

УДК 621.39

**О. М. Ткаченко, Н. О. Біліченко,
О. Д. Феферман, С. В. Хрущак**

Вінницький національний технічний університет
вул. Хмельницьке шосе, 95, 21021 Вінниця, Україна

LSF-вокодер на основі векторного квантування

Досліджено модель LSF-вокодера з векторним квантуванням параметрів. Порівняно різні способи розбиття вектора LSF на підвектори, для чого використано міри спектрального спотворення кодованого сигналу.

Ключові слова: *кодування мови, вокодер, LSF, LPC, векторне квантування, спектральне спотворення, зважена Евклідова відстань.*

Вступ

При низькошвидкісному та середньошвидкісному кодуванні мови для представлення короткочасної спектральної інформації про сигнал широко використовуються параметри, що отримали назву коефіцієнтів лінійного прогнозування (LPC — Linear Predictive Coefficients). При практичному застосуванні LPC, як правило, перетворюються на лінійні спектральні частоти (LSF — Linear Spectral Frequencies), які, завдяки певним властивостям [1], є менш чутливими до каналних завад і більш зручними для використання в кодових книгах.

У [2] розглянуто методику розробки кодових книг з урахуванням специфіки української мови на основі скалярного квантування. Але в [3] показано, що векторне квантування дозволяє отримати менше спотворення мовного сигналу порівнянно зі скалярним квантуванням за умов однакової ємності інформаційних параметрів. Тому було вирішено на базі наявного фонетичного матеріалу за методом K -середніх побудувати векторну кодову книгу та розробити модель вокодера на основі векторного квантування LSF-параметрів.

Метою даної статті є зменшення обсягу даних, необхідного для передавання мовних сигналів за умови збереження придатної якості їхнього звучання при відновленні. Для досягнення поставленої мети досліджено вокодер, побудований на основі векторного квантування. За основу було взято вокодер лінійного прогнозування, розглянутий у [4].

Постановка задачі

При використанні LPC кожні 20 мс цифрового мовного сигналу кодуються десятьма коефіцієнтами лінійного прогнозування та двома додатковими параметрами: періодом основного тону та коефіцієнтом підсилення. Таким чином, загальна швидкість, яка необхідна для передачі кодованого сигналу каналом зв'язку, складає 4800 біт/с. У сучасних системах зв'язку звичайно замість LPC використовують лінійні спектральні частоти LSF, які є більш стійкими до каналних завад, а також легше піддаються перетворенню, ніж коефіцієнти лінійного прогнозування. При цьому, для зменшення швидкості до каналу зв'язку передаються не самі значення LSF, а їхні індекси в таблиці — кодовій книзі, тобто відбувається скалярне квантування параметрів. Після квантування з десяти 8-бітних коефіцієнтів отримуються набори біт розмірністю 3, 4, 4, 4, 4, 3, 3, 3, 3, 3 біта відповідно.

Скалярне квантування, хоча і дозволяє зменшити обсяг даних, що передається, але не є самим ефективним способом представлення інформації про сигнал. Векторне квантування LSF є набагато ефективнішим і дозволяє досягти кращої якості мовного сигналу при незмінних інших параметрах. Для розробки векторної кодової книги необхідно дослідити такі питання:

- 1) які міри повинні використовуватись для оцінювання спотворень сигналу що вносяться за рахунок кодової книги;
- 2) як порівнювати вектори між собою, з огляду на різну вагу елементів;
- 3) на скільки векторів повинен розбиватись набір коефіцієнтів LSF, яка розмірність векторів, скільки елементів має бути в кодовій книзі для кожного з векторів.

Вимірювання спектрального спотворення

Вибір правильного способу вимірювання похибки, що вноситься в кодований сигнал за рахунок кодової книги, є одним із найважливіших питань при проектуванні векторного квантизатора. Базуючись на значенні похибки, можна визначити мінімальну кількість векторів, необхідну для отримання придатної якості відновленого звукового сигналу.

Важливою властивістю LSF є локалізована спектральна чутливість, яка полягає в тому, що зміна одного LSF-параметра спотворює спектр вихідного сигналу тільки у вузькому діапазоні частот навколо зміненого параметра. Дана властивість робить ці параметри придатнішими за інші для використання в кодових книгах. Вектор LSF може квантуватись окремими частинами, при цьому похибка буде привноситись тільки в певну область спектра. На рис. 1 зображено зміну спектра сигналу при внесенні похибки в сьомий елемент вектора LSF. Зміна помітна тільки в області 2800 Гц і ніяк не впливає на інші області спектра.

У роботі [2] для оцінювання якості створених кодових книг використовувалася нормована енергія квадрату похибки:

$$e = \frac{E_e}{E_s} = \frac{\sum_{i=1}^N (s_{2i} - s_{1i})^2}{\sum_{i=1}^N s_{1i}^2}, \quad (1)$$

де s_1 та s_2 — відліки сигналів, синтезованих відповідно з LPC і LSF; N — кількість відліків.

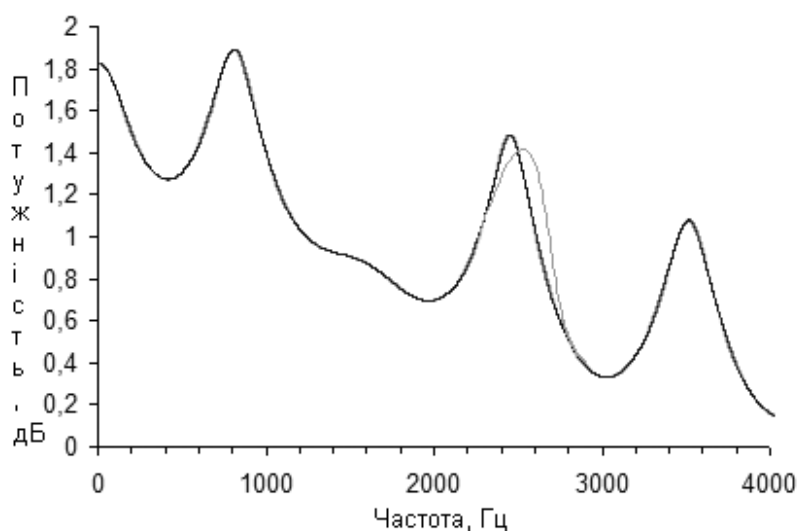


Рис. 1. Зміна спектральної потужності при зміні одного коефіцієнта LSF

Але в [3] було доведено, що для вимірювання спотворення вхідного звукового сигналу доцільно використовувати сумарне спектральне спотворення (SD — Spectral Distortion) на всьому діапазоні частот. Спектральне спотворення для i -го фрейму у децибелах, визначається за формулою:

$$D_i^2 = \frac{1}{F_B} \int_0^{F_B} \left[10 \log_{10}(P_i(f)) - 10 \log_{10}(\hat{P}_i(f)) \right]^2 df, \quad (2)$$

де F_B — частота дискретизації в герцах, а $P_i(f)$ та $\hat{P}_i(f)$ — спектральні потужності для i -го фрейму, що визначаються, виходячи зі співвідношень:

$$P_i(f) = 1 / |A_i(\exp(j2\pi f / F_B))|^2, \quad (3)$$

$$\hat{P}_i(f) = 1 / |\hat{A}_i(\exp(j2\pi f / F_B))|^2, \quad (4)$$

де $A_i(z)$ і $\hat{A}_i(z)$ — неквантизований та квантизований LSF-поліноми, що відпові-

дають i -му фрейму.

Спектральне спотворення обчислюється на частотах від 0 до 4 кГц для всіх фреймів тестових даних, та знаходиться його середнє значення. Графік розподілу спектрального спотворення представлено на рис. 2.

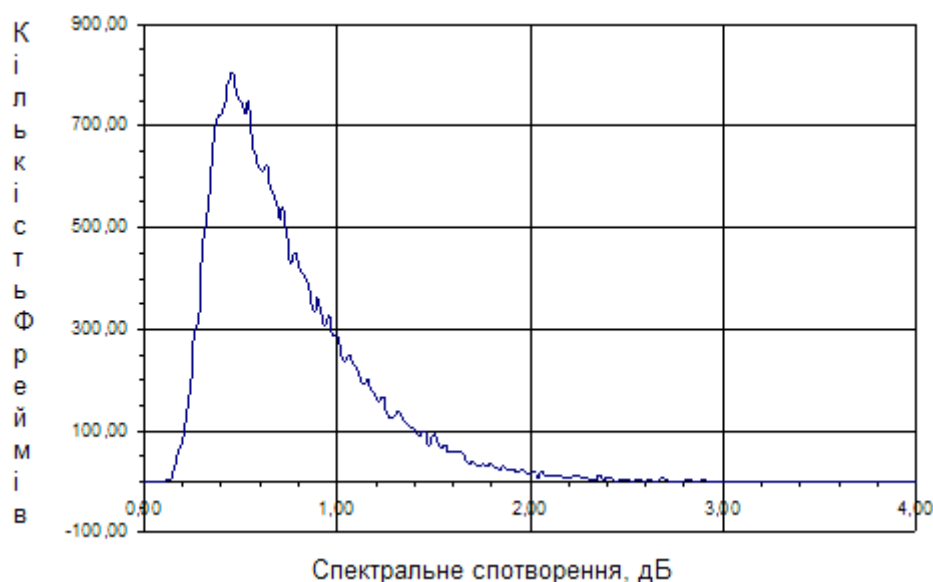


Рис. 2. Розподіл спектрального спотворення для квантизатора з розбиттям 5×5 підвекторів

У системах низькошвидкісного ущільнення звукового сигналу середнє значення спектрального спотворення не повинно перевищувати 1 дБ. Окремо виділяють фрейми, що мають спотворення більше 2 дБ; їх розподіляють на два типи: ті, що мають спектральне спотворення в діапазоні 2–4 дБ, і ті, що мають спектральне спотворення більше, ніж 4 дБ. Тобто, при побудові кодової книги необхідно дотримуватись таких умов:

- 1) середнє спектральне викривлення повинно бути близько 1 дБ;
- 2) фреймів, що мають спектральне спотворення в діапазоні 2–4 дБ, повинно бути менше ніж 2 %;
- 3) відсутні фрейми, у яких спектральне спотворення більше 4 дБ.

Зважене вимірювання відстані LSF

Для побудови векторної кодової книги та вибірки з неї необхідно обрати міру для оцінювання відстані між двома векторами LSF. Складність вимірювання полягає в тому, що кожен коефіцієнт має різну вагу. Коефіцієнт, який би врівноважував значення елементів вектора LSF, має вигляд:

$$\omega_i = [P(f_i)]^r, \quad (5)$$

де f_i — значення i -го коефіцієнта LSF у герцах; $P(f_i)$ — спектральна потужність LSF, представлена як функція частоти f_i ; r — емпірична константа, яка контролює

відносні ваги, що надаються різним LSF, вона визначається експериментально та має значення $r = 0,15$.

Тоді вираз для Евклідової відстані між двома векторами f_i та \hat{f}_i буде мати вигляд:

$$d(f, \hat{f}) = \sum_{i=1}^{10} [w_i(f_i - \hat{f}_i)]^2. \quad (6)$$

Аталом та Палівалом у [3] було запропоновано ввести додатковий коефіцієнт зваження — c_i , який базується на тому, що людське вухо більш чутливе до різниці між низькими частотами, ніж між високими. Експериментально значення c_i було встановлено таким:

$$c = \begin{cases} 1,0 & \text{для } 1 \leq i \leq 8, \\ 0,8 & \text{для } i = 9, \\ 0,4 & \text{для } i = 10. \end{cases} \quad (7)$$

Враховуючи коефіцієнт зваження, Евклідова відстань між двома векторами буде мати вигляд:

$$d(f, \hat{f}) = \sum_{i=1}^{10} [c_i w_i(f_i - \hat{f}_i)]^2. \quad (8)$$

Було проведене дослідження, на скільки покращується якість ущільнення при використанні зваженої Евклідової відстані (8). За критерій оцінки було взято середнє спектральне відхилення. Результати дослідження наведено в табл. 1. Можна побачити, що зважене вимірювання відстані дає певне покращення якості сигналу відносно звичайної Евклідової міри. Це відбувається за рахунок того, що векторний квантизатор із використанням зваженої Евклідової метрики при виборі вектора з кодової книги надає перевагу елементам LSF з меншими значеннями частоти. Надалі буде розглядатися квантизатор тільки з використанням зваженої Евклідової метрики (8).

Таблиця 1. Дослідження векторних кодових книг з використанням зваженої міри та без неї

	Тип розбиття	Кількість векторів	Середнє спектральне спотворення, дБ	2–4 дБ, %	> 4 дБ, %
Зважена міра	3×3×4	256×256×512	0,834041	1,78045	0,003879
Звичайна міра	3×3×4	256×256×512	0,942977	1,958883	0,007758
Зважена міра	5×5	4096×4096	0,746784	1,357642	0,003879
Звичайна міра	5×5	4096×4096	0,882325	1,214119	0,003879

Розбиття вектора LSF

Для того, щоб визначити число компонентів, на які необхідно розбивати вектор LSF, було досліджено декілька видів розбиття. При цьому використовувалась різна кількість підвекторів, які зберігаються в кодовій книзі, що дає змогу визначити оптимальне співвідношення якості сигналу до необхідної швидкості передачі даних для квантизатора. Результати дослідження наведено в табл. 2.

Усі квантизатори тестувались на записах, надиктованих десятьма дикторами. Запис тестового матеріалу відбувався в студійних умовах. Надиктовані тексти кодувалися з частотою дискретизації 8 кГц, по 16 біт на відлік. Загальна кількість проаналізованих фреймів становила 25780 (майже 9 хвилин). Слід зазначити, що під час попередньої обробки зі звукових файлів було видалено паузи, що покращило отримані результати.

Вектори в кодовій книзі будувались з урахуванням особливостей української мови. Для порівняння також надаються результати кодування з використанням стандартних кодових книг.

Таблиця 2. Результати дослідження різних видів розбиття вектора LSF

	Тип розбиття	Кількість векторів	Середнє спектральне спотворення, дБ	2–4 дБ, %	> 4 дБ, %
стандартна	3×3×4	256×256×512	0,978686	3,553142	0,003879
досліджувана	3×3×4	256×256×512	0,834041	1,78045	0,003879
стандартна	5×5	4096×4096	0,871322	2,199379	0,007758
досліджувана	5×5	4096×4096	0,746784	1,214119	0,003879
досліджувана	5×5	512×512	1,087711	5,919317	0,015516
досліджувана	5×5	1024×1024	0,953129	3,630721	0,003879
досліджувана	4×6	4096×4096	1,021776	6,45299	0,256102
досліджувана	4×6	2048×8192	1,782519	3,92519	0,744995
досліджувана	6×4	8192×2048	0,840095	1,993794	0

Висновки

Найкращі результати за критерієм мінімального середнього спотворення були досягнуті при використанні розбиття на два підвектори по 5 елементів у кожному, кодова книга містила по 4096 значень для кожного з підвекторів. При цьому для опису одного фрейму використовувалося 24 біти. Цілком придатними виявилися також результати розбиття на два підвектори розмірністю 6 і чотири елементи відповідно. У цьому варіанті фрейми із спотворенням більше 4 дБ взагалі відсутні. Варіант розбиття на три підвектори є також достатньо цікавим з огляду на те, що в цьому випадку вимоги до пам'яті мінімальні (3584 комірки замість 40960 і 57344 комірок у двох попередніх випадках відповідно). Використання зваженої Евклідової відстані (8) в усіх варіантах дозволяло зменшити середнє спектральне спотворення, а також кількість фреймів із спотворенням вище 2 дБ і 4 дБ. Використання вокодера на основі векторного квантування дозволило зменшити обсяг

даних, що передається до каналу зв'язку, для одного фрейму з 30-ти до 24-х біт, порівняно зі скалярним квантуванням.

Отримані результати можуть застосовуватися при побудові низькошвидкісних вокодерів для систем службового зв'язку зі збереженням необхідного рівня якості мовного сигналу та мінімальними вимогами до пам'яті.

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

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

3. *Palival K., Atal S.* Efficient Vector Quantization of LPC Parameters at 24 Bits/Frame // IEEE Transactions on Speech and Audio Processing. — 1993. — Vol. 1, N 2. — P. 3–14.

4. *Ткаченко О.М., Феферман О.Д., Хрущак С.В.* Вокодер LSF зі швидкістю 1600 біт/с // Вісник ВПІ. — 2006. — № 1. — С. 78–86.

5. *Lahouti F., Khandani A.K.* Quantization of LSF Parameters Using Trellis Modeling // IEEE Transactions on Speech and Audio Processing. — 2003. — Vol. 11, N 5. — P. 400–412.

6. *Рабинер Л.Р., Шафер Р.В.* Цифровая обработка речевых сигналов / Пер. с англ. — М.: Радио и связь, 1981. — 496 с.

7. *Lahouti F., Fazel A.R., Safavi-Naeini A.H.* Single and Double Frame Coding of Speech LPC Parameters Using a Lattice-Based Quantization Scheme. — Department of Electrical and Computer Engineering, Technical report, 2004. — 22 p.

Надійшла до редакції 11.01.2007