Jonathan Mooser
Paper Review:
HUMAN PERCEPTION AND COMPUTER EXTRACTION OF MUSICAL BEAT STRENGTH
George Tzanetakis, et al.
2/3/2005

In *Human Perception and Computer Extraction Of Musical Beat Strength,* George Tzantetakis investigates the musical concept of beat strength as a qualitative property of music. The research consists of two fundamental phases. The first is a psychological study based on the hypothesis that untrained subjects will generally agree on whether a given piece of music has a stronger or weaker beat. Having established that, the authors propose and test an algorithm that automatically measures relative beat strengths of audio signals. The results are fairly convincing and may ultimately be a valuable component of a larger music analysis system.

The human subject phase of the research involved asking 32 participants to listen to 50 recordings and rate the beat strength of each on a scale of one to five. A graph of the results showed a clear trend. There was general agreement amongst subjects as to which pieces of music had stronger or weaker beats and their observations corresponded well to the categories established by the researchers beforehand. It is interesting to note that the authors did not provide a definition of beat strength, either in its instructions to the subjects or as part of the paper. They argue that a formal definition is unnecessary. As long as humans can universally recognize it, beat strength is a valid way to characterize music.

The question now is whether a machine can be trained to automatically recognize beat strength. Building on some of their previously published research, the authors propose an algorithm based on the concept of a "Beat Histogram." The Beat Histogram takes an audio signal as input, uses a wavelet transform to break it up into frequency bands, then uses autocorrelation to find periodicities in the signal. The result is a histogram that gives the relative strength of different BPM tempos within the sample.

Beat strength is derived from the Beat Histogram in two ways. The first is to simply calculate the sum of all of the histogram values and the second is to find the ratio of the histogram's highest peak to its average value. Both of these methods return results that correlate fairly well with the observations of the human subjects, suggesting that either could reasonably used as a beat strength detection algorithm.

The authors further speculate that the cognitive recognition of beat strength is based on similar kinds of signal processing, implying that they may have uncovered its fundamental nature. I would be interested in seeing more research in this direction. One possible experiment would be to create noise at various frequencies and engineer the signal to have the properties that the program looks for. The result would not be musical, per se, but would be recognized by their algorithm as having a strong beat. If human subjects found those same sounds to have a strong beat, we could assert that the proposed method truly taps into the root of beat strength. If not, then real music must have some more subtle characteristics that provide humans with the experience of a strong beat.

Overall, the results are convincing. The human subject phase demonstrates that beat strength is a very real part of the music listening experience and the algorithms they test are good enough to be potentially useful for some applications. The results are far from perfect, but shows significant promise for future research.