

О. М. Ткаченко, к. т. н., доц.

МЕТОД ВЕКТОРНОГО ДЕЛЬТА-КВАНТУВАННЯ ПАРАМЕТРІВ МОВЛЕННЕВОГО СИГНАЛУ

У статті запропоновано метод векторного квантування LSF-параметрів мовленнєвого сигналу з прогнозуванням подальшого значення. Основна ідея методу полягає в тому, що замість квантування дійсного вектора LSF кодується різниця між дійсним та прогнозованим значеннями. Такий підхід дозволяє зменшити динамічний діапазон вхідних величин і відповідно похибку квантування. Розроблено ітераційну процедуру побудови кодових книг для реалізації запропонованого методу. Наведено експериментальні результати апробації методу для різних швидкостей кодування мовленнєвого сигналу.

Ключові слова: мовленнєвий сигнал, параметризація, кодові книги, роздільне та багатоетапне векторне квантування, вектори LSF-параметрів, похибка кластеризації, коефіцієнт кореляції, спектральне спотворення.

Вступ

Помічена Г. Муром закономірність зростання продуктивності процесорів удвічі кожні два роки дійсно підтверджувалася на практиці останні 40 років [1]. Приблизно такими ж темпами зростали й обсяги мультимедійної інформації, що передається та зберігається в комп'ютерних системах. Проте існування фундаментальних обмежень, пов'язаних з атомарною природою речовини та швидкістю розповсюдження світла, змусили у 2007 р. Г. Мура визнати, що закон незабаром перестане діяти. Водночас аналогічних обмежень на зростання обсягів інформації наразі не виявлено.

Отже, існує проблема недостатньої продуктивності комп'ютерних систем, призначених для оброблення мультимедійної, зокрема, голосової інформації. Зменшення обсягів даних, необхідних для коректного відновлення цифрового мовленнєвого сигналу, дозволяє суттєво підвищити ефективність роботи таких систем.

На сьогодні найвищий ступінь ущільнення забезпечують параметричні методи з подальшим квантуванням отриманих параметрів. На етапі параметризації сигнал розбивають на ділянки (фрейми) однакової довжини, як правило, по 10 – 30 мс, та для кожної з них обчислюють деякий вектор параметрів. У сучасних системах ущільнення мовленнєвих сигналів стандартом де-факто стало обчислення параметрів на основі моделі лінійного прогнозування десятого порядку. На етапі квантування отриманий вектор параметрів замінюють так званим квантованим вектором – найближчим до вхідного вектора представника репрезентативної сукупності параметрів, що зберігаються в кодових книгах (КК). У роботі [2] було показано:

1. Як параметри квантування доцільно використовувати лінійні спектральні частоти (LSF), які обчислюють за значеннями коефіцієнтів лінійного прогнозування (LPC).

2. Оптимальне квантування повного LSF-вектора є обчислювально складним, тому його розбивають на два або три підвектори, кожен з яких квантується окремо. Альтернативним варіантом є багатоетапне квантування повного вектора за допомогою КК меншого розміру, коли на кожному наступному етапі квантується похибка, що залишилася після квантування на попередньому етапі.

3. Похибку квантування для i -го фрейму оцінюють за спектральним спотворенням SD_i :

$$SD_i^2 = \frac{1}{F_s} \int_0^{F_s} [10 \log_{10}(P_i(f)) - 10 \log_{10}(\hat{P}_i(f))]^2 df,$$

де $P_i(f) = 1/|A_i(\exp(j2\pi f / F_s))|^2$, $\hat{P}_i(f) = 1/|\hat{A}_i(\exp(j2\pi f / F_s))|^2$ – спектральні щільності відповідно початкового і квантованого векторів для i -го фрейму, $A_i(z)$, $\hat{A}_i(z)$ – початковий та квантований LPC-поліноми, що відповідають i -му фрейму, F_s – частота дискретизації.

4. Для досягнення транспарентності середнє значення похибки квантування SD_{av} за всіма фреймами має складати приблизно 1 дБ; кількість фреймів, для яких $SD > 2\text{дБ}$, не повинна перевищувати 2%; фреймів, для яких $SD > 4\text{дБ}$, не повинно бути взагалі.

5. Зазначених умов досягають під час розбиття вектора LSF-параметрів на два підвектори та використання для квантування вхідного вектора 24 бітів на один фрейм.

Подальшому зменшенню обсягу даних для опису спектральної інформації було присвячено кілька робіт [3, 4]. Проте задача зменшення бітової швидкості, необхідної для передавання мовленнєвої інформації, за умов збереження транспарентності квантування, а також виконання наявних обмежень за обсягами пам'яті, складністю обчислень та затримкою кодування, залишається актуальною.

У роботі запропоновано метод ущільнення мовленнєвої інформації, який дозволяє зменшити обсяги даних для опису спектральної інформації до 20 бітів на один фрейм при збереженні придатної якості відновленого мовленнєвого сигналу.

Метод векторного дельта-квантування з прогнозуванням

Будь-який метод ущільнення ґрунтується на використанні надлишковості, що властива мовленнєвому сигналу. Ця надлишковість зберігається й після переходу до параметричного опису сигналу. Кількісною мірою надлишковості слугує коефіцієнт кореляції. У таблиці 1 наведено значення коефіцієнта кореляції між окремими параметрами LSF у межах фрейму, отримані для тренувальної послідовності з 90000 векторів.

Таблиця 1

Коефіцієнти кореляції LSF усередині фрейму

№	1	2	3	4	5	6	7	8	9	10
1	1,00	0,45	-0,05	-0,07	-0,17	-0,27	-0,28	-0,28	-0,24	-
2		1,00	0,57	0,23	0,11	-0,02	0,03	0,06	-0,04	-
3			1,00	0,56	0,32	0,39	0,38	0,38	0,19	-
4				1,00	0,58	0,46	0,44	0,26	0,19	-
5					1,00	0,62	0,42	0,24	0,01	-
6						1,00	0,62	0,45	0,20	-
7							1,00	0,63	0,25	-
8								1,00	0,58	-
9									1,00	-
10										1,00

У таблиці 2 наведено значення коефіцієнта кореляції між відповідними LSF-параметрами для послідовних фреймів.

Коефіцієнти кореляції LSF для сусідніх фреймів

№	1	2	3	4	5	6	7	8	9	10
	0,68	0,73	0,76	0,81	0,86	0,84	0,82	0,81	0,76	0,75

Як можна побачити з таблиць, спостерігаємо доволі високу кореляцію як між окремими параметрами всередині фрейму, так і між компонентами вектора LSF для сусідніх фреймів. Можливі такі шляхи використання цих двох видів кореляції:

- перехід від скалярного квантування (СК) параметрів до векторного квантування (ВК);
- застосування методів прогнозування з подальшим видаленням прогнозованих значень з метою зменшення динамічного діапазону вхідних величин.

Існує принципова можливість застосування обох підходів для використання як міжфреймової кореляції, так і кореляції всередині фрейму. Проте СК LSF-параметрів із прогнозуванням усередині фрейму, хоча й вимагає мінімальних обчислювальних витрат, значно поступається за ефективністю ВК, що виражається у збільшенні спектрального спотворення [3]. Водночас квантування LSF-параметрів для послідовних фреймів пов'язано з внесенням додатково затримки, що є неприпустимим для деяких застосувань. Отже, доцільним видається варіант ВК параметрів LSF для окремих фреймів із подальшим прогнозуванням значень для послідовних фреймів. Саме цей підхід, названий автором векторним дельта-квантуванням з прогнозуванням (ВДКП), запропоновано у цій роботі.

В основу ВК за допомогою КК покладено просту ідею: замість передачі дійсного значення параметрів вхідного мовленнєвого сигналу обчислюється та передається індекс вектора з КК, який є найближчим до вхідного. При цьому в КК зберігаються найбільш репрезентативні значення параметрів, отримані в результаті кластеризації тренувальної послідовності вхідних векторів LSF. На відміну від цього, за методом ВДКП у КК зберігаються значення параметрів, отримані як результат кластеризації різниці (дельти) e_i дійсного l_i та прогнозованого значення \tilde{l}_i LSF. За умов вдалого прогнозу, динамічний діапазон вхідних величин, які використовують для створення КК, і відповідно, похибка кластеризації мають зменшитися, що, у свою чергу, призведе до зменшення спектрального спотворення у процесі квантування.

Отже, у процесі створення КК вхідна послідовність складатиметься з векторів $e_i = l_i - \tilde{l}_i$. При використанні авторегресійної моделі лінійного прогнозування першого порядку прогнозоване значення $\tilde{l}_i(m)$ m -го компонента LSF для i -го фрейму обчислюють за формулою:

$$\tilde{l}_i(m) = \alpha(m)\hat{l}_{i-1}(m) + \beta(m), \quad (1)$$

де параметри $\alpha(m)$ та $\beta(m)$ обчислюють, виходячи з умови мінімізації середньоквадратичної похибки прогнозування:

$$\alpha(m) = \frac{COV(l_i(m), \hat{l}_{i-1}(m))}{VAR(l_i(m))}, \quad (2)$$

$$\beta(m) = E(l_i(m)) - \alpha(m)E(\hat{l}_{i-1}(m)), \quad (3)$$

де E , COV , VAR позначають відповідно математичне сподівання, коваріацію та дисперсію, $\hat{l}_{i-1}(m)$ – квантоване значення m -го компонента вектора LSF на $i-1$ -му фреймі; $\hat{l}_{i-1}(m) = l_{i-1}(m)$.

Процедуру квантування LSF-параметрів за методом ВДКП представлено на рис. 1.

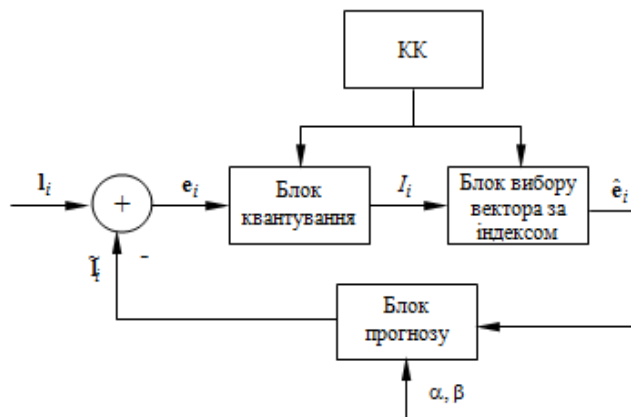


Рис. 1. Структурна схема квантування за методом ВДКП

Створення кодових книг за методом ВДКП

Параметри α та β обчислюють для тренувальної послідовності векторів LSF. Це означає неможливість безпосереднього застосування формул (2) та (3) у процесі створення КК, оскільки значення $\hat{l}_{i-1}(m)$ обчислюють лише під час квантування, яке, у свою чергу, може відбуватися лише за наявності КК.

Отже, практичне застосування ВДКП потребує реалізації ітераційної процедури створення КК.

Крок 1. Для $i = 1, 2, \dots, N$, $m = 1, 2, \dots, M$ присвоїти $\hat{l}_{i-1}(m) = l_{i-1}(m)$,

Крок 2. За формулами (2) та (3) обчислити параметри α , β .

Крок 3. За формулою (1) обчислити прогнозовані значення параметрів LSF $\tilde{\mathbf{l}}_i$.

Крок 4. Обчислити значення векторів похибки $\mathbf{e}_i = \mathbf{l}_i - \tilde{\mathbf{l}}_i$.

Крок 5. Для отриманих векторів похибки \mathbf{e}_i за методом k -середніх створити КК \mathbf{Y} .

Крок 6. Використовуючи параметри α , β та створену КК \mathbf{Y} , провести квантування тренувальної послідовності векторів \mathbf{l}_i й отримати квантовані значення векторів $\hat{\mathbf{l}}_i$.

Крок 7. Обчислити похибку квантування. Якщо вона змінилася на достатньо малу величину порівняно з попередньою ітерацією, завершити процедуру. Якщо ні, повернутися до кроку 2.

Отже, процедура створення КК складається з трьох частин: обчислення параметрів прогнозу, квантування, а також безпосередньо кластеризації. Створення початкової КК доцільно виконувати за вдосконалим методом k -середніх [5]. На наступних ітераціях можна застосувати класичний алгоритм k -середніх, використовуючи наявні вектори \mathbf{Y} для ініціалізації центрів. Схематично процедуру створення КК представлено на рис. 2:

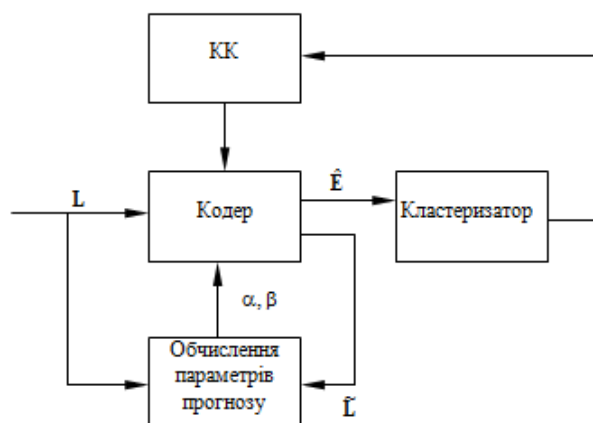


Рис. 2. Структурна схема створення КК для методу ДВКП

Розглянуту вище процедуру створення КК можна використовувати як для роздільного, так і для багатоетапного ВК. В останньому випадку процедура створення КК може потребувати додаткових ітерацій, зумовлених необхідністю узгодження між собою окремих частин КК [6]. Проте підвищення продуктивності, що виражається у зменшенні спектрального спотворення в процесі квантування, є достатнім обґрунтуванням додаткових часових витрат на стадії створення КК, яка виконується один раз на підготовчому етапі.

Експериментальні результати

Для експериментального дослідження розробленого методу було використано загальнодоступну частину англійського акустичного корпусу TIMIT [7]. Тренувальна послідовність складалася з 90000 векторів LSF, отриманих на основі моделі лінійного прогнозування десятого порядку. Тестова послідовність складалася з 15000 векторів LSF, що відрізнялися від векторів тренувальної послідовності. Довжина фрейму складала 20 мс. Вимірювання відстані відбувалося за зваженою Евклідовою метрикою з використанням ваг розрядів, обчислених за спектральною чутливістю [8].

За основу для порівняння було обрано схему роздільного векторного квантування (РВК) з розбиттям 10-вимірної вектора LSF на два підвектори розмірністю п'ять. Таке ж розбиття використовувалося при роздільному векторному квантуванні з прогнозуванням (РВДКП). При використанні багатоетапного квантування процес кодування відбувався у два етапи, розміри обох КК були однакові. Моделювали два варіанти пошуку: послідовний, коли для пошуку на другому етапі використовувався один найближчий вектор, отриманий на першому етапі, і пошук по дереву, коли на другому етапі використовували N_1 векторів, отриманих на першому етапі. Відповідні схеми позначено як БЕВДКП з ПП та БЕВДКП з ДП.

У таблицях 3 – 5 наведено результати, отримані при використанні для квантування вхідного вектора 24, 22 та 20 бітів на один фрейм відповідно. Продуктивність квантування оцінювали за спектральним спотворенням.

Таблиця 3

Продуктивність квантування спектральної інформації 24 бітами на фрейм

Схема квантування	SD, дБ	% перевищень	
		SD > 2 дБ	SD > 4 дБ
РВК	1,18	1,96	0,00
РВДКП	0,93	1,45	0,00
БЕВДКП з ПП	0,94	1,45	0,00
БЕВДКП з ДП, $N_1 = 10$	0,92	1,27	0,00
БЕВДКП з ДП, $N_1 = 30$	0,90	1,12	0,00

Таблиця 4

Продуктивність квантування спектральної інформації 22 бітами на фрейм

Схема квантування	SD, дБ	% перевищень	
		SD > 2 дБ	SD > 4 дБ
РВК	1,34	6,28	0,06
РВДКП	1,06	3,56	0,03
БЕВДКП з ПП	1,06	3,57	0,03
БЕВДКП з ДП, $N_1 = 10$	1,01	2,23	0,01
БЕВДКП з ДП, $N_1 = 30$	0,98	2,04	0,00

Таблиця 5

Продуктивність квантування спектральної інформації 20 бітами на фрейм

Схема квантування	SD, дБ	% перевищень	
		SD > 2 дБ	SD > 4 дБ
РВК	1,51	15,23	0,16
РВДКП	1,22	7,11	0,10
БЕВДКП з ПП	1,23	7,13	0,11
БЕВДКП з ДП, $N_1 = 10$	1,10	3,63	0,05
БЕВДКП з ДП, $N_1 = 30$	1,08	3,42	0,03

Як можна побачити з таблиць, умови транспарентності (див. п. 4 вступу) виконуються при використанні для опису спектральної інформації 20 бітів на один фрейм сигналу.

Слід зазначити, що обчислювальна складність квантування за методом ВДКП практично збігається з РВК. Лише за застосування пошуку по дереву кількість операцій збільшувалася в N_1 разів, що зумовлено особливостями схеми пошуку, а не методу ВДКП. Можливість зменшення обчислювальної складності пошуку найближчого вектора розглянуто в [9].

На рис. 3 представлено похибку квантування як функцію від номера ітерації у процесі створення КК. Похибку обчислюють як значення відстані між вхідним вектором і найближчим вектором КК та усереднюють за всіма фреймами. Як можна побачити, уже на п'ятій ітерації похибка квантування змінюється лише на 0,7%.



Рис. 3. Залежність похибки квантування від номеру ітерації

Висновки

Розроблений метод ВДКП дозволяє підвищити продуктивність квантування параметрів мовленнєвого сигналу завдяки зменшенню динамічного діапазону вхідних векторів, що використовуються для створення КК і, відповідно, зменшенню похибки кластеризації. Це

досягається за рахунок ускладнення підготовчого етапу, який потребує ітераційної процедури створення КК. Експериментальна перевірка розробленого методу показала, що умови транспарентності виконуються для двохетапного ВДКП з пошуком по дереву при використанні 20 бітів на один фрейм для опису спектральної інформації.

СПИСОК ЛІТЕРАТУРИ

1. Родоначалник [Електронний ресурс] / Ю. Ревич // Компьютерра-Онлайн. – 2012. – Режим доступу до журн.: <http://www.computerra.ru/xterra/homo/21659/>.
2. Paliwal K. K. Efficient vector quantization of LPC parameters at 24 bits/frame / K. K. Paliwal, B. S. Atal. // IEEE Transaction on Speech and Audio Processing. – 1993. – No. 2, vol. 1. – P. 3 – 14.
3. Lahouti F. Quantization of LSF parameters using a trellis modeling / F. Lahouti, A. K. Khandani // IEEE Transaction Speech and Audio Processing. – Sep. 2003. – No. 5, vol. 11. – P. 400 – 412.
4. Hai Le Vu. Efficient Distance Measure for Quantization of LSF and Its Karhunen–Loeve Transformed Parameters / Hai Le Vu and Laszlo Lois // IEEE Transactions on speech and audio processing. – Nov. 2000. – No. 6, vol. 8. – P. 744 – 746.
5. Ткаченко О. М. Метод кластеризації на основі послідовного запуску k-середніх з обчисленням відстаней до активних центроїдів / О.М. Ткаченко, Н. О. Біліченко, О. Ф. Грійо Тукало, О. В. Дзись // Реєстрація, зберігання і обробка даних. – 2012. – № 1, Т. 14. – С. 25 – 34.
6. Chu W. C. Speech Coding Algorithms: Foundation and Evolution of Standardized Coders / Wai C. Chu // NY.: John Wiley & Sons, Inc. – 2003 – 558 p.
7. Fisher William M. The DARPA Speech Recognition Research Database: Specifications and Status / William M. Fisher, George R. Doddington, Kathleen M. Goudie-Marshall // Proceedings of DARPA Workshop on Speech Recognition. – Feb. 1986. – P. 93 – 99.
8. Gardner W. R. Theoretical analysis of the high-rate vector quantization of LPC parameters / W. R. Gardner, B. D. Rao // IEEE Transaction Speech and Audio Processing. – Sep. 1995. – Vol. 3. – P. 367 – 381.
9. Ткаченко О. М. Двоетапна стратегія пошуку у векторних кодових книгах для ущільнення мовлення / О. М. Ткаченко, О. Ф. Грійо Тукало // Вісник Вінницького політехнічного інституту. – 2011. – № 6. – С. 194 – 201.

Ткаченко Олександр Миколайович – к. т. н., доцент кафедри обчислювальної техніки, тел. 59-84-13, e-mail: ant@vstu.vinnica.ua.

Вінницький національний технічний університет.