Short-Time Fourier Transform

Chia-Ping Chen

Department of Computer Science and Engineering
National Sun Yat-sen University
Kaohsiung, Taiwan 800

Spring 2012

Short-Time Fourier Transform

short-time Fourier transform

The **short-time Fourier transform** (STFT) is defined as

$$X[n, \lambda) = \sum_{m=-\infty}^{\infty} x[n+m]w[m]e^{-j\lambda m},$$

where $w[n]$ is a window sequence which determines the portion of the signal to be analyzed.

$X[n, \lambda)$ is a function of two variables, the discrete time index n and the continuous frequency λ.

STFT as discrete-time Fourier transform

Given n, $X[n, \lambda)$ can be interpreted as the DTFT of the time-shifted sequence $x[n+m]$, as viewed through the window $w[m]$.

- Note that the window $w[m]$ has a stationary origin, and different portion of the signal is viewed through the window as n changes.
- The existence of $X[n, \lambda)$ is guaranteed if $w[m]$ is of finite duration.

recovering signal from STFT

Using inverse DTFT, we have

$$x[n+m]w[m] = \frac{1}{2\pi} \int_0^{2\pi} X[n, \lambda) e^{j\lambda m} d\lambda.$$

It follows that

$$x[n] = \frac{1}{2\pi w[0]} \int_{-\pi}^{\pi} X[n, \lambda) d\lambda.$$

Example

An example of STFT is given in Example 10.9.

$w[m]$: Hamming window of length 400

$x[n]$: linear chirp signal

linear filtering interpretation

Changing variable $m' = n + m$, we can re-write STFT as

$$X[n, \lambda) = \sum_{m'=-\infty}^{\infty} x[m']w[-(n - m')]e^{j\lambda(n-m')}$$
$$= x[n] * h_\lambda[n], \quad h_\lambda[n] = w[-n]e^{j\lambda n}.$$

Thus, the STFT given λ can be interpreted as linear filtering.

Note that the frequency response of the filter is

$$H_\lambda(e^{j\omega}) = W(e^{j(\lambda-\omega)}).$$

For common windows $w[n]$, the duration is finite and $W(e^{j\omega})$ is generally lowpass. So $H_\lambda(e^{j\omega})$ is a bandpass filter centering around λ.

windowing effect

Let the DTFT of $x[n]$ be $X(e^{j\omega})$. Consider the extreme case where $w[m] = 1$, $\forall m$, i.e., no window. We have

$$X[n, \lambda) = \sum_{m=-\infty}^{\infty} x[n+m]e^{-j\lambda m} = \sum_{m'=-\infty}^{\infty} x[m']e^{j\lambda(n-m')} = X(e^{j\lambda})e^{j\lambda n}.$$

That is, $X[n, \lambda)$ is the same as $X(e^{j\lambda})$ except for a linear phase. The DTFT of $X[n, \lambda)$ is

$$X(e^{j\omega})H_\lambda(e^{j\omega}) = X(e^{j\omega})W(e^{j(\lambda-\omega)}).$$

Taking the inverse DTFT, we have

$$X[n, \lambda) = \frac{1}{2\pi}\int_0^{2\pi} e^{j\theta n}X(e^{j\theta})W(e^{j(\lambda-\theta)})d\theta.$$

sampling in time and frequency

Let the window be finite-duration

$$w[m] = 0, \quad \text{outside } 0 \leq m \leq L-1.$$

Consider $\lambda_k = {}^{2\pi k}/_{N}$, with $N \geq L$. Define (DFT)

$$X[n,k] = X[n,\lambda_k) = \sum_{m=0}^{L-1} x[n+m]w[m]e^{-j(2\pi/N)km}.$$

Thus,

$$X[n,k] = x[n] * h_k[n], \quad h_k[n] = w[-n]e^{j(2\pi/N)kn},$$

which is equivalent to a bank of N filters.

Using IDFT, we have

$$x[n+m]w[m] = \frac{1}{N}\sum_{k=0}^{N-1} X[n,k]e^{j(2\pi/N)km}, \quad 0 \le m \le L-1.$$

Therefore,

$$x[n+m] = \frac{1}{Nw[m]}\sum_{k=0}^{N-1} X[n,k]e^{j(2\pi/N)km}, \quad 0 \le m \le L-1.$$

Since using $X[n_0,k]$ we are able to reconstruct up to $x[n_0+L-1]$, the time index can be sampled as well.

Define the sampled time-dependent Fourier transform as

$$\begin{aligned} X[rR,k] &= X[rR, 2\pi k/N) \\ &= \sum_{m=0}^{L-1} x[rR+m]w[m]e^{-j(2\pi/N)km} \\ &\triangleq X_r[k], \quad -\infty < r < \infty,\ 0 \le k \le N-1. \end{aligned}$$

Note that $X_r[k]$ is simply the DFT of the windowed signal segment

$$x_r[m] = x[rR+m]w[m], \quad -\infty < r < \infty,\ 0 \le m \le N-1.$$

Figure 10.15 illustrates the sampling in frequency and time.

frame convolution

Suppose $x[n]$ is causal, $h[n]$ is causal and of finite-duration P. Let the N-point DFT of $h[n]$ be $H[k]$. Suppose $R = L$, so adjacent windows are non-overlapping. We have

$$x_r[m] = x[rL + m], \quad 0 \leq m \leq L - 1$$

and

$$x[n] = \sum_{r=0}^{\infty} x_r[n - rL].$$

Let

$$Y_r[k] = H[k]X_r[k], \quad 0 \leq k \leq N - 1.$$

Taking the inverse DFT, and from the circular convolution theorem, we have

$$y_r[m] = \sum_{l=0}^{N-1} x_r[l] h[((m-l))_N], \quad 0 \le m \le N-1.$$

If $N \ge L + P - 1$, then $y_r[m]$ will be identical to the linear convolution of $h[m]$ and $x_r[m]$. It follows that

$$y[n] = \sum_{r=0}^{\infty} y_r[n - rL]$$

is the convolution of $x[n]$ and $h[n]$.