**Карпов Олег Миколайович. Методи аналізу і розпізнавання складних сигналів в автоматизованих системах мовного діалогу: дисертація д-ра техн. наук: 05.13.06 / Національний авіаційний ун-т. - К., 2003**

|  |  |
| --- | --- |
| |  | | --- | | **Карпов О.М.** Методи аналізу і розпізнавання складних сигналів в атоматизованих системах мовного діалогу. - Рукопис.  Дисертація на здобуття вченого ступеня доктора технічних наук за спеціальністю 05.13.06 -автоматизовані системи управління і прогресивні інформаційні технології. - Національний авіаційний університет. Київ. 2003.  Дисертація присвячена питанням побудови моделей аналізу складних сигналів і їх спектрів, методів визначення параметрів систем, що формують ці сигнали, розробки алгоритмів розпізнавання мовних сигналів і розробка на цій основі - інформаційних технологій побудови систем розпізнавання мови.  Для рішення поставлених задач у роботі запропоновані нові методи аналізу сигналів і їх спектрів, що розглядають їх як функції багатьох змінних з невідомим порядком і складом компонент, що входять до них і невідомими параметрами компонент. Методи рішення виходять із теореми Колмогорова A.М. про те, що функцію багатьох змінних можна уявити як суму, добуток і суперпозицію одновимірних функцій. У даній роботі розглядається уявлення багатовимірних функцій у базисі елементарних функцій виду: для параметричного уявлення спектра і частотної функції мовотворного тракту - резонансні функції другого порядку, в тому числі функції запізнюючіх ланок і експоненційні функції; для непараметричного уявлення спектра - параболи другого порядку і локон Аньєзі - крива 3-го порядку. Для резонансних функцій і запізнюючих ланок задача вирішена в часовій і частотній областях. Параметри мовного тракту визначені методами: послідовного вилучення складових сигналу - CET і послідовних наближень - SAT для акустичної і біомеханічної систем.  Алгоритми розпізнавання реалізовані як послівне, фонемне і сегментно-складове розпізнавання з застосуванням згладжуючих сплайнів, сплайн-ідентифікації і сплайн-синтезу. Реалізован базовий алгоритм розпізнавання на основі екстремальних функцій, який досліджен по словнику 800 слів і виявив стабільну надійність розпізнавання слів не менше 98% на одного диктора. Словарь еталонів, в загальному випадку, утримує по одної реалізації кожного слова. Добре разрізняються слова типу “бочка - дочка”. | |
| |  | | --- | | Результати вирішення науково-технічної проблеми по створенню теоретичних основ аналізу і розпізнаванню мовних сигналів - функцій багатьох змінних як суперпозиції функцій меньшого числа змінних і розробки методів реалізації при обробці мовних сигналів зводяться до наступного:  1. Вперше запропонований і реалізований метод побудови частотної функції мовотворного тракту людини шляхом послідовного вилучення компонент у часовій і частотній областях (Consecutive Extraction Transform - CET) і послідовних наближень -SAT ( Sequential Approximation Transform). Розроблені алгоритми і схеми аналізу в прямому і логарифмічному видах при розв’язку задач визначення механізму формування мовних сигналів у часовій і частотній областях. Дано опис мовного сигналу в базисі функцій CET- і SAT- перетворень при розв’язку задач структурної апроксимації спектрів у різних базисах.  2. Запропоновано і реалізовано 5 методів формування функції сегментації: верифікацією послідовності параметрів; диференціюванням функцій у частототних смугах; ідентифікацією відрізків мови символами Г, С, У, Ш по групових ознаках; виділенням із обвідних сигналів у частотних смугах сигналів в діапазоні частот (312) Гц; по методу СЕТ, що у сукупності дозволяє розв’язати у повному обсязі задачу сегментації.  3. Для структурної апроксимації параметричних функцій у виді СЕТ- і SAT-перетворень сигналів реалізовано сплайн-перетворення для фонемних і сегментно-складових послідовностей, що дозволяє розв’язати задачу опису і синтезу функцій параметрів при розпізнаванні, що визначає послідовність переходу від однієї категорії інформації до іншої, так наприклад, від параметричної до фонетичної, від символьної до словесної, при цьому мінімізуються часові витрати розпізнавання мови.  4. Теоретично й експериментально обгрунтована модель адаптивного аналізу основного тону (ОТ), яка заснована на дробових номерах коефіцієнтів Фур'є, що дозволяє здійснити синхронізацію процесу розпізнавання.  5. Розроблено обчислювальні схеми, алгоритми і програмне забезпечення сплайн-синтезу і сплайн-ідентифікації мовних сигналів.  6. Обгрунтована і реалізована оптимальна послідовність порівняння мовних параметрів при розпізнаванні, що дозволяє побудувати ієрархію розпізнавання у виді подвійкового дерева розподілу словника еталонів на підсловарі.  7. Для аналізу і наступного об'єднання різноманітних категорій інформації про мовний сигнал розроблена концепція проектування класів програмних систем із застосуванням передпроцесорних засобів мови програмування СІ і об'єктно-орієнтованого програмування, що дозволяє генерувати і модифікувати програмні засоби. Розроблені алгоритми і програмне середовище розпізнавання мови, протестовані і реалізовані для значного об’єму словників на мові DELPHI.  8. Результати вирішення даної науково-техничної проблеми впроваджені у організаціях різних галузей науки і техніки та учбовому процесі. | |