Прогнозування спектральної обвідної мовленнєвих сигналів в просторі лінійних спектральних параметрів найвищого розщеплення

Павлов О.І., Герасименко К.В., Аполонов Є.В.

Національний технічний університет України «Київський політехнічний інститут»,
Міжнародний науково-навчальний центр інформаційних технологій та систем
op-mail@ukr.net

Анотація

Наведений аналіз спектральних параметрів (СП) і процесів їх перетворення, які застосовуються для прогнозування спектральної обвідної (СО) мовленнєвих сигналів в Рекомендації *G*.729. Запропонована методика порівняння характеристик процесів прогнозування СО в різних альтернативних еквівалентних просторах СП. Представлені результати експериментальних досліджень прогнозування СО в просторах класичних лінійних СП (ЛСП) і ЛСП найвищого розщеплення (ЛСП-НР). Показано, що перехід до просторів ЛСП-НР зменшує похибку прогнозування СО.

1. Вступ

Під час розробки та перевірки методів ефективного кодування СО застосовують різноманітні техніки стиснення даних, до яких відносяться моделювання, квантування, прогнозування, інтерполяція. Яскравим прикладом такого підходу може бути Рекомендації *G*.729, де основний параметричний опис СО здійснюється в просторі коефіцієнтів лінійного прогнозування (ЛП), квантування і прогнозування — в просторі класичних лінійних спектральних частот (ЛСЧ), інтерполяція — в просторі класичних лінійних спектральних проекцій.

Проте, в роботах [1—11] було розроблено новий метод ЛСП-НР і показано, що він має цілу низку переваг перед класичним методом ЛСП, який є лише першим етапом перетворень методу ЛСП-НР.

Зауважимо, що у вказаних роботах абревіатура ЛСП-НР розшифровувалася як «ЛСП найвищої ***регресії***», однак, більш точним слід вважати термін «ЛСП найвищого ***розщеплення***», який і буде вживатися далі.

Було показано, що параметри ЛСП-НР мають значно меншу обчислювальна складність [3, 4], не містять методичної похибки їх оцінювання [1, 2], їх застосування зменшує похибку в операціях квантування [9] і інтерполяції [8].

Метою даної роботи є дослідження і порівняння характеристик процесів прогнозування параметрів СО в просторах класичних ЛСП і ЛСП-НР.

2. Екстраполяція класичних ЛСП в Рекомендації *ITU-T G*.729

2.1. Класичні вектори *LPC, LSP* і *LSF*

Для кодування СО в Рекомендації *G*.729 [12, 13] на кожному поточному кадрі аналізу (кадрі з індексом m) спочатку, за допомогою методу ЛП, проводиться розрахунок компонент вектору *LPC* $a_i^{(m)}$, $i = 1,...,10$, на підставі яких, за класичним методом ЛСП, оцінюються компоненти вектору *LSP* $x_i^{(m)}$, $i = 1,...,10$. Ці компоненти, в свою чергу, перетворюються в відповідні компоненти вектору *LSF* $\omega_i^{(m)}$, $\omega_i = \arccos(x_i)$, $i = 1,...,10$. Останні розглядаються як залежні ***композитні*** величини, що утворюються в результаті процесу ковзного середнього (КС) через суму зважених обілених ***елементарних*** (некомпозитних) *LSF* $l_i^{(m)}$, $i = 1,...,10$, і використовуються разом з ними в операціях векторного квантування та векторного прогнозування СО.

2.2. Модель композитних і елементарних *LSF*

Прийнята в Рекомендації *G*.729 теоретична модель утворення композитних компонент вектору *LSF* $\omega_i^{(m)}$, $i = 1,...,10$ наведена на рис. 1 і задається формулою КС

$$\omega_i^{(m)} = p_{i,0} l_i^{(m)} + \sum_{k=1}^{K} p_{i,k} l_i^{(m-k)}, \; i = 1,...,10, \qquad (1)$$

де $K = 4$ — порядок провісника КС- моделі, $p_{i,k}$ — коефіцієнти лінійної регресії композитних *LSF* $\omega_i^{(m)}$ на обілені елементарні *LSF* $l_i^{(m)}$, $p_{i,0} = 1 - \sum_{k=1}^{K} p_{i,k}$.

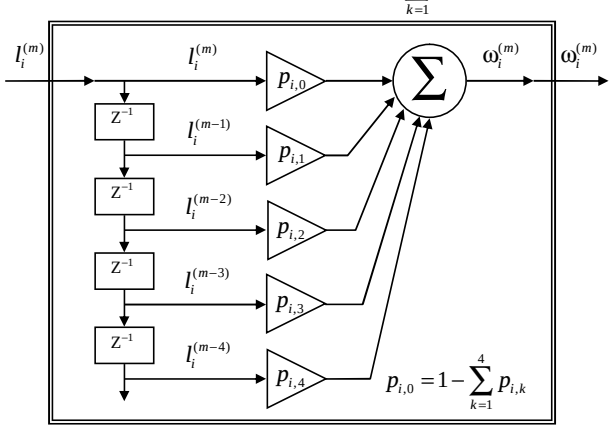

Рисунок 1: Теоретична модель КС- процесу утворення композитних *LSF* з обілених елементарних *LSF*

Наперед невідомими при цьому є як значення обілених елементарних *LSF* $l_i^{(m)}$, так і значення *оптимальних* коефіцієнтів КС $p_{i,k}$.

2.3. Вагові коефіцієнти векторного КС- провісника

Процедура оцінювання коефіцієнтів $p_{i,k}$ теоретичної КС- моделі, рис. 1, на практиці безпосередньо не виконується. Замість цього використовуються *рекомендовані* набори *квантованих* коефіцієнтів $\hat{p}_{i,k}$ за фіксованою матричною кодовою книгою, яка в Рекомендації G.729 містить дві кодових матриці розміром 10 x 4, що задають два відповідних КС- провісники. Який з двох провісників слід обрати — з'ясовується в передавачі під час випробування кожного з них за методом «аналіз через синтез» (шляхом оцінювання елементарних *LSF*, їх квантування та відновлення композитних *LSF*). Номер кращого провісника передається до приймача.

2.4. Оцінювання елементарних *LSF*

Оцінювання елементарних *LSF* $l_i^{(m)}$, $i = 1,...,10$ на практиці здійснюється за зворотною КС- моделлю, рис. 2, де вони, фактично, є *нормованими похибками прогнозування* поточних композитних *LSF* $\omega_i^{(m)}$ за *квантованими* попередніми похибками $\hat{l}_i^{(m-k)}$ такого прогнозування при використанні *квантованих* коефіцієнтів $\hat{p}_{i,k}$, і описується формулами

$$l_i^{(m)} = \left(\omega_i^{(m)} - \sum_{k=1}^{4} \hat{p}_{i,k}\hat{l}_i^{(m-k)}\right) \Big/ \hat{p}_{i,0}\,,\ \hat{l}_i^{(m)} = \mathrm{q}(l_i^{(m)})\,,\ i = 1,...,10\,. \quad (2)$$

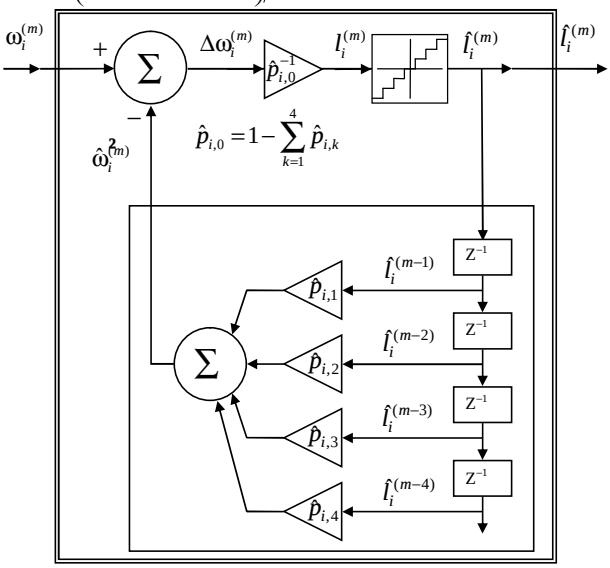

Рисунок 2: Схема оцінювання елементарних *LSF* в Рекомендації *G*.729

2.5. Квантування елементарних *LSF*

Процес квантування вектору елементарних *LSF* $l_i^{(m)}$, $i = 1,...,10$, який показаний на рис. 2 у вигляді одного блоку, здійснюється в передавачі за допомогою дворівневої схеми векторного квантування.

На першому рівні використовується векторний квантувач 10-го порядку, який працює з кодовою книгою L1, що має 128 кодових векторів. На другому рівні теж використовується векторний квантувач 10-го порядку, але він працює з двома розщепленими кодовими книгами L2 та L3 5-го порядку (кожна з яких має по 32 кодових вектори).

2.6. Пересування квантованих елементарних *LSF*

Квантований вектор елементарних *LSF* $\hat{l}_i^{(m)}$, $i = 1,...,10$ підлягає (в разі необхідності) операції *пересування своїх компонент* для запобігання гострих резонансів в квантованому фільтрі синтезу мовленнєвого сигналу. В результаті чого його посунуті компоненти (елементарні *LSF* $\hat{l}_i^{(m)}$) розміщуються так, що сусідні компоненти мають відстань одна від одної не менше за мінімальну відстань *J*:

for $i = 2,...,10$

$$if\,(\hat{l}_{i-1} > \hat{l}_i - J)\,\{\,\hat{l}_{i-1} = (\hat{l}_i + \hat{l}_{i-1} - J)/2;\ \ \hat{l}_i = (\hat{l}_i + \hat{l}_{i-1} + J)/2\,\}\,\cdot$$

end

Така операція пересування виконується двічі. Спочатку для *J* = 0.0012, потім для *J* = 0.0006.

2.7. Відновлення квантованих композитних *LSF*

Квантовані елементарні *LSF* $\hat{l}_i^{(m)}$, $i = 1,...,10$ після їх пересування використовуються для відновлення однакових (в передавачі і приймачі) оцінок квантованих композитних *LSF* $\hat{\omega}_i^{(m)}$, $i = 1,...,10$ за тими самими (які застосовувалися під час аналізу) квантованими коефіцієнтами КС $\hat{p}_{i,k}$, рис. 3, а саме:

$$\hat{\omega}_i^{(m)} = \hat{p}_{i,0}\hat{l}_i^{(m)} + \sum_{k=1}^{K} \hat{p}_{i,k}\hat{l}_i^{(m-k)}\,,\ i = 1,...,10\,, \quad (3)$$

де $\hat{\omega}_i^{(m)}$, $\hat{l}_i^{(m)}$, $\hat{p}_{i,k}$ — квантовані: композитні *LSF*, елементарні *LSF* та коефіцієнти КС,. $\hat{p}_{i,0} = 1 - \sum_{k=1}^{K} \hat{p}_{i,k}$.

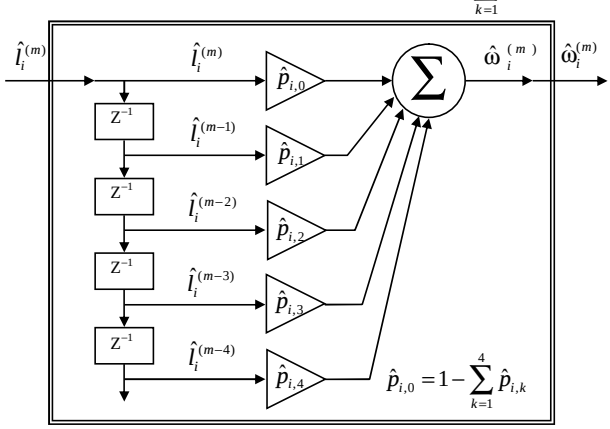

Рисунок 3: Схема відновлення квантованих композитних *LSF* з квантованих елементарних *LSF* за допомогою КС-процесу з квантованими коефіцієнтами

Початкові значення квантованих елементарних *LSF* $\hat{l}_i^{(k)}$, $i = 1,...,10$ для всіх $k < 0$ задаються на першому кадрі при запуску процесу відновлення як $\hat{l}_i^{(k)} = i\pi/11$, $i = 1,...,10$, і є рівномірно розміщеними на інтервалі $]0;\pi[$.

2.8. Перевірка на стійкість відновлених *LSF*

Після обчислення квантованих композитних *LSF* $\hat{\omega}_i^{(m)}$, $i = 1,...,10$ відповідний до них фільтр синтезу перевіряється на стабільність. Це робиться таким чином:

- композитні *LSF* $\hat{\omega}_i^{(m)}$ розташовуються за зростанням;
- якщо $\hat{\omega}_1^{(m)} < 0{,}005$ то $\hat{\omega}_1^{(m)} = 0.005$;
- якщо $\hat{\omega}_{i+1}^{(m)} - \hat{\omega}_i^{(m)} < 0.0391$, то $\hat{\omega}_{i+1}^{(m)} = \hat{\omega}_i^{(m)} + 0.0391$;
- якщо $\hat{\omega}_{10}^{(m)} > 3.135$, то $\hat{\omega}_{10}^{(m)} = 3.135$.

2.9. Висновки

Прогнозування СО кожного поточного кадру мовленнєвого сигналу здійснюється в просторі класичних *LSF* за допомогою КС- провісника з фіксованими наборами коефіцієнтів, який обчислює зважену суму квантованих (і скоригованих шляхом пересування) похибок такого прогнозування (обілених елементарних *LSF*) на попередніх 4 кадрах.

Використання моделі у вигляді КС- провісника ускладнює порівняння характеристик процесів прогнозування СО в різних альтернативних просторах СП через відсутність простих методів оцінювання оптимальних КС- коефіцієнтів.

3. Методика порівняння екстраполяційних характеристик класичних ЛСП і ЛСП-НР

3.1. Зв'язок АР-, КС- та АРКС- моделей

Відомо, що задану АР-, КС- або АРКС- модель, яка має скінчену кількість параметрів можна апроксимувати у вигляді будь-якої з двох інших моделей. АРКС- та КС-процеси можна подати за допомогою однієї АР- моделі, в загальному випадку з нескінченою розрядністю. Цей факт дуже важливий, бо дозволяє обирати будь-яку з трьох моделей і отримувати задовільну апроксимацію при достатньо великому порядку моделі. Оскільки більшість ефективних алгоритмів оцінювання параметрів моделей розроблено для АР- моделі, то саме АР- модель великого порядку обирають як перший етап оцінювання параметрів КС- та АРКС- моделей, [14].

В нашому випадку можна записати таке різницеве рівняння: $\omega_n = -\sum_{k=1}^{p} a_k \omega_{n-k} + \sum_{k=0}^{q} b_k l_{n-k} = \sum_{k=0}^{\infty} h_k l_{n-k}$, де ω_n — послідовність на виході каузального фільтру ($h_k = 0$ при $k < 0$), який формує дані, що спостерігаються, l_n — вхідна збуджуюча послідовність. Не звужуючи розгляду можна вважати $b_0 = 1$. Тоді системна функція такого фільтру є $H(z) = B(z)/A(z)$, де $A(z) = 1 + \sum_{k=1}^{p} a_k z^{-k}$, $B(z) = 1 + \sum_{k=1}^{q} b_k z^{-k}$, $H(z) = 1 + \sum_{k=1}^{\infty} h_k z^{-k}$. Нулі поліномів $A(z)$ і $B(z)$ мають знаходитися всередині одиничного кола на z-площині для того, щоб каузальний фільтр був стійким і мінімально-фазовим. Якщо треба подати таку АРКС(p, q) модель у вигляді АР(∞) моделі, тобто $B(z)/A(z) = 1/C(z)$, де $C(z) = 1 + \sum_{k=1}^{\infty} c_k z^{-k}$, то коефіцієнти c_k (АР параметри еквівалентної АР(∞) моделі) будуть визначатися так:

$$c_n = \begin{cases} 1, & n = 0; \\ -\sum_{k=1}^{q} b_k c_{n-k} + a_n, & 1 \le n \le p; \\ -\sum_{k=1}^{q} b_k c_{n-k}, & n > p; \end{cases} \qquad (4)$$

з початковими умовами $c_{-1} = ... = c_{-q} = 0$. І навпаки, якщо треба відновити параметри АРКС(p, q) моделі, еквівалентної до відомої АР(∞) моделі, то спочатку знаходять значення коефіцієнтів b_k (КС параметрів АРКС(p, q) моделі), розв'язуючи матричне рівняння

$$\begin{pmatrix} c_p & c_{p-1} & \dots & c_{p-q+1} \\ c_{p+1} & c_p & \dots & c_{p-q+2} \\ \dots & \dots & \dots & \dots \\ c_{p+q-1} & c_{p+q-2} & \dots & c_p \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_q \end{pmatrix} = -\begin{pmatrix} c_{p+1} \\ c_{p+2} \\ \dots \\ c_{p+q} \end{pmatrix} \qquad (5)$$

відносно b_k, яке відповідає (4) при $p + 1 \le k \le p + q$. Оскільки матриця коефіцієнтів c_k в рівнянні (5) є матрицею Тепліца, то для його розв'язання можна використати алгоритм Левінсона – Дурбіна. Значення коефіцієнтів a_k (АР параметрів АРКС(p, q) моделі) відновлюють за допомогою згортки $a_n = c_n + \sum_{k=1}^{q} b_k c_{n-k}$, $1 \le n \le p$, після обчислення коефіцієнтів b_k.

3.2. Досліджувана модель спектральних параметрів

Для порівняння характеристик процесів прогнозування СО в різних альтернативних еквівалентних просторах СП була обрана АР- модель їх утворення. На рис. 4 наведений приклад застосування такої моделі у вигляді схеми відновлення композитних *LSF* $\omega_i^{(m)}$.

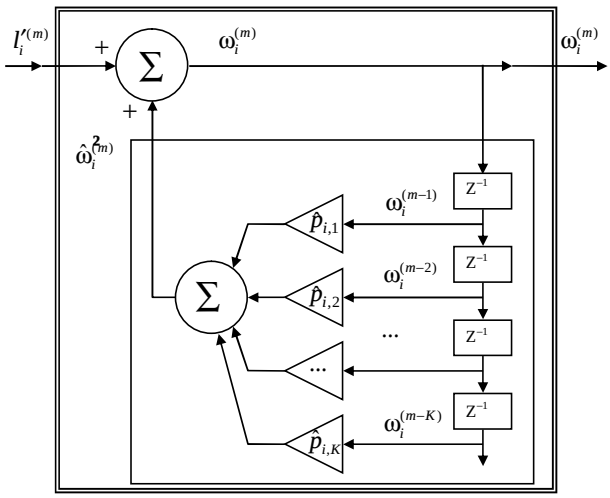

Рисунок 4: Досліджувана теоретична АР- модель утворення СП на прикладі схеми відновлення композитних *LSF*

В схемі на рис. 4 використовується АР- провісник *K*-го порядку з *квантованими* АР- коефіцієнтами $\hat{p}_{i,k}$. Сигналом збудження такої схеми є відповідні обілені похибки АР- прогнозування $l_i'^{(m)}$. Оцінювання їх значень здійснюється за схемою аналізу, рис. 5, в два етапи. На першому етапі аналізу методом ЛП обчислюються матриці *оптимальних неквантованих* АР- коефіцієнтів $p_{i,k}$ (а не $\hat{p}_{i,k}$, які показано на рис. 5) для всієї навчальної вибірки досліджуваних СП.

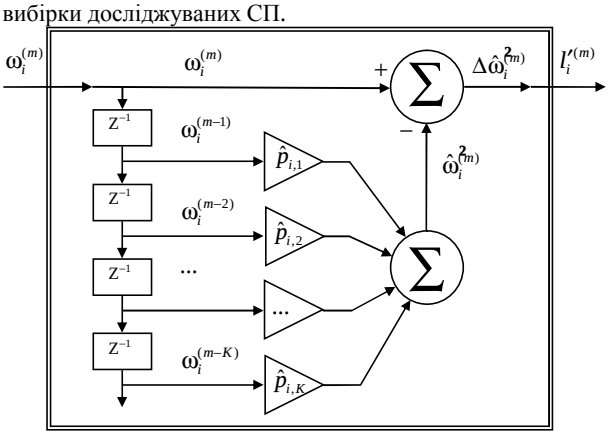

Рисунок 5: Схема оцінювання обілених похибок ЛП СП на прикладі композитних *LSF*

По завершенню першого етапу всі отримані матриці *оптимальних неквантованих* АР- коефіцієнтів $p_{i,k}$ об'єднуються в навчальну вибірку АР- коефіцієнтів, на підставі якої, методом кластерного аналізу (за допомогою *LBG*- алгоритму [15]), синтезується матрична кодова книга с заданою кількістю кодових матриць. Елементи отриманих кодових матриць використовуються в якості *квантованих* АР- коефіцієнтів $\hat{p}_{i,k}$ на другому етапі аналізу похибки АР- прогнозування $l_i'^{(m)}$ (повторно для всієї навчальної вибірки досліджуваних СП), що саме і показано на рис. 5.

3.3. Міра спотворень

Відповідно до схеми на рис. 5 для *i*-компоненти вектору $\omega_i^{(m)}$, $i = 1,...,10$ на *n*-м кадрі аналізу обчислюються:

1. Поточне прогнозоване значення *i*-компоненти:

$$\hat{\omega}_{i,m} = \sum_{k=1}^{K} p_{i,k}\omega_{i,m-k} \ . \tag{6}$$

2. Поточна похибка прогнозування *i*-компоненти:

$$l'_{i,m} = \omega_{i,m} - \hat{\omega}_{i,m} \ . \tag{7}$$

3. Енергія похибки прогнозування *i*-компоненти:

$$\xi_i^{(n)} = \sum_{m=m_{0,n}}^{m_{1,n}} l'^2_{i,m} \ . \tag{8}$$

4. Енергія *i*-компоненти:

$$E_i^{(n)} = \sum_{m=m_{0,n}}^{m_{1,n}} \omega^2_{i,m} \ . \tag{9}$$

5. Відносна енергія похибки прогнозування *i*-компоненти:

$$\alpha_i^{(n)} = \frac{\xi_i^{(n)}}{E_i^{(n)}} = \sum_{m=m_{0,n}}^{m_{1,n}} l'^2_{i,m} \Bigg/ \sum_{m=m_{0,n}}^{m_{1,n}} \omega^2_{i,m} \ , \tag{10}$$

$$10\lg\alpha_i^{(n)} = 10\lg\left(\sum_{m=m_{0,n}}^{m_{1,n}} l'^2_{i,m} \Bigg/ \sum_{m=m_{0,n}}^{m_{1,n}} \omega^2_{i,m} \right). \tag{11}$$

Для всієї навчальної вибірки (тобто за всіма кадрами аналізу) обчислюються:

6. Максимальна відносна енергія похибки прогнозування *i*-компоненти:

$$\alpha_{i,\max} = \max_n \left\{ \alpha_i^{(n)} \right\} . \tag{12}$$

7. Мінімальна відносна енергія похибки прогнозування *i*-компоненти:

$$\alpha_{i,\min} = \min_n \left\{ \alpha_i^{(n)} \right\} . \tag{13}$$

8. Середня відносна енергія похибки прогнозування *i*-компоненти:

$$\alpha_{i,ave} = \frac{1}{n}\sum_n \alpha_i^{(n)} = \frac{1}{n}\sum_n \left(\sum_{m=m_{0,n}}^{m_{1,n}} l'^2_{i,m} \Bigg/ \sum_{m=m_{0,n}}^{m_{1,n}} \omega^2_{i,m} \right), \tag{14}$$

$$10\lg\alpha_{i,ave} = 10\lg\frac{1}{n}\sum_n \left(\sum_{m=m_{0,n}}^{m_{1,n}} l'^2_{i,m} \Bigg/ \sum_{m=m_{0,n}}^{m_{1,n}} \omega^2_{i,m} \right). \tag{15}$$

9. Стандартне відхилення відносної енергії похибки прогнозування *i*-компоненти:

$$\alpha_{i,\mathrm{std}} = \sqrt{\left(\sum_n \left(\alpha_i^{(n)} \right)^2 \right) - \left(\frac{1}{n}\sum_n \alpha_i^{(n)} \right)^2} \ . \tag{16}$$

Середня відносна енергія похибки прогнозування *i*-компоненти (15) є основною мірою спотворень, яка використовується під час всіх подальших досліджень.

3.4. Методика досліджень

Кореляційні властивості часових послідовностей векторів СП досліджуються шляхом розрахунку АР-коефіцієнтів ЛП кожної часової траєкторії, що утворюється з компонент цих векторів.

Дослідження виконуються за однаковою методикою для послідовностей векторів СП, що кодують форми СО на кожному кадрі одного й того ж мовленнєвого сигналу в чотирьох еквівалентних спектральних просторах:

- *LSF* — просторі класичних (першого ступеня розщеплення) лінійних спектральних частот (ЛСЧ).
- *LSP* — просторі класичних (першого ступеня розщеплення) лінійних спектральних проекцій (ЛСПр).
- *FHI* — просторі лінійних спектральних частот найвищого ступеня розщеплення (ЛСЧ-НР).
- *PHI* — просторі лінійних спектральних проекцій найвищого ступеня розщеплення (ЛСПр-НР).

Реалізується коваріаційний підхід оцінювання кореляційної матриці вхідних даних на кожному кроці аналізу.

Перша частина досліджень виконуються з метою з'ясування впливу на похибку ЛП СП таких факторів, як:

- розміру кадру аналізу (кількості векторів СП),
- порядку прогнозування (кількості векторів СП).

За результатами першої частини досліджень обираються оптимальні налаштування методу ЛП для другої частини досліджень — оцінювання погіршення похибок ЛП після квантування отриманих оптимальних АР- коефіцієнтів.

4. Результати експериментальних досліджень

4.1. Вплив порядку прогнозування СП

При проведенні дослідження впливу порядку ЛП на відносні похибки прогнозування було обрано розмір кадру аналізу у 100 векторів (що відповідає 100 СО, обчисленим через кожні 10 мс, тобто 1000 мс мовленнєвого сигналу).

Зсув кожного наступного кадру аналізу відносно попереднього виконувався на 10 векторів СП, що відповідає 100 мс мовленнєвого сигналу.

Під час досліджень для кожного кадру аналізу векторів СП обраховувалася відносна похибка прогнозування (10), а за результатами обробки всіх кадрів (для даних умов була отримана вибірка з 8858 кадрів) відбиралося найбільше (12) і найменше (13) значення, а також обчислювалось середнє значення (15) і стандартне відхилення (16).

Результати досліджень для векторів ЛСЧ і ЛСЧ-НР наведено на рис. 6.

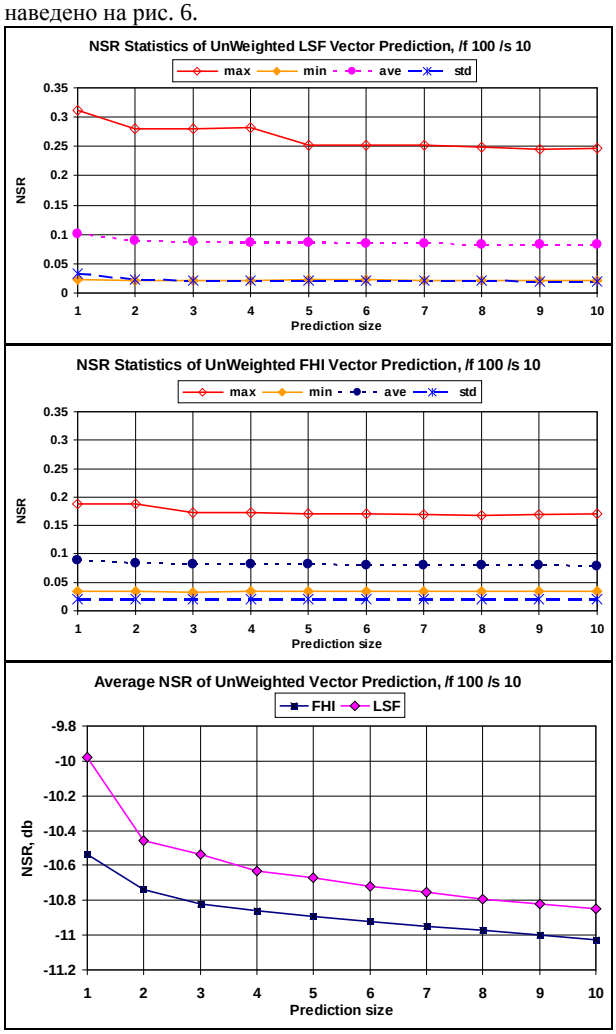

Рисунок 6: Статистики та середні значення відносних похибок прогнозування векторів ЛСЧ і ЛСЧ-НР при різних значеннях порядку прогнозування

Можна бачити, що для класичних ЛСЧ з ростом порядку прогнозування зменшується середнє значення відносної похибки, але цей процес є дуже повільним і не існує чіткого оптимального значення порядку прогнозування. Аналогічна картина процесу спостерігається і для ЛСЧ-НР, але при менших середніх значеннях відносної похибки прогнозування і меншим розкидом даних.

На рис. 6 видно, що *прогнозування векторів ЛСЧ-НР забезпечує меншу похибку прогнозування для будь-якого порядку прогнозування*.

Такі самі дослідження були виконані для класичних ЛСПр та ЛСПр-НР. Відповідні результати подано на рис. 7.

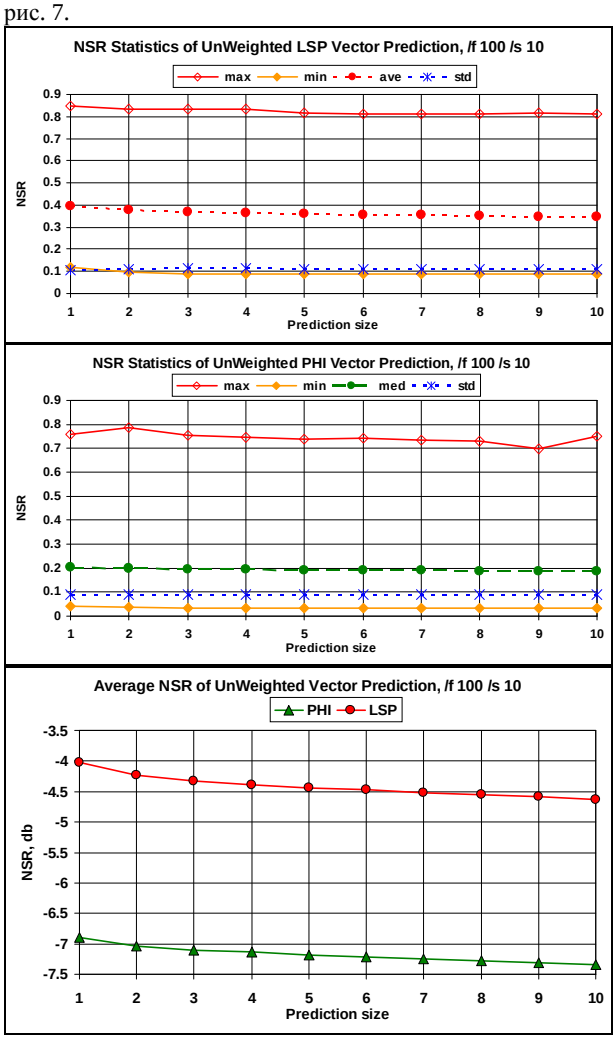

Рисунок 7: Статистики та середні значення відносних похибок прогнозування векторів ЛСПр і ЛСПр-НР при різних значеннях порядку прогнозування

Можна бачити, що для лінійних спектральних проекцій, рис. 7, характер процесу такий самий, як і для лінійних спектральних частот, рис. 6. При збільшенні порядку прогнозування дуже повільно зменшується середнє значення відносної похибки прогнозування і не існує чіткого оптимального значення порядку прогнозування. Порівняння характеристик процесів

прогнозування векторів ЛСПр і ЛСПр-НР приводить до висновку, що прогнозування ЛСПр-НР є суттєво кращим.

Результати досліджень для 4 альтернативних просторів СП зведені до табл. 1.

Таблиця 1: Вплив порядку прогнозування

Розмір порядку прогнозу-вання	Нормовані спотворення, $10\lg\alpha_{i,ave}$, дБ,			
	ЛСЧ-НР	ЛСЧ	ЛСП-НР	ЛСП
1	-10,5376	-9,97933	-6,8947	-4,0285
2	-10,7386	-10,4558	-7,0353	-4,2329
3	-10,8221	-10,5383	-7,1016	-4,3336
4	-10,8582	-10,6284	-7,1409	-4,3865
5	-10,8922	-10,6674	-7,1770	-4,4336
6	-10,9204	-10,7191	-7,2109	-4,4770
7	-10,9498	-10,7528	-7,2444	-4,5152
8	-10,9740	-10,7918	-7,2721	-4,5496
9	-11,0017	-10,8199	-7,3046	-4,5893
10	-11,0274	-10,8504	-7,3360	-4,6255

При порівняння графіків на рис. 6 і рис. 7, а також даних табл. 1 видно, що з 4 альтернативних СП найбільшу перевагу мають вектори ЛСЧ-НР.

За отриманими результатами можна зробити висновок, що в подальшому, для проведення інших досліджень, можна використовувати порядок прогнозування від 2 до 10, що буде відповідати 20—100 мс тривалості попереднього мовленнєвого сигналу. Також можна схвалити компромісний вибір 4 порядку прогнозування, як це зроблено в рекомендації *G*.729.

4.2. Вплив розміру кадру аналізу

За результатами попередніх досліджень були обрані умови проведення поточних експериментів, а саме: 4-й порядок прогнозування і відсутність операції зважування кадрів аналізу з векторами СП вікном Хеммінга, зсув кожного наступного кадру аналізу відносно попереднього на 10 векторів СП.

Досліджувався вплив розміру кадру аналізу, який змінювався від 10 векторів СП (100 мс мовленнєвого сигналу) до 2000 і більше. Результати для векторів *FHI* подано на рис. 8.

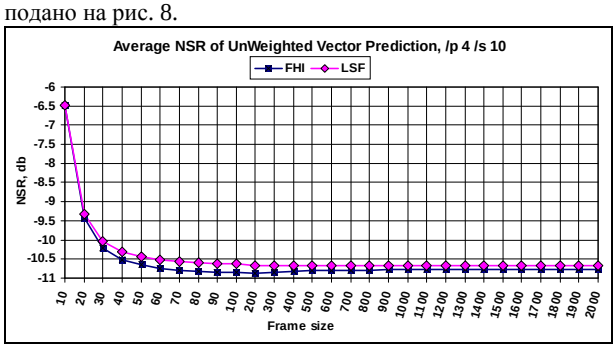

Рисунок 8: Середні значення відносних похибок прогнозування векторів ЛСЧ і ЛСЧ-НР при різних розмірах кадрів аналізу

Для векторів ЛСЧ-НР, рис. 8, існує пологий оптимум щодо вибору розміру кадру аналізу. Мінімум середнього значення відносної похибки прогнозування при цьому припадає на розміри кадру у 200—300 векторів СП (2000—3000 мс).

Те саме можна сказати і відносно класичних векторів ЛСЧ, рис. 8. Мінімальне середнє значення відносної похибки прогнозування спостерігається для АР-коефіцієнтів, які обчислюються за кадрами розміру 200—300 векторів СП.

Порівняння середніх значень відносних похибок прогнозування, рис. 8, приводить до висновку, що вектори ЛСЧ-НР і тут мають перевагу.

Аналогічні дослідження були виконані для векторів класичних ЛСПр і ЛСПр-НР, рис. 9.

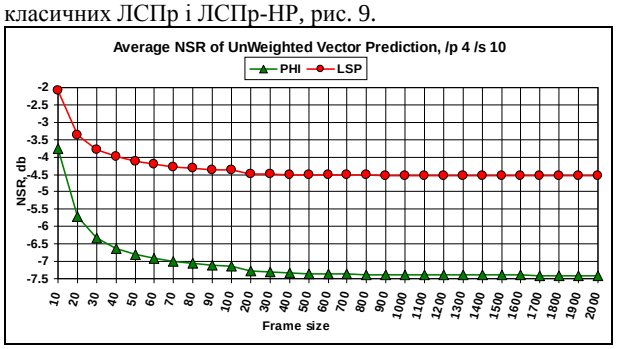

Рисунок 9: Середні значення відносних похибок прогнозування векторів ЛСПр і ЛСПр-НР при різних розмірах кадрів аналізу

На рис. 9 видно, що для лінійних спектральних проекцій характер процесу залишається таким самим, але оптимум (мінімум середнього значення відносної похибки прогнозування) суттєво зсувається в сторону збільшення розміру кадру аналізу. Для векторів класичних ЛСП оптимум спостерігається при розмірі кадру в 4000 векторів, а для векторів ЛСПр-НР — ще більше. Але зміни середнього значення відносної похибки в інтервалі розмірів кадру від 200 до 4000 настільки незначні, що можна обирати будь-який розмір кадру аналізу зі вказаного інтервалу, тим більше, що отримані оптимальні матриці АР- коефіцієнтів все одно будуть усереднюватися в процесі подальшого кластерного аналізу і синтезу матричної кодової книги квантованих АР- коефіцієнтів.

Порівняння процесів для векторів класичних ЛСПр і ЛСПр-НР, рис. 9, знов таки свідчить про перевагу ЛСПр-НР перед класичними ЛСПр.

4.3. Вплив квантування АР- коефіцієнтів

Результати попередніх досліджень було покладено в основу в поточного експерименту, а саме: 4-й порядок прогнозування для всіх типів векторів СП, розміри кадрів аналізу — 200 векторів СП для класичних ЛСЧ і ЛСЧ-НР (2000 мс мовленнєвого сигналу), 4000 векторів СП для класичних ЛСПр і ЛСПр-НР (40 с мовленнєвого сигналу), відсутність операції зважування кадрів аналізу з векторами СП вікном Хеммінга, зсув кожного наступного кадру аналізу відносно попереднього на 10 векторів СП.

Отримані при таких фіксованих параметрах аналізу значення АР- коефіцієнтів для кожного кадру векторів СП були об’єднані в відповідні навчальні вибірки. Після цього для кожної навчальної вибірки був виконаний

кластерний аналіз (матричне квантування) з метою створення матричних кодових книг різного розміру і подальшого використання їх кодових матриць (центроїдів відповідних кластерів) в якості обмеженого набору можливих квантованих АР- коефіцієнтів при повторному прогнозуванні відповідних досліджуваних векторів СП.

Навчальна вибірка неквантованих матриць АР-коефіцієнтів розглядалася як така, яка забезпечує потенційне (мінімальне можливе) середнє значення відносної похибки прогнозування, а кодові матриці різних за розмірами кодових книг — як такі, які будуть погіршувати потенційну похибку.

Результати досліджень впливу процесу квантування АР- коефіцієнтів на прогнозування векторів класичних ЛСЧ і ЛСЧ-НР подано на рис. 10, з якого видно, що процес квантування впливає на похибку прогнозування очікуваним чином: для кодової книги з 8848 кодових матриць (що дорівнює розміру самої навчальної вибірки, тобто нічого не квантується, а використовуються обчислені оптимальні неквантовані АР- коефіцієнти) забезпечує потенційну похибку, яка є найменшою. Далі, при використанні кодової книги з одним кодовим вектором (будь яке значення оптимальних неквантованих АР- коефіцієнтів замінюється на це єдине значення кодового вектору) дає найбільше середнє значення відносної похибки прогнозування. Збільшення розміру кодової книги вдвічі, тобто за схемою 1—2—4—8—16, приводить до зменшення середнього значення відносної похибки прогнозування, яка прагне до потенційного значення на нескінченості.

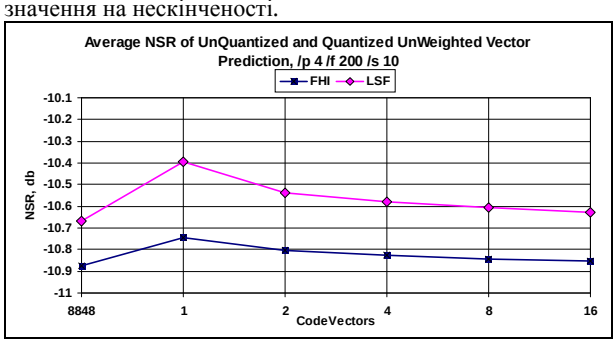

Рисунок 10: Середні значення відносних похибок прогнозування векторів ЛСЧ і ЛСЧ-НР при різних розмірах квантованої кодової книги

Аналогічні дослідження були проведені і для лінійних спектральних проекцій, результати яких наведені на рис. 11. Можна бачити, що вплив процесу квантування АР- коефіцієнтів залишається таким самим і для цих векторів. Порівняння характеристик на рис. 11 свідчить, що проекції вищого ступеня розщеплення так саме мають перевагу перед класичними проекціями і в цьому випадку.

Результати досліджень впливу квантування АР-коефіцієнтів на процес прогнозування СП для 4 альтернативних просторів СП зведені до табл. 2.

Порівнюючи середні значення відносної похибки для досліджуваних альтернативних СП 4 типів, можна зробити висновок, що вектори ЛСЧ-НР є найбільш переважними для використання їх в якості СП, які кодують СО мовленнєвого сигналу і застосовуються для її прогнозування.

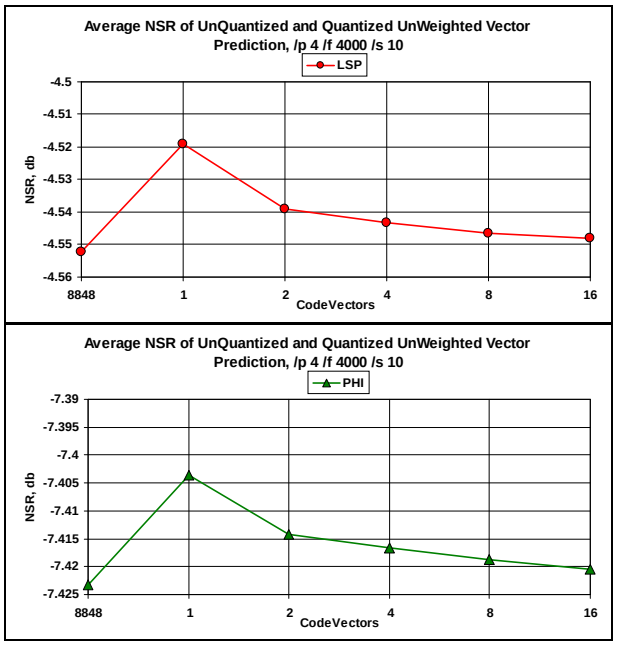

Рисунок 11: Середні значення відносних похибок прогнозування векторів ЛСПр і ЛСПр-НР при різних розмірах квантованої кодової книги

Таблиця 2: Результати прогнозування за квантованими АР- коефіцієнтами

Розмір квантованої кодової книги	Нормовані спотворення, $10\lg\alpha_{i,ave}$, дБ,			
	ЛСЧ-НР	ЛСЧ	ЛСП-НР	ЛСП
8848	-10,8730	-10,6705	-7,4232	-4,5522
1	-10,7436	-10,3970	-7,4036	-4,5192
2	-10,8014	-10,5372	-7,4143	-4,5390
4	-10,8269	-10,5802	-7,4167	-4,5434
8	-10,8412	-10,6076	-7,4187	-4,5466
16	-10,8527	-10,6268	-7,4205	-4,5480

5. Висновки

Проведено дослідження властивостей 4-х типів векторів кодування СО мовленнєвого сигналу з метою з'ясування впливу на похибку їх ЛП таких параметрів, як:

- розміру кадру аналізу,
- порядку прогнозування,
- розміру кодової книги при векторному квантуванні навчальної вибірки АР- коефіцієнтів.

1. Збільшення порядку ЛП зменшує середнє значення відносної похибки прогнозування. Явного оптимуму не існує. Цілком влучним можна вважати прогнозування 4-го порядку, як це зроблено в рекомендації G.729.
2. Накладання вагової функції на кадр векторів СП для розрахунку АР-коефіцієнтів з метою отримання навчальної вибірки для подальшого кластерного аналізу приводить до погіршення потенційних (найменш можливих) середніх значень відносної похибки прогнозування і є небажаним.

3. Розмір кадру аналізу впливає на відносну похибку ЛП так, що існує дуже пологий оптимум, який для векторів *LSF* і *FHI* становить близько 200-300 векторів у кадрі аналізу, а для векторів *LSP* і *PHI* — близько 4000.
4. Кластеризація навчальної вибірки АР-коефіцієнтів приводить до погіршення оцінок відносної похибки в порівнянні з потенційною оцінкою.
5. Збільшення розміру квантованої кодової книги (збільшення кількості кодових матриць, які можуть бути застосовані при прогнозуванні) приводить до того, що відносна похибка починає зменшуватися і наближатися до потенційної.
6. Дослідження показали безперечну перевагу ЛСПр-НР і ЛСЧ-НР (*PHI* і *FHI*) в порівнянні з класичними ЛСПр і ЛСЧ (*LSP* і *LSF*).
7. Найкращі результати отримано для векторів *FHI* — лінійних спектральних частот найвищого ступеня розщеплення.
8. Потенційний виграш, що очікується при прогнозуванні СО зі застосуванням СП найвищого ступеня розщеплення і кодової книги розміру 2 становить: в просторі ЛСЧ-НР по відношенню до класичного простору ЛСЧ — 0,264 дБ, в просторі ЛСПр-НР по відношенню до класичного простору ЛСПр — 2,875 дБ, в просторі ЛСЧ-НР по відношенню до простору ЛСП-НР — 3,387 дБ.

6. Посилання

[1] Павлов О.И. Прямое П-преобразование в линейном предсказании речи // Радиоэлектроника. — 2000. — N12. — С. 53 — 66. (Изв. высш. учеб. заведений).

[2] Павлов О.И. Упрощение реализации метода линейных спектральных пар (частот) в линейном предсказании речи // Труды 3-й Международной конференции “Цифровая обработка сигналов и ее применение”, Т.3, стр. 128 — 132, Москва, 2000.

[3] Павлов О.И. Быстрый алгоритм и графическое представление прямого преобразования в методе линейных спектральных частот высшего порядка // Труды 3-й Международной конференции “Цифровая обработка сигналов и ее применение”, Т. 3, стр. 132 — 136, Москва, 2000.

[4] Павлов О.И. Алгоритм быстрого прямого П-преобразования и особенности его математического аппарата // Радиоэлектроника. — 2001. — N2.— С. 61 — 73. (Изв.высш.учеб.заведений).

[5] Павлов О.И. Обратное П-преобразование в линейном предсказании речи // Радиоэлектроника. — 2001. — N1. — С. 61 — 73. (Изв. высш. учеб. заведений).

[6] Павлов О.И. Алгоритм быстрого обратного П-преобразования // Радиоэлектроника. — 2001. — N8.— С. 67 — 77. (Изв.высш.учеб.заведений).

[7] Павлов О.И. Свойства линейных спектральных частот высших порядков // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2001. Вып. 117. С. 62 — 64

[8] Павлов О.И. Межкадровая интерполяция спектральной огибающей речевого сигнала в пространстве линейных спектральных частот наивысшей регрессии // Радиоэлектроника. — 2008. — N4.— С. 56 — 69. (Изв.высш.учеб.заведений).

[9] Павлов О.І., Стасевич П.А., Тертичний Г.М. Оцінка ефективності кодування спектральної обвідної мовленнєвих сигналів в просторах лінійних спектральних параметрів найвищої регресії методом кластерного аналізу // Праці 9-ої Всеукраїнської міжнародної конференції “Оброблення сигналів і зображень та розпізнавання образів”, стор. 189—192 , Київ, 2008.

[10] Павлов О.І. Розділ 6. Кодування спектральної обвідної мовленнєвих сигналів методом ЛСП-НР // Розробити апаратні засоби, системне та проблемно-орієнтоване прикладне програмне забезпечення високотехнологічних електронних виробів мовленнєвої інформатики, здійснити їх впровадження в навчальний процес, виробництво та в базову комп’ютерно-телекомунікаційну платформу «Образний комп’ютер» (ОК_2008_2_ЦЕНТР_КПІ) : звіт про НДР № 205 (заключний) / РТФ НТУУ «КПІ» МОН України; кер. О.І.Рибін; викон.: О.І.Павлов [та ін.]. — К., 2008. — 139 с. — № ДР 0208U010175.

[11] Павлов О.І. Розділ 5. Інтеграція базового модуля усномовних інформаційних технологій у сучасні комп’ютерні системи // Розроблення програмно-апаратних засобів базового модуля усномовної комп’ютерної технології, що вбудовується в сучасні комп’ютерні системи, створення на їх основі високотехнологічних електронних виробів широкого застосування та здійснення заходів для їх впровадження у виробництво (ОК_2009_2) : звіт про НДР (заключний) / МННЦ ІТіС НАН та МОН України; кер. Т.К.Вінцюк; викон.: О.І.Павлов [та ін.]. — К., 2010. — 149 с. — № ДР 0109U004244.

[12] ITU-T Recommendation G.729. General Aspects of Digital Transmission Systems. Coding of Speech at 8 Kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear Prediction (CS-ACELP). // Telecommunication Standardization Sector of the International Telecommunication Union, Geneva, 19th of March 1996.

[13] ITU-T Recommendation G.729 – Annex A: Reduced Complexity 8 Kbit/s CS-ACELP Speech Codec / Coding of Speech at 8 Kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear Prediction (CS-ACELP) / Series G: Transmission Systems and Media. Digital transmission systems – Terminal equipments – Coding of analogue signals by methods other than PCM // Telecommunication Standardization Sector of the International Telecommunication Union, Geneva, 8th of November 1996.

[14] Марпл.-мл. С.Л. Цифровой спектральный анализ и его приложения: Пер. с англ. — М.: Мир, 1990. — 584 с., ил.

[15] Дж. Макхоул, С. Рукос, Г. Гиш. Векторное квантование при кодировании речи // ТИИЭР, 1985 , т.73, №11, с. 19 — 61.