

Lecture 24: Pitch Tracking

Mark Hasegawa-Johnson

ECE 401: Signal and Image Analysis

- 1 Review: Autocorrelation and Covariance
- 2 Autocorrelation and power spectrum of a windowed complex exponential
- 3 Autocorrelation and power spectrum of a windowed cosine
- 4 Autocorrelation of a windowed complex tone
- 5 Pitch tracking
- 6 Summary

Outline

- 1 Review: Autocorrelation and Covariance
- 2 Autocorrelation and power spectrum of a windowed complex exponential
- 3 Autocorrelation and power spectrum of a windowed cosine
- 4 Autocorrelation of a windowed complex tone
- 5 Pitch tracking
- 6 Summary

Review: Autocorrelation and Covariance

- Autocorrelation:

$$r_x[n] = \sum_{m=-\infty}^{\infty} x[m]x^*[m-n]$$

- If the signal is stationary and zero-mean, then autocorrelation is a scaled estimate of covariance:

$$\rho_x[m, m-n] = E [(x[m] - \mu_x[m])(x[m-n] - \mu_x[m-n])^*]$$

- Autocorrelation is flipped convolution:

$$r_x[n] = x[n] * x^*[-n]$$

- Power spectrum is the DTFT of convolution:

$$\mathcal{F}\{r_x[n]\} = |X(\omega)|^2$$

Example: White Noise

- For example, **white noise** is a signal whose samples are completely uncorrelated with each other:

$$\begin{aligned} E[r_x[n]] &= E[x[m]x^*[m-n]] \\ &= \delta[n] \end{aligned}$$

- This is called white noise because its Fourier transform has equal energy at all frequencies, like white light:

$$\begin{aligned} E[|X(\omega)|^2] &= \mathcal{F}\{\delta[n]\} \\ &= 1 \end{aligned}$$

Outline

- 1 Review: Autocorrelation and Covariance
- 2 Autocorrelation and power spectrum of a windowed complex exponential
- 3 Autocorrelation and power spectrum of a windowed cosine
- 4 Autocorrelation of a windowed complex tone
- 5 Pitch tracking
- 6 Summary

Autocorrelation of a windowed complex exponential

Suppose that $x[n] = w[n]e^{j\omega_0 n}$. What is its autocorrelation?

$$r_x[n] = \sum_{m=-\infty}^{\infty} w[m]e^{j\omega_0 m} w^*[m-n]e^{-j\omega_0(m-n)}$$

$$\begin{aligned} r_x[n] &= e^{j\omega_0 n} \sum_{m=-\infty}^{\infty} w[m]w^*[m-n] \\ &= e^{j\omega_0 n} r_w[n] \end{aligned}$$

- The autocorrelation of a windowed complex exponential is a windowed complex exponential!
- The autocorrelation is windowed by $r_w[n]$, the autocorrelation of $w[n]$

Power spectrum of a windowed complex exponential

$$x[n] = w[n]e^{j\omega_0 n} \Rightarrow r_x[n] = e^{j\omega_0 n} r_w[n]$$

What is its power spectrum?

$$\begin{aligned} X(\omega) &= W(\omega - \omega_0) \\ |X(\omega)|^2 &= |W(\omega - \omega_0)|^2 \end{aligned}$$

Outline

- 1 Review: Autocorrelation and Covariance
- 2 Autocorrelation and power spectrum of a windowed complex exponential
- 3 Autocorrelation and power spectrum of a windowed cosine**
- 4 Autocorrelation of a windowed complex tone
- 5 Pitch tracking
- 6 Summary

Autocorrelation of a windowed cosine

- Suppose that $x[n] = w[n] \cos(\omega_0 n)$. What is its autocorrelation?

$$r_x[n] = \sum_{m=-\infty}^{\infty} w[m] \cos(\omega_0 m) w^*[m-n] \cos(\omega_0(m-n))$$

- WARNING: AUTOCORRELATION IS NOT LINEAR! For DTFT we used the relationship $\cos(\omega n) = \frac{1}{2}(e^{j\omega n} + e^{-j\omega n})$ to make the problem easy, but that won't work for autocorrelation, because it's not linear.
- ... but we can use a trig identity to help us:

$$\cos(a) \cos(b) = \frac{1}{2} (\cos(a+b) + \cos(a-b))$$

$$\sum_m w[m] w^*[m-n] \cos(\omega_0 m) \cos(\omega_0(m-n)) =$$

$$\sum_m w[m] w^*[m-n] \frac{1}{2} (\cos(\omega_0(2m-n)) + \cos(\omega_0 n))$$

Autocorrelation of a windowed cosine

$$x[n] = w[n] \cos(\omega_0 n)$$

$$\begin{aligned} r_x[n] &= \sum_{m=-\infty}^{\infty} w[m] \cos(\omega_0 m) w^*[m-n] \cos(\omega_0(m-n)) \\ &= \sum_m w[m] w^*[m-n] \frac{1}{2} (\cos(\omega_0(2m-n)) + \cos(\omega_0 n)) \\ &\approx \frac{1}{2} \cos(\omega_0 n) r_w[n] \end{aligned}$$

The last relationship is only approximate, because

$\sum_m w[m] w^*[m-n] \cos(\omega_0(2m-n))$ is not exactly zero, but the approximation becomes better as the window length gets longer.

Power spectrum of a windowed cosine

$$x[n] = w[n] \cos(\omega_0 n) \Rightarrow r_x[n] \approx \frac{1}{2} \cos(\omega_0 n) r_w[n]$$

$$X(\omega) = \frac{1}{2} W(\omega - \omega_0) + \frac{1}{2} W(\omega + \omega_0)$$

$$|X(\omega)|^2 \approx \frac{1}{4} |W(\omega - \omega_0)|^2 + \frac{1}{4} |W(\omega + \omega_0)|^2$$

Again, the second line is only approximate: It ignores the cross-terms involving the product of $W(\omega - \omega_0)W(\omega + \omega_0)$. We can usually say that the cross-terms are approximately zero.

Autocorrelation and power spectra of pure tones

- Autocorrelation of an exponential is an exponential:

$$\begin{aligned}x[n] = w[n]e^{j\omega_0 n} &\Rightarrow r_x[n] = r_w[n]e^{j\omega_0 n} \\ &\Rightarrow |X(\omega)|^2 = |W(\omega - \omega_0)|^2\end{aligned}$$

- Autocorrelation of a cosine is approximately a cosine:

$$\begin{aligned}x[n] &= w[n] \cos(\omega_0 n) \\ &\Rightarrow r_x[n] \approx \frac{1}{2} r_w[n] \cos(\omega_0 n) \\ &\Rightarrow |X(\omega)|^2 \approx \frac{1}{4} |W(\omega - \omega_0)|^2 + \frac{1}{4} |W(\omega + \omega_0)|^2\end{aligned}$$

Autocorrelation eliminates phase!

- What about $x[n] = w[n]e^{j(\omega_0 n + \theta)}$?

$$\begin{aligned} r_x[n] &= \sum_{m=-\infty}^{\infty} w[m]e^{j(\omega_0 m + \theta)} w^*[m-n]e^{-j(\omega_0(m-n) + \theta)} \\ &= e^{j\omega_0 n} r_w[n] \end{aligned}$$

- What about $x[n] = w[n] \cos(\omega_0 n + \theta)$?

$$\begin{aligned} r_x[n] &= \sum_{m=-\infty}^{\infty} w[m]w^*[m-n] \cos(\omega_0 m + \theta) \cos(\omega_0(m-n) + \theta) \\ &\approx \frac{1}{2} \cos(\omega_0 n) r_w[n] \end{aligned}$$

- Autocorrelation eliminates the phase!

Autocorrelation squares the magnitude!

- What about $x[n] = Aw[n]e^{j(\omega_0 n + \theta)}$?

$$\begin{aligned} r_x[n] &= \sum_{m=-\infty}^{\infty} Aw[m]e^{j(\omega_0 m + \theta)} A^* w^*[m - n]e^{-j(\omega_0(m-n) + \theta)} \\ &= |A|^2 e^{j\omega_0 n} r_w[n] \end{aligned}$$

- What about $x[n] = Aw[n] \cos(\omega_0 n + \theta)$?

$$\begin{aligned} r_x[n] &= \sum_m Aw[m]A^* w^*[m - n] \cos(\omega_0 m + \theta) \cos(\omega_0(m - n) + \theta) \\ &\approx \frac{|A|^2}{2} \cos(\omega_0 n) r_w[n] \end{aligned}$$

- Autocorrelation squares the magnitude, and eliminates phase!

Autocorrelation and power spectra of pure tones

- Autocorrelation of an exponential is an exponential:

$$\begin{aligned}x[n] = w[n]Ae^{j(\omega_0 n + \theta)} &\Rightarrow r_x[n] = r_w[n]|A|^2 e^{j\omega_0 n} \\ &\Rightarrow |X(\omega)|^2 = |A|^2 |W(\omega - \omega_0)|^2\end{aligned}$$

- Autocorrelation of a cosine is approximately a cosine:

$$\begin{aligned}x[n] &= w[n]A \cos(\omega_0 n + \theta) \\ \Rightarrow r_x[n] &\approx \frac{|A|^2}{2} r_w[n] \cos(\omega_0 n) \\ \Rightarrow |X(\omega)|^2 &\approx \frac{|A|^2}{4} |W(\omega - \omega_0)|^2 + \frac{|A|^2}{4} |W(\omega + \omega_0)|^2\end{aligned}$$

Outline

- 1 Review: Autocorrelation and Covariance
- 2 Autocorrelation and power spectrum of a windowed complex exponential
- 3 Autocorrelation and power spectrum of a windowed cosine
- 4 Autocorrelation of a windowed complex tone**
- 5 Pitch tracking
- 6 Summary

Complex Tone

- Remember that any signal with a period of N_0 , windowed by $w[n]$, can be written as

$$x[n] = \sum_{k=-\lfloor \frac{N_0-1}{2} \rfloor}^{\lfloor \frac{N_0-1}{2} \rfloor} X_k w[n] e^{j \frac{2\pi kn}{N_0}}$$

- To find $r_x[n]$ we have to be careful, because autocorrelation is not a linear operator.

Autocorrelation of a complex tone

$$\begin{aligned} r_x[n] &= \sum_{m=-\infty}^{\infty} x[m]x^*[m-n] \\ &= \sum_k \sum_{\ell} X_k X_{\ell}^* e^{j\frac{2\pi\ell n}{N_0}} \sum_{m=-\infty}^{\infty} w[m]w^*[m-n] e^{j\frac{2\pi(k-\ell)m}{N_0}} \\ &\approx \sum_k |X_k|^2 r_w[n] e^{j\frac{2\pi kn}{N_0}} \end{aligned}$$

Again, the last line is only approximate because

$\sum_{m=-\infty}^{\infty} w[m]w^*[m-n] e^{j\frac{2\pi(k-\ell)m}{N_0}}$ is not *exactly* zero for $k \neq \ell$, but it's pretty close.

Power spectrum of a complex tone

$$x[n] = \sum_k X_k w[n] e^{j \frac{2\pi kn}{N_0}} \Rightarrow X(\omega) = \sum_k X_k W\left(\omega - \frac{2\pi kn}{N_0}\right)$$

$$r_x[n] \approx \sum_k |X_k|^2 r_w[n] e^{j \frac{2\pi kn}{N_0}} \Rightarrow |X(\omega)|^2 \approx \sum_k |X_k|^2 \left| W\left(\omega - \frac{2\pi kn}{N_0}\right) \right|^2$$

Again, the last line is only approximate because it ignores the cross-terms caused by the tails of one harmonic multiplied by another harmonic.

Why is this useful? Answer: Autocorrelation eliminates phase!



$$x[n] = \sum_k X_k w[n] e^{j \frac{2\pi kn}{N_0}}$$

... where each coefficient has an arbitrary phase

$$X_k = |X_k| e^{j\theta_k}.$$



$$r_x[n] \approx \sum_k |X_k|^2 r_w[n] e^{j \frac{2\pi kn}{N_0}}$$

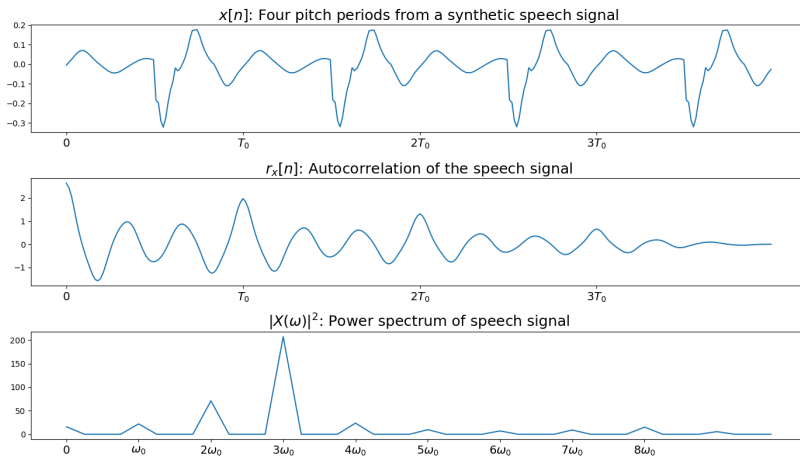
with coefficients that are all positive real numbers!

- An impulse train has positive real coefficients. Does this mean that $r_x[n]$ looks like an impulse train?

$$x[n] = \sum_{p=-\infty}^{\infty} \delta[n - pN_0] \Rightarrow X_k = \frac{2\pi}{N_0} \text{ for all } k$$

Autocorrelation eliminates phase

Here is an example speech waveform plus its autocorrelation and power spectrum:



Autocorrelation of complex tones

Autocorrelation squares the magnitude, and eliminates phase.

$$x[n] = \sum_k X_k w[n] e^{j \frac{2\pi kn}{N_0}}$$

$$r_x[n] \approx \sum_k |X_k|^2 r_w[n] e^{j \frac{2\pi kn}{N_0}}$$

$$|X(\omega)|^2 \approx \sum_k |X_k|^2 \left| w \left(\omega - \frac{2\pi kn}{N_0} \right) \right|^2$$

Try the quiz!

Go to the course webpage, and try the quiz!

Outline

- 1 Review: Autocorrelation and Covariance
- 2 Autocorrelation and power spectrum of a windowed complex exponential
- 3 Autocorrelation and power spectrum of a windowed cosine
- 4 Autocorrelation of a windowed complex tone
- 5 Pitch tracking**
- 6 Summary

The “Duplex theory of pitch perception”

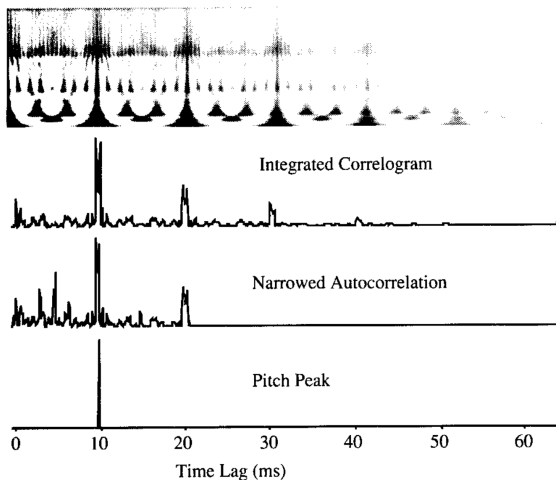
In 1951, Licklider proposed that our brains use autocorrelation to estimate pitch. Your ear first bandpass filters the speech in a series of different sub-bands, producing signals that we call $x(\tau)$ which are electromechanically transduced onto neurons. Then:

- ① One bank of neurons delays $x(\tau)$ to different delays, $x(\tau - T_0)$
- ② Another bank of neurons multiplies and integrates:

$$R_x(T_0) = \int_{-\infty}^t x(\tau)x(\tau - T_0)d\tau$$

- ③ The neuron with the highest autocorrelation is the “winner,” and determines what pitch frequency you hear

Here's a picture of Licklider's idea, as implemented by Slaney and Lyon (1990). They call the top plot a "correlogram:" The y-axis is sub-band center frequency, the x-axis is T_0 . The "integrated correlogram" integrates across sub-band center frequency to compute the signal autocorrelation.

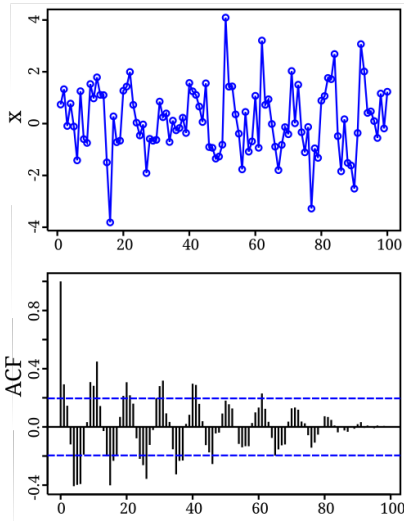


The SIFT pitch tracking algorithm

Markel (1972) proposed a pitch tracking algorithm called SIFT that was the basis of most pitch tracking algorithms since then:

- 1 Lowpass filter the signal to eliminate high-frequency noise
- 2 Compute the autocorrelation, and find its first peak
- 3 That's the pitch period!

Autocorrelation can even find the pitch of a sine wave embedded in lots of noise, like the one at right



https://commons.wikimedia.org/wiki/File:Acf_new.svg

Outline

- 1 Review: Autocorrelation and Covariance
- 2 Autocorrelation and power spectrum of a windowed complex exponential
- 3 Autocorrelation and power spectrum of a windowed cosine
- 4 Autocorrelation of a windowed complex tone
- 5 Pitch tracking
- 6 Summary**

Summary

- Complex exponential, $x[n] = Ae^{j(\omega_0 n + \theta)}$:

$$r_x[n] = |A|^2 e^{-j\omega_0 n} r_w[n] \Rightarrow |X(\omega)|^2 = |A|^2 |W(\omega - \omega_0)|^2$$

- Complex tone, $x[n] = \sum_k X_k w[n] e^{j\frac{2\pi kn}{N_0}}$:

$$r_x[n] \approx \sum_k |X_k|^2 r_w[n] e^{j\frac{2\pi kn}{N_0}}$$

$$\Rightarrow |X(\omega)|^2 \approx \sum_k |X_k|^2 \left| W\left(\omega - \frac{2\pi kn}{N_0}\right) \right|^2$$