

Speech signal representation

Time Domain Representation

Time domain representation of speech signal is the most common way of representing speech signal. In this representation, the amplitude of the signal is plotted against time. The time domain representation of speech signal is also known as waveform representation of speech signal.

What is a signal?

A signal is a function that conveys information about a phenomenon. In the context of speech processing, the phenomenon is speech.

$$s : Time \rightarrow Pressure$$

What is sampling?

Sampling is the process of converting a continuous signal into a discrete signal. T is the sampling period, which is the time interval between two consecutive samples.

Sampling period is calculated as

$$T = \frac{1}{f_s}$$

where f_s is the sampling frequency, which is the number of samples per second.

Nyquist Theorem

Nyquist theorem states that the sampling frequency should be at least twice the highest frequency component of the signal. If the sampling frequency is less than twice the highest frequency component of the signal, then the signal is said to be under-sampled and aliasing occurs.

Example:

$$\begin{aligned} f_s &= 8000Hz \\ f_{max} &= 4000Hz \end{aligned}$$

In this case, the signal is not under-sampled because the sampling frequency is twice the highest frequency component of the signal.

Amplitude Quantization

Amplitude quantization is the process of converting the continuous amplitude values of the signal into discrete amplitude values. The number of bits used to represent the amplitude values is called the bit depth. The bit depth determines the number of discrete amplitude values that can be represented.

The amplitude of values are the values of the signal.

Periodicity

Periodicity is the property of a signal that repeats itself after a certain interval of time.

Periodic Signal

A signal is periodic if it remains unchanged after a certain interval of time.

$$x(t) = x(t + T)$$

where T is the period of the signal and t is the time.

How to check if a signal is periodic?

$$x(n) = \cos\left(\frac{\pi}{10}n\right)$$

In this case, the signal is periodic because it repeats itself after every 20 samples. The steps to check if a signal is periodic are:

1. Find the period of the signal
2. Check if the signal repeats itself after the period
3. If the signal repeats itself after the period, then the signal is periodic
4. If the signal does not repeat itself after the period, then the signal is aperiodic

Windowing

Split the signal into quasi stationary segments and apply Fourier Transform to each segment.

Types of windowing

1. Sine window: The amplitude of the signal is multiplied by a sine function.
2. Hamming window: The amplitude of the signal is multiplied by a Hamming function.
3. Hann: The amplitude of the signal is multiplied by a Hann function.

What is a frame

A frame is a segment of the signal that is analyzed at a time.

Overlap-Add Method

Overlap-Add method is used to reconstruct the signal from the frames. The frames are overlapped and added to reconstruct the signal.

What are window length and hop length?

Window length is the length of the frame and hop length is the distance between two consecutive frames.

Zero padding

Zero padding is the process of adding zeros to the signal to increase the length of the signal.

Time domain features

Time domain features are the features that are extracted from the time domain representation of the signal.

Root Mean Square

Root Mean Square (RMS) is the square root of the mean of the square of the signal.

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$$

Autocorrelation

Autocorrelation is the correlation of a signal with itself. It is used to find the periodicity of the signal.

Zero Crossing Rate

Zero Crossing Rate is the rate at which the signal changes its sign. Meaning, the number of times the signal crosses the zero axis.

Continuous time fourier series

Periodic signal can be represented as a sum of sinusoids.

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t}$$

where c_n is the coefficient of the sinusoid and ω_0 is the fundamental frequency.

Discrete time fourier series

Discrete time fourier series is used to represent a periodic discrete signal as a sum of sinusoids.

Discrete Fourier Series (DFS)

$$\tilde{x}(n) \rightarrow_{DFS} X(k) = \sum_{n=0}^{N-1} \tilde{x}(n) e^{-j2\pi kn/N}$$

where $X(k)$ is the DFS of the signal and N is the length of the signal.

Finite duration signal

A signal is said to be finite duration if it is non-zero for a finite duration of time.

Idk, take a finite duration signal and repeat it. Then use DFS to represent the signal.

Simmetry properties of DFS

If $x(n)$ is a real signal, then $X(k)$ is a sequence with the following properties:

1. $X(k)$ is conjugate symmetric, meaning $X(k) = X^*(N-k)$
2. $X(k)$ is real, meaning $X(k) = X^*(-k)$

Spectrum of a signal

Frequency distribution of the signal's power. Represent the DFT complex coefficients. It's in dB.

Spectrogram

A spectrogram is a visual representation of the spectrum of frequencies of a signal as it varies with time. It's a 2D representation of the signal.

Wideband and Narrowband signals

Wide band means short window, narrow band means long window.

In particular, a wideband signal has a wide range of frequencies, while a narrowband signal has a narrow range of frequencies.

Cepstrum

Cepstrum is the inverse Fourier transform of the log magnitude of the Fourier transform of a signal. It is used to analyze the periodicity of the signal.

Cepstrum is good because machines are good at analyzing this kind of representation.

Mel filter bank

Mel filter bank is a set of filters that are used to extract the Mel frequency cepstral coefficients (MFCCs) from the signal. The Mel filter bank is used to extract the Mel frequency cepstral coefficients (MFCCs) from the signal.

Discrete Cosine Transform (DCT)

Discrete Cosine Transform (DCT) is a type of Fourier transform that is used to convert a signal into a set of coefficients. The DCT is used to extract the Mel frequency cepstral coefficients (MFCCs) from the signal.