

Секція:

Приладобудування

УДК 612.78

Суль В. – ст. гр. ПМпм-62

Тернопільський державний технічний університет імені Івана Пулюя

**ФОНЕМНА СЕГМЕНТАЦІЯ МОВНОГО СИГНАЛУ З
ВИКОРИСТАННЯМ ВЕЙВЛЕТ-ПЕРЕТВОРЕННЯ**

Науковий керівник: к.м.н., доцент Гевко О.В.

Найважливішим етапом опрацювання мови в процесі розпізнавання, є виділення інформативних ознак, які однозначно характеризують мовний сигнал. Для цього в системах автоматичного розпізнавання мови здійснюють сегментацію мовного сигналу відповідно до фонетичної транскрипції. Методи математичного опрацювання мовних сигналів (кепстральний аналіз, приховане марківське моделювання) базуються на спектральному аналізі. Проте, при використанні перетворення Фур'є початковий сигнал замінюється на періодичний, з періодом що дорівнює тривалості досліджуваного зразка, і при зміні параметрів з часом дає усереднені коефіцієнти для досліджуваного зразка. У мовному сигналі інформативним є зміни його частотно-часових характеристик. Тому, для його дослідження потрібні базисні функції, здатні виявляти в аналізованому сигналі як частотні, так і часові характеристики. Виходячи з цих міркувань доречно застосувати вейвлет-перетворення.

В результаті вейвлет-перетворення мовного сигналу, було виявлено що він складається із фонем, для яких характерні квазістаціонарні ділянки, що перемежуються ділянками з порівняно швидкими змінами спектральних характеристик сигналу (міжфонемні переходи, переходи мова-пауза). Отримані розклади використано для сегментації мовного сигналу, тобто було виділено ділянки сигналу, що відповідають його окремим структурним одиницям. За такі одиниці вибрано фонемні, і завдання сегментації звелось до виявлення міжфонемних переходів.

На міжфонемних переходах сигнал зазнає значних змін відразу на кількох масштабах дослідження, і, відповідно, характеризується зростанням вейвлет-коефіцієнтів, в той час як, на стаціонарних ділянках фонем вейвлет-коефіцієнти змінюються незначно. Таким чином, знаходження міжфонемних меж зведено до встановлення моментів зростання значень вейвлет-коефіцієнтів.

Розкладання за вейвлетами мовного сигналу довжиною N відліків є сумою[1]:

$$f(t) = \sum_{k=0}^{\frac{N}{2^n}-1} s_{jk} \varphi_{jk} + \sum_{j=1}^n \sum_{k=0}^{\frac{N}{2^j}-1} d_{jk} \psi_{jk} \quad (1)$$

де: s_{jk} - коефіцієнти апроксимації; d_{jk} - коефіцієнти деталізації; φ_{jk} та ψ_{jk} - масштабовані і зміщені версії масштабної функції (скейлінг-функції) φ і базисного („материнського“) вейвлета ψ ; n - рівень деталізації; $j, k \in Z$.

Література:

1. Дремін І.М., Іванов О.В., Нечитайло В.А. Вейвлети и их использование. // Успехи физических наук, т. 171, №5 с. 465-500, 2001.