

МЕТОДЫ АНАЛИЗА ВОКАЛЬНЫХ ЗВУКОВ РЕЧИ НА ОСНОВЕ ФОНЕМЛЕТ-ФУНКЦИЙ

Федоров Е. Е., доктор технических наук, доцент

Донецкая академия автомобильного транспорта, Донецк, Украина

На сегодняшний день активно разрабатываются интеллектуальные компьютерные системы распознавания речевых образов, используемые для набора текста, расшифровки стенограмм, голосового управления и др.

Важную роль для этих систем играет проблема построения эффективных методов, которые обеспечивают высокую скорость обучения модели распознавания, а также высокую вероятность, адекватность и скорость распознавания речевых сигналов. Современные системы распознавания образцов речевого сигнала используют следующие подходы к обучению коннекционистских и генеративных моделей [1–4]:

1. Выбор в качестве опорных образцов всех обучающих образцов, при этом весовые коэффициенты модели фиксированы. Недостаток — длительное распознавание.

2. Настройка весовых коэффициентов модели, при этом опорные образцы отсутствуют. Недостаток — длительное обучение и низкая вероятность распознавания в случае трудно разделимых классов.

3. Формирование опорных образцов в виде настройки векторов весовых коэффициентов модели. Недостаток — низкое качество распознавания в случае трудно разделимых классов.

4. Настройка весовых коэффициентов модели и формирование опорных образцов в виде статистических оценок множества обучающих образцов. Недостаток — длительное обучение и низкая вероятность распознавания в случае трудно разделимых классов.

5. Настройка весовых коэффициентов модели и выбор в качестве опорных образцов тех обучающих образцов, которые попадают на границы классов. Недостаток — длительное обучение и низкая вероятность распознавания в случае трудно разделимых классов.

Таким образом, все указанные подходы обладают недостатками.

Для систем распознавания образцов речевого сигнала также актуальна проблема создания методов формирования системы признаков. В работах [5–6] для определения признаков сигнала часто применяются вейвлеты. Однако эти вейвлеты не всегда учитывают особенности речевых сигналов. Таким образом, возникает необходимость в разработке методов, которые используют оконные функции, ориентированные на особенности вокальных фонем.

Целью работы является создание методов создания опорных образцов и классификация тестовых образцов колебаний вокальных фонем на основе семейства фонемлет-функций.

Для решения поставленной задачи в работе осуществляется:

- создание семейства фонемлет-функций на основе образцов колебаний вокальных фонем;
- создание опорных образцов колебаний вокальных фонем, представленных ограниченной непрерывной функцией, на основе семейства фонемлет-функций;
- классификация тестовых образцов колебаний вокальных фонем, представленных ограниченной непрерывной функцией, на основе семейства фонемлет-функций;
- восстановление фонемлета;
- дискретизация семейства фонемлет-функций;
- создание опорных образцов колебаний вокальных фонем, представленных дискретной функцией, на основе дискретизированного семейства фонемлет-функций;
- классификация тестовых образцов колебаний вокальных фонем, представленных дискретной функцией, на основе дискретизированного семейства фонемлет-функций.

Выводы:

1. В работе впервые вводится определение фонемлета, который описывает колебание вокальной фонемы.
2. На основе предложенного фонемлета создается семейство фонемлет-функций, которые являются разновидностью оконных функций. В отличие от вейвлет-функций, фонемлет-функции ориентированы на конкретные участки вокальных фонем определенного диктора, которые находятся в заданном звуковом окружении. Фонемлет-функции ориентированы на сдвиг и масштабирование колебаний вокальных фонем по времени и амплитуде. Использование фонемлет-функций позволяет уменьшить количество опорных образцов участков вокальных фонем — вместо хранения подмножества опорных образцов в процессе распознавания изменяются четыре параметра, которые характеризуют фонемлет-функцию (два параметра для сдвига и два для масштабирования). Это приводит к более компактному представлению попарных образцов и увеличению скорости распознавания за счет уменьшения количество сопоставлений тестового образца с опорными образцами.
3. Для практических приложений предлагается дискретизация семейства фонемлет-функций и восстановление фонемлета.
4. Создание опорных образцов колебаний вокальных фонем и распознавание тестовых образцов колебаний вокальных фонем осуществляется посредством норм пространств Лебега.
5. Разработанные методы могут использоваться для решения широкого круга задач, связанных с биометрической идентификацией человека и распознаванием речи.

Литература

1. Sivanandam S. N. Introduction to Neural Networks using Matlab 6.0 / S. N. Sivanandam, S. Sumathi, S. N. Deera — New Delhi: The McGraw-Hill Comp., Inc., 2006. — 660 p.
2. Haykin S. Neural networks: a comprehensive foundation / S. Haykin. — Delhi: Pearson Education Ptc. Ltd., 2005. — 823 с.
3. Rabiner L. R. Fundamentals of speech recognition / L. R. Rabiner, B. H. Jang. — Englewood Cliffs, NJ: Prentice Hall PTR, 1993. — 507 p.
4. Винцюк Т. К. Анализ, распознавание и интерпретация речевых сигналов / Т. К. Винцюк. — К. : Наук.шва думка, 1987. — 261 с.
5. Daubechies I. Ten lectures on Wavelets / I. Daubechies. — Philadelphia, PA: SIAM, 1992. — 464 p.
6. Chui C. K. An introduction to wavelets / C. K. Chui. — New York: Academic Press, 1992. — 412 p.

Анотація

У статті уведено поняття фонемлета, що описує коливання вокальної фонемі, визначене сімейство фонемлет-функцій, які враховують зрушення та масштабування коливання вокальної фонемі за часом й амплітудою, розробляються методи створення опорних зразків і методи розпізнавання тестових зразків коливань вокальних фонем на основі сімейства фонемлет-функцій. Пропоновані методи призначені для рішення широкого кола задач, пов'язаних з біометричною ідентифікацією людини та розпізнаванням мовлення диктора.

Ключові слова: фонемлет-функція, обробка мовленого сигналу, коливання вокальної фонемі, формування опорних зразків, розпізнавання тестових зразків, дискретизація сімейства фонемлет-функцій, відновлення сигналу.

Аннотация

В статье введено понятие фонемлета, который описывает колебание вокальной фонемі, определено семейство фонемлет-функций, которые учитывают сдвиг и масштабирование колебания вокальной фонемі по времени и амплитуде, разрабатываются методы создания опорных образцов и методы распознавания тестовых образцов колебаний вокальных фонем на основе семейства фонемлет-функций. Предлагаемые методы предназначены для решения широкого круга задач, связанных с биометрической идентификацией человека и распознаванием речи диктора.

Ключевые слова: фонемлет-функция, обработка речевого сигнала, колебание вокальной фонемі, формирование опорных образцов, распознавание тестовых образцов, дискретизация семейства фонемлет-функций, восстановление сигнала.

Abstract

In article the concept fonemlet which describes fluctuation of a vocal phoneme is entered, the family of fonemlet-functions which consider shift and scaling of fluctuation of a vocal phoneme on time and amplitude is defined, methods creation of support patterns and methods of recognition of test patterns of fluctuations of vocal phonemes on the basis of family of fonemlet-functions are developed. Offered methods are intended for the decision of a wide range of the problems connected with biometric identification of the person and recognition of speech of the speaker.

Keywords: fonemlet-function, processing of a speech signal, fluctuation of a vocal phoneme, formation of support patterns, recognition of test patterns, sampling of family of fonemlet-functions, signal restoration.