

# ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА КОМП'ЮТЕРНА ТЕХНІКА

---

УДК 621.39

О. М. Ткаченко, канд. техн. наук, доц.;

С. Л. Козлов, студ.

## АДАПТИВНЕ ПОДАННЯ ПАРАМЕТРІВ У КОДОВИХ КНИГАХ ДЛЯ УЩІЛЬНЕННЯ МОВЛЕННЄВИХ СИГНАЛІВ

*Запропоновано метод адаптивного подання параметрів у кодових книгах, який відноситься до методів роздільного квантування для ущільнення мовленнєвих сигналів. Проведено експериментальне дослідження запропонованого методу.*

### Вступ

Наслідком розвитку сучасних технологій цифрового зв'язку, таких як стільникова та IP-телефонія, є необхідність розробки високоефективних методів і алгоритмів кодування мовлення, що реалізують високий ступінь ущільнення зі збереженням достатнього рівня якості звучання. Серед причин, що зумовлюють необхідність підвищення рівня ущільнення мовленнєвих сигналів, можна виділити такі:

- обмежена пропускна здатність цифрових каналів зв'язку;
- необхідність криптографічного захисту інформації;
- необхідність компактного збереження інформації.

У теперішній час найвищий рівень ущільнення (найменшу швидкість передачі) забезпечують методи на основі лінійного прогнозування (LPC — Linear Prediction Coding), які в процесі безпосередньої реалізації дозволяють зменшити швидкість передачі даних до 4800 біт/с. На практиці лінійні коефіцієнти прогнозування зазвичай подають у вигляді LSF (лінійних спектральних частот), що надає додаткові можливості покращення якості ущільненого сигналу. Для подальшого підвищення рівня ущільнення використовуються квантизатори, що квантують LSF за допомогою певних наборів значень, які називають кодовими книгами. Це дозволяє отримати швидкість передачі даних 2400 біт/с та менше. Проте подальше зменшення швидкості передавання призводить до суттєвого погіршення якості відновленого звуку [1].

*Метою роботи є підвищення ефективності передавання мовленнєвих сигналів через розробку методу адаптивного квантування. Квантизатор, що реалізує запропонований метод, є різновидом роздільного квантизатора, проте позбавлений його недоліків.*

Для досягнення поставленої мети необхідно розв'язати такі задачі: проаналізувати існуючі методи квантування мовленнєвих сигналів; розробити математичну модель адаптивного квантизатора; провести експериментальні дослідження.

### Розробка методу адаптивного кодування

В ході практичної реалізації систем зв'язку параметри сигналу (коефіцієнти лінійного прогнозування, лінійні спектральні частоти) безпосередньо не передаються в канали зв'язку, оскільки несуть надлишкову інформацію. Зазвичай цієї надлишковості намагаються позбутися для ущільнення даних. Це досягається передачею не параметрів мовленнєвого сигналу, а індексів їх наборів, що містяться в певних таблицях — кодових книгах [2].

Чим менше індексів міститиме книга, тим більшим буде ступінь ущільнення інформації. Тому є сенс зберігати в кодових книгах лише найбільш репрезентативні набори параметрів, тобто такі, за допомогою яких можливо найкращим чином закодувати весь спектр наборів параметрів. Таким чином, в процесі кодування сигналу після отримання певного вектора параметрів у кодовій книзі,

проводиться пошук найбільш схожого вектора, індекс якого передається в канал зв'язку. Декодуючи за цим індексом, обирається вектор з тієї ж кодової книги, на основі якого відтворюється мовленнєвий сигнал. Для розрахунку значень, що ввійдуть у кодові книги, використовуються методи кластеризації [1].

Кодові книги поділяють на скалярні та векторні. Скалярні книги містять набори значень для кожного елемента вектора параметрів окремо. Вони є простішими в обчисленнях і вимагають менше місця для зберігання. Векторні книги містять набори векторів параметрів. Обчислювальні витрати і витрати пам'яті з використанням векторних кодових книг набагато перевищують витрати використання скалярних, проте вони забезпечують більший ступінь ущільнення.

Скалярний квантизатор  $Q$  розміром  $N$  — це перетворення множини дійсних чисел  $x \in R$  у скінченну множину  $Y$ , що містить  $N$  скалярних значень, званих центроїдами,  $Q : R \rightarrow Y$ , де  $(y_1, y_2, y_3) \in Y$  [3].

Множину  $Y$  називають кодовою книгою квантизатора. Операцію квантування можемо записати у вигляді

$$Q(x) = y_i, x \in R, i = 1, \dots, N. \quad (1)$$

У більшості випадків розмір квантизатора скінченний, а отже, доцільно характеризувати квантизатор кількістю двійкових розрядів, яка необхідна для опису індексу квантованого значення в кодовій книзі. Для цього зручно користуватись поняттям роздільної здатності квантизатора [4]

$$r = \log_2 N. \quad (2)$$

Векторний квантизатор  $Q$  розмірності  $K$  та розміру  $N$  — це перетворення  $K$ -мірного вектора  $X$  з Евклідового простору  $R^K$  в скінченний набір  $Y$ , що містить  $N$   $K$ -мірних векторів, званих кодовими векторами або центроїдами,  $Q : R^K \rightarrow Y$ ,

де  $X = [x_1, x_2, \dots, x_K]$ ;  $(Y_1, Y_2, \dots, Y_N) \in Y$ ;  $Y_i = [y_{i1}, y_{i2}, \dots, y_{iK}]$ ;  $i = 1, 2, \dots, N$ .

Оскільки векторний квантизатор є узагальненням скалярного, то його роздільна здатність визначається за аналогією до скалярного.

Мірою спотворення називається невід'ємна величина  $d(X, Q(X))$ , що оцінює спотворення в процесі квантування

$$d(X, Q(X)) = \begin{cases} 0, & Q(X) = X; \\ > 0, & Q(X) \neq X. \end{cases} \quad (3)$$

Існують різні шляхи оцінювання спотворення, але в більшості прикладних реалізацій використовують середньоквадратичну міру відстані завдяки її простоті [4]

$$d^2(X, Y_i) = \sum_{k=1}^K (x_k - y_{ik})^2. \quad (4)$$

Крім того, на практиці використовують зважену міру відстані, вперше запропоновану в роботі [5]. Показано, що використання зваженої міри відстані дозволяє зменшити роздільну здатність з 26 біт до 24 біт зі збереженням необхідного рівня якості.

У більшості випадків немає можливості реалізувати векторний квантизатор з прийнятною роздільною здатністю, оскільки він вимагатиме великих витрат пам'яті (на зберігання кодових книг) і часу (на пошук векторів, найближчих до вхідного). Для мінімізації витрат часу та пам'яті на практиці використовують субоптимальне квантування, коли кодові книги розбивають на дві або більше частини меншої розмірності [4]. Залежно від способу формування вихідного вектора з кількох підвекторів розрізняють багатоетапне та роздільне квантування [5]. Роздільний квантизатор характеризується простотою реалізації та високою швидкістю роботи, але має низку недоліків, що пов'язані з залежностями між параметрами LSF.

Суть першого недоліку полягає у тому, що підвектори у кодових книгах вибираються незалежно і їх сумісність не враховується. Тому через параметри, що знаходяться на межах підвекторів, може виникати спотворення ущільненого сигналу. Цього недоліку можна позбутися ускладненням процедури пошуку, але результат у більшості випадків не виправдовує обчислювальні витрати.

Як правило, у кожній книзі зберігаються послідовні компоненти векторів. При цьому кожний на-

ступний параметр в наборі LSF повинен мати значення, більше за попереднє. А отже, хоча пошук у кодових книгах проходить незалежно, для кожної наступної кодової книги кількість фактичних претендентів буде зменшуватись. В цьому полягає другий недолік.

Для виправлення зазначених недоліків пропонується новий підхід до розміщення параметрів кодового вектора у кодових книгах.

Нехай квантизатор розмірністю  $2K$  складається з двох кодових книг  $Y^A$  та  $Y^B$  розміром  $N$ , відповідно операції квантування позначимо  $Q^A$  та  $Q^B$ . Необхідно квантувати вхідний вектор  $x$ . Представимо його квантованими векторами  $y^A$  та  $y^B$  відповідно для кожної книги.

$$y_i^A = Q_A(x_{2i}), i \in [1, K]. \quad (5)$$

З наведеної формули видно, що перша кодова книга складається з парних елементів вхідного вектора. Квантований вектор з другої книги обчислюється на основі отриманих квантованих значень для першого вектора за формулою

$$y_i^B = Q_B\left(\frac{x_{2i-1} - t_i}{y_i^A - t_i}\right); i \in [1, K]; \quad (6)$$

$$t_j = \begin{cases} 0, & j = 1; \\ y_{j-1}^A, & j \neq 1. \end{cases} \quad (7)$$

Друга кодова книга буде складатись з відношень різниці непарного елемента та попереднього йому, на різницю двох сусідніх елементів. При цьому, у випадку першого елемента, значення попереднього вважається нулем.

Під час декодування значення вихідного вектора будуть обчислюватись таким чином:

$$x_{2i} = y_i^A, i \in [1, K]; \quad (8)$$

$$x_{2i-1} = y_i^B (y_i^A - t_i) + t_i, i \in [1, K]. \quad (9)$$

Слід зазначити, що квантування другою кодовою книгою здійснюється на основі результатів квантування першою кодовою книгою. Отже за достатніх розмірів другої кодової книги непарні елементи відновленого вектора будуть близькими до значень вхідного незалежно від навчального матеріалу. Таким чином, можна говорити про адаптацію системи до мовця.

Ілюстрацію цього ефекту наведено на рис. 1. На ньому показано графіки вхідного (суцільна лінія) та відновленого векторів (пунктирна лінія). Можна побачити, що значення, 4-го, 6-го та 8-го компонентів векторів значно відрізняються. У той самий час точки, що відповідають непарним елементам векторів, досить близькі для вхідного та вихідного векторів. Тобто, незважаючи на те, що в процесі квантування першою кодовою книгою не знайшлося задовільного вектора, ситуацію було певною мірою виправлено за рахунок закладеної можливості адаптації.

Процедура створення кодових книг за адаптивного квантування є достатньо простою. Аналогічно до процедури створення кодових книг за роздільного квантування, вона зводиться до виконання кластеризації. Єдиною відмінністю є етап підготовки даних, на якому необхідно отримати набори парних елементів (перша кодова книга) та відповідних відношень (друга кодова книга) з навчального матеріалу.

Легко бачити, що процедура пошуку кодових векторів дещо ускладниться. Як і за роздільного квантування вона складатиметься з двох етапів, але тепер вони будуть відрізнятись:

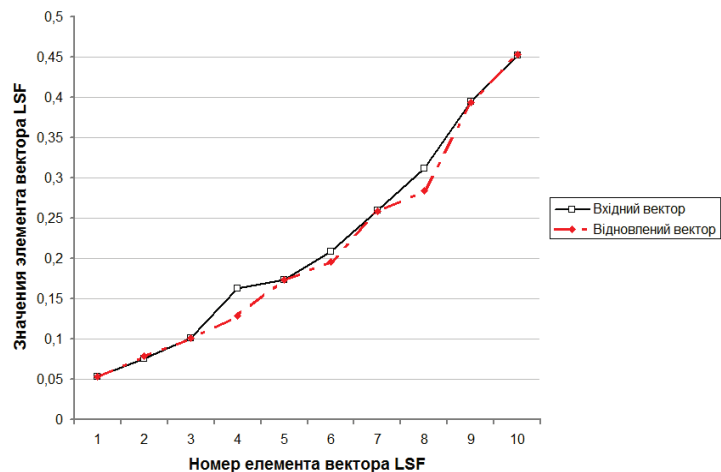


Рис. 1. Графічне зображення вхідного та відновленого векторів за адаптивного квантування

- пошук найкращого кодового вектора у першій книзі за зваженим критерієм відстані;
- пошук найкращого кодового вектора у другій книзі за середньоквадратичним критерієм відстані.

### Проведення експериментальних досліджень

Основою для проведення тестів використовувалась демонстраційна версія фонетичної бази даних ТІМІТ (Texas Instruments and Massachusetts Institute of Technology), що складається з фраз, надиктованих американською англійською мовою людьми різної статі. Цей матеріал був випадковим чином розбитий на дві частини: 10-хвилинний навчальний набір для створення кодових книг та набір тривалістю 3,5 с для проведення тестів.

Створення кодових книг відбувалось таким чином. Навчальний фонетичний матеріал був поданий у вигляді наборів векторів, елементами яких є параметри LSF. Отримані набори за допомогою середовища Matlab було розділено на частини відповідно до типу квантизатора (роздільне квантування, адаптивне квантування), після чого кластеризовано за допомогою стандартної реалізації методу *k*-середніх, який входить до пакету Matlab. Оцінювання результатів проводилося за спектральним спотворенням *SD* [5].

Для проведення порівняння було обрано межі роздільної здатності від 20 біт до 24 біт. В цих межах протестовано можливі конфігурації кодових книг для роздільного та адаптивного квантування зі зваженим пошуком. Результати наведено в табл. 1 і 2.

Порівнюючи результати експериментів, можемо зазначити, що для однакових конфігурацій кодових книг значення спектрального спотворення зменшилось у середньому на 0,033 дБ, а процент фреймів зі спотворенням 2...4 дБ зменшився на 0,19 %. Фрейми зі спотворенням, більшим за 4 дБ, відсутні.

Таблиця 1

Результати роздільного квантування зі зваженим пошуком

Конфігурація кодової книги	Роздільна здатність	Середнє SD	Процент фреймів з SD в межах 2...4 дБ
4096×4096	24	0,868525	1,53535
4096×2048	23	0,91315	2,213675
2048×4096	23	0,92265	1,9974
4096×1024	22	0,945025	3,2027
1024×4096	22	0,94935	2,1785
2048×2048	22	0,96545	2,726425
1024×2048	21	0,97905	2,55965
2048×1024	21	0,98245	3,72105
1024×1024	20	1,020625	4,177125

Таблиця 2

Результати адаптивного квантування в межах 20...24 біт

Конфігурація кодової книги	Роздільна здатність	Середнє SD	Процент фреймів з SD в межах 2...4 дБ
8192×2048	24	0,791575	1,2062
4096×4096	24	0,814525	1,395475
8192×1024	23	0,831625	1,360325
4096×2048	23	0,845225	1,5236
2048×4096	23	0,879875	2,3351
4096×1024	22	0,88425	1,806375
2048×2048	22	0,909425	2,436375
4096×512	21	0,92995	1,9322
2048×1024	21	0,945675	2,6893
4096×256	20	0,9738	2,2127
1024×1024	20	1,020775	3,81045
2048×512	20	0,98925	3,105

Використання асиметричних кодових книг, як, наприклад,  $8192 \times 2048$  дозволить підвищити якість ущільненого сигналу (спектральне спотворення зменшиться на 0,076, кількість фреймів зі спотворенням 2...4 дБ зменшиться на 21 %) порівняно з кодовою книгою  $4096 \times 4096$  тієї ж роздільної здатності для роздільного квантування, при цьому обчислювальні затрати збільшаться приблизно на 25 %.

На рис. 2 наведено розподіл спектрального спотворення кодових книг  $4096 \times 4096$  для адаптивного квантування (суцільна лінія) та роздільного квантування (пунктирна лінія). З рисунку видно, що за адаптивного квантування збільшується кількість фреймів з малим спектральним спотворенням.

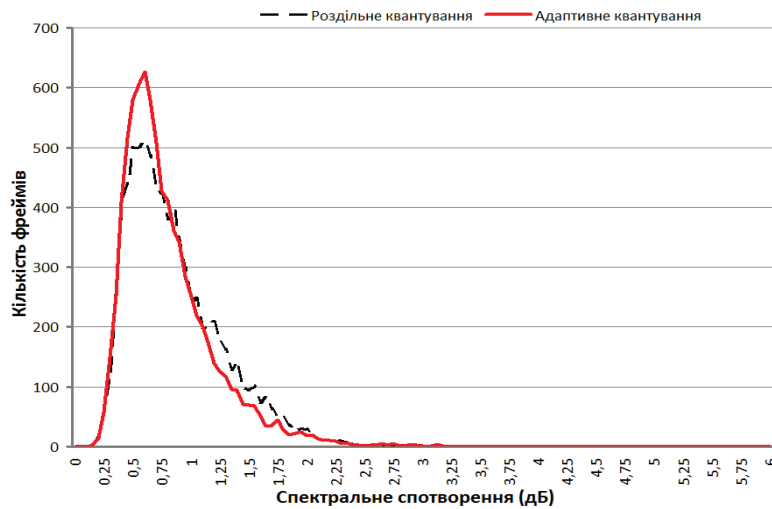


Рис. 2. Розподіл спектрального спотворення кодових книг  $4096 \times 4096$  для адаптивного та роздільного квантування

## Висновки

Запропоновано метод адаптивного подання мовленнєвих сигналів, що полягає у специфічному поданні параметрів мовленнєвого сигналу у кодових книгах. Для забезпечення адаптивності, пошук у другій кодовій книзі здійснюється з урахуванням значень, квантованих першою кодовою книгою.

Результати експериментального дослідження адаптивного квантизатора зі застосуванням запропонованого методу показали, що значення середнього спектрального спотворення зменшується на 0,033 дБ порівняно з роздільним квантуванням.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Ткаченко О. М. Методи кластеризації даних в цифровій обробці мовленнєвих сигналів. [Електронний ресурс] / О. М. Ткаченко, О. Д. Феферман // Наукові праці ВНТУ. — 2010. — № 1. — Режим доступу : [http://www.nbu.gov.ua/e-journals/VNTU/2010\\_1/2010-1.files/uk/10omtoss\\_ua.pdf](http://www.nbu.gov.ua/e-journals/VNTU/2010_1/2010-1.files/uk/10omtoss_ua.pdf).
2. Дж. Макхоул Векторное квантование при кодировании речи / Дж. Макхоул, С. Русос. Г. Гиш // ТИИЭР. — 1985. — 73 с.
3. K. K. Paliwal Quantization of LPC parameters in Speech Coding and Synthesis / K. K. Paliwal, W. B. Kleijn. — Amsterdam : Elsevier. — 1995. — P. 443—466.
4. Wai C. Chu Speech Coding Algorithms – Foundation and evolution of Standardized Coders / Wai C. Chu Wiley // Interscience. — USA, 2003. — 585 p.
5. K. K. Paliwal Efficient Vector Quantization of LPC Parameters at 24 Bits/Frame / K. Paliwal, S. Atal // IEEE Transactions on Speech and Audio Processing. — 1993. — Vol. 1, No. 2. — P. 3—14.

Рекомендована кафедрою обчислювальної техніки

Стаття надійшла до редакції 3.11.10  
Рекомендована до друку 11.11.10

**Ткаченко Олександр Миколайович** — доцент кафедри обчислювальної техніки;

**Козлов Сергій Леонідович** — студент Інституту інформаційних технологій та комп'ютерної інженерії.

Вінницький національний технічний університет, Вінниця