Alec Goldfarb

Music Signal Processing

10/16/18

Def. Pitch: that attribute of auditory sensation in terms of which sounds may be ordered on a scale from low to high. Related to frequency of periodic sound, in terms of which sounds may be judged as dull (low) or sharp (high). It is phase insensitive, stable over a wide range of amplitudes and durations. In 'classical' music theory, fundamental frequencies are usually what we track in pitch space.

Place theory: some intellectual forebearers are Aristoxenos, Safi al-Din, Marin Mersenne, Galileo Galilei. Mersenne "First confirmed experimentally the laws of strings, according to which frequency varies inversely with the length of a string, proportionally to the square root of its tension, and inversely with the square root of its weight per unit length. This done, he stretched strings long enough to count the vibrations and, halving their lengths repeatedly, he derived the frequencies of every note of the scale." (10) DuVerny in 1693 was the first to offer a theory and model tonotopic projection to the brain based on a metaphor to steel string behavior. A theory to explain the superimposed vibrations of strings evolved throughout the 18/19th centuries, relying on concept of linearity (implying principle of superposition)- this is how pitch and period came to be mapped. Through Fourier, pitch has to be related to fundamental frequency- thus, pitch is an abstraction (many-to-one mapping). Helmholtz viewed the basilar membrane as a fourier transformer: "Sound is analyzed within the cochlea by the basilar membrane (BM) considered as a bank of radially taut strings, each loosely coupled to its neighbors. Resonant frequencies are distributed from high (base) to low (apex), and thus a sound undergoes a spectral analysis, each locus responding to partials that match its characteristic frequency." This theory is not accepted (but still used in certain areas) today because it does not account for missing fundamentals. Pattern Matching: we scan a waveform for a pattern, then match it to a template- "Templates are indexed by pitch, and the one that gives the best match indicates the pitch." Whatever pattern of 'partial pitches' (partials in the sound), the F0 template is recognized.

Here is possibly the fundamental contrast between time and place: Is it more reasonable to assume that the ear counts vibrations, or contains calibrated resonators?

Time theory: explains how rate/frequency is measured by a listener (the ear "counting" vibrations), so it just follows the temporal envelope of waveforms- period/interval between peaks = pitch. Ear seems devoid of 'counters', so this must happen in the brain. Can be traced to Anaxagoras, and the telephone theory of Rutherford: the ear just transmits vibrations to the brain like a telephone. Unclear because we're not sure if stead-

state nerve spikes can fire fast enough to track high frequency pitches, say something that's 9k: "Steady-state discharge rates in the auditory nerve are limited to about 300 spikes per second, but the pattern of instantaneous probability can carry time structure that can be measured up to 3 to 5 kHz in the cat (Johnson 1980). The limit is lower in the guinea pig, higher in the barn owl (9 kHz, Ko¨ppl 1997), and unknown in humans." (23) Auto-correlation is a theoretical self comparison when the brain slides a waveform against a copy of itself= "shift and compare"- where's maximum self-similarity?