

УДК 621.396

МЕТОДИКА РОБАСТНОЇ ОЦІНКИ КЕПСТРУ МОВНОГО СИГНАЛУ

Максим Кузнецов

Центр Інформаційно-аналітичних досліджень
стратегічних програм

Анотація: Розглядається методика одержання новітнього результату в сфері аналізу мовних сигналів – залежностей від частоти кепстру сигналу квантилів варіаційних рядків – порядкових статистик кепстральної функції мовного сигналу.

Summary: The methodology is presented in this article covers the newest result in speech processing field – the variational series fractals depending on speech signal cepstrum frequency. This effect was named the order statistics of speech signal cepstrum.

Ключові слова: Мовний сигнал, кепстральний аналіз, порядкові статистики, статистичний кепстр.

Вступ

Застосування кепстрального аналізу має місце в багатьох напрямках галузі дослідження мови – виділення кепстральних зрізів фонем: системи тон/шум, що працюють по принципу аналізу розподілення потужності звукового коливання по спектру частот; системи відокремлення дзвінких та глухих звуків мови та ін. Апарат кепстрального аналізу застосовується, також, для фільтрації перемножених сигналів та близького завдання нелінійної фільтрації згорнутих сигналів. Такій підхід використовується під час аналізу та обробки мовних сигналів, у рамках моделі утворення мовного сигналу, що складається зі згортки трьох функцій, котрі характеризують висоту основного тону, форму голосових зв'язок та конфігурацію мовного тракту. Основною стратегією багатьох систем стиснення смуги частот мови та автоматичного розпізнавання звуків мови є виділення та роздільне описання збудження та імпульсної характеристики мовного тракту. Таким чином, численні системи обробки мови частково базуються на виконанні операції розгортки сигналу.

Кепстральний аналіз знайшов своє застосування у вокодерній техніці тощо. Застосування його доцільне для виділення частоти основного тону мовного сигналу та подальшого синтезування мовного сигналу. Ці можливості кепстрального аналізу використовуються, у тому числі, як частина голосового усередненого образу при розв'язанні завдання ідентифікації джерела мовного сигналу.

I Постановка завдання

Завданням цього дослідження є демонстрація новітнього результату у царині дослідження кепстру складних сигналів – порядкових статистик кепстру мовного сигналу, отриманого шляхом застосування стійкого до помилок та аномальних перешкод методу, із урахуванням значень багатовимірних функцій щільності розподілу ймовірності кепстру.

II Основна частина

Процес утворення голосового сигналу складний. Повітряний струмінь з легень розсовує голосові зв'язки у поперечному напрямку, однак, під впливом м'язів, голосові зв'язки повертаються у зімкнутий стан. Цей процес є циклічним. Вже на цьому етапі голос характеризується висотою, силою та тембром. Останні дві характеристики змінюються під час проходження повітря мовоутворюючим трактом залежно від параметрів резонаторів – порожнин. Вважається, що основним органом, що формує інтенсивність мовного сигналу та утворює звуки мови є надставна труба периферійного голосового апарату, а форма порожнин практично повністю характеризує передавальну функцію мовного тракту. Різноманітні рухи активних артикуляторів (друга група органів надставної труби) забезпечують формування складних резонаторів, в результаті чого утворюються ті чи інші звуки мови – фонемі. Система фонем різних мов містить неоднакову кількість звуків, однак вона завжди скінченна та не перевищує декількох десятків. Зіставлення цього числа з безперервними рухами активних органів вимови, що утворюють нескінченну різноманітність форм резонаторів, забезпечує широку можливість варіювання фізичних параметрів акустичних сигналів, які відповідають тим чи іншим дискретним фонематичним одиницям. Таким чином, мовний сигнал утворюється шляхом квазіперіодичної модуляції потоку повітря, що видихається легеньми.

Сфера застосування кепстрального аналізу в аспекті аналізу та синтезу мови є дуже широкою. Таким чином, важливо вести дослідження якостей кепстру не тільки з метою вдосконалення існуючих алгоритмів кепстрального аналізу, а й з метою одержання нових знань щодо якостей кепстральної функції.

Аудіосигнал є випадковим процесом, його акустичні (електричні) характеристики змінюються в часі безперервно. Це є сукупність випадкових функцій. Статистичні якості аудіосигналів характеризуються, перш за все, законами розподілу: миттєвих значень та рівнів звукових тисень (напружень) у часі; звукових тисень відповідно до частоти сигналу; тривалістю пауз; тривалістю безперервного існування рівнів, що є визначеними. Але найбільш повний опис статистичних якостей величини X надає функція розподілу $G(x)$, в це поняття вкладається сенс вірогідності події $X \langle x$, де x - деяке поточне значення випадкової величини. Похідна від $G(x)$ має назву функції щільності розподілення ймовірності (ЩРЙ) $g(x)$ випадкової величини x . Вона відбиває ймовірність того, що випадкова величина опиниться у межах від x до $x + \Delta x$, де $\Delta x \rightarrow 0$. Залежності $g(x)$ та $G(x)$ встановлюють закон розподілу випадкової величини X .

Потенційні характеристики систем дослідження складних сигналів (мовних, гідроакустичних, біомедичних, сейсмічних, ультразвукових) взагалі та, як окремий випадок – систем кепстрального аналізу, можуть бути забезпечені лише при повному урахуванні значень багатовимірних функцій ЩРЙ інтенсивності кепстру сигналів на частотах аналізу (1).

$$g_{\Sigma}([A_i, f_i]) = \prod_{i=1}^n g(A_i, f_i); \quad (1)$$

де: g_{Σ} – сумарна багатовимірна функція ЩРЙ (сукупність даних про досліджуваній сигнал), $[A_i, f_i]$ - матриця значень амплітуди A_i інтенсивності кепстру сигналу на кожній дискретній частоті (сачоті) f_i кепстру сигналу $x(t)$, що досліджується, $g(A_i, f_i)$ - багатовимірна функція ЩРЙ амплітуд

A_i на частотах f_i в кожному перерізі (відлік сигналу) сегменту аналізу складного сигналу, $\prod_{i=1}^n$ – добуток всіх значень g на сегменті аналізу.

Саме підхід (1) надасть найбільш повний та достовірний опис складових випадкового, нестационарного процесу, яким є мовний сигнал. Цей підхід в царині дослідження мови традиційно вважається перспективним, але надскладним. Порівняння значень функцій ЩРЙ сигналу, що досліджується, у такий спосіб на практиці неможливо, бо не реалізовано метод, котрий враховуватиме весь масив чисел, що описують значення функцій ЩРЙ. Тому зазвичай проводять усереднення значень багатовимірних функцій ЩРЙ. Після чого, оперуючи поняттями середнього, дисперсії, асиметрії функцій ЩРЙ, ексцесу, тобто параметрами, яки характеризують сумарну багатовимірну функцію ЩРЙ – виконують дослідження кепстру мовного сигналу.

Усереднення значень сумарної багатовимірної функції ЩРЙ g_{Σ} веде до «згладжування» впливу характеризуючих особливостей будови мовотворчого тракту людини. Тому, будь-які параметри та складові мовного сигналу, що подаватимуться на дослідження, будуть в значній мірі неінформативні, збіднені, що проявлятиметься при розгляді їх від одного квазістационарного сегмента аналізу до наступного. У випадку дослідження мовних сигналів різних дикторів такий підхід (усереднення) до вирішення завдання матиме наслідком значну схожість параметрів (кепстру) сигналів різних дикторів. У той же час, відповідно до застосованого методу усереднення, можливе отримання протилежних висновків дослідників щодо однієї й тієї же мовної реалізації.

Враховуючи вищезазначене, можна зробити висновок, що існує необхідність покращення алгоритмів функціонування систем кепстрального аналізу складних сигналів у сенсі збереження природного інформаційного наповнення складного сигналу, що відбиватиметься на його кепстрі.

Оскільки існує складність із урахуванням величезних об'ємів значень функцій ЩРЙ, потрібно зауважити, що з високою мірою точності функція ЩРЙ кепстру характеризується порядковими статистиками – квантилями (квартилями, децилями, п'ятилями, процентилями) аргументу у зоні визначення функції із одночасним спрощенням завдання статистичного порівняння. Під спрощенням слід розуміти, що виконується статистичне порівняння не значень функції ЩРЙ статистичних кепстрів, а статистичне порівняння векторів порядкових статистик (рис. 1). Такий підхід гарантуватиме повноту відображення характеризуючої кепстр мовного сигналу інформації.

На рис. 1 надано графічне представлення функції ЩРЙ (нормовано до 1), функції розподілу (ФР), та квантильне розбиття статистичних інтервалів.

Виконання завдання підвищення інформативності результатів кепстрального аналізу мовних сигналів виконуватиметься шляхом отримання порядкових статистик кепстру. Ці дані міститимуть повну інформацію щодо характерних параметрів будови джерела утворення сигналу відповідно до якостей порядкових статистик, оскільки виконується перехід до використання векторів порядкових статистик, ніякого усереднення параметрів кепстру не застосовується, інформаційне наповнення мовного сигналу акумулюватиметься в статистичних залежностях квантилів кепстру [3].

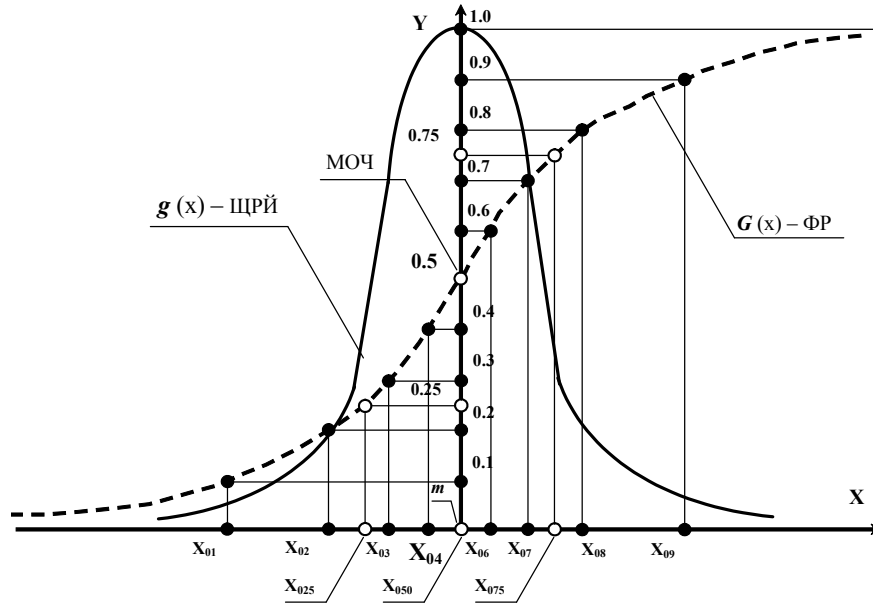


Рисунок 1 – Графічне представлення функції щільності розподілу ймовірності $g(x)$, функції розподілу $G(x)$ (ФР), та значень аргументу функції $G(x) - x_{01} \dots x_{09}$

Оскільки квантилі варіаційних рядків абсолютних значень дискретних відліків поточних кепстрів сигналів, як порядкові статистики вибіркового розподілу ймовірностей абсолютних значень дискретних відліків поточних кепстрів для кожної частоти дискретного кепстру сигналу розподіляються за нормальним законом незалежно від негаусовського характеру багатовимірних функцій розподілу вибірових значень кепстрів сигналів, що досліджуються, виконується «нормалізація» розподілу випадкових значень кепстру. Ця якість істотно впливатиме на інформаційне наповнення кепстральної функції сигналу, що, в свою чергу значно підвищить ефективність функціонування систем кепстрального аналізу складних сигналів, в тому числі систем ідентифікування джерела складних сигналів за критеріями згоди, котрі орієнтовані на статистичну ідентифікацію нормально розподілених випадкових величин [2, 3].

Сигнали, що досліджувалися в роботі – це студійні записи на дисках професійних дикторів (жінок та чоловіків російською та англійською мовами). Диски є доступними в продажу. Характеристики мовного сигналу в оригіналі – формат запису .mp3; 192 кБіт/сек; $F_{\text{дискр.}}$ становить 44,1 кГц; стерео. Сигнал із такими параметрами потребує попередньої обробки.

Особливості якостей кепстру потребують вирахування кепстральної функції не для довільної мовної реалізації, а для деякої мовної одиниці, тому було прийнято рішення використовувати для отримання порядкових статистик кепстру мовного сигналу слово «and». З вихідного сигналу було взято 62 таких слова що (з розрахунку одне слово на сегмент аналізу) надасть достатній статистичний матеріал для отримання залежностей квантилів від частоти кепстру сигналу та для оцінки отриманих статистичних даних (рис. 2). Зазвичай, у кепстральному аналізі використовуються більш короткі сегменти, довжина котрих відповідає інтервалу квазістаціонарності мовного сигналу, оскільки апарат кепстрального (гомоморфного) аналізу використовується для фонемного аналізу. В роботі прийнято, що розмір сегменту аналізу сигналу (інтервал стаціонарності) становить 6614 відліків або 0,3 с., відповідно до значення $F_{\text{аєнєд.}} = 22050$ Гц (моно). Таким чином, на сегменті аналізу, що нараховує 6614 відліків, необхідно

розглядати 6614 значень A_i на 6614 значеннях f_i та 6614 гістограм багатовимірної функції щільності розподілу ймовірності $g(A_i / f_i)$. Кількість сегментів аналізу мовного сигналу зазвичай велика та залежить від довжини мовної реалізації.

Методику оцінювання параметрів кепстру реалізовано в системі MatLAB, тому виконуємо зміну формату сигналу стосовно Windows PCM, що відповідає вимогам до акустичних сигналів, які оброблятимуться в середовищі MatLAB. Оскільки на ефективність функціонування методики не впливає тип реалізації (стереофонічна або монофонічна), то в той же час здійснюємо перетворення до монофонічної реалізації. Тобто отримуємо наступні характеристики сигналу – формат запису .wav; 705 кБіт/сек; $F_{\text{дискр.}}$ становить 22050 Гц; 16 біт; моно. Ці перетворення виконувалися за допомогою редактора звукових файлів Cool Edit Pro v.2.00.

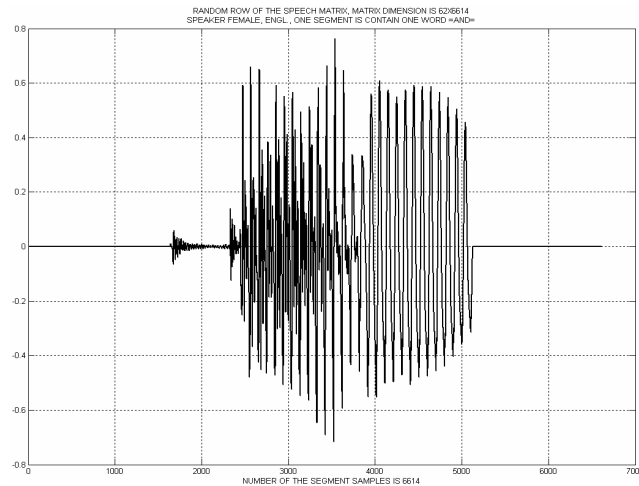


Рисунок 2 – Довільний сегмент аналізу мовного сигналу диктора із числом відліків 6614

Рішення поставленого завдання досягається тим, що в заданій послідовності взаємопов'язаних операцій здійснюють перетворення сигналу, що досліджується (рис. 3).

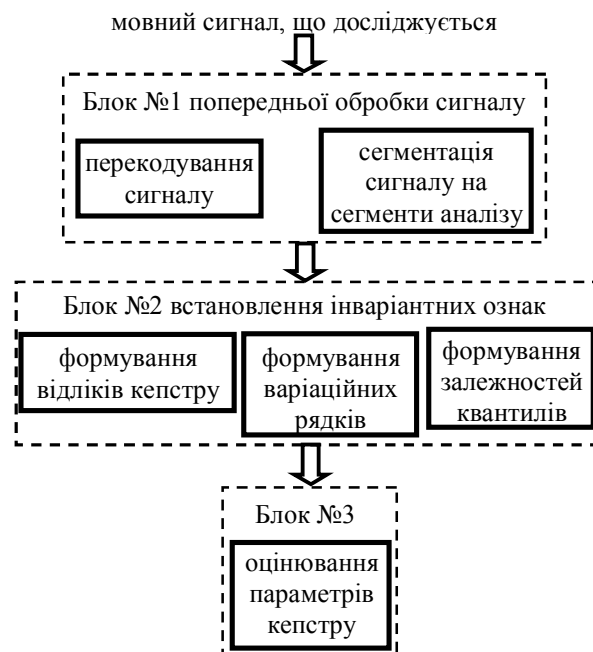


Рисунок 3 – Блок-схема методики робастної оцінки параметрів кепстру мовного сигналу

У першому блоці попередньої обробки сигналу виконується перекодування та сегментація сигналу на сегменти аналізу однакової тривалості. Функціоналом другого блоку передбачено вирахування реального кепстру речовинного сигналу $x(t)$ шляхом застосування зворотного перетворення Фур'є до логарифму комплексної спектральної щільності Фур'є сигналу.

$$x(t) = x_c(t) + x_s(t);$$

$$x(t) \xrightarrow{\text{FFT}} X_F(f) \xrightarrow{\text{abs}} A_x(f) \xrightarrow{\ln} \ln(A_x(f) + 10^{-5}) \xrightarrow{F^{-1}} y(t) = \hat{\text{cep}}[x(t)];$$

де $x(t)$ - довільний речовинний сигнал; $x_c(t)$ - парна компонента сигналу; $x_s(t)$ - непарна компонента сигналу.

Пряме та зворотне перетворення Фур'є (БПФ):

$$X_F(f) = \int_{-\infty}^{\infty} x(t) \exp[-j2\pi ft] dt; \quad x(t) = \int_{-\infty}^{\infty} X_F(f) \exp[j2\pi ft] df.$$

В результаті виконання розглянутих вище операцій перетворення мовного сигналу на сегменті для кожного дискретного значення частоти його поточного кепстру може бути отримана сукупність значень модулів $x_i, i = 1 \dots n$, за всіма сегментами аналізу. Ця сукупність незалежних випадкових величин може вважатися вибіркою об'єму n (кількість сегментів аналізу) з генеральної сукупності, функцією розподілу, що характеризується $G(x)$ [2]. В даних умовах проведення експерименту зміна початку відліку (використання правила помилкового нуля), масштабу і ладу (перемішування, зміна порядку нумерації) значень $x_i, i = 1 \dots n$, ніяк не повинні позначатися на результатах статистичного аналізу в аспекті спотворення даних.

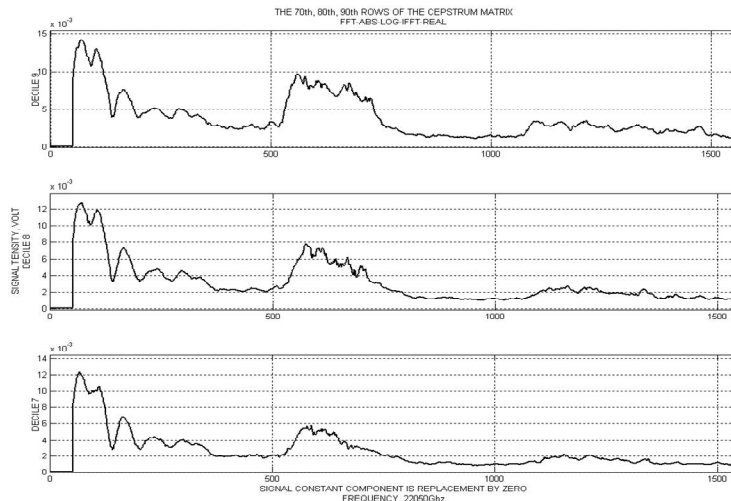


Рисунок 4 – Реальний кепстр мовного сигналу на сегменті залежності квантилів №№ 7, 8, 9

Для проведення будь-якої обробки значень $[x_i]$ та оцінювання статистик вибірки можна сформувати варіаційний ряд $[\hat{x}_{i/n}]$ значень $[x_i]$ для кожної частоти кепстру сигналу:

$$\hat{x}_{1/n} \leq \hat{x}_{2/n} \leq \dots \leq \hat{x}_{i/n} \leq \dots \leq \hat{x}_{n-1/n} \leq \hat{x}_{n/n},$$

розташувавши їх в лад зростання, де індекс елементів варіаційного ряду $[\hat{x}_{i/n}]$ запишемо у виді i/n , тобто $\hat{x}_{1/n} = \min_i[x_i], \hat{x}_{n/n} = \max_i[x_i]$.

Для кожного варіаційного ряду $[\hat{x}_{i/n}]$ за значенням номерів i/n елементів цього ряду визначають вибіркові квантили $\hat{x}_\alpha: \hat{x}_{j/4}, j = 1, 2, 3$ – вибіркові квантили ($\hat{x}_{0,25}, \hat{x}_{0,5}, \hat{x}_{0,75}$); $\hat{x}_{j/10}, j = 1 \dots 9$ –

вибіркові децилі $(\hat{x}_{0,1}, \hat{x}_{0,2}, \dots, \hat{x}_{0,9})$; $\hat{x}_{j/20}, j = 1 \dots 19$ – вибіркові п'ятилі $(\hat{x}_{0,05}, \hat{x}_{0,1}, \hat{x}_{0,15}, \hat{x}_{0,20}, \dots, \hat{x}_{0,95})$; $\hat{x}_{j/100}, j = 1 \dots 99$ – вибіркові процентилі $(\hat{x}_{0,01}, \hat{x}_{0,02}, \dots, \hat{x}_{0,99})$, які ділять весь діапазон $[\hat{x}_{1/n}, \hat{x}_{n/n}]$ зміни значень варіаційного ряду $[\hat{x}_{i/n}]$ відповідно на 4, 10, 20 і 100 нерівних за величиною підінтервалів, в які потрапляє однакова кількість значень ряду [2, 3]. Після чого виконується формування залежності виділених квантилів від дискретних значень частоти кепстру сигналу (рис. 4).

У третьому блоці (оцінці) виконується оцінка значень залежностей кепстру з метою подальшого використання цих даних відповідно до завдань, що вирішуються апаратом кепстрального аналізу в окремих випадках.

У зв'язку з тим, що квантилі варіаційних рядків як порядкові статистики вибіркового розподілу ймовірностей поточних кепстрів для кожної частоти дискретного кепстру сигналу розподіляються за нормальним законом (незалежно від негаусовського характеру багатовимірних функцій розподілу вибірових значень кепстрів сигналів, що досліджуються), істотно підвищується ефективність функціонування систем кепстрального аналізу складних сигналів, в тому числі систем ідентифікації джерела складних сигналів за критеріями згоди, котрі орієнтовані на статистичну ідентифікацію нормально розподілених випадкових величин [2, 3].

Постійна складова отриманого кепстру піддається усуненню шляхом заміни на нульові компоненти перших 45 Гц кепстру, це дає можливість спостерігати вигляд кепстру без спотворень, що спричиняє дуже велике значення амплітуди постійної складової. На рис. 4 наведено три децилі сьомий, восьмий та дев'ятий. Спостерігаються складові кепстру в смугах від 50 до 200 Гц, від 500 до 750 Гц, та від 1100 до 1500 Гц., котрі вважаються трьома функціями, що характеризують висоту основного тону, функцію голосових зв'язок та функцію мовного тракту [1].

Виходячи з якостей порядкових статистик, отриманий кепстр є статистично вірним та стабільним, стабільність передається в достатній кількості статистичних, не усереднених даних. Отриманий таким чином кепстр практично не змінює свій образ від сегмента до сегмента аналізу мовного сигналу диктора та має наявні відмінності від диктора до диктора. Тобто, в аспекті питання ідентифікування джерела мовного сигналу, незважаючи на малу інформативність (порівняно з голосовим статистичним спектральним образом) (рис. 5), отримані залежності квантилів варіаційних рядків кепстру можливо використовувати в системах ототожнення. На рис. 5 наведено зображення спектрального статистичного образу того ж мовного сигналу на сегменті аналізу, квантиль № 9.

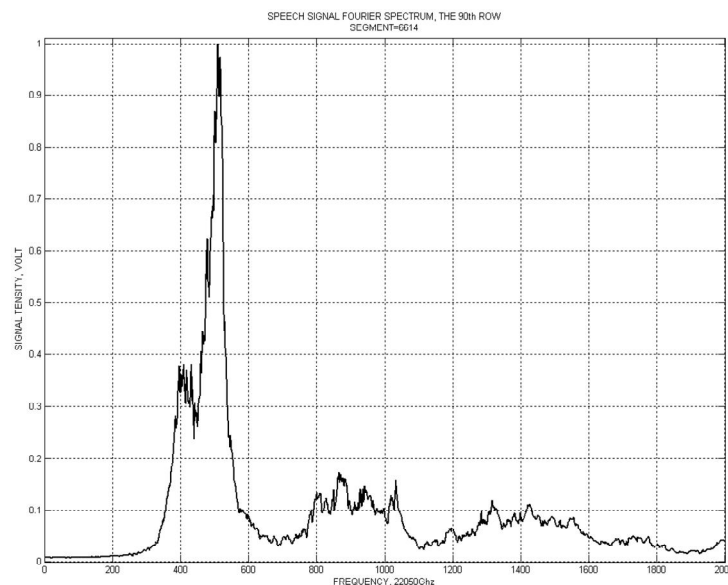


Рисунок 5 – Спектр Фур'є сигналу на сегменті аналізу

Взагалі кепстри складних сигналів за своєю природою є функціями із дуже вузькою смугою (порівняно зі спектрами) із вкрай обмеженою інформативністю. Кепстри, що отримані за допомогою класичних методик [1], на додачу, ще й усереднені, що не може не позначитися на можливостях щодо їхнього застосування.

Висновки

Результатом дослідження є новітні дані (отримані експериментальним шляхом, котрі характеризують реальний кепстр мовного сигналу без втрат характеризуючих складових мовного сигналу) щодо порядкових статистик $\hat{x}_\alpha(f_k)$ кепстру мовного сигналу. Кепстр у цьому випадку представлено статистичними неусередненими даними – графіками залежностей розподілу квантилів кепстру мовного сигналу. Методика є стійкою до помилок у каналах зв'язку, реєстрації та відтворення складних сигналів завдяки тому, що шумові складові залежностей розміщені по краях варіаційного ряду, ці квантилі не використовуються для дослідження кепстру [3].

Комплексне використання перерахованих відмічених особливостей та пов'язаних з ними позитивних ефектів дозволяють забезпечити підвищення ефективності систем кепстрального аналізу мовних (та інших складних) сигналів.

Література: 1. Голд Б., Рейдер Ч. *Цифровая обработка сигналов. Пер. с англ., под ред. А. М. Трахтмана. М., «Сов. радио», 1973, 368с.* 2. Дейдвуд Дж. *Порядковые статистики: Пер. с англ. – М.: Мир, 1989.-540с.,ил.* 3. Патент України на винахід № 88208. *Спосіб кепстральної ідентифікації сигналів / В. Л. Селетков., М. В. Кузнецов. Опубліковано бюл. «Промислова власність» № 13 від 10.07.2009 р.*

УДК 651.928

ОХОРОНА АДВОКАТСЬКОЇ ТАЄМНИЦІ В КРИМІНАЛЬНОМУ СУДОЧИНСТВІ, ЇЇ ОБ'ЄКТ ТА СУБ'ЄКТИ

Роман Луценко

Національний технічний університет України «КПІ»

Анотація: Розглядаються питання, пов'язані з особливостями правового регулювання адвокатської таємниці в кримінальному судочинстві України. Визначені об'єкт та суб'єкт адвокатської таємниці, часові межі її збереження.

Summary: Some aspects of defence attorney secrets legal regulations in criminal procedure of Ukraine are discussed. An object and a subject of a defence attorney secret, as well as the secrecy time limits are defined.

Ключові слова: Адвокатська таємниця, конфіденційність, інформація, професійна таємниця, конфіденційна інформація.

І Вступ

Україна ратифікувала Міжнародний пакт про громадянські і політичні права 1966 року і Європейську конвенцію про захист прав людини і основоположних свобод 1950 року у 1973 і 1997 роках відповідно. І взяла на себе зобов'язання забезпечити кожну людину, права якої порушено, ефективними засобами правового захисту (ст. 3 Пакту і ст. 13 Конвенції) [1, 2].

В свою чергу стаття 55 Конституції України закріпила право кожної особи на правовий захист, а ст. 59 – право на правову допомогу, визначаючи, що ці дії покладаються саме на адвокатуру [3]. В межах Кримінально-процесуального кодексу України (далі КПКУ), а саме в ст. 21, зазначено, що підозрюваному, обвинуваченому і підсудному забезпечується право на захист [4].

Звернення особи за юридичною допомогою в рамках кримінального судочинства до адвоката передбачає встановлення довіри й відкритості між ними, належного і якісного виконання адвокатом своїх професійних обов'язків. Правова основа довіри між адвокатом і його клієнтом – це закріплення інституту адвокатської таємниці, дотримання якого вимагають як міжнародні стандарти, так і законодавство України.

Конституційною основою професійних таємниць в Україні є ст. 32 Конституції України, яка містить заборону втручатись в особисте і сімейне життя, а також передбачає, що збирання, зберігання, використання та поширення конфіденційної інформації про особу без її згоди не допускається, крім випадків, визначених законом, і лише в інтересах національної безпеки, економічного добробуту та прав людини.