Lecture 5: Spectograms

ECE 401: Signal and Image Analysis

University of Illinois

2/2/2017

Outline

On-Board Practice

Can I have 3 volunteers to come try this one on the board? Thanks!

$$x(t) = \sum_{k=-\infty}^{\infty} X_k e^{jk\Omega_0 t} = \begin{cases} 1 & -\frac{T_0}{4} \leq t \leq \frac{T_0}{4} \\ 0 & \frac{T_0}{4} < t < \frac{3T_0}{4} \\ x(t - T_0) & \text{always} \end{cases}$$

1. Plot $x(t)$
2. What is X_0?
3. What is X_k for $k \neq 0$?

Outline

Spectral Analysis of Quasi-Periodic Signals

- Voiced speech is periodic. But the fundamental frequency, $\Omega_0(t)$, changes over time.
- Spectral amplitudes $X_k(t)$ tell us which vowel is being produced.

$$x(t) = \sum_{k=-\infty}^{\infty} X_k(t) \exp\left(jk\Omega_0(t)t\right)$$

Spectrogram = Quasi-Periodic Analysis

In order to find the spectrum, $X_k(t)$, at each time...

1. Excise one frame of speech, that starts at time t:

$$x(t,\tau) = \begin{cases} x(t+\tau) & 0 \le \tau < T \\ 0 & \text{otherwise} \end{cases}$$

2. Pretend that one frame is a single period from some perfectly periodic longer signal.
3. Short-time Fourier transform (STFT) $X_k(t) =$ Fourier series analysis of each frame:

$$X(t,k) = \frac{1}{T}\int_0^T x(t+\tau)e^{-jk2\pi k\tau/T}d\tau$$

4. Spectogram = STFT converted to dB

$$S(t,k) = 10\log_{10}|X_k(t)|^2$$

Spectrogram: Practical Issues

1. How long is each frame?
   1. $T = T_0$: pitch synchronous analysis
   2. $T > T_0$: narrowband spectrogram
   3. $T < T_0$: wideband spectrogram
2. Digital spectrogram is computed using the discrete time Fourier series (DFS) or discrete Fourier transform (DFT)

Outline

Frequency Resolution

The answer to "how long should each frame be" is intimately connected with the idea of frequency resolution.

$$X(t,k) = \frac{1}{T}\int_0^T x(t+\tau)e^{-jk2\pi k\tau/T}d\tau$$

Has the property that the separation between two neighboring frequency bins, k and $k+1$, is

$$\Delta F = \frac{1}{T}, \quad \Delta\Omega = \frac{2\pi}{T}$$

Pitch-Synchronous Spectrogram

Suppose that $x(t)$ is really periodic with period T_0, so that

$$x(t) = \sum_{\ell=-\infty}^{\infty} X_\ell e^{\frac{j2\pi\ell(t)t}{T_0}}$$

Suppose that we perform analysis using period $T \approx T_0$:

$$X(t,k) = \frac{1}{T}\int_0^T x(t+\tau)e^{-\frac{j2\pi k\tau}{T}}d\tau$$

Then

- GOOD: If we're right, and $T = T_0$, then $X(t,k) = X_k$ exactly!
- BAD: If we're mistaken by a small amount (up to $\pm 70\%$, roughly), then the numbers $X(t,k)$ don't tell us anything about what T_0 actually was.
- WHY: $\Delta\Omega = \frac{2\pi}{T} \approx \Omega_0$. We can separate tones that are Ω_0 apart, but smaller variations are invisible.

Narrowband Spectrogram

$$x(t) = \sum_{\ell=-\infty}^{\infty} X_\ell e^{\frac{j2\pi\ell(t)t}{T_0}}, \quad X(t,k) = \frac{1}{T}\int_0^T x(t+\tau)e^{-\frac{j2\pi k\tau}{T}}d\tau$$

Suppose we choose $T \gg T_0$, say, $T \approx 2T_0$. Then

- GOOD: If $T = 2T_0$ exactly, then

$$X(t,k) = \begin{cases} X_{k/2} & k \text{ even} \\ 0 & k \text{ odd} \end{cases}$$

- GOOD: If we don't know T_0 exactly, then we choose T large enough so that $T \geq 2T_0$. Then we can estimate T_0 by seeing how many of the bins $|X(t,k)|$ have large amplitude.
- WHY: $\Delta\Omega = \frac{2\pi}{T} \ll \Omega_0$, so we can measure changes that are small relative to Ω_0.

Wideband Spectrogram

- Suppose we want to measure the spectrum $X(F)$ as a function of frequency (its timbre, or vowel quality); we want our measurement to be independent of the pitch period.
- This can be done by choosing $T \ll T_0$. For example, if $T = 0.5T_0$, then (not exactly, but approximately):

$$X(t,k) \approx X_{2k} + 0.5\left(X_{2k-1} + X_{2k+1}\right)$$

- Even if $x(t)$ is **not periodic at all**, $X(t,k)$ still tells us how much energy the signal has in the frequency band $\frac{k-0.5}{T} \le F \le \frac{k+0.5}{T}$.
- We can still measure the pitch period, T_0, by measuring the variation of $X(t,k)$ in the time domain:

$$X(t+T_0,k) = X(t,k)$$

Outline

Digital Spectrogram

A digital spectrogram is computed by first sampling the signal at F_s samples/second:

$$x[n] = x(n/F_s)$$

Then we compute the discrete time Fourier series:

$$X(t,k) = \frac{1}{N}\sum_{n=0}^{N-1} x[tF_s + n]e^{\frac{-j2\pi kn}{N}}$$

...and convert to decibels:

$$S(t,k) = 10\log_{10}|X(t,k)|^2$$

Discrete Time Fourier Series

The DTFS (discrete time Fourier series) is just like the CTFS (continuous time Fourier series) except that (1) we use a sum instead of an integral, (2) the number of frequency-domain samples is the same as the number of time-domain samples.

$$X_k = \frac{1}{N}\sum_{n=0}^{N-1} x[n]e^{-jk\omega_0 n}, \quad \omega_0 = \frac{2\pi}{N}$$

$$x[n] = \sum_{k=0}^{N-1} X_k e^{jk\omega_0 n}$$

Frequency Resolution of a Digital Spectrogram

$$\Delta f = \frac{1}{N} \; \frac{\text{cycles/sample}}{\text{frequency bin}}$$

$$\Delta \omega = \frac{2\pi}{N} \; \frac{\text{radians/sample}}{\text{frequency bin}}$$

$$\Delta F = \frac{F_s}{N} \; \frac{\text{Hertz}}{\text{frequency bin}}$$

$$\Delta \Omega = \frac{2\pi F_s}{N} \; \frac{\text{radians/second}}{\text{frequency bin}}$$