

Міністерство освіти і науки України
Вінницький національний технічний університет

О. М. Ткаченко, С. В. Хрущак

**УЩІЛЬНЕННЯ ПАРАМЕТРІВ
МОВЛЕННЄВОГО СИГНАЛУ
НА ОСНОВІ ВЕКТОРНОГО КВАНТУВАННЯ**

Монографія

Вінниця
ВНТУ
2014

УДК 004.934 + 004.627

ББК 32.973.202

T48

Рекомендовано до друку Вченою радою Вінницького національного технічного університету Міністерства освіти і науки України (протокол № 11 від 26.06.2013 р.)

Рецензенти:

В. М. Лисогор, доктор технічних наук, професор

С. І. Перевозніков, доктор технічних наук, професор

Ткаченко, О. М.

T48 Ущільнення параметрів мовленнєвого сигналу на основі векторного квантування : монографія / О. М. Ткаченко, С. В. Хрущак. – Вінниця : ВНТУ, 2014. – 116 с.

ISBN 978-966-641-566-3

У монографії розглянуто питання ущільнення мовленнєвих сигналів на основі векторного квантування параметрів сигналу. Вдосконалено методи ущільнення параметрів сигналу за рахунок структуризації кодових книг, дельта-ущільнення. Розроблено методику та програмні засоби для дослідження запропонованих методів ущільнення. Книга розрахована на науковців, аспірантів та інженерів, які займаються розробкою комп'ютерних систем ущільнення, передавання та зберігання мовленнєвих сигналів.

УДК 004.934 + 004.627

ББК 32.973.202

ISBN 978-966-641-566-3

© О. Ткаченко, С. Хрущак, 2014

ЗМІСТ

ВСТУП	5
РОЗДІЛ 1 АНАЛІЗ ІСНУЮЧИХ СИСТЕМ УЩІЛЬНЕННЯ МОВЛЕННЄВИХ СИГНАЛІВ	7
1.1 Аналіз комп'ютерних систем ущільнення мовленнєвих сигналів	7
1.2 Ущільнення із кодуванням форми сигналу	9
1.3 Параметричні методи ущільнення мовленнєвих сигналів	13
1.3.1 Аналіз методу лінійного прогнозування	13
1.3.2 Інші форми подання коефіцієнтів лінійного прогнозування	16
1.3.3 Застосування кодових книг для підвищення ступеня ущільнення	18
1.3.4 Векторне квантування параметрів мовленнєвого сигналу ...	20
1.3.5 Багатоетапне векторне квантування	22
1.4 Базова структура комп'ютерної системи передавання мовленнєвих сигналів	24
РОЗДІЛ 2 ТЕОРЕТИЧНІ ЗАСАДИ ДЛЯ ПОБУДОВИ ЗАСОБІВ ВЕКТОРНОГО КВАНТУВАННЯ ПАРАМЕТРІВ МОВЛЕННЄВИХ СИГНАЛІВ НА ОСНОВІ МАЖОРУВАННЯ ВЕКТОРІВ	27
2.1 Оцінювання відстаней між векторами	27
2.2 Структуризація векторних кодових книг	31
2.2.1 Структуризація векторних кодових книг на основі відношення мажорювання	31
2.2.2 Мажорювання векторів відстаней	34
2.3 Прискорений пошук у структурованій кодовій книзі	36
2.4 Метод дельта-ущільнення параметрів мовленнєвого сигналу ..	39
2.4.1 Дельта-ущільнення з використанням алгоритму Вітербі	39
2.4.2 Метод дельта-ущільнення з прогнозуванням	47
РОЗДІЛ 3 СТРУКТУРНО-ФУНКЦІОНАЛЬНА ОРГАНІЗАЦІЯ КОМП'ЮТЕРНОЇ СИСТЕМИ ОБРОБЛЕННЯ МОВЛЕННЄВИХ СИГНАЛІВ НА БАЗІ ВПОРЯДКОВАНИХ КОДОВИХ КНИГ	51
3.1 Загальні вимоги до комп'ютерної системи передавання та зберігання мовленнєвих сигналів	51

3.2 Архітектура комп'ютерної системи передавання мовленнєвих сигналів на базі кодових книг	53
3.3 Особливості побудови підсистеми визначення голосової активності	61
3.4 Методика створення мовленнєвого пакета	66
3.5 Особливості побудови підсистеми зберігання мовленнєвих даних	69
РОЗДІЛ 4 ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ МЕТОДІВ ТА РЕАЛІЗАЦІЯ КОМП'ЮТЕРНОЇ СИСТЕМИ УЩІЛЬНЕННЯ МОВЛЕННЄВИХ СИГНАЛІВ	73
4.1 Програмні засоби для дослідження методів прискореного пошуку векторів у кодовій книзі	73
4.2 Умови тестування розроблених методів	76
4.3 Експериментальні дослідження запропонованих методів ущільнення мовленнєвих сигналів у комп'ютерних системах	77
4.3.1 Експериментальні дослідження запропонованої міри та різних видів розбиття кодової книги	77
4.3.2 Експериментальні дослідження прискореного пошуку векторів в кодовій книзі	78
4.3.3 Експериментальні дослідження методу дельта-ущільнення мовленнєвих сигналів	80
4.3.4 Дослідження методів побудови екстраполюючої функції	83
4.3.5 Дослідження методу визначення голосової активності	86
4.4 Реалізації вокодера з використанням дельта-ущільнення для передачі та зберігання мовленнєвих сигналів у комп'ютерних системах	87
4.5 Оцінювання розбірливості синтезованого мовленнєвого сигналу	92
ВИСНОВКИ	95
ЛІТЕРАТУРА	96
Додаток А Лістинг програми для дослідження методів ущільнення мовленнєвих сигналів на основі векторного квантування	104

ВСТУП

Актуальним напрямком розвитку сучасних комп'ютерних систем є оброблення, зберігання та передавання мультимедійної інформації, зокрема мовленнєвих сигналів. При цьому одним з найбільш важливих питань є обсяг даних, необхідних для забезпечення якісного відтворення мовленнєвого сигналу. Ущільнення мовленнєвого сигналу зменшує обсяг даних, необхідних для його опису та відповідно апаратні витрати на його зберігання та передавання. Це дозволяє збільшувати кількість користувачів мовленнєвих сервісів мережі без збільшення апаратних витрат.

Створення кодерів з високим ступенем ущільнення дозволить використовувати їх у побудові систем голосових повідомлень, автоматичних мовленнєвих довідкових служб, систем передавання мовленнєвих сигналів по низькошвидкісних комп'ютерних мережах, недорогих систем безпроводного зв'язку. Тому світовий ринок мультимедійних систем розвивається під безпосереднім технологічним впливом вчених та розробників мовленнєвих кодеків.

Слід відзначити, що основні зусилля в розробці методів і систем оброблення, передавання та зберігання цифрових мовленнєвих сигналів належать високорозвиненим країнам – США, Японії, Франції та Великобританії. Кінцеві результати досліджень формуються у вигляді національних стандартів і рекомендацій міжнародних організацій, таких як, ITU-T і групи MPEG.

Значний внесок у розвиток систем ущільнення мовлення внесли Л. Р. Рабінер, Б. С. Атал, К. К. Палівал, Ф. Ітакура, Ж. Джін, Дж. Макхоул, Т. К. Вінцюк.

Проте дослідження у цьому напрямку не втрачають своєї актуальності, оскільки залишилось чимало нерозв'язаних задач. Відсутні методи ущільнення мовленнєвих сигналів, які враховували б кореляцію між кадрами мовленнєвого сигналу, а також методи, які б використовували впорядковані кодові книги в процесі векторно-го квантування параметрів мовленнєвого сигналу. Таким чином

актуальною є науково-технічна задача розробки програмно-апаратних засобів ущільнення даних, необхідних для подальшого передавання або відтворення мовленнєвого сигналу без втрати прийнятної якості декодованого звуку.

Метою монографії є підвищення продуктивності оброблення мовленнєвих сигналів в комп'ютерній системі за рахунок розробки програмно-апаратних засобів ущільнення даних, необхідних для подальшого передавання, зберігання та відтворення мовленнєвого сигналу без втрати прийнятної якості декодованого мовлення.

Застосування розроблених програмних засобів дозволило зменшити кількість мовленнєвої інформації, необхідної для подальшої передачі або відтворення у комп'ютерній системі при збереженні прийнятної якості вихідного мовленнєвого сигналу.

Автори будуть вдячні за відгуки на монографію та пропозиції щодо розвитку подальших досліджень.

РОЗДІЛ 1

АНАЛІЗ ІСНУЮЧИХ СИСТЕМ УЩІЛЬНЕННЯ МОВЛЕННЄВИХ СИГНАЛІВ

1.1 Аналіз комп'ютерних систем ущільнення мовленнєвих сигналів

Відомо, що для мовленнєвого сигналу характерна висока ступінь надлишковості, що забезпечує його високий ступінь завадостійкості. Цифрова обробка мовленнєвих сигналів дозволяє значно зменшити цю надлишковість і скоротити обсяг даних, необхідних для опису сигналу, що, в свою чергу, дозволяє зменшити витрати на канали зв'язку [1].

У загальному випадку ущільнення (англ. encoding) мовленнєвих сигналів – це процедура перетворення оцифрованих мовленнєвих даних в представлення за допомогою якомога меншої кількості біт, при збереженні якості сигналу, прийнятної для конкретних застосувань цього методу ущільнення [1, 2]. Структура комп'ютерної системи ущільнення мовленнєвих сигналів може бути представлена у вигляді [3], показаному на рис. 1.1.

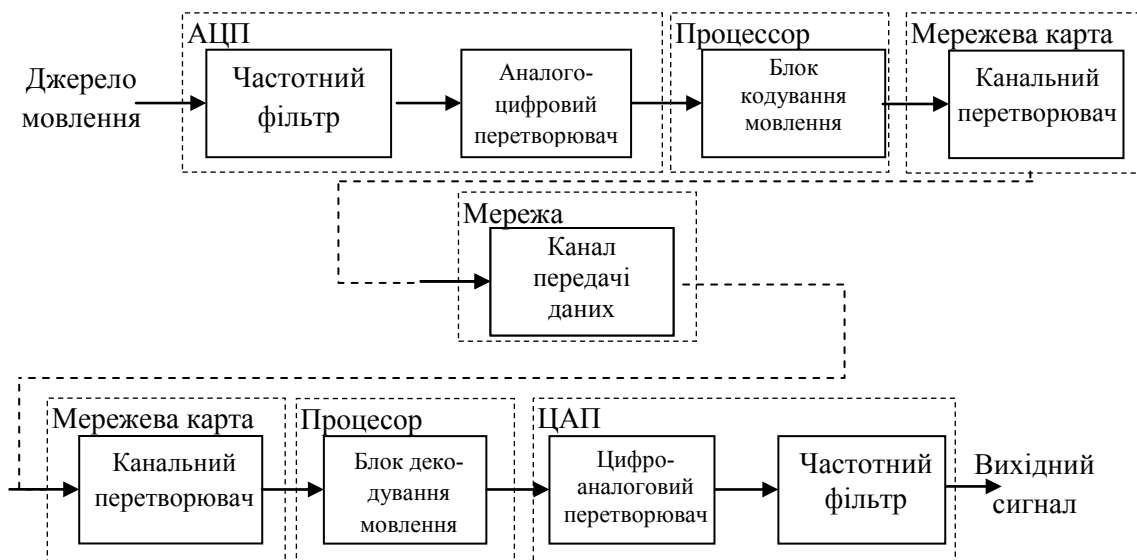


Рисунок 1.1 – Загальна структура комп'ютерної системи передавання, оброблення та зберігання мовленнєвого сигналу

Мовленнєвий сигнал спочатку проходить фільтрацію та оцифровується за допомогою аналого-цифрового перетворювача. В більшості випадків, використовується частота дискретизації 8 КГц та 16 біт для опису одного відліку сигналу. В результаті отримуємо швидкість вхідного потоку даних $V_r = 128$ Кбіт/с. Отриманий оцифрований сигнал надходить на вхід блока кодування сигналу, результатом роботи якого є кодований мовленнєвий сигнал. Бітова швидкість цього сигналу в більшості випадків значно менша за бітову швидкість вхідного потоку [3, 4]. Далі, в загальному випадку, кодований мовленнєвий сигнал обробляється за допомогою каналного перетворювача, який підвищує завадостійкість сигналу під час передачі бітового потоку по каналу зв'язку [5, 6].

Далі перетворення сигналу відбуваються у зворотному порядку – каналний декодер обробляє дані, що надходять з каналу зв'язку, щоб відновити кодований сигнал, який далі передається на декодер мовленнєвого сигналу. Декодер генерує вихідний мовленнєвий сигнал з бітовою швидкістю оригінального сигналу. Отриманий цифровий сигнал перетворюється за допомогою цифроаналогового перетворювача в аналоговий сигнал [7].

Основна мета розробки систем ущільнення мовленнєвих сигналів – це або отримати якомога вищу якість сигналу (якомога менше спотворення) A при заданій бітовій швидкості V , або пониження бітової швидкості сигналу V при заданій якості вихідного сигналу. В обох варіантах вводяться додаткові обмеження, найбільш важливими з яких є: величина затримки сигналу τ , обчислювальна складність C , обсяги пам'яті M_c , стійкість до збоїв в каналах зв'язку S [8, 9]:

- | | | | |
|----|----------------------|----|----------------------|
| 1) | $V \rightarrow \min$ | 2) | $A \rightarrow \max$ |
| | $A \geq A_0$ | | $V = \text{const}$ |
| | $\tau \leq \tau_0$ | | $\tau \leq \tau_0$ |
| | $C \leq C_0$ | | $C \leq C_0$ |
| | $M_c \leq M_0$ | | $M_c \leq M_0$ |
| | $S \leq S_0$ | | $S \leq S_0$ |

Також можна виділити характеристики, які є бажаними для усіх систем ущільнення мовленнєвих сигналів або вокодерів [10–12]:

1. Низька бітова швидкість, що необхідна для опису сигналу. Чим менша швидкість необхідна для передачі мовленнєвого сигналу, тим

ефективніше буде працювати система і тим меншою буде її вартість. Зменшення обсягу даних, необхідних для опису сигналу, призводить до погіршення якості сигналу, що накладає деякі обмеження на систему. Тому на практиці вибирають таку швидкість та якість, які будуть прийнятними в конкретних умовах [7, 8].

2. Висока якість [12] відтвореного сигналу. Існує багато способів оцінювання якості декодованого сигналу – натуральність, розбірливість, оцінювання за шкалою MOS (середня експертна оцінка), впізнаваність диктора, передача емоційного стану диктора та інші [8, 12].

3. Надійність для різних дикторів та мов. Методи та моделі, які застосовуються у вокодері, повинні однаково добре працювати для різних дикторів – чоловіків, жінок та дітей та для різних мов [11].

4. Стійкість до каналних перешкод. Ця властивість є критичною для систем цифрового зв'язку, де каналні завади здійснюють негативний вплив на кінцеву якість сигналу [7].

5. Невелика затримка кодування. При кодуванні-декодуванні сигналу в канал зв'язку неминуче вноситься затримка. Велике значення затримки створює значні проблеми при спілкуванні у реальному часі, тому для систем цифрового зв'язку вважається прийнятною затримка до 200 мс [8].

6. Ущільнення немовленнєвих сигналів. У системах загального користування таких, як, наприклад, VoIP, важливою вимогою до вокодера є можливість достовірно передавати інші звуки, окрім мовлення, наприклад, музику [11].

7. Низька обчислювальна складність та обсяг пам'яті. Для того, щоб система ущільнення мовленнєвих сигналів була практично цінною, вартість її реалізації повинна бути низькою, це стосується вимог до пам'яті та обчислювальних витрат, необхідних для ущільнення мовлення [7, 10].

1.2 Ущільнення із кодуванням форми сигналу

Методи кодування форми сигналу мають за мету скопіювати форму вихідного сигналу. Найбільш простим серед таких методів є рівномірне квантування – метод, в якому оцифровані відліки сигналу зберігаються без перетворень [7]. При оцифруванні сигналу рівні та інтервали квантування вибираються рівномірно:

$$x_i - x_{i-1} = \Delta, \quad (1.1)$$

$$\hat{x}_i - \hat{x}_{i-1} = \Delta, \quad (1.2)$$

де x та \hat{x} – відповідно звичайні та квантовані значення сигналу; Δ – крок квантування [10].

Зазвичай використовується двійкове подання сигналу зі знаком, коли крайній лівий розряд вважається знаковим, тоді рівні квантування пов'язані з кодовими словами співвідношенням [7]

$$\hat{x}(n) = (\Delta / 2) \cdot \text{sign}(c(n)) + \Delta c(n), \quad (1.3)$$

де $\text{sign}(c(n)) = 1$, якщо перший розряд $c(n)=0$, та $\text{sign}(c(n)) = -1$, якщо знаковий розряд $c(n) = 1$.

Вираз для співвідношення сигнал/шум при B -розрядному квантуванні сигналу з максимальною амплітудою X_{\max} буде мати вигляд [11]

$$SNR = \frac{\sigma_x^2}{\sigma_e^2} = \frac{\sigma_x^2}{\Delta^2 / 12} = \frac{3 \cdot 2^{2B} \sigma_x^2}{X_{\max}^2}, \quad (1.4)$$

де σ_x – дисперсія вхідного сигналу; σ_e – дисперсія помилки квантування.

З цього виразу випливає, що у випадку, коли дисперсія вхідного сигналу складає лиш половину від тієї, на яку розрахований квантизатор, то це призведе до погіршення сигналу на 6 дБ. У той же час відомо, що дисперсія на невокалізованих сегментах мовленнєвого сигналу може бути на 20–30 дБ меншою, ніж дисперсія на вокалізованих, тобто короткочасне відношення сигнал/шум на невокалізованих сегментах може бути значно меншим ніж на вокалізованих. Тому для того, щоб отримати постійну відносну похибку квантування, використовують логарифмічно розподілені рівні квантування [7, 8]. Реалізацією такого підходу є імпульсно-кодова модуляція (Pulse Code Modulation, PCM). При логарифмічному квантуванні декодований сигнал буде мати вигляд

$$\hat{x}(n) = \exp(\log|x(n)| + \varepsilon(n)) \cdot \text{sign}(x(n)) = x(n) \exp(\varepsilon(n)), \quad (1.5)$$

де $\varepsilon(n)$ – похибка квантування відліку сигналу. Якщо $\varepsilon(n)$ є досить малим, то це рівняння можна апроксимувати у вигляді

$$\hat{x}(n) \approx x(n)(1 + \varepsilon(n)) = x(n) + \varepsilon(n) \cdot x(n) = x(n) + f(n), \quad (1.6)$$

де $f(n) = \varepsilon(n) x(n)$. Таким чином, оскільки $\varepsilon(n)$ та $x(n)$ є некорельованими, співвідношення сигнал/шум матиме вигляд [7]

$$SNR = \frac{1}{\sigma_e^2}. \quad (1.7)$$

Тобто співвідношення сигнал/шум не залежить від потужності сигналу, а залежить тільки від кроку квантування. Використовувати таке квантування на практиці неможливо, оскільки динамічний діапазон є нескінченним, і, таким чином, необхідна нескінченна кількість рівнів квантування. Тому використовуються дві модифікації РСМ з близькими до логарифмічних рівнями квантування – μ -law РСМ (Північна Америка та Японія) та a -law РСМ (Європа, та більшість інших країн світу), в яких телефонна розмова попередньо квантована з розрядністю 12 або 13 бітів, кодується з частотою 8 КГц й ущільнюється до 8 бітів на кожен відлік. Таким чином, для передачі мовленнєвого сигналу тільки в один бік необхідна швидкість 64 Кбіт/с з початкових 96 Кбіт/с [9, 13].

У методі ущільнення, що заснований на μ -законі (μ -law), функція перетворення відліків $y(n)$ буде мати вигляд

$$y(n) = F(x(n)) = X_{\max} \frac{\log(1 + \mu \frac{x(n)}{X_{\max}})}{\log(1 + \mu)} \text{sign}(x(n)). \quad (1.8)$$

Для $\mu = 0$ рівняння (1.8) перетворюється в рівність $y(n) = x(n)$, однак для більших значень μ і $|x(n)|$ залежність перетворюється в логарифмічну. Співвідношення сигнал/шум при цьому матиме вигляд [11]

$$SNR = 6B + 4,77 - 20 \log_{10}(\ln(1 + \mu)) - 10 \log_{10}(1 + (\frac{X_{\max}}{\mu \sigma_x})^2 + \sqrt{2}(\frac{X_{\max}}{\mu \sigma_x})). \quad (1.9)$$

Таким чином, квантування за μ -законом дозволяє отримати постійне співвідношення сигнал/шум у широкому діапазоні дисперсій вхідного сигналу.

Для A -закону функція перетворення матиме вигляд:

$$y(n) = \begin{cases} X_{\max} \frac{A \frac{x(n)}{X_{\max}}}{1 + \log(A)}, & \text{при } 0 \leq \frac{x(n)}{X_{\max}} \leq \frac{1}{A}; \\ X_{\max} \frac{1 + A \frac{x(n)}{X_{\max}}}{1 + \log(A)} \text{sign}(x(n)), & \text{при } \frac{1}{A} \leq \frac{x(n)}{X_{\max}} \leq 1. \end{cases} \quad (1.10)$$

Існує також модифікація РСМ, яка суміщає дискретизацію сигналу з його ущільненням – метод DPCM кодує значення сигналу як різницю між попереднім та поточним значенням. У середньому це дозволяє зменшити кількість бітів, необхідних для передачі одного відліку, приблизно на 25 % [10].

Наступним кроком у зменшенні обсягу даних для опису мовленнєвого сигналу є спроба забезпечити рівномірну відносну точність вимірювання амплітуди сигналу. Для цього динамічний діапазон амплітуди розбивають на вісім логарифмічних піддіапазонів, в кожному з яких значення амплітуди кодують п'ятьма розрядами. Таким чином результати ущільнення підлаштовуються відповідно до енергії сигналу так, щоб слабкі сигнали квантувалися малими ступенями квантування, а сильні сигнали – великими. Завдяки безперервному підлаштуванню кроку квантування до поточної потужності мовленнєвого сигналу, розрядність шкали квантування в ADPCM вдалося понизити до чотирьох біт і одержати ущільнення зі швидкістю передачі 32 Кбіт/с і якістю, близькою до РСМ [11–12].

Перетворення мовленнєвого сигналу методами кодування форми хвилі забезпечує прийнятну для комерційного застосування якість на швидкостях від 24 Кбіт/с і вище. Подальше зменшення інформаційної ємності сигналу з використанням цього підходу вважається неефективним, тому в більшості сучасних систем телефонії (мобільний зв'язок, IP-телефонія) використовуються параметричні вокодери.

1.3 Параметричні методи ущільнення мовленнєвих сигналів

1.3.1 Аналіз методу лінійного прогнозування

Метод лінійного прогнозування використовується як основа в більшості сучасних алгоритмів ущільнення мовленнєвих сигналів. Цей метод є одним з найефективніших при ущільненому представленні мовленнєвого сигналу з метою його низькошвидкісної передачі. Важливість методу визначається високою точністю отримуваних оцінок та відносною простотою обчислень [8, 14].

Основною ідеєю цього методу є те, що відрізок мовленнєвого сигналу можливо з деякою похибкою обрахувати на основі попередніх відрізків. У середині одного відрізка вагові коефіцієнти, що використовуються для отримання лінійної послідовності, обраховуються, виходячи з мінімізації середньоквадратичної похибки. Ці вагові коефіцієнти називаються коефіцієнтами лінійного прогнозування. В основі методу лінійного прогнозування лежить авторегресійна модель, а саме: припущення, що мовленнєвий сигнал може бути змодельовано як авторегресійний сигнал. Лінійне прогнозування являє собою обчислення параметрів авторегресії з початкового відрізка сигналу. Також лінійне прогнозування дає оцінку спектру сигналу, оскільки параметри авторегресійної моделі визначають спектр потужності сигналу. Отже, лінійне прогнозування може розглядатися як задача ідентифікації систем, в якій параметри авторегресійної моделі отримуються з початкового мовленнєвого сигналу (рис. 1.2). Вхідний білий шум $x[n]$ пропускається через синтезатор авторегресійного процесу, в результаті чого утворюється авторегресійний сигнал $s[n]$ з параметрами авторегресії – \hat{a}_i [15–16].

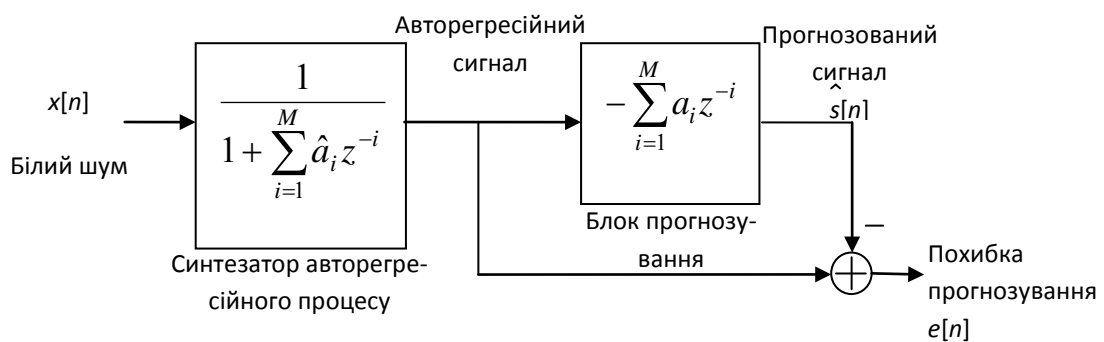


Рисунок 1.2 – Лінійне прогнозування в задачах ідентифікації систем

Блок лінійного прогнозування використовується для прогнозування сигналу $s[n]$ на основі M попередніх відліків згідно з виразом [8]

$$\hat{s}[n] = \sum_{i=1}^M a_i \cdot s[n-i], \quad (1.11)$$

де a_i – коефіцієнти лінійного прогнозування LPC, які являють собою оцінки параметрів авторегресійної функції; M – порядок моделі прогнозування. Таким чином, прогнозування базується на лінійній комбінації M попередніх відліків сигналу. Похибка прогнозування являє собою різницю між вихідним та прогнозованим відрізками сигналу і обчислюється згідно з формулою:

$$e[n] = s[n] - \hat{s}[n]. \quad (1.12)$$

Задача ідентифікації полягає у визначенні коефіцієнтів авторегресійної функції \hat{a}_i з набору вхідних даних $s[n]$. Критерієм при цьому є мінімізація середньоквадратичної похибки прогнозування J , яка обчислюється, виходячи зі співвідношення (1.13),

$$J = E\{e^2[n]\} = E\left\{\left(s[n] - \sum_{i=1}^M a_i s[n-i]\right)^2\right\}, \quad (1.13)$$

де $E\{\}$ – математичне очікування [8].

Функція J є функцією другого порядку від коефіцієнтів LPC. Графічно ця залежність є параболоподібною $M + 1$ вимірною поверхнею з M рівнями свободи. Ця поверхня має єдиний мінімум, отже для того, щоб знайти набір коефіцієнтів LPC, який мінімізує середньоквадратичну похибку, необхідно прирівняти часткову похідну від функції J по a_i до нуля:

$$\frac{\partial J}{\partial a_k} = 2E\left\{\left(s[n] - \sum_{i=1}^M a_i s[n-i]\right)s[n-k]\right\} = 0, \quad (1.14)$$

де $k = 1, 2, \dots, M$.

Якщо рівність (1.14) виконується, тоді знайдений набір коефіцієнтів LPC – a_i є шуканими параметрами авторегресії \hat{a}_i [8].

Необхідним параметром при описі сигналу є коефіцієнт підсилення моделі. Коефіцієнт підсилення моделі (G) можна визначити шляхом узгодження енергії сигналу та лінійно спрогнозованих відліків.

Якщо повні енергії сигналу позначити як $[W(0)]$, то коефіцієнт підсилення можна отримати виходячи з виразу

$$G = R_k(0) - \sum_{i=1}^M a_i W_k(i). \quad (1.15)$$

Також, окрім коефіцієнтів лінійного прогнозування та коефіцієнта підсилення, використовуються такі параметри: класифікатор вокалізованих і невокалізованих звуків та період основного тону для вокалізованих сегментів [8,10]. Мовленнєвий сигнал швидко змінює параметри в часі, тому коефіцієнти повинні обраховуватись на коротких часових сегментах, в більшості випадків вибирають сегменти по 10–30 мс [7, 8, 17].

Мовленнєвий сигнал може бути синтезований за параметрами лінійного прогнозування різними шляхами, найпростішим з яких є використання для синтезу системи, що описується тими ж параметрами, що й застосовувались при аналізі [7]. На рис. 1.3 зображено структурну схему такого синтезатора. На цій схемі генератор, залежно від значення періоду основного тону, працює як джерело імпульсів збудження на вокалізованих сегментах, формуючи імпульси на початку кожного періоду основного тону, або як генератор білого шуму на невокалізованих сегментах сигналу, формуючи некорельований рівномірно розподілений випадковий процес з одиничною дисперсією та нульовим середнім значенням [14].

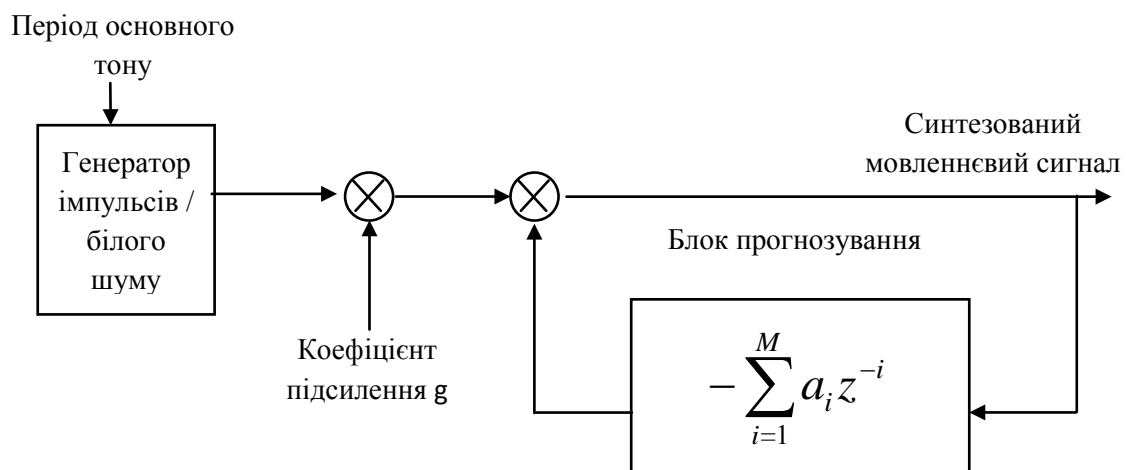


Рисунок 1.3 – Структурна схема синтезатора на основі лінійного прогнозування

Коефіцієнт підсилення визначає повну амплітуду збудження. Відлік синтезованого мовленнєвого сигналу визначається, виходячи зі співвідношення

$$\hat{s}[n] = \sum_{i=1}^M a_i \cdot \hat{s}[n-i] + g \cdot u[n], \quad (1.16)$$

де $u[n]$ – сигнал збудження.

До недоліків вокодера LPC можна віднести те, що коефіцієнти лінійного прогнозування виявляються дуже нестійкими до каналних завад в системах зв'язку. Для параметрів LPC немає методу визначення чи є даний набір коефіцієнтів стійким, тому якщо в процесі квантування чи інтерполяції цих параметрів виникне спотворення певного коефіцієнта, то це призведе до повної втрати фрейму [15]. Також одним з головних недоліків вокодера є наявність призвуків і синтетичності голосу, що спричинено тим, що сигнал збудження формується тільки за допомогою періодичних імпульсів.

1.3.2 Інші форми подання коефіцієнтів лінійного прогнозування

Враховуючи недоліки, які властивості параметрам LPC, у сучасних вокодерах для передачі параметрів голосового тракту використовують похідні від них коефіцієнти. Найбільш поширеними є RC (Reflection coefficients), LAR (Log-Area Ratios) та LSP (Line Spectral Pair) [18].

RC-коефіцієнти (коефіцієнти відображення) отримуються під час обчислення коефіцієнтів LPC за допомогою алгоритму Левінсона–Дарбіна, тому не потребують додаткових обчислень. Дані коефіцієнти дозволяють дуже легко визначати стійкість вихідного LPC фільтра, але для того, щоб вихідний набір параметрів можна було декодувати, необхідно, щоб виконувалась умова $|RC_m| < 1$ для кожного з коефіцієнтів [17, 18].

У [16] показано, що RC-коефіцієнти, значення за модулем яких близьке до одиниці, є дуже чутливими до помилок квантизації порівняно з коефіцієнтами з меншими значеннями. Тому Вішенейтоном і Макхоулом [18] запропоновані похідні від них коефіцієнти LAR, які

мають сталу чутливість до завад. Ці коефіцієнти виражаються співвідношенням

$$LAR_j = \frac{1}{2} \ln \left(\frac{1 + RC_j}{1 - RC_j} \right). \quad (1.17)$$

Ці параметри розширюють діапазон RC-коефіцієнтів до діапазону $(-1; 1)$, так що однакові помилки квантизації спричинять в середньому однакове спотворення сигналу.

В [16] Ітакура запропонував використовувати для передачі параметрів лінійного прогнозування коефіцієнти LSP (Line Spectral Pair), які, як доведено в [19, 20], є найбільш ефективними для подальшої квантизації та передачі та більш стійкими до різного роду перешкод.

Коефіцієнти LSP, або за іншою термінологією – LSF (Line Spectral Frequencies), є коренями поліномів [21–23]:

$$P(z) = P_{p+1}(z) = A_p(z) + z^{-(p+1)} A_p(z^{-1}); \quad (1.18)$$

$$Q(z) = Q_{p+1}(z) = A_p(z) - z^{-(p+1)} A_p(z^{-1}). \quad (1.19)$$

Корені поліномів $P(z)$ та $Q(z)$ x_j та y_j лежать на одиничному колі, тому

$$x_j = \cos \omega_{2j-1} \pm i \sin \omega_{2j-1}, \quad (1.20)$$

$$y_j = \cos \theta_{2j} \pm i \sin \theta_{2j}, \quad (1.21)$$

$$1 \leq j \leq p/2. \quad (1.22)$$

Набори $\{\omega_{2j-1}\}$ та $\{\theta_{2j}\}$ за визначенням є коефіцієнтами LSP. Для цих коефіцієнтів властива перемешованість значень [24]

$$0 < \omega_1 < \theta_2 < \omega_3 < \dots < \omega_{p-1} < \theta_p < \pi. \quad (1.23)$$

Це робить їх менш чутливими до каналних помилок порівняно з LPC, а також значно спрощує перевірку системи на стійкість. Саме ця властивість надає можливість відстежити спотворення коефіцієнтів при передачі і виправити їх за допомогою інтерполяції.

1.3.3 Застосування кодових книг для підвищення ступеня ущільнення

На практиці зручно передавати не самі коефіцієнти, і їх індекси в таблиці значень LSP – кодовій книзі. Саме тому цей підхід застосовується у вокодерах, розроблених починаючи з 1994 року. Кодові книги представляють собою репрезентативний набір значень вхідних коефіцієнтів LSP, які знаходяться на передавальній та приймальній сторонах. На передавальній стороні для коефіцієнтів LSP, що отримані в результаті ущільнення, знаходяться репрезентативні значення з кодової книги, індекси яких передаються на приймальну сторону, де можна з них відновити вихідні коефіцієнти. Такий підхід дозволяє значно зменшити кількість інформації, що передається в канал зв'язку [25–27].

Виділяють два основні види кодових книг: скалярні та векторні. Скалярні кодові книги містять набір квантованих значень для кожного окремого коефіцієнта LSP. Важливою перевагою скалярного квантування є відносно невисока обчислювальна складність, але обсяги інформації, що необхідно передавати для кожного фрейму, залишаються досить великими. Як показали дослідження [26, 27] при використанні скалярних кодових книг, для опису параметрів одного фрейму достатньо 32–40 біт інформації без значної втрати якості.

При векторному квантуванні набір коефіцієнтів LSP розглядається як один вектор, і векторна кодова книга містить квантовані значення всього набору коефіцієнтів. Такий підхід є набагато ефективнішим і дозволяє досягти меншого спотворення сигналу при тій самій бітовій швидкості, порівняно зі скалярним квантуванням. Математично векторне квантування Q масиву M розмірністю N є відображення вектора x у M -розмірному евклідовому просторі R^M у кінцевий набір Y , який містить N M -вимірних точок, які названі кодовими векторами чи кодовими словами [28]:

$$Q: R^M \rightarrow Y, \quad (1.24)$$

де $x = [x_1, x_2, \dots, x_M]^T$ – вектор, що квантується; $y_i = [y_{i1}, y_{i2}, \dots, y_{iM}]^T$; $i = 1, \dots, N$ – набір у який відбувається відображення; Y – кодовий вектор у кодовій книзі квантизатора.

Саме відображення буде мати вигляд

$$Q(x) \rightarrow y_i, i = 1, \dots, N. \quad (1.25)$$

При кодуванні параметрів мовленнєвого сигналу відбувається заміна вектора коефіцієнтів на індекс найближчого квантованого значення з кодової книги, отже кодування (E) можна описати як відображення з R^M до набору індексів $I = \{1, 2, \dots, N\}$, а декодування (D) – відображення набору індексів I у набір кодових векторів Y :

$$E : R^M \rightarrow I, \quad (1.26)$$

$$D : I \rightarrow R^M. \quad (1.27)$$

Для точного опису параметрів фрейму векторний квантизатор повинен використовувати якомога більше бітів інформації, а це означає, що векторна кодова книга повинна містити значну кількість векторів. Відповідно ускладнюється процедура створення та зберігання такої кодової книги, а також процедура пошуку найближчого вектора при квантуванні. Таким чином, вираш у швидкості передачі при векторній квантизації досягається ціною підвищення обчислювальної складності і зростання обсягів пам'яті, необхідної для збереження кодових книг [29].

Зважаючи на вказані причини, останнім часом багато робіт [26, 29–31] присвячено пошуку субоптимальних алгоритмів векторного квантування, які дозволяють досягти компромісу між швидкістю передачі та ефективністю процесу ущільнення мовленнєвого сигналу. Так, наприклад, Морія і Хонда [29] запропонували гібридний векторно-скалярний квантизатор, обробляючи коефіцієнти за допомогою векторного квантизатора на першому кроці, та за допомогою скалярного – на другому. Шохамом [26] запропонований каскадний векторний квантизатор для квантизації LPC. Але жоден запропонований квантизатор не забезпечує достатньої якості сигналу, тому Бішну Аталом та Кульдіп Палівалом [29] було запропоноване розділене векторне квантування (split vector quantization). Цей підхід базується на тому, що набір з LSP коефіцієнтів розбивається на кілька підвекторів, які потім квантуються окремо. В [30, 31] досліджено, що розбиття 10 лінійних спектральних пар, на підвектори 5×5 коефіцієнтів, при розмірі кодової книги підвекторів 4096×4096 забезпечує виконання вимог, сформованих вище.

Розділене векторне квантування дозволяє значно зменшити обсяг векторної кодової книги (в 4096 разів) та зменшити обчислювальні

витрати на квантизацію вектора, та при цьому отримати кращу якість кодованого сигналу порівняно зі скалярним квантуванням.

1.3.4 Векторне квантування параметрів мовленнєвого сигналу

Прикладом реалізації векторної квантизації параметрів є метод CELP, запропонований Етолом і Шредером в 1984 р., що базується на принципі аналізу через синтез. Суть цього підходу полягає в тому, що на передавальній стороні окрім обчислення параметрів мовленнєвого сигналу, що використовуються для ущільнення, здійснюється синтез сигналу з метою порівняння отриманого декодованого сигналу з оригінальним та оптимізація параметрів опису мовленнєвого сигналу з метою мінімізації помилки кодування. Цей метод дозволяє отримати високу якість кодованого сигналу на швидкостях від 4,8 до 16 Кбіт/с. На його базі розроблено апаратуру для систем мобільного зв'язку, яка не поступається за якістю іншим аналогам, але є значно дешевшою [28, 29, 32, 33].

Для опису голосового тракту в CELP використовується модифікація методу лінійного прогнозування з розміром фрейму 10–30 мс. В свою чергу кожен фрейм ділиться на чотири підфрейми по 60 відліків, для кожного з яких виконується аналіз мовленнєвого сигналу і виділяються параметри LPC-моделі. Далі ці параметри кодуються в бітовий потік та передаються в канал зв'язку. Для моделі CELP сигнал на виході формується, виходячи із співвідношення [29]

$$s(n) = \sum_{m=1}^M a(m)s(n-m) + x(n), \quad (1.28)$$

де M – порядок моделі; $s(n)$ – послідовність відліків мовленнєвого сигналу; $a(m)$ – коефіцієнти лінійного прогнозування, що характеризують властивості голосового тракту; $x(n)$ – послідовність, що породжує сигнал помилки голосового тракту.

Однією з головних відмінностей від FS 1015 LPC є те, що для формування сигналу збудження використовується векторна кодова книга, яка містить фіксований набір визначених імпульсів або «білий шум». На передавальній стороні знаходиться індекс найкращої послідовності для збудження сигналу, а на приймальній стороні вибирається послідовність сигналів з кодової книги, яка відповідає цьому індексу. Обрана послідовність підсилюється відповідно до коефіцієнта підсилення та пропускається через два фільтри: фільтр синтезу періоду

ЛІТЕРАТУРА

1. Игнатов В. А. Теория информации и передачи сигналов / В. А. Игнатов. – М. : Советское радио, 1979. – 280 с.
2. Куприянов М. С. Цифровая обработка сигналов / М. С. Куприянов, Б. А. Матюшкин. – С-Пб. : Питер, 1998. – 416 с.
3. Прокис Дж. Цифровая связь / Дж. Прокис. – М. : Радио и связь. – 2000. – 800 с.
4. Kleijn W. B. Wideband Speech Coding. Speech Coding and Synthesis / W. B. Kleijn, K. K. Paliwal. – The Netherlands : Elsevier Science, 1995. – P. 289–310.
5. Ворсано Д. Н. Кодирование речи в цифровой телефонии / Д. Н. Ворсано // Сети и системы связи. – 1996. – № 8. – С. 24–27.
6. Быховский М. А. Пионеры информационного века. История развития теории связи. / М. А. Быховский. – М. : Техносфера, 2006. – 376 с.
7. Рабинер Л. Р. Цифровая обработка речевых сигналов / Л. Р. Рабинер, Р. В. Шафер. – М. : Радио и связь, 1981. – 496 с.
8. Chu W. C. Speech Coding Algorithms – Foundation and Evolution of Standardized Coders / W. C. Chu. – New Jersey : Wiley, 2003. – 553 p.
9. Ткаченко О. М. Вокодер LSF зі швидкістю 1600 біт/с / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Вісник вінницького політехнічного інституту. – 2006. – № 1. – С. 78–86.
10. Сапожков М. А. Вокодерная связь / М. А. Сапожков, В. Г. Михайлов. – М. : Радио и связь, 1983. – 248 с.
11. Вокодерная телефония. Методы и проблемы / ред. А. А. Пирогова. – М. : Связь, 1974. – 246 с.
12. Gray A. H. Quantization and bit allocation in speech processing / A. H. Gray, J. D. Markel // IEEE Trans. Acoust., Speech, Signal Processing. – 1986. – V. ASSP-24. – P. 459–473.
13. Ткаченко О. М. Адаптивне визначення голосової активності при дельта-ущільненні мовленнєвих сигналів / О. М. Ткаченко, Л. В. Крупельницький, С. В. Хрущак // Наукові праці Вінницького національного технічного університету. – 2011. – № 3. – Режим доступу:

http://www.nbuu.gov.ua/e-journals/vntu/2011_3/2011-3.files/uk/11omtoss_ua.pdf.

14. Itakura F. Line spectrum representation of linear predictive coefficients of speech signals / F. Itakura // *Acoust. Soc. Amer.* – 1985. – V. 57. – P. 535.

15. Viswanathan R. Quantization properties of transmission parameters in linear predictive systems / R. Viswanathan, J. Makhoul. // *Speech, Signal processing. IEEE Trans. Acoust.* – 1985. – V. ASSP-23. – P. 309–323.

16. Kang G. S. Application of line-spectrum pairs to low-bit-rate speech encoders / G. S. Kang and L. J. Fransen // *Speech, Signal Processing: proc. IEEE Int. Conf. Acoust.* – FL : Tampa, 1985. – P. 244–247.

17. Иванов В. Н. Вычисление линейных спектральных частот / В. Н. Иванов // *Электросвязь.* – 1997. – № 6. – С. 25–27.

18. Soong F. K. Line spectrum pair (LSP) and speech data compression / F. K. Soong and B. H. Juang // *Acoust. Speech. Sig. Processing: proc. IEEE Int. Conf.* – CA : San Diego, 1987. – P. 1.10.1–1.10.4.

19. Palival K. Efficient Vector Quantization of LPC Parameters at 24 Bits/Frame / K. Palival, S. Atal // *IEEE Transactions on Speech and Audio Processing.* – 1993. – V. 1, № 2. – P. 3–14.

20. Soong F. K. Optimal quantization of LSP parameters / F. K. Soong and B. H. Juang // *Speech Signal Processing: proc. IEEE Inf. Conf. Acoust.* – FL : Tampa, 1988. – P. 394–397.

21. Lahouti F. Single and Double Frame Coding of Speech LPC Parameters Using a Lattice-based Quantization Scheme / F. Lahouti, A. R. Fazel, A. H. Safavi-Naeini // *Department of Electrical and Computer Engineering: technical report.* – 2004. – 22 p.

22. Shoham Y. Cascaded likelihood vector coding of the LPC information / Y. Shoham // *Speech, Signal Processing: proc. IEEE Int. Conf. Acoust.* – Scotland : Glasgow, 1989. – P. 160–163.

23. Ткаченко О. М. Дослідження моделі вокодера на основі векторного квантування / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак

// Автоматика 2006 : матеріали XIII міжнародної конференції з автоматичного управління – Вінниця : ВНТУ, 2007. – С. 487–491.

24. Ткаченко О. М. Модель вокодера на основі LSP / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Контроль і управління в складних системах (КУСС-2005) : тези доповідей VII Міжнародної науково-технічної конференції. – Вінниця : УНІВЕРСУМ-Вінниця, 2005. – С. 265.

25. Makhoul J. Vector quantization in speech coding / J. Makhoul, S. Houcos, H. Gish // Proc. IEEE. – 1985. – № 73. – P. 1551–1558.

26. Paliwal K. K. Quantization of LPC parameters in Speech Coding and Synthesis / K. K. Paliwal, W. B. Kleijn. – Amsterdam : Elsevier, 1995. – P. 443–466.

27. Soong F. Line spectrum pair (LSP) and speech data compression / F. Soong, B. Juang // Signal Processing: proc. Acoustic Speech. – San Diego, 1984. – P. 1.10.1–1.10.4.

28. Soong F. Optimal quantization of lsp parameters / F. Soong, B. Juang // Speech, Signal Processing: proc. Conf. Acoustic. – New York, 1988. – P. 394–397.

29. Paliwal K. K. Efficient Vector Quantization of Line Spectral Frequencies Using the Switched Split Vector Quantiser / K. K. Paliwal, S. Stephen // School of Microelectronic Engineering: technical report. – 2004. – 4 p.

30. Ткаченко О. М. Вокодер LSF на основі векторного квантування / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // XIII Міжнародна науково-технічна конференція з автоматичного управління (Автоматика–2006) : тези доповідей – Вінниця : УНІВЕРСУМ-Вінниця, 2006. – С. 442.

31. Ткаченко О. М. Ефективне векторне квантування LSF-параметрів при ущільненні мовних сигналів / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Інформаційні технології та комп'ютерна інженерія. – 2007. – № 1. – С. 124–129.

32. Ткаченко О. М. Ущільнення мовних сигналів із застосуванням векторного квантування в системах цифрового зв'язку / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Методи та засоби кодування, захисту й ущільнення інформації : тези доповідей I Між-

народної науково-практичної конференції. – Вінниця : ВНТУ, 2007. – С. 108–109.

33. LSF-вокодер на основі векторного квантування / Н. О. Біліченко, О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Реєстрація, зберігання і обробка даних. – 2007. – № 1. – С. 35–41.

34. Гольдштейн Б. С. IP-телефонія / Б. С. Гольдштейн, А. В. Пинчук, А. Л. Суховицкий. – М. : Радио и связь, 2006. – 336 с.

35. Росляков А. В. IP-телефонія / А. В. Росляков, М. Ю. Самсонов, И. В. Шиббаева. – М. : Радио и связь, 2003. – 115 с.

36. Гольдштейн Б. С. Протокол SIP : справочник. / Б. С. Гольдштейн, А. А. Зарубин, В. В. Саморезов. – СПб. : Питер, 2005. – 275 с.

37. Соколов О. В. Организация и возможности сетей IP-телефонии / О. В. Соколов, Н. В. Слободской. // Voip-Exchange. – 2006. – № 7. – С. 54–67.

38. Платов М. Что важно знать об IP-телефонии / М. Платов // Системный администратор. – 2005. – № 5. – С. 12–16.

39. Гольдштейн Б. С. Протокол SIP: учебное пособие / Б. С. Гольдштейн, В. Ю. Гойхман, Д. Н. Онучина. – СПб. : ГОУВПО СПбГУТ, 2008. – 75 с.

40. Гольдштейн Б. С. Сети NGN. Оборудование IMS: учебное пособие / Б. С. Гольдштейн, В. Ю. Гойхман, Ю. В. Столповская. – СПб. : ГОУВПО СПбГУТ, 2010. – 57 с.

41. Harte L. IPTV testing: Service quality monitoring, analyzing, and diagnostics for IP television systems and services / L. Harte // West Vance Street Fuquay Varina. – 2008. – № 27. – P. 65–67.

42. Olsson U. Towards the all-IP vision / U. Olsson // Ericsson Review. – 2005. – № 1. – P. 82–85.

43. Darling P. Telstra's Next Generation Network / P. Darling // Telecommunications Journal of Australia. – 2006. – № 1. – P. 56–78.

44. Ткаченко О. М. Розробка кодових книг для вокодера на основі LSP / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Оптико-електронні інформаційно-енергетичні технології. – 2005. – № 2. – С. 219–223.

45. Tkachenko A. N. Codebooks research for Ukrainian speech compression / A. N. Tkachenko, O. D. Feferman, S. V. Hruschak // Інтернет-Освіта-Наука-2006 : збірник матеріалів V Міжнародної конференції ІОН-2006. – Вінниця : УНІВЕРСУМ-Вінниця, 2006. – Т. 2. – С. 689–691.

46. Маршалл А. Неравенства: теория мажоризации и ее приложения / А. Маршалл, И. Олкин. – М. : Мир, 1983. – 576 с.

47. Ткаченко А. Н. Сжатие речевых сигналов с применением теории мажоризации / А. Н. Ткаченко, О. Д. Феферман, С. В. Хрущак // Инфокоммуникационные технологии в науке, производстве и образовании : материалы конференции. – Ставрополь : СевКавГТУ, 2008. – С. 167–169.

48. Швидкий пошук при векторному квантуванні лінійних спектральних частот / Н. О. Біліченко, О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Реєстрація, зберігання і обробка даних. – 2008. – С. 37–47.

49. Ткаченко О. М. Ефективний пошук векторів у кодовій книзі / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Проблеми інформатизації та управління. – 2007. – № 22. – С. 121–127.

50. Метод прискореного пошуку векторів у кодовій книзі / О. Д. Азаров, О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Современные информационные и электронные технологии : труды VIII международной научно-практической конференции. – Одесса : ОНПУ, 2007. – С. 28.

51. Ущільнення мови із застосуванням структурованих кодових книг / О. Д. Азаров, О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Современные информационные и электронные технологии : труды девятой международной научно-практической конференции – Одесса : ОНПУ, 2008. – С. 28.

52. Ткаченко О. М. Effective codebook vector search / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Інтернет–освіта–наука-2008 : збірник матеріалів VI міжнародної конференції. – Вінниця : УНІВЕРСУМ-Вінниця, 2008. – Т. 2. – С. 298.

53. Ткаченко О. М. Підвищення ефективності пошуку векторів у кодовій книзі / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Об-

роблення сигналів і зображень та розпізнавання образів : праці ІХ Всеукр. міжнар. конф. – К., 2008. – С. 193–195.

54. Ефективний пошук векторів у кодовій книзі в процесі ущільнення мовних сигналів / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак, Д. В. Монастирський // Контроль і управління в складних системах : матеріали ІХ Міжнародної конференції. – Вінниця, 2008. – С. 20.

55. Ткаченко О. М. Дельта-ущільнення мовленнєвих сигналів з прогнозуванням індексів кодової книги / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Вісник Вінницького політехнічного інституту. – 2009. – № 1. – С. 85–88.

56. Коваленко И. Н. Теория вероятностей и математическая статистика : учебное пособие / И. Н. Коваленко, А. А. Филиппова. – М. : Высшая школа, 1982. – 256 с.

57. Ткаченко О. М. Створення векторних кодових книг для ущільнення мови / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Методи та засоби кодування, захисту й ущільнення інформації : збірник тез доповідей ІІ МНТК. – Вінниця, 2009. – С. 146–147.

58. Гурский Е. И. Теория вероятностей с элементами математической статистики : учебное пособие для вузов / Е. И. Гурский. – М. : Высшая школа, 1971. – 328 с.

59. Хинчин А. Я. Избранные труды по теории вероятностей / А. Я. Хинчин. – М. : ТВП, 1995. – 556 с.

60. Ткаченко О. М. Використання методу дельта-ущільнення для вокодера MELP / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Методи та засоби кодування, захисту й ущільнення інформації : збірник тез доповідей ІІ МНТК. – Вінниця, 2009. – С. 148–149.

61. Пат. 32410 Україна, МПК (2006) G10L 21/00 G10L 19/00. Спосіб ущільнення мовних сигналів / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак ; заявник і власник патенту Вінницький національний технічний університет. — № 200800956 ; заявл. 28.01.08 ; опубл. 12.05.08, Бюл. № 9.

62. Пат. 38505 Україна, МПК (2009) G10L 21/00 G10L 19/00. Спосіб дельта-ущільнення мовних сигналів / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак ; заявник і власник патенту Вінницький національ-

ний технічний університет. — № 200810041 ; заявл. 04.08.08 ; опубл. 12.01.09, Бюл. №1.

63. Соколов Н. А. Качество обслуживания трафика речи в сети NGN / Н. А. Соколов // Мир связи. – 2006. – № 7. – С. 21–27.

64. Lane J. Best-effort network layer packet reordering in support of multipath overlay packet dispersion / J. Lane, A. Nakao // Global telecommunications: proc. – New Orleans, 2008. – P. 2457–2462.

65. Контроль показателей качества обслуживания с учетом перехода к сети связи следующего поколения / Б. С. Гольдштейн, М. А. Маршак, Е. Д. Мишин, Н. А. Соколов // Техника Связи. – 2009. – № 1. – С. 14–21.

66. Хинчин А. Я. Математические методы теории массового обслуживания / А. Я. Хинчин // Труды Математического института им. В. А. Стеклова. – 1955. – Т. 49. – 122 с.

67. Ramirez J. Voice Activity Detection. Fundamentals and Speech Recognition System Robustness / J. Ramirez, J. M. Gorriz, J. C. Segura // Robust Speech Recognition and Understanding. – 2007. – P. 1–23.

68. Benyassine A. A silence compression scheme for use with G.729 optimized for V.70 digital simultaneous voice and data application: ITU-T recommendation G.729 annex B / A. Benyassine, E. Shlomot, and H.-Y. Su. // IEEE Communications Magazine. – 2006. – V. 35. – P. 64–73.

69. Хрущак С. В. Визначення голосової активності з використанням адаптивного порогу // Методи та засоби кодування, захисту й ущільнення інформації : збірник тез доповідей III МНТК. – Вінниця, 2011. – С. 178–179.

70. Özer H. A geometric algorithm for voice activity detection in nonstationary Gaussian noise / H. Özer and S. G. Tanyer // IEEE Transactions on Speech and Audio Processing. – 2000. – V. 8. – P. 478–482.

71. C. Intensity Fluctuations in Telephone Traffic / C. Palm // Ericsson Tech. – 1943. – V. 1, № 44. – P. 56–64.

72. Перцева Л. В. Качество передачи речи при использовании низкоскоростных кодеков на абонентских линиях / Л. В. Перцева // Электросвязь. – 1987. – № 8. – С. 48–64.

73. А. с. 12396. Комп'ютерна програма аналізу алгоритмів ущільнення мови / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак. – № 14842 ; заявл. 10.01.05 ; опубл. 03.03.2005. – 5 с.

74. А. с. 24757. Комп'ютерна програма структуризації векторних кодових книг для застосування в системах цифрового зв'язку / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак. – № 25094 ; заявл. 18.03.08 ; опубл. 17.06.2008. – 5 с.

75. А. с. 26051. Комп'ютерна ущільнення мовних сигналів із застосуванням структурованих кодових книг / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак. – № 26051 ; заявл. 02.07.08 ; опубл. 13.10.2008. – 5 с.

76. Ткаченко О. М. Метод дельта-ущільнення мовних сигналів / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Інформаційні технології та комп'ютерна інженерія. – 2008. – № 1(11). – С. 8–13.

77. Ткаченко О. М. Дельта-ущільнення мовленнєвих сигналів із застосуванням методу визначення голосової активності / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Укробраз–2010 : матеріали Міжнародної науково-практичної конференції. – К., 2010. – С. 193–198.

78. ГОСТ Р 50840-95. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости. Введ. 97-01-01. – М : Издательство стандартов, 1995. – 234 с.

79. ГОСТ 16600-72. Передача речи по трактам радиотелефонной связи. Требования к разборчивости речи и методы артикуляционных измерений. Введ. 74.19.01. – М : Издательство стандартов, 2007. – 76 с.

Наукове видання

**Ткаченко Олександр Миколайович
Хрущак Сергій Вікторович**

**УЩІЛЬНЕННЯ ПАРАМЕТРІВ
МОВЛЕННЄВОГО СИГНАЛУ
НА ОСНОВІ ВЕКТОРНОГО КВАНТУВАННЯ**

Монографія

Редактор С. Малішевська

Оригінал-макет підготовлено С. Хрущак

Підписано до друку 15.04.2014 р.
Формат 29,7×42¼. Папір офсетний.
Гарнітура Times New Roman.
Друк різнографічний. Ум. др. арк. 6,7
Наклад 300 (1-й запуск 1–75) прим. Зам № В2014-15

Вінницький національний технічний університет,
КІВЦ ВНТУ,
21021, м. Вінниця, Хмельницьке шосе, 95,
ВНТУ, ГНК, к. 114.
Тел. (0432) 59-85-32.

Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.

Віддруковано ФОП Барановська Т. П.
21021, м. Вінниця, вул. Порики, 7.
Свідоцтво суб'єкта видавничої справи
серія ДК № 4377 від 31.07.2012 р.