

Виокремлення квазістаціонарних ділянок мовного сигналу на основі псевдоінваріантів

Пелешко Д. Д., Пелех Ю. М., Киричук В. І.

Національний університет «Львівська політехніка» вул. С. Бандери 12, Львів, Україна, 79013
e-mail: peleshko@polynet.lviv.ua

Proposed methods of building a quasistationary areas and pseudo-invariants of speech signal based on the singular value decomposition of really determined square matrices to solve some problems of speech signals processing. Feature of these methods is their independence from the model voice tract.

ВСТУП

Задача первинного поділу мовного сигналу на квазістаціонарні ділянки є однією із ключових задач аналізу, перетворення та синтезу мови.

Успішне вирішення якої визначає ефективність розв'язку проблеми в цілому.

Метою даної роботи є розробка методів сегментації (поділу на квазістаціонари) мовного сигналу на базі використання спектральної метрики в топологічному просторі, які не залежать від моделі мовотворення.

Запропоновано методіку орієнтовану на одновимірні сигнали. Проте при незначних доопрацюваннях може застосовуватись для сигналів будь-якої розрядності.

1. ПОБУДОВА МАТРИЦІ ОПЕРАТОРА І ЇЇ СИНГУЛЯРНИЙ РОЗКЛАД

Нехай визначений на часовому проміжку $\tau = [0; T]$ дискретний мовний сигнал $x(t)$ задано у вигляді топології Γ , елементами якої виступають елементарні ділянки X_i ,

$$x(t) = \bigcup_{i=1}^n X_i, X_i \in \eta. \quad (1)$$

де η - скінченне диз'юнктивне покриття заданої на просторі сигналу топології Γ .

Сигнал є нормованим за схемою мінімаксного множника

$$K = \frac{1}{\sup_{i \in [1, n]} x(t)}. \quad (2)$$

Для кожного елемента X_i покриття η будемо оператор $\nabla_i : \mathbb{R}^l \times \mathbb{R}^l \rightarrow \mathbb{R}^l$

$$\nabla_i = \left(\delta_{i(p,q)} = \frac{x(t_q)}{x(t_p)} \middle| x(t_q), x(t_p) \in X_i \right)_{p=1, q=1}^{l, l}. \quad (3)$$

Оператор ∇_i є квадратною додатно визначеною матрицею, до якої можна застосувати операцію сингулярного розкладу (розклад SVD, [2])

$$\nabla_i = U_i \Sigma_i V_i^*, \quad (4)$$

де U_i , V_i – унітарна матриця порядку $l \times l$; V_i^* - спряжено-транспонована матриця до V_i ; Σ_i - діагональна матриця порядку $l \times l$, яка складається із сингулярних чисел $\{\sigma_{i,q} | q = 1..l\}$ таких, що $\sigma_{i,1} \geq \sigma_{i,2} \geq \dots \geq \sigma_{i,l} \geq 0$. Оскільки ∇_i є матрицею дійсних чисел, то $V_i^* = V_i^T$.

2. Виділення псевдоінваріантів мовного сигналу

Для вирішення завдання пошуку псевдоінваріантів $\{y_i \mid i = 1..n\}$ для кожного елемента покриття χ розглянемо рівняння

$$\nabla_i y_i = x_i \rightarrow y_i = \nabla_i^{-1} x_i, \quad (5)$$

де y_i – l -вимірний вектор, який виступає псевдоінваріантом для i -го елемента даного покриття.

Оскільки ∇_i є виродженою матрицею ($\det(\nabla_i) = 0$), то обернена матриця ∇_i^{-1} не існує. Взамін розглянемо псевдо обернену матрицю ∇_i^+ , яка за сингулярним розкладом (4) визначається так

$$\nabla_i^+ = V_i \Sigma_i^+ U_i^T. \quad (6)$$

Оскільки матриця $\nabla_i^+ \nabla_i$ не є виродженою, то псевдоінваріант i -го елемента покриття χ буде визначатись так [1, 2]

$$y_i = \nabla_i^+ x_i + (1 - \nabla_i^+ \nabla_i) r_i, \quad (7)$$

де $(1 - \nabla_i^+ \nabla_i)$ – оператор проектування на ядро оператора ∇_i ; r_i – з точності до розмірності випадковий вектор.

Зазначимо, що пошук псевдо оберненої матриці може бути здійснений за псевдо обертанням Мура–Пенроуза [1, 2], яке узагальнює поняття оберненої матриці.

3. Виділення квазістаціонарних ділянок мовного сигналу

Суть задачі визначення квазістаціонарних ділянок X_i' полягає у побудові нового представлення сигналу $x(t)$

$$x(t) \leftrightarrow \bigcup_{i=1}^m X_i'. \quad (8)$$

де X_i' – квазістаціонарна ділянка (надалі квазістаціонар), яка виступає об'єднанням послідовних елементарних ділянок X_i , m – кількість квазістаціонарних ділянок X_i' .

У просторі елементів X_i покриття η можна породити метричний простір за допомогою евклідової метрики

$$\forall i, j \in [1, n]: d(X_i, X_j) = \sqrt{\sum_{z=1}^l (y_{i,z} - y_{j,z})^2}, \quad (9)$$

Завдяки метриці (9) визначимо умову належності елементарної ділянки X_j до квазістаціонарну X_i'

$$\forall j \in [I_i, I_i + m_i - 1]: d(X_i, X_j) \leq \varepsilon, \quad (11)$$

де $\varepsilon \in \mathbb{R}^{1,+}$ – похибка віднесення елементарної ділянки X_j до квазістаціонару X_i' . Алгоритм визначення m_i полягає у наступному. Якщо умова (11) між ділянкою, яка визначає початок стаціонару і ділянкою, яка слідує після наступної, що входить в квазістаціонар, виконується, то збільшується значення m_i , а остання ділянка вважається такою, що включена в квазістаціонар. В протилежному випадку починається побудова нового квазістаціонару.

ВИСНОВКИ

Експериментально доведено доцільність і ефективність використання даного методу.

Розроблений метод побудови квазістаціонарних ділянок може ефективно доповнювати інші методи сегментації, які базуються на вибраних моделях мовотворення.

Основною областю використання інваріантів є задачі реконструкції сигналів.

Інваріанти є корисними при розв'язуванні задач класифікації математичних об'єктів, оскільки вони відображають внутрішні властивості об'єкта досліджень.

ЛІТЕРАТУРА

- [1] Алберт А., Регрессия, псевдоинверсия и рекуррентное оценивание, пер. с англ., главная редакция физико-математической литературы из-ва «Наука», 1977, М., 224 с.
- Роджер Пенроуз, A generalized inverse for matrices. Proceedings of the Cambridge Philosophical Society 51, 406-413 (1955).
- Є.Г. Жиляков, Є.І. Прохоренко, А.В. Болдишев, А.А. Фірсова, М.В. Фатова, Сегментація мовних сигналів на основі аналізу особливостей розподілу часток енергії за частотним інтервалам, Вісник НТУ "ХПІ". Тематичний випуск: Інформатика і моделювання. – Харків: НТУ "ХПІ". – 2011. – № 17. – С. 44 – 50.
- Сорокин В.Н., Цыплихин А.И., Сегментация и распознавание гласных, Информ. процессы. — 2004, т. 4, № 2. — С. 202—220.
- Wall, Michael E., Andreas Rechtsteiner, Luis M. Rocha, "Singular value decomposition and principal component analysis". In D.P. Berrar, W. Dubitzky, M. Granzow. A Practical Approach to Microarray Data Analysis. Norwell, MA: Kluwer. — 2003. — С. 91–109.

