EECS225d Audio Signal Processing in Humans and Machines **Spring 1995**

Lecture 19: 3 March 1995

Lecturer: Ben Gold *Scribe: Daishi Harada*

We continue our discussion of pitch perception.

19.1 Virtual Pitch Shifting

Consider the triad formed by the frequencies $800+20n$ Hz, $1000+20n$ Hz, and $1200+20n$ Hz for $0 \leq n \leq 10$. We find that this causes a single pitch to be perceived, of frequencies $\{200, 204, \ldots, 241\}$ Hz, respectively, and hence we might conclude that the ear computes the "best fit" fundamental given that the triad consists of consecutive harmonics, under some error metric. Let us illustrate this with the following simplified example.

> **Example 19.1** We shall compute the predicted perceived pitch when the ear is presented with the frequencies 220Hz and 420 Hz simultaneously, using absolute error as our metric of fitness. If we suppose that it is immediately apparent that we wish to consider the two frequencies as approximations of the first and second harmonics of some pitch, then we find that we wish to minimize $|220 - x| + |420 - 2x|$. The solution to this is $x = 210$.

If this phenomena is tested on pitch perception algorithms based on one of periodicity or place, we find that both behave qualitatively correctly, but with a certain amount of bias.

19.2 Comb Filtered Noise

Consider the output of a comb filter whose input is white noise. Recollect that the frequency response of a basic comb filter can be either $a\sin(b\omega)$ or $a\cos(b\omega)$ depending on its configuration. Note that maximum response occurs at the frequencies $(\pi/b)(n+1/2)$ and $(\pi/b)n$ for the two basic types, which corrsponds to peaking at odd numbered harmonics and all harmonics, respectively. A pitch is perceived by people when presented with this stimulus. We find that the pitch perception algorithms do not perform well in general. However, there appears to be a frequency for the fundamental (i.e., $\pi/2b$ or π/b for some b) below which the place algorithm performs acceptably, while the periodicity algorithm performs poorly. Above this frequency, the performance of the algorithms are reversed.

19.3 Ambiguous Pitches

Consider a signal $z(x) = \sum_{i=0}^{\infty} A_i \sin(\omega_i t)$ where $A_i = \exp(-\frac{1}{\sigma^2}(\mu - \omega_i)^2)$ and $\omega_i = 2^{i+x} f$ for fixed σ, μ, and $f \ll \mu$. An example of what this signal may look like for fixed x is given in Figure 19.1. Suppose we let x vary continuously between 0 and 1, and repeat this process. Noting that $z(1) = \sum_{i=1}^{\infty} A_i \sin(\omega_i t) \approx z(0)$ since A_0 small, the transition from $x = 1$ to $x = 0$ will be perceived as continuous. In fact, this stimulus causes the perception of a continually rising pitch. This result is due to Risset, while the discretized version was presented by Shepard. Further analysis may be found in (Meddis and Hewitt, 1991).

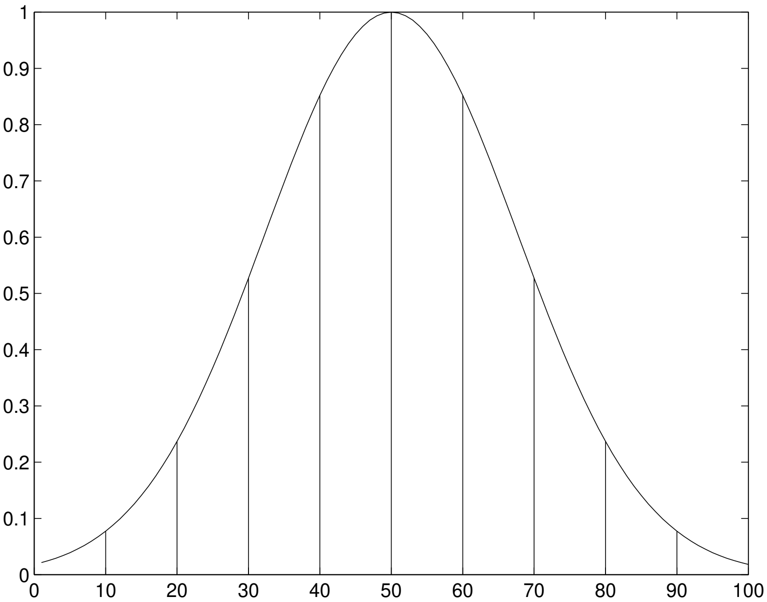

Figure 19.1: Amplitude/frequency plot for a signal causing pitch ambiguity.

There are other more simple ways in which ambiguity of perceived pitch for human beings can be achieved. For example, intervals of pure tones can be ambiguous: suppose subjects are presented with 199×9 Hz and 199×10 Hz simultanously. Then it was found that the perceived pitch varied from 145 Hz to 220 Hz.

References

Meddis and Hewitt (1991), Virtual Pitch and Phase Sensitivity, in *Journal of the Acoustical Society of America*, pp.2866–2882, Vol. 89(6), June.