Chương 1

Xử lý tín hiệu âm thanh

1.1 Phân tích phổ tín hiệu âm thanh

1.1.1 Biến đổi Fourier rời rạc

$$DFT(x[n]) = X[k] = \sum_{n=1}^{N} x[n]e^{-j2\pi kn/N}, k = 0, 1, \dots N - 1 \tag{1.1}$$

trong đó:

- n: chỉ số mẫu (theo thời gian)
- k: chỉ số tần số
- $\omega_k = 2\pi k/N$: tần số (đơn vị rad/s)
- $f_k = f_s k/N$: tần số (đơn vị Hz), với f_s là tần số lấy mẫu (sampling rate).

Biểu đồ phổ biên độ - pha. Biên độ (đơn vị dB)

$$mX[k] = 20 \log_{10} |X[k]| \tag{1.2}$$

$$aX[k] = \mathrm{unwrap}(\measuredangle X[k]) \tag{1.3}$$

Bản chất của biến đổi Fourier rời rạc. DFT chiếu tín hiệu $x[n]$ lên bộ cơ sở tạo ra bởi các hàm tuần hoàn trong không gian phức

$$s_k^*[n] = e^{-j2\pi kn/N} = \cos(2\pi kn/N) - j\sin(2\pi kn/N) \tag{1.4}$$

$$X[k] = \sum_{n=1}^{N} x[n]s_k^*[n] = \langle x, s_k^* \rangle \tag{1.5}$$

DFT của một số tín hiệu đặc biệt
Tính chất của DFT

- Tuyến tính: $DFT(ax_1[n] + bx_2[n]) = aX_1[k] + bX_2[k]$.
- Dịch chuyển: $DFT(x[n - n_0]) = e^{-j2\pi kn_0/N}X[k]$.

- DFT độ dài N tuần hoàn theo chu kì N.
- DFT của tín hiệu thực có phần thực là hàm chẵn, phần ảo là hàm lẻ, biên độ là hàm chẵn, pha là hàm lẻ.
- DFT của tín hiệu thực, chẵn có phần thực là hàm chẵn, phần ảo bằng 0, biên độ là hàm chẵn, pha bằng $n\pi$.
- Bảo toàn năng lượng:

$$\sum_{n=-N/2}^{N/2} |x[n]|^2 = \frac{1}{N} \sum_{k=-N/2}^{N/2} |X[k]|^2 \tag{1.6}$$

- Điền thêm 0 vào cuối tín hiệu → nội suy hệ số $X[k]$.
- Zero-phase:

Biến đổi DFT ngược

$$x[n] = \frac{1}{N} \sum_{i=1}^{N} X[k] s_k[n] = \frac{1}{N} \sum_{i=1}^{N} X[k] e^{j2\pi kn/N} \tag{1.7}$$

1.2 Biến đổi Fourier thời gian ngắn (STFT)

Áp dụng biến đổi DFT trên từng đoạn tín hiệu (cửa sổ - frame) có độ dài N, cách nhau H đơn vị (hop-size):

$$X_\ell[k] = \sum_{n=-N/2}^{N/2} w[n] x[n + \ell H] e^{-j2\pi kn/N} \tag{1.8}$$

trong đó ℓ là số hiệu cửa sổ (frame index), $\ell = 0, 1, \ldots$, còn $w[n]$ là hàm cửa sổ (ví dụ: cửa sổ Hamming).

Hàm cửa sổ

Biến đổi STFT ngược

1.3 Chỉnh sửa, tạo hiệu ứng

1.4 Mã hóa - giải mã, nén - giải nén