

АНАЛІЗ СПОСОБІВ ПОБУДОВИ СИСТЕМ РОЗПІЗНАВАННЯ МОВИ

Автоматичне розпізнавання мови - технологія, що дозволяє ідентифікувати слова мовця в мікрофон або по телефону. Завдання автоматичного розпізнавання мови дотепер не має якісного рішення, що забезпечувало б взаємодію з ЕОМ, що обумовлено складною частотно-часовою структурою мовного сигналу.

Процес розпізнавання мови включає наступні етапи: запис і оцифрування мови, обчислення часових і спектральних характеристик зразків, класифікацію часових фреймів по фонетичних категоріям, підбор відповідних категорій слів, вимір ступеня довіри до ухваленого рішення.

Найпростіший підхід до розпізнавання - зіставлення вхідного образу з еталонами, що зберігаються в пам'яті (рис. 1.). Розпізнаваний образ (невідомого класу) рівняється з еталонами кожного класу. Класифікація ґрунтується на заздалегідь обраному критерії подоби, тобто якщо вхідний образ краще відповідає еталону і-го класу, чим будь-якому іншому, те вхідний образ класифікується як приналежний і-ому класу.

Для порівняння необхідно зробити виміру на вхідному образі. Подання мовних елементів через систему ознак складно. Навіть фонема не є стаціонарною ділянкою мови, її складно виділити з мовного потоку. Сильний вплив роблять сусідні звуки.

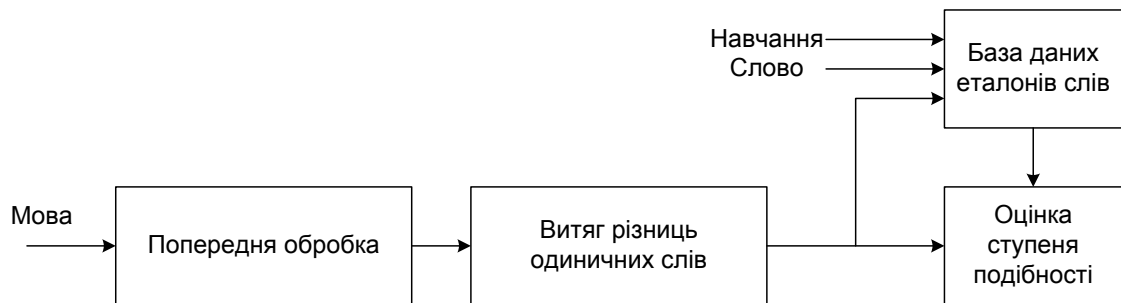


Рис. 1. Структура системи, що розпізнає, з використанням еталонів слів

Структура системи, що розпізнає, з використанням еталонів фонем наведена на рис.2. Розпізнавання слова починається з розкладання по тимчасових інтервалах фонем. БД еталонів фонем формується в процесі навчання. Потім, використовуючи фонем як параметри розрізнення, обчислюються ступені подоби з еталонами, що зберігаються в базі даних для всіх фонем.

По максимальному ступені подоби невідомих фонем еталонам виконується їхня оцінка або ідентифікація параметрів. Слідом за розпізнаванням фонем оцінюється окреме слово на основі баз даних (таблиця відповідності слів і сполучень фонем). При розпізнаванні з використанням подібних фонемних одиниць досить підготувати базу даних з декількох десятків фонем основних видів. Подальше розпізнавання безлічі слів уже виконується досить легко.

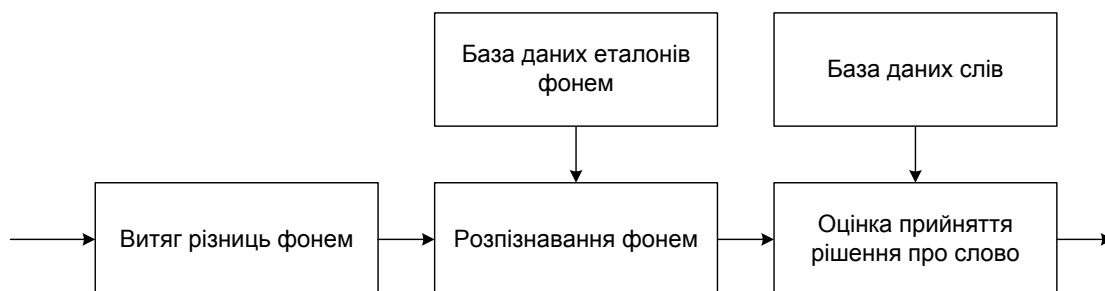


Рис. 2. Структура системи розпізнавання - фонемний підхід

Основна мета розпізнавання фраз – розуміння значеннєвого змісту фрази, тому цю процедуру називають розумінням мови.

Узагальнена схема системи розпізнавання мови представлена на рис. 3.

При розпізнаванні використовуються дві групи методів:

1. Параметричні, спрямовані на математичне перетворення мовного сигналу з виділенням і стабілізацією основних інформативних ознак:

- перетворення Фур'є,
- цифрова фільтрація,
- вейвлет-аналіз і інш.

2. Лінгвістичні, ціл'ю яким контекстна обробка висловлення (динамічне програмування, сховане марковське моделювання).

Вихідною інформацією для лінгвістичної обробки служать результати параметризації мовного сигналу.

При проектуванні системи класифікації мовних сигналів необхідно:

- поставити завдання й визначити кількості еталонних класів;
- виконати аналіз мовних сигналів з метою виділення ознак, що характеризують об'єкти розпізнавання;
- розробити алгоритм навчання, ціл'ю якого сформулювати й занести в базу знань системи еталони класів;
- ухвалити рішення про приналежність до певного класу інформації на вході.

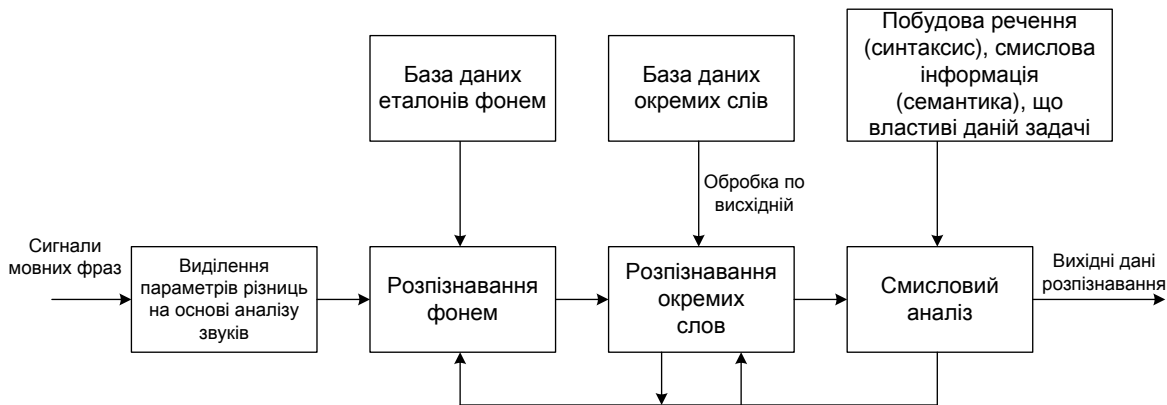


Рис. 3. Розпізнавання фраз

Основні методи розпізнавання мови.

Розпізнавання на основі ознак припускає, що кожний об'єкт, що підлягає класифікації, можна охарактеризувати вектором ознак і безліччю значень ознак. Критерій відбору ознак заснований на важливості їх для характеристик образу або важливості їх впливу на якість розпізнавання. Систему ознак вибирають інтуїтивно.

Рішення, що вимірювати, є суб'єктивним і залежить від конкретного розв'язуваного завдання - у системах розпізнавання мови використовуються вейвлет-коефіцієнти, спектральний опис, коефіцієнти лінійного проорокування й інші параметри. Подання мовних елементів через систему ознак складно. Навіть фонема не є стаціонарною ділянкою мови, її складно виділити з мовного потоку. Сильний вплив роблять сусідні звуки.

Завдання розпізнавання формулюється як завдання статистичного ухвалення рішення.

Нейромережевий підхід до розпізнавання. Нейромережа - набір елементів (нейронів), що виконують операції над вхідними даними, які мають властивості класифікації. В нейромережі пам'ять не локалізована, а розподілена по всій структурі.

Одна з переваг подібної архітектури - "життєздатність", тобто вихід з ладу n нейронів не порушить функціонування системи. Нейромережі вирішують завдання: генерації, навчання й тестування нейромереж, визначення значимості вхідних параметрів, контрастування нейромережі, пошук мінімального вирішального набору вхідних даних.