

## ВИКОРИСТАННЯ ПСЕВДОБЕРТАННЯ НЕСИМЕТРИЧНОЇ МАТРИЦІ САМОПОДІБНОСТІ ДЛЯ ПОБУДОВИ КВАЗІСТАЦІОНАРНИХ ДІЛЯНОК МОВНОГО СИГНАЛУ

Пелешко Д.Д., Пелех Ю.М., Рашкевич М., Пелешко М.

*Національний університет «Львівська політехніка», Україна, Львів, вул. С. Бандери 12,  
e-mail: dpeleshko@gmail.com*

Статистичні та статистично-мережеві методи (приховані марківські процеси, змішані гаусівські моделі і ін.) є одними із найбільш популярних сучасних методів обробки мовних сигналів. Визначальним для них є формулювання моделі і організації вірогідного попереднього навчання (наприклад у [5]) для визначення параметрів цієї моделі. Саме в останньому і полягає основна слабкість цих методів. Відсутність якісних наборів для попереднього визначення параметрів моделі призводить до того, що такі моделі є нечутливими до випадкових чи тимчасових мовних викидів, які можуть виникати у мовному сигналі внаслідок різноманітних причин. У випадку статистичних методів, наприклад при використанні прихованих марківських моделей без нейронних мереж, цих проблем немає. Проте якість роботи методу в цілому, наприклад при розпізнаванні мови, є залежною від якості стаціонарних чи квазістаціонарних ділянок, що забезпечує правильне визначення можливих станів марківського процесу та імовірностей переходу між станами й спотворення стану [4]. Тому важливим завданням в обробці мовних сигналів є задача виділення стаціонарних чи квазістаціонарних ділянок.

Основним завданням роботи є розробка методів виділення квазістаціонарних ділянок на основі побудови метричних просторів характеристичних векторів елементарних ділянок. Метод виділення цих векторів базується на вирішенні лінійної виродженої системи на основі методу Мура-Пенроуза [1]. Вироджена квадратна матриця-оператор лінійної системи визначається як матриця відносних несиметричних мір конвергенції, псевдобертання якої здійснюється за її сингулярним розкладом.

Для вирішення основного завдання часовий проміжок  $[0; T]$ , якому визначено мовний сигнал  $x(t)$  розіб'ємо на  $l$  елементарних ділянок, які не перетинаються між собою. На кожній з цих ділянок будемо оператор  $\nabla_i$  як один із видів матриці подібності, а саме як матрицю відносних несиметричних мір конвергенції (матриця включення), елементи якої є несиметричними мірами подібності об'єктів [6]:

$$\forall i \in [1; l]: \nabla_i = \begin{pmatrix} 1 & \sigma_{1,2} & \dots & \sigma_{1,l} \\ (\sigma_{1,2})^{-1} & 1 & \dots & \sigma_{2,l} \\ \dots & \dots & \dots & \dots \\ (\sigma_{1,l})^{-1} & (\sigma_{2,l})^{-1} & \dots & 1 \end{pmatrix}, \sigma_{n,m} = \frac{x(t_m)}{x(t_n)}, n, m \in [1, l]; t_i \in T_i .$$

Тоді для вирішення завдання виділення векторів, які виступатимуть характеристиками елементарної ділянки мовного сигналу для кожного елемента покриття  $\chi$  розглянемо рівняння  $\nabla_i y_i = x_i$ , де  $x_i = \{x(t_p) \in X_i \mid p = 1..l\}$ ;  $y_i = (y_{i,1}, \dots, y_{i,l})$  –  $l$ -вимірний вектор амплітудних значень мовного сигналу  $x(t)$  на  $i$ -му інтервалі. З цього рівняння очевидно, що вектор  $y_i = (y_{i,1}, \dots, y_{i,l})$  буде визначатись так:  $y_i = \nabla_i^{-1} x_i$ .

Оскільки матриця  $\nabla_i$  є виродженою, то знаходження вектора  $y_i$  пропонується за [2]

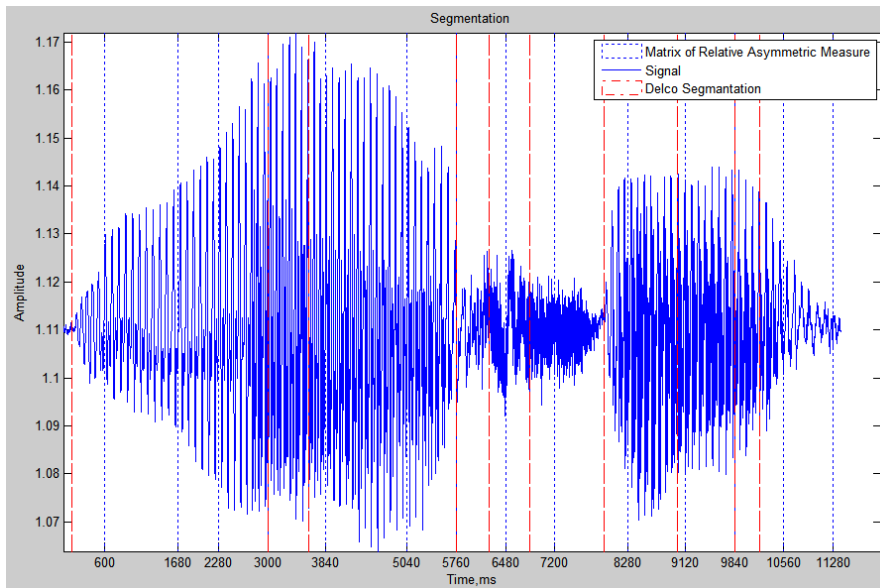
$$y_i = \nabla_i^+ x_i + (1 - \nabla_i^+ \nabla_i) r_i,$$

Де  $r_i$  – випадковий вектор розмірності  $l$ ;  $\nabla_i^+$  – узагальнена обернена матриця Мура-Пенроуза (псевдобернена до  $\nabla_i$  матриця [1, 3]), яка визначається за сингулярним розкладом матриці  $\nabla_i$  так [1]

$$\nabla_i^+ = V_i \Sigma_i^+ U_i^T,$$

де  $U_i, V_i$  – унітарні матриці порядку  $l \times l$  сингулярного розкладу матриці  $\nabla_i$ ;  $\Sigma_i^+$  – матриця порядку  $l \times l$ , яка є псевдо оберненою до діагональної матриці  $\Sigma_i$  сингулярного розкладу матриці  $\nabla_i$ . Оскільки матриця  $\Sigma_i$  є також виродженою, то матриця  $\Sigma_i^+$  отримується з  $\Sigma_i$  шляхом заміни усіх ненульових сингулярних чисел на відповідно обернені до них  $1/\sigma_{i,q}$ .

В ітераційному процесі знаходження за  $y_i^{j+1}$  випадковий вектор  $r_i^{j+1}$  – визначався за нев'язкою:  $r_i^{j+1} = \|x_i - \nabla_i y_i^j\|_l$ , тут  $\|\cdot\|_l$  –  $l$ -норма [3].



Результати сегментації мовного сигналу (слово миша) за алгоритмом DELCO та за методом на основі матриці відносних асиметричних мір конвергенції

З вектора  $y_i = (y_{i,1}, \dots, y_{i,l})$ , сформуємо такі набори даних:  $\Lambda = \{\Lambda_i \mid \Lambda_i = (y_{1,i}, \dots, y_{l,i})\}_{i=1..l}$ . Якщо вибрати метрику Хаусдорфа  $\rho_H(\Lambda_i, \Lambda_j)$ , то за нею можна визначити близькість двох елементів  $\Lambda_i$  і  $\Lambda_j$ .

За метрикою  $\rho_H(\Lambda_i, \Lambda_j)$  квазістаціонар формується за правилом – будь-яка множина  $\{X_n, X_m\}$  послідовних елементів покриття  $\eta$  утворює квазістаціонар  $Y_i$ , якщо для будь-яких двох елементів цієї множини відхилення за метрикою  $\rho_H(\Lambda_i, \Lambda_j)$  не перевищує наперед заданого порогу.

**Висновки.** Розроблений метод сегментації мовного сигналу на квазістаціонарні ділянки апробовано на прикладі первинного поділу слова “миша” (тривалість слова – 1,03 секунди, частота дискретизації 11 025 Гц, довжина елементарної ділянки –  $l = 120$  відліків), хвильове представлення якого наведено на рисунку.

У якості експертної оцінки порівняльних результатів сегментації обрано результати сегментації цього ж сигналу за алгоритмом DELCO із пороговим значенням 1.6 при тому самому розмірі елементарної ділянки. За розробленим методом виділено 13 сегментів, а за методом DELCO – 10. При цьому, у трьох точках результати сегментації повністю співпадають. На решті частини сигналу результати є близькими з різними мірами близькості. На основі цих результатів можна констатувати, що розроблений метод у порівнянні з алгоритмом DELCO є більш чутливим до змін енергії в часі. Так перевага отриманого методу може бути дуже корисною в задачах, ефективне розв’язання яких є залежним від змін енергетичних характеристик мовного сигналу. Наприклад у задачах кодування для процесів передачі мовних сигналів телекомунікаційними каналами.

#### ЛІТЕРАТУРА

1. Penrose R. A generalized inverse for matrices. Proceedings of the Cambridge Philosophical Society 51, 406-413 (1955).
2. Rashkevych Y., Peleshko D., Kovalchuk A., Kupchak M., Pelekh Y. Speech signal pseudoinvariants. Computer Science and Information Technologies: Materials of the Vith International Scientific and Technical Conference CSIT 2011. - Lviv: Publishig House Vezha&Co, 2011. 2011, c.21-22.
3. Алберт А. Регрессия, псевдоинверсия и рекуррентное оценивание/ А.Алберт// пер.с англ. – М.: Наука. – 1977. - 224с.
4. Добрушкін Г. О., Данилов В. Я. Основні підходи до розпізнавання мовленнєвої інформації/ Г. О. Добрушкін, В. Я. Данилов// Інформаційні технології та комп’ютерна техніка. Вісник Вінницького політехнічного інституту. – 2009. – № 4. – С. 50-64.
5. Дорохин О.А., Старушко Д.Г., Федоров Е.Е., Шелепов В.Ю. Сегментация речевого сигнала. «Искусственный интеллект». – 2000. – №3. – С. 450- 458.
6. Матрица мер конвергенции [Електронний ресурс] – Режим доступу до журналу: [http://ru.wikipedia.org/wiki/Матрица\\_мер\\_конвергенции](http://ru.wikipedia.org/wiki/Матрица_мер_конвергенции).