення огинаючої складової голосового сигналу.



## Бібліографія

1. Рабинер Лоренс. Цифровая обработка речевых сигналов : пер. с англ. / Л. Рабинер, Р. Шафер ; [под ред. М. В. Назарова, Ю. Н. Прохорова]. – М. : Радио и связь, 1981. – 496 с.
2. Максимов Е.М. Актуальные задачи речевой акустики / Е.М. Максимов, Ю.Н. Ромашкин, С.А. Лопатина // Речевые технологии. – 2008. – №2. – С. 66-71.
3. Вокодерная телефония. Методы и проблемы / Под ред. А.А. Пирогова - М.: Связь, 1974. - 246 с.
4. Ворсано Д. Н. Кодирование речи в цифровой телефонии // Сети и системы связи. - 1996. - №8. - С. 24-27.
5. Дьяконов В. MATLAB. Обработка сигналов и изображений. Специальный справочник.- С-Пб.: Питер, 2002. - 608с.
6. Зюко А.Г., Банкет В.Л., Лехан В.Ю. Методы низкоскоростного кодирования при цифровой передаче речи // Зарубежная радиоэлектроника. - 1986. - №11.-С. 53-70.
7. Коротаев Г.А. Анализ и синтез речевого сигнала методом линейного предсказания // Зарубежная радиоэлектроника. -1990. - №3. - С. 31-52.
8. Коротаев Г.А. Некоторые аспекты линейного предсказания при анализе речевого сигнала // Зарубежная радиоэлектроника. -1991. -№ 7. -С. 13-31.
9. Куприянов М.С., Матюшкин Б.А., - Цифровая обработка сигналов, С-Пб, Питер, 1998. - 416с.
10. Маркел Дж., Грэй А. Линейное предсказание речи - М.: Связь, 1980. - 308с.
11. Минин С. А. Кодер АДИКМ аппаратуры цифровой спутниковой связи // Электросвязь.- 1992. - №11. - С. 32-44.

12. Назаров Х.З., Прохоров Ю.Н. Методы цифровой обработки и передачи речевых сигналов - М.: Связь, 1982. - 236с.
13. Перцева Л. В. - Качество передачи речи при использовании низкоскоростных кодеков на абонентских линиях // Электросвязь. -1987.-№8.- С. 48-64.
14. Прокис Дж. Цифровая связь / Под ред. Д.Д. Кловского - М.: Радио и связь, 2000. - 800с.
15. Прохоров Ю.Н.. Статистические модели и рекуррентное предсказание речевых сигналов - М., Радио и связь, 1986. - 316 с.
16. Шелухин О.И., Лукьянцев Н.Ф. Цифровая обработка и передача речи, М., Радио и связь, 2000. - 328с.
17. Шульгин В.И. Основы теории передачи информации – Учебное пособие. -Харьков: Нац. аэро-косм. ун-т. «Харьк. авиац. ин-т», 2003. - 102с.
18. J. L. Flanagan, R. M. Golden, "Phase Vocoder," Bell System Technical Journal, November 1966, 1493-1509. (<http://www.ee.columbia.edu/~dpwe/e6820/papers/FlanG66.pdf>)
19. Mark Dolson, "The phase vocoder: A tutorial," Computer Music Journal, vol. 10, no. 4, pp. 14 -- 27, 1986.

# ДОДАТКИ

УДК 519.218:612.78

**А. Іскра, Р. Небожук, Л. Дедів**

(Тернопільський національний технічний університет імені Івана Пулюя)

## **МЕТОД ОЦІНЮВАННЯ ПЕРІОДУ ОСНОВНОГО ТОНУ ГОЛОСОВИХ СИГНАЛІВ ДЛЯ МЕДИЧНИХ ДІАГНОСТИЧНИХ СИСТЕМ**

В праці [1] розглядається задача медичної діагностики стану органів голосового апарату за результатами належного опрацювання голосових сигналів методами, які визначаються адекватною фізичній природі таких сигналів та поставленій задачі математичною моделлю голосових сигналів. Проводиться обґрунтування вибору математичної моделі голосових сигналів, зокрема вокалізованих фрикативних звуків, у вигляді періодично корельованого випадкового процесу. В працях [2,3] проведено опрацювання голосових сигналів синфазним методом, необхідність та коректність застосування якого визначається обґрунтованою у праці [1] математичною моделлю таких сигналів у вигляді періодично корельованого випадкового процесу. На першому етапі застосування синфазного методу опрацювання голосових сигналів проводиться обчислення оцінок стаціонарних компонент шляхом формування вибірок значень окремої реалізації голосового сигналу, взятих через період корельованості. При цьому необхідно знати інтервал існування та наближене значення періоду корельованості сигналу.

В працях [2,3] припущено, що значення періоду корельованості буде близьким або рівним значенню періоду основного тону (ПОТ), і відповідно задача пошуку періоду корельованості зведеться до задачі визначення ПОТ. Найпростішим методом знаходження ПОТ є піковий метод [4], що ґрунтується на оцінюванні часової структури голосового сигналу. Також, відомим є кепстральний метод обчислення основного тону [4]. В праці [4] описується метод визначення ПОТ, що ґрунтується на принципах лінійного передбачення. Однак, розглянутим методам притаманні недоліки, зокрема низька точність значення ПОТ.

Найпоширенішим методом знаходження ПОТ є автокореляційний. Початкова оцінка ПОТ визначається місцезнаходженням максимального значення автокореляції в межах визначеного інтервалу. Однак відомо, що на точність обчислення ПОТ за автокореляційною функцією можуть впливати обертони, присутні в сигналі, та гармоніки з вищою амплітудою. Для вирішення проблеми пропонується обчислити перетворення Фур'є від кореляційної функції. Гармоніки, з яких складається функція кореляції, перетворяться на спектральні піки, рознесені по частотам. Таким чином вирішується проблема розділення основного тону та обертонів і, відповідно, одержується оцінка спектральної густини потужності сигналу.

Оскільки ПОТ є змінним для різних дикторів, необхідно проводити обчислення періоду корельованості для кожної реалізації сигналу.

### **Література**

1. Драган, Я. Обґрунтування математичної моделі фрикативного звуку у вигляді періодично корельованого випадкового процесу / Я. Драган, Є. Яворська, В. Дозорський // Вісник Тернопільського національного технічного університету. – Тернопіль : ТНТУ ім. І. Пулюя, 2010. – Т15. – №3. – С. 159–164.
2. Дозорський, В. Синфазний метод статистичного опрацювання фрикативних звуків для задач діагностики голосового апарату / В. Дозорський // Вісник Сумського державного університету. Технічні науки. – Суми : видавництво СумДУ, 2012. – № 3. – С. 16–21.
3. Драган, Я. Метод опрацювання фрикативних звуків для діагностики захворювань органів голосового апарату на ранніх стадіях / Я. Драган, В. Дозорський // Вісник Національного університету “Львівська політехніка”. Комп'ютерні науки та інформаційні технології. – Львів : НУЛП, 2011. – № 694. – С. 376–382.
4. Рабинер Лоренс. Цифровая обработка речевых сигналов : пер. с англ. / Л. Рабинер, Р. Шафер ; [под ред. М. В. Назарова, Ю. Н. Прохорова]. – М. : Радио и связь, 1981. – 496 с.