

УДК 621.391

**О. К. Юдін** — д-р техн. наук, проф.  
Національний авіаційний університет  
orcid.org/0000-0001-5098-7796  
e-mail: yak333@ukr.net;

**Р. В. Зюбіна** — аспірант  
Національний авіаційний університет  
orcid.org/0000-0002-8654-6981  
e-mail: sbt\_nau @ukr.net

## КЛАСИФІКАЦІЯ МЕТОДІВ ІДЕНТИФІКАЦІЇ ЧАСТОТИ ОСНОВНОГО ТОНУ

### Вступ

У процесі інформатизації сучасного суспільства наукоємні технології пропонують все більш прогресивні методи та моделі обробки та перетворення інформаційних сигналів. Використання сучасних операційних систем (ОС) та прикладного програмного забезпечення (ПЗ) для систем управління, мобільних пристроїв дозволяє використовувати голосові команди для керування складними інформаційними системами, органайзерами, браузером і навіть робототехнікою, що значно підвищує ефективність автоматизованих систем та дає економію часу для вирішення складних прикладних завдань. Впровадження систем управління складними процесами на базі ідентифікації корисного аудіо-сигналу є актуальною науково-прикладним завданням.

Голос — це набір певних звуків, які утворюються в гортані людини під тиском повітря формуючи інформаційне повідомлення у вигляді складного гармонійного сигналу. Значна роль у формуванні характеристик аудіо-сигналу, відводиться акустичним властивостям фізіологічної системи реалізації повідомлень тієї чи іншої мовної групи, яка відповідає за основні голосові (аудіо) параметри характеристик повідомлення, такі як частота основного тону сигналу, інформаційні фрейми та властивості частотного діапазону, гучність, періодичність і темп мовлення, інтонацію тощо. Крім того, акустична специфіка системи реалізації повідомлень певної мовної групи, відповідає за логічні та лінгвістичні наголоси в словах та реченнях.

Таким чином, напрямами **подальших наукових досліджень авторів є**: по-перше, ідентифікація корисного аудіо-сигналу на тлі завад, а по-друге, визначення базових характеристик голосу, що можуть характеризувати психофізичний стан людини або оператора системи управління інформаційною системою. Перше завдання достатньо відоме, тому найбільш цікавим напрямом досліджень повинні бути критичні умови впровадження нових методів та моделей ідентифіка-

ції, а саме великий рівень штучних або природних завад різних класів, висока ймовірність ідентифікації, час на прийняття рішень тощо.

Друге завдання практично не досліджувалось. Науковий пошук повинен стосуватись нових методів, що надають можливість визначити основні характеристики голосу та на їх базі створити психоемоційний портрет людини, а також отримати інформацію про стан людини, її відношення до певних речей, реакцію на ті чи інші запитання.

За допомогою багатокритеріального аналізу аудіо-повідомлення необхідно визначити правдивість відповідей оператора системи управління або взагалі людини, що ідентифікується.

**Метою роботи є** — аналіз основних аудіо-характеристик голосу людини, систематизація й організація процедури структурного представлення вокалізованого сигналу, а також класифікація наявних методів ідентифікації частоти основного тону, визначення переваг та недоліків у процесі їх реалізації.

Актуальність використання звукових сигналів у техніці зумовлена специфічними характеристиками голосової системи, що робить голос кожної людини унікальним і завдяки цьому дає можливість використовувати його у системах біометрії.

Залежно від співвідношення амплітуд частотних складових спектра, звук приймає різне відтворення і сприймається системою ідентифікації, як тон або шум. Якщо розглядати спектральне представлення сигналу, можна встановити залежності у вигляді чітко виражених піків амплітуд гармонійних складових спектру, що системно повторюються з деяким інтервалом частот. Зрозуміло, що кожна така складова відтворює один тон сигналу та повинна сприйматися, як індивідуальна інформаційна характеристика повідомлення. У випадку, коли піки амплітуд не мають чітко вираженого характеру (гладка функція спектру) і визначити піки практично неможливо, тоді такий звуковий сигнал сприймається системою як шум.

Як відомо слухова система людини здатна розрізняти будь-які сигнали в зоні низьких частот (до 8–10 кГц), у результаті чого роздільна здатність слухової системи погіршується під час переходу до області більш високих частот (до 10–12 кГц).

Таким чином, основним і найважливішим параметром голосового сигналу, у розрізі науково-прикладних завдань: кодування, компресії, ідентифікації та верифікації є частота основного тону.

**Основний тон** — це тон (гармонійна складова) створений акустичною системою, коливання якого відбуваються з найменшою частотою.

Амплітуда основного тону визначається природною частотою коливання системи, крім того основний тон вважається першою гармонією спектрального представлення сигналу.

**Обертон**ом називається синусна складова періодичного коливання складної форми з більш високою частотою ніж основний тон. Якщо, частоти всіх обертонів кратні частоті основного тону, то такі обертони називаються *кратними гармоніками* більш високих частот. У випадку, коли частоти залежать від частоти основного тону більш складним чином, то такі обертони називаються *некратними* основному тону.

**Спектром аудіо-сигналу** називається сукупність основного тону та обертонів, які є складовими сформованого складного звуку (інформаційного сигналу або корисного повідомлення).

Піки в частотній області, що виникають через квазіперіодичні коливання голосових зв'язок, називають **періодом основного тону** [1].

**Голос** людини — це сума синусоїдальних коливань, що відтворюють інформаційне повідомлення. Залежно від психоемоційного стану особи частота основного тону та обертонів може змінюватись, однак існує базова частота основного тону — найінформативніший частотний діапазон коливань голосу людини, так для чоловіків він становить 130 Гц, а для жінок — 260 Гц.

Як відомо, мовний сигнал поділяється на два типи, а саме **вокалізований і шумовий**. Основною різницею між ними є те, що вокалізований сигнал виникає як результат імпульсів збудження мовного тракту та коливання голосових зв'язок. Імпульси які зумовлюють появу вокалізованого сигналу мають чітко виражений період основного тону в частотній області, до них відносяться всі голосні букви «а», «о», «у» та ін. Шумовий сигнал (не вокалізований), у свою чергу, виникає в результаті звуження певних місць мовного тракту при проходженні через нього повітряного потоку. До не вокалізованих належать фрикативні звуки, що виникають у разі утворення турбулентного широкосмугового шу-

му («с», «ш»), і вибухові, які формуються утворенням у тракті перепони, на шляху повітря, з подальшим вивільненням стиснутого в області за перепопою повітря («ч», «п», «к»).

**Вокалізований гармонійний сигнал** — діапазон частот, що складається з гармонійних складових голосового сигналу, тобто діапазону, на якому чітко простежується наявність основного тону [2].

На сьогодні при визначенні частоти основного тону виникає низка проблем, а саме:

- складність реалізації алгоритмів визначення частоти основного тону;
- низька ймовірність визначення;
- помилки у процесі визначення;
- низька стійкість алгоритму до зовнішніх змін.

**Аналіз існуючих методів визначення частоти основного тону голосового сигналу**

Класифікуємо основні методи оцінки частоти основного тону:

- методи, що працюють у часовій області;
- методи, що працюють у частотній області;
- методи, основані на структурі слухового апарату.

Відомі такі, методи як метод максимальної правдоподібності та метод визначення частоти основного тону з надвисокою роздільною здатністю.

**Метод максимальної правдоподібності:** сутність цього методу полягає в тому, щоб зіставити кадр у частотній області з усіма можливими відрізками імпульсів, для знаходження максимальної кореляції.

**Визначення частоти основного тону з надвисокою роздільною здатністю** [6]: основна ідея даного методу полягає у використанні функції кореляції двох суміжних сегментів спектру, коли вони знаходяться один від одного на визначеному періоді або кратні йому. Метод дозволяє кількісно оцінити ступінь подібності між двома суміжними сегментами за допомогою лінійної інтерполяції. Однак, автори вважають за доцільне зупинитись на більш цікавих методах та моделях з точки зору аналітичного апарату та критеріях якості прийняття рішень, що до ідентифікації базових параметрів інформаційного аудіо-сигналу.

### 1. Методи, що працюють у часовій області

**Метод амплітудної селекції (метод переходу через нуль)**

В амплітудному методі на стаціонарній ділянці вокалізованого сигналу, при низькому рівні шумів, форма мовного коливання майже точно повторюється на кожному черговому періоді основного тону. Відстань між глобальними максимумами мовного сигналу можна вважати приблизно рівною періоду основного тону.

Основні труднощі в реалізації алгоритмів амплітудної селекції полягають в необхідності придушення локальних помилкових максимумів. Цю проблему можна вирішити шляхом підвищення порогу прийняття рішення в схемі пошуку максимумів. Однак, при цьому збільшується ймовірність пропуску істинного максимуму. Важливо зазначити, що як пропуск істинного максимуму, так і втрата можуть привести до істотних викривлень синтезованого звуку.

Підвищити надійність визначення періоду основного тону можна, додавши другий канал амплітудної селекції, що виділяє положення мінімумів мовного сигналу.

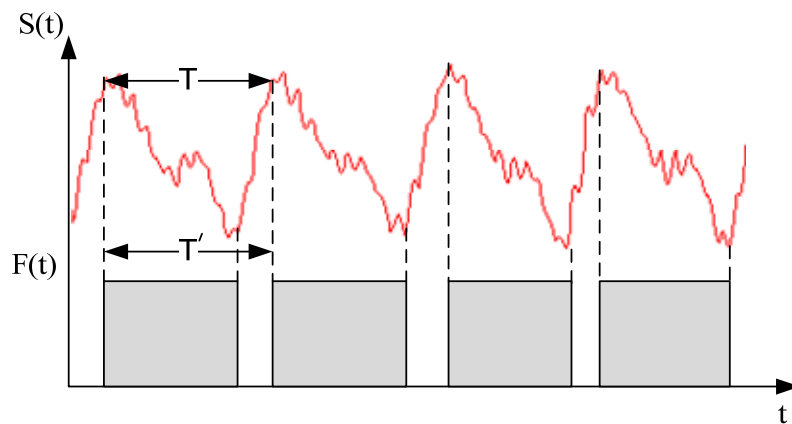


Рис. 1. Метод амплітудної селекції

**Метод автокореляції**

Цей метод працює на основі аналізу двох сигналів, один із яких є зміщеною копією першого.

$$y(n) = \sum_{i=1}^N U(t)U(t+t_0).$$

Якщо сигнал періодичний, то можна визначити його автокореляційну функцію  $y(n)$ , а якщо сигнал гармонійний, функція автокореляції матиме піки на кратних частотах відносно основного тону.

Нехай мовний сигнал представлений у вигляді послідовності відліків  $S_i, i = 0, 1, 2, \dots, N$ . Для вокалізованих звуків можна вважати що:  $S_n \approx S_{n-T}$ , де  $T$  — період основного тону, виражений у кількості відліків. В якості його оцінки в момент часу  $n$  оберемо значення  $k$ , яке мінімізує функцію:

$$L(n, k) = \sum_{i=0}^{N-1} (S_{n+i} - S_{n-k+i})^2$$

Припустимо, що енергія мовного сигналу не змінюється на ділянці квазістаціонарності. Тоді оцінка періоду основного тону повинна максимізувати кореляційну функцію.

Головною перевагою використання методу амплітудної селекції є простота реалізації, а основним недоліком — низька точність і нестійкість визначення частоти основного тону (навіть при відносно невеликому рівні шумів). Окрім того, цей метод коректно використовувати у тих випадках, коли амплітуда корисного сигналу значно перевищує амплітуду завади.

На рис. 1 показано вхідний мовний сигнал  $S(t)$  з деяким періодом  $T$ . Визначивши максимальні і мінімальні значення амплітуди мовного сигналу  $S(t)$ , можна сформувати сигнал  $F(t)$ , період якого дорівнює  $T'$  і є наближеним до  $T$ .

$$R(n, k) = \sum_{i=0}^{N-1} S_{n+1} S_{n-k+1}.$$

Також існують модифікації даного методу, у яких за основу взято обчислення взаємної кореляційної функції:

$$\tilde{R}(n, k) = \sum_{i=0}^{N-1} S_{n+1} F(S_{n-k+1}).$$

Підібравши зручну функцію  $F(x)$  можна спростити алгоритм, зробивши його придатним для апаратної реалізації. Модифікуємо цільову функцію

$$L(n, k) = \sum_{i=0}^{N-1} (S_{n+1} - \alpha_k S_{n-k+1})^2.$$

Параметр  $\alpha_k$  — це коефіцієнт підсилення. Легко показати, що для зсуву  $n$  оптимальне значення  $\alpha_k$  обчислюють за формулою

$$\alpha_k = \frac{\sum_{i=0}^{N-1} S_{n+1} S_{n-k+1}}{\sum_{i=0}^{N-1} S_{n-k+1}^2}$$

Тоді як оцінку періоду основного тону в момент часу  $n$  треба вибрати таке значення  $k$ , яке максимізує функцію:

$$M(n, k) = \frac{\left( \sum_{i=0}^{N-1} S_{n+i} S_{n-k+i}^2 \right)}{\sum_{i=0}^{N-1} S_{n-k+i}^2}.$$

Цей метод більш ефективний для роботи в області низьких та середніх частот.

Завдяки обмеженому діапазону частот, він використовується в додатках для розпізнавання мови. Однак залежно від довжини фрейму, метод автокореляції вимагає більших обчислювальних потужностей і він виявляється нестійким при використанні амплітудної модуляції.

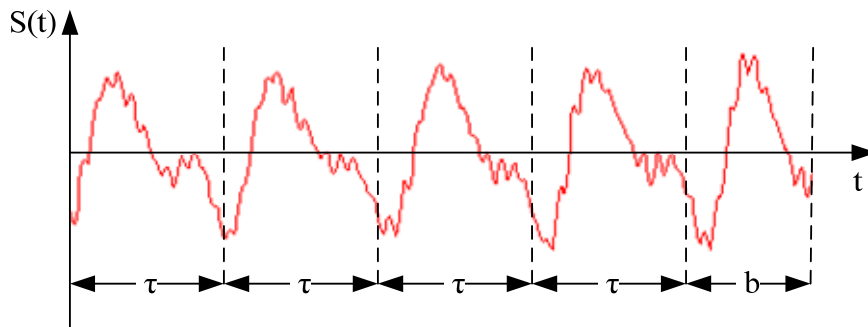


Рис. 2. Сигнал розбитий на  $N$  відрізків довжиною  $T$  плюс один відрізок довжиною  $b$

Тепер додаються такі сегменти (попередньо) відповідно до виразу:

$$S(t, \tau) = \begin{cases} \frac{1}{N+1} \sum_{i=0}^N S(t+n\tau) \\ \frac{1}{N} \sum_{i=0}^N S(t+n\tau) \end{cases},$$

де  $0 \leq t \leq b$ ;  $b \leq t \leq \tau$ .

Ці сегменти додаються послідовно, доки  $\tau = \tau_0$ . Отже, в результаті всіх дій, ми можемо сформувати цільову функцію:

$$V(\tau) = (N+1) \sum_{t=0}^b S^2(t, \tau) + N \sum_{t=0}^b S^2(t, \tau)$$

і отримати  $\tau_0$ .

Проте  $V(t)$  не є одновершинною і має багато піків, як це зазвичай буває з цільовими функціями методу максимальної правдоподібності. Якщо  $V(t)$  досягає максимуму при  $\tau = \tau_0$ , то його піки також будуть дорівнювати  $\tau = m\tau_0$ , у випадку коли  $m$  ділиться на  $N$ . Таким чином  $V(t)$  буде мати майже періодичні локальні максимуми [3].

### Методи адаптивної фільтрації

При використанні цього методу можна реалізувати багато підходів.

**Вузькосмуговий фільтр:** вхідний сигнал передається на вузькосмуговий фільтр.

### Метод максимальної правдоподібності

Цей метод дає змогу знайти оптимальний шлях для визначення періоду основного тону в часовій області.

Припустимо, що існує сигнал  $S(t)$  з відомою довжиною  $T$  і невідомим періодом  $\tau_0$ .

Розіб'ємо даний сигнал на  $N$  сегментів довжиною  $\tau$ , за умови ( $1 < \tau \leq T$ ).

$$T = N\tau + b,$$

де  $0 \leq b < \tau$  (рис. 2).

Нефільтрований і фільтрований сигнали потрапляють на розпізнавальну детекторну схему (рис. 3). Після цього сигнал подається назад для контролю центральної частоти смугового фільтра, змушуючи фільтр звести все до частоти вхідного сигналу [4].

**Оптимальний гребінчастий фільтр** [5]: цей метод призначений для визначення гребінчастого фільтра який мінімізує його вхідний сигнал. Порог гребінчастого фільтра має адаптивно налаштовуватись до частоти вхідного сигналу.

**Контур адаптивного підсилення** (показаний на схемі рис. 4).

### 2. Методи, що працюють в частотній області

Типовий аналіз сигналів полягає у розбитті сигналу на маленькі фрейми, що кратні вікну, і використанні швидкого перетворення Фур'є (STFT). Якщо сигнал періодичний, то за допомогою перетворення Фур'є визначаються піки, кратні частоті основного тону. Такі методи описують шляхи виявлення максимумів, які відповідають частоті основного тону. Однак у цього методу є недолік — STFT поділяє діапазон частот аудіо-сигналу на набори рівновіддалених частотних каналів. Оскільки сприйняття частоти основного тону людини в основному логарифмічне, то це означає що низькі частоти, можуть, простежуватись не так точно, як високі.

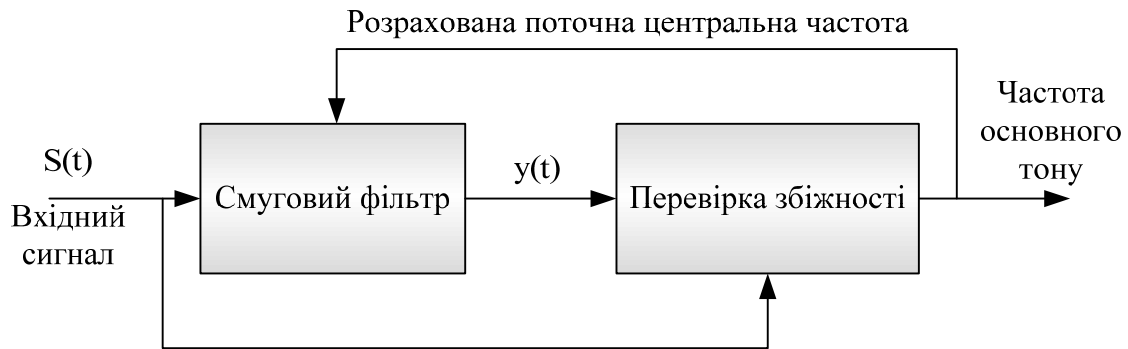


Рис. 3. Детектор частоти основного тону оснований на схемі смугового фільтра

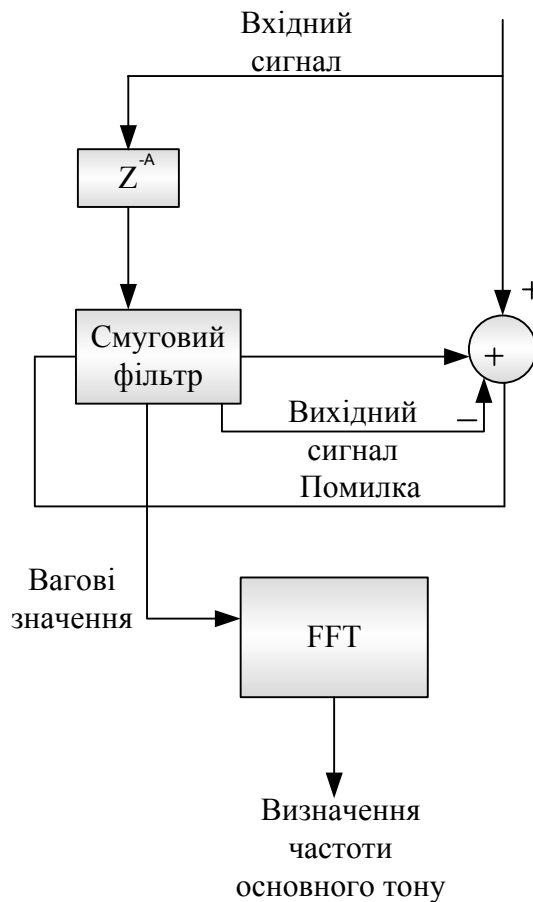


Рис. 4. Контур адаптивного підсилення для визначення частоти основного тону

**Спектральний гармонічний метод:** частота основного тону може бути визначена шляхом знаходження частоти гармонійних складових з максимальною амплітудою та обчислення найбільшого спільного дільника цих частот для гармонічних складових. Цей дільник може бути визначений шляхом внесення запису в частотну гістограму для кожної гармоніки.

Гармоніка з найбільшою амплітудою являє собою найбільший спільний дільник, і отже, є основною частотою. До переваг використання цього методу відносять: простоту реалізації, стійкість до адитивних і мультиплікативних за-

вад, можливість регулювання деяких характеристик (рис. 5).

**Метод кепстрального аналізу:**

Термін «кепстр» введений Богертом і є в даний час загальноприйнятим для позначення зворотного перетворення Фур'є логарифма спектра потужності сигналу; термін «комплексний кепстр» означає, що застосовується комплексний логарифм. Кепстральний аналіз застосовують для сигналів, що являють собою згортку двох тимчасових функцій, причому таких, що після перетворення їх в спектр утворюють на осі  $q$  імпульси, які не перекриваються.

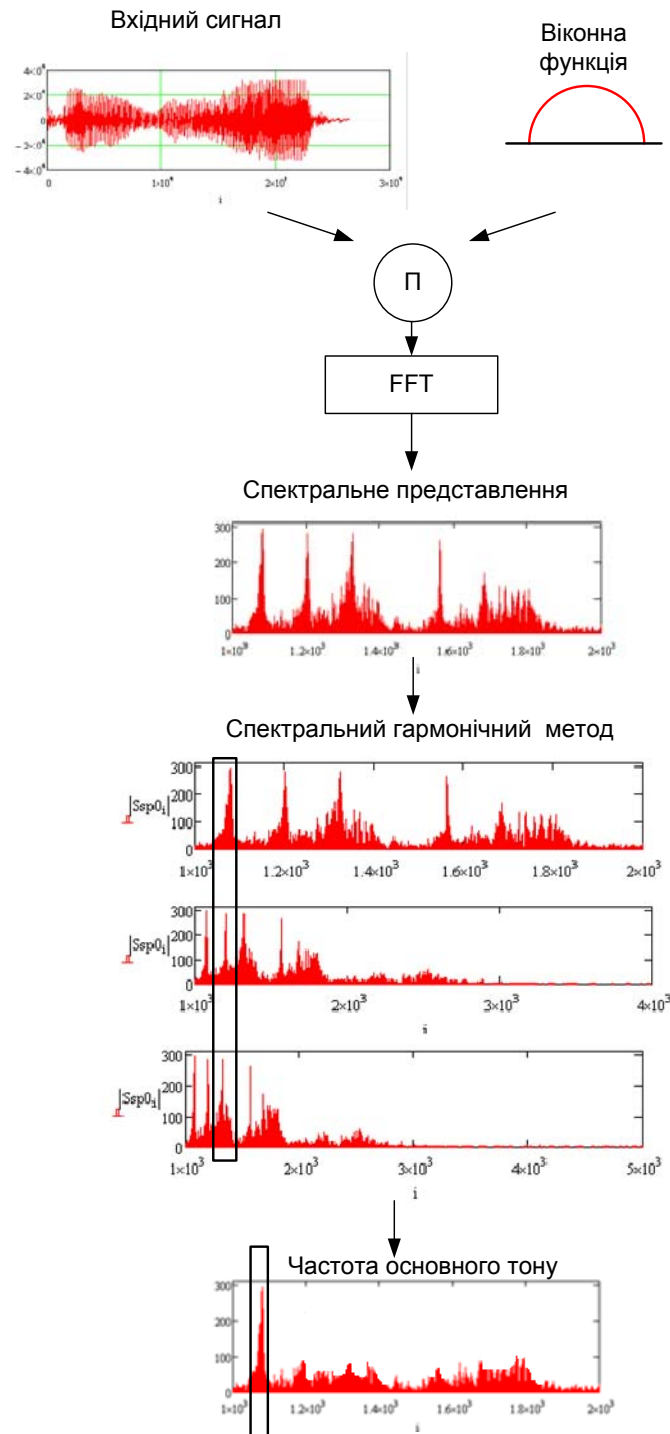


Рис. 5. Функціональна схема спектрального гармонічного методу

Якщо існує спектр гармонійного сигналу, то він має бути періодичним, і повторне використання швидкого перетворення Фур'є дасть можливість побачити піки, що повторюються з певним періодом і відображають період основного тону.

Метод визначення частоти основного тону на базі кепстрального аналізу полягає в аналізі окремих обчислених купстрів (рис. 6). Якщо пік у кепстрі перевищує поріг, то сегмент класифікується як вокалізований, а координата піку дозво-

ляє оцінити період основного тону. У випадку, коли пік не перевищує визначений поріг прийняття рішення, то сегмент відноситься до невокалізованих. Зважаючи на параметри збудження мови людини, кепстр обраховують кожні 10–20 мегасекунд. Однак такий метод визначення вокалізованих сегментів не дуже точний, так як існування піків у кепстрі залежить від багатьох чинників, включаючи довжину вікна і структуру самого сигналу.

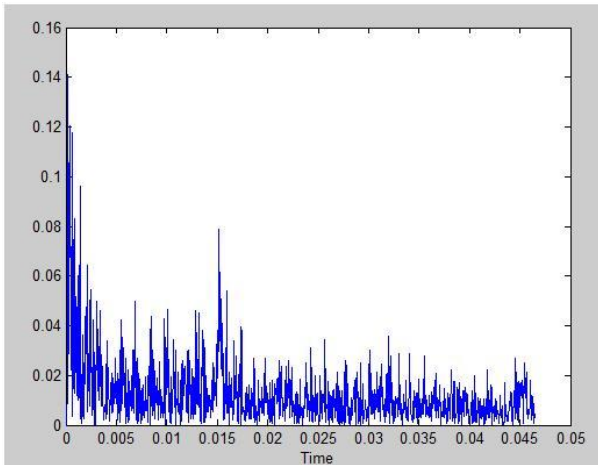


Рис. 6. Керстр вокалізованого сигналу

Від характеристик віконної функції залежить якість визначення періоду основного тону, тому бажано обирати його максимально вузьким.

У випадку правильно вибраної довжини вікна на вході, положення і амплітуда імпульсу кепстра забезпечують хорошу оцінку періоду основного тону, а також можливість визначення співвідношення тон/шум

Але цей метод має ряд значних недоліків:

- використання додаткових методів визначення порогу прийняття рішення;
- робота в нереальному масштабі часу;

– необхідність застосування віконних функцій і операцій згладжування.

**Метод заснований на вейвлет-перетворенні**

Вейвлет-перетворення — інтегральне перетворення, яке являє собою згортку вейвлет-функції з сигналом.

Вейвлет-перетворення переводить сигнал із часового уявлення в частотно-часове.

В основі вейвлет-перетворень, у загальному випадку, лежить використання двох безперервних, взаємозалежних сигналів інтегрованих по незалежній змінній функції (рис. 7). Переваги та недоліки вейвлет-перетворення:

- вейвлет-перетворення ні чим не поступається перетворенню Фур’є;
- базиси можуть бути добре локалізованими як за частотою, так і за часом. При виділенні в сигналах добре локалізованих різномасштабних процесів можна розглядати тільки ті масштабні рівні розкладання, які являють інтерес;
- базиси, на відміну від перетворення Фур’є, мають досить багато різноманітних базових функцій, властивості яких орієнтовані на вирішення різних завдань. Базисні вейвлети можуть мати і кінцеві, і нескінченні носії, що реалізуються функціями різної гладкості;
- складність реалізації.

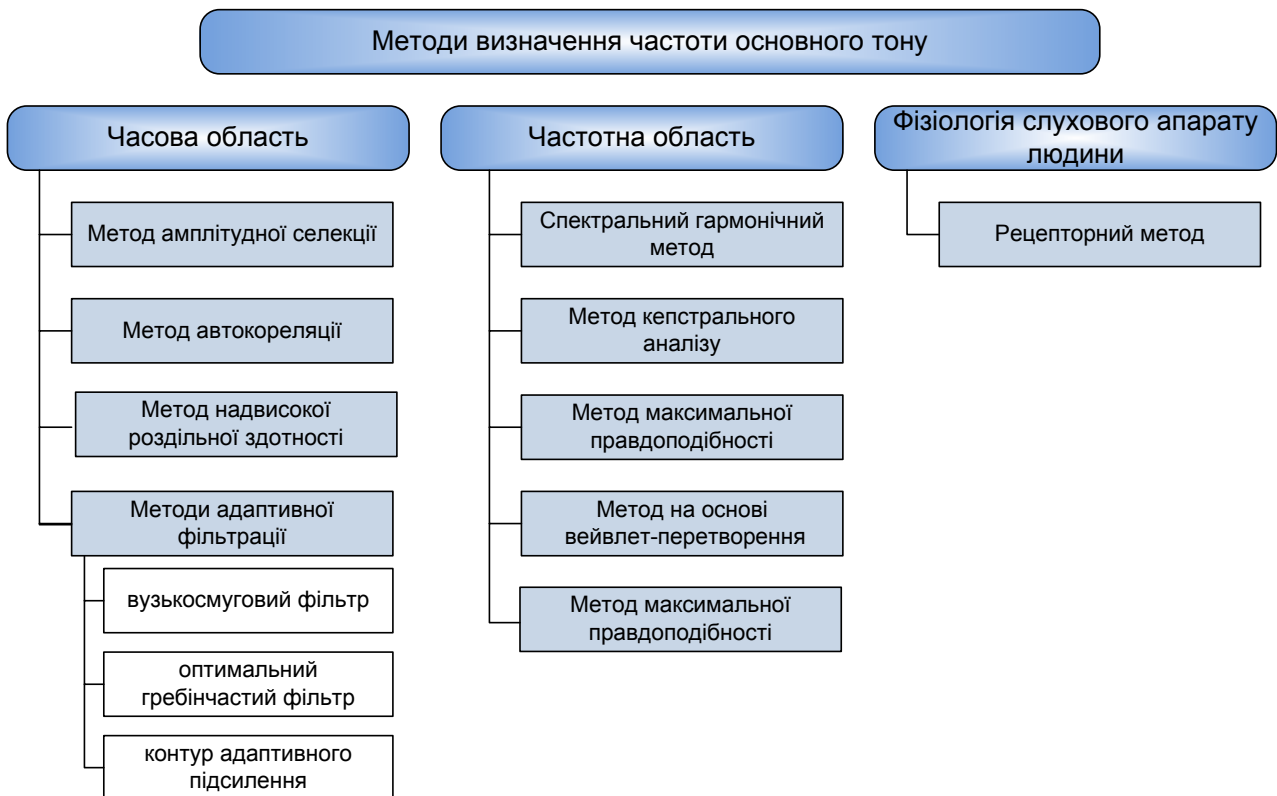


Рис. 7. Класифікаційна схема методів визначення частоти основного тону

Вейвлет-перетворення базується на застосуванні безперервного або дискретного вейвлет-перетворення. Слід відзначити позитивну сторону даного методу: для генерованого еталонного чотирьох гармонійного сигналу, а саме те, що відносна похибка оцінки частоти основного тону і формантних частот методом вейвлет-перетворення не перевищувала 0,38 %. Недоліки цього методу: необхідність коригування вікна перетворення під кожну оцінювану частоту; складність алгоритму реалізації; великі обчислювальні втрати.

**Рецепторний метод визначення частоти основного тону:** цей метод визначення частоти основного тону розробили М. Слейн і Р. Ліон на базі теорії розробленої Дж. Ліклайдером [7].

Такий метод дозволяє створити точну модель функціонування слухової системи людини в процесі визначення частоти основного тону.

Рецепторний метод дає можливість точно визначити висоту комплексного гармонійного і агармонійного подразника, а також демонструє хороші результати в умовах наявності шумової складової.

#### Висновки

Проведено аналіз основних аудіо-характеристик голосу людини, систематизовано методи і процедури структурного представлення вокалізованого сигналу, а також проведено класифікацію наявних методів ідентифікації частоти основного тону, визначено переваги та недоліки у процесі їх реалізації.

У статті визначено основні характеристики голосу людини. Проведено детальний аналіз методів визначення частоти основного тону.

Юдін О. К., Зюбіна Р. В.

### КЛАСИФІКАЦІЯ МЕТОДІВ ІДЕНТИФІКАЦІЇ ЧАСТОТИ ОСНОВНОГО ТОНУ

*Значна роль у формуванні голосових характеристик інформаційних сигналів відводиться акустичній системі представлення певної мовної групи, яка відповідає за основні голосові параметри, такі як частота основного тону сигналу, інформаційні фрейми та властивості частотного діапазону, гучність, періодичність і темп мовлення, інтонацію тощо. Визначено основні критерії, що характеризують особливості голосу. Проаналізовано основні характеристики аудіо-сигналів сформованих людиною певної мовної групи. Проведено аналіз методів і моделей ідентифікації частоти основного тону, що в подальшому дало можливість визначити період основного тону. В сучасних дослідженнях, при визначенні частоти основного тону, виникають такі проблеми, як складність реалізації методів, низька ймовірність та помилки у процесі визначення; низька стійкість методів до зовнішнього впливу завад. Проведено класифікацію методів визначення частоти основного тону за площиною їх реалізації.*

**Ключові слова:** ідентифікація аудіо-сигналу, частота основного тону, період основного тону, обертон, спектр аудіо-сигналу, частотна область, часова область.

Розроблено їх класифікацію та класифікацію наявних методів ідентифікації частоти основного тону. Визначено, недосконалість моделей знаходження частоти основного тону, що, в свою чергу, ставить нові завдання в процесі дослідження ідентифікації аудіо-сигналів за голосовими особливостями людини.

#### ЛІТЕРАТУРА

1. Аль-Келані Ф. Дослідження характеристик мовного сигналу в задачах розпізнавання, 2004.

2. Ашихмін А. В. Підвищення точності та швидкості обчислення миттєвого спектра гармонійних сигналів за допомогою детектора основного тону, 2008.

3. A. M. Noll. Pitch determination of human speech by the harmonic product spectrum, the harmonic sum spectrum, and maximum likelihood estimate Proceedings of the Symposium on Computer Processing in Communications, April, 1969.

4. Widrow B., Stearns S. D., Burgess J. C. Adaptive signal processing edited by bernard widrow and samuel d. stearns // The Journal of the Acoustical Society of America, 1986. — Т. 80. — № 3. — С. 991–992.

5. Moorer J. A. 1973. The optimum comb method of pitch period analysis of continuous digitized speech AIM-207. Stanford: Stanford Artificial Intelligence Laboratory.

6. Medan Y., Yair E., Chazan D. Super resolution pitch determination of speech signals //IEEE transactions on signal processing, 1991. — Т. 39. — № 1. — С. 40–48.

7. Licklider J. C. R. A duplex theory of pitch perception // The Journal of the Acoustical Society of America, 1951. — Т. 23. — № 1. — С. 147–147.



Yudin O. K., Ziubina R. V.

## CLASSIFICATION OF THE FUNDAMENTAL FREQUENCY IDENTIFICATION METHODS

*A significant role in the formation of the voice characteristics of the information signals given to the speaker system represents a language group, which is responsible for basic voice parameters such as pitch frequency of the signal, information frames and properties frequency range, volume, frequency and rate of speech, intonation, etc. The main criteria features of voice were identified. The basic characteristics of audio signals generated by a person of a certain language group were analyzed. The methods and models to identify the fundamental frequency, which further gave the opportunity to determine the period of the fundamental tone were analyzed. In the current research, the determination of the fundamental frequency, there are problems such as the complexity of the implementation methods, and low probability of error in determination process; low resistance methods to external interference. Methods of determination of fundamental frequency with the plane of their implementation were classified/*

**Key words:** identification audio signal, the pitch frequency, the period of the fundamental tone, the overtone spectrum of the audio signal, frequency range, time domain.

Юдин А. К., Зюбина Р. В.

## КЛАССИФИКАЦИЯ МЕТОДОВ ИДЕНТИФИКАЦИИ ЧАСТОТЫ ОСНОВНОГО ТОНА

*Значительная роль в формировании голосовых характеристик информационных сигналов отводится акустической системе представления определенной языковой группы, которая отвечает за основные голосовые параметры, такие как частота основного тона сигнала, информационные фреймы и свойства частотного диапазона, громкость, периодичность и темп речи, интонацию и тому подобное. Определены основные критерии, характеризующие особенности голоса. Проанализированы основные характеристики аудио-сигналов сформированных человеком определенной языковой группы. Проведен анализ методов и моделей идентификации частоты основного тона, что в дальнейшем дало возможность определить период основного тона. В современных исследованиях, при определении частоты основного тона возникают такие проблемы, как сложность реализации методов, низкая вероятность и ошибки в процессе определения; низкая устойчивость методов к внешнему воздействию помех. Проведена классификация методов определения частоты основного тона с плоскостью их реализации.*

**Ключевые слова:** идентификация аудио-сигнала, частота основного тона, период основного тона, обертона, спектр аудио-сигнала, частотная область, временная область.

Стаття надійшла до редакції 28.02.2017 р.

Прийнято до друку 29.02.2017 р.

Рецензент — д-р техн. наук, проф. В. В. Бараннік