

## СПЕКТРАЛЬНИЙ АНАЛІЗ МОВНОГО СИГНАЛУ

У системах обробки мови аналоговий мовний сигнал надходить на вхід мікрофона, з виходу якого знімається електричний сигнал. Далі сигнал піддається дискретизації за часом і квантуванням по амплітуді.

У процесі квантування виникають спотворення (помилки квантування), які, по суті, означають втрату інформації.

Вихідна інформація представлена у вигляді залежності амплітуди від часу (звичайно це .wav файли).

Отримана послідовність цифрових даних надалі піддається обробці, з метою визначення частотного діапазону й інших характеристик сигналу, на основі яких його можна відтворити. Оскільки сигнал звичайно зашумлений, найпростішим способом видалення шуму є обнуління тих значень сигналу, які менше деякого граничного значення.

Завдання цифрової фільтрації - змінити співвідношення потужностей частот у звуці.

ЦФ виконує перетворення послідовності вхідних цифрових сигналів з використанням: підсумовування сигналів, розгалуження, множення на постійний коефіцієнт, затримки на один інтервал дискретизації (D).

Часова форма подання сигналу, тобто зміни сигналу залежно від часу, дозволяє визначити амплітуду, енергію, потужність і тривалість. Моделі сигналів у вигляді функції часу використовуються для аналізу форми сигналів. Складні сигнали можна представити у вигляді системи базисних функцій  $\varphi_k(t)$

$$f(t) = \sum_k c_k \varphi_k(t), \quad t \in [t_1, t_2] \quad (1)$$

де  $[t_1, t_2]$  – інтервал існування сигналу.

При обраному наборі базисних функцій сигнал  $f(t)$  повністю визначається сукупністю безрозмірних коефіцієнтів  $c_k$ . Такі сукупності чисел називають дискретними спектрами сигналів. Базисна функція  $\varphi_k(t) = e^{pt}$ , де  $p = \pm j\omega$  використовується в перетворенні Фур'є.

Крім часових характеристик сигналу важливі і його частотні властивості. Для їх дослідження використовуються частотні подання функції у вигляді спектра. Спектральне подання сигналу - розкладання його на кінцеву або нескінченну суму гармонійних сигналів. Знання частотних властивостей сигналу дозволяє вирішувати завдання ідентифікації сигналу (визначення його найбільш інформативних параметрів), фільтрації (виділення корисного сигналу на тлі перешкод), вибору частоти дискретизації безперервного сигналу, тому що цей параметр є визначальним для апаратури обробки.

Сукупність синусоїдальних складових складного звуку, заданих за допомогою амплітуд і частот цих складових представляють акустичний спектр. Для спектрального аналізу сигналу використовується дискретне перетворення Фур'є (ДПФ) і швидке перетворення Фур'є (ШПФ), що представляє процедуру прискореного ДПФ.

Відповідно до теореми Котельникова, довільний сигнал, спектр якого не містить частот вище  $F_b$  Гц, може бути повністю відновлений, якщо відомі відлікові значення цього сигналу, узяті через рівні проміжки часу  $1/(2 \cdot F_b)$  с.

Реальний мовний сигнал має кінцеву тривалість, при поданні в частотній області його спектр необмежений. Тому сигнал сегментують на ділянки порядку 10 мс, на яких він вважається стаціонарним.

Мовний сигнал має ряд особливостей, які необхідно враховувати: властивості сигналу не постійні на обраному для аналізу відрізку довжиною в слово, це нестационарний випадковий процес; складність форми сигналу (мова нагадує скоріше шум, чим регулярний сигнал).

Для подолання цих труднощів дискретний випадковий процес оцифрованого мовного сигналу вважається стаціонарним на інтервалі порядку 10 мс, тому що параметри голосового тракту на цьому інтервалі значно не змінюються. Це обґрунтований експериментально часовий інтервал.

Основне завдання обробки сигналу складається в обчисленні по вхідному сигналу сукупності параметрів (ознак), які містять інформацію про сигнал, яка використовується при синтезі й розпізнаванні.

Звичайно визначають наступні параметри сигналу:

- частоту основного тону для формування траєкторії періоду основного тону;
- короткочасну енергію для синтезу траєкторії короткочасної енергії;
- коефіцієнти лінійного прококування (КЛП) для побудови траєкторії передатної функції мовоутворюючого тракту;
- формантні частоти для відтворення траєкторії формантних частот.