

Haojun Wang
haojunwa@usc.edu

Week 5 Review

Review on A Positive-Evidence Model for Rhythmical Beat Induction

A Positive-Evidence Model for Rhythmical Beat Induction is a rule-based model to predict downbeats of music pieces with the assumption that the positive evidence (accent onsets) is primarily responsible for downbeat induction.

Downbeats are “strong beats” when people tap the hands or feet while listening to the music. Several rule-based models have already been proposed to induct the downbeat. (i) NEG model only considers the rest and unaccent onsets. (ii) POS model only consider onsets information. (iii) HYBRID model is a combination of above two. With the variation of weights, this model can be adjusted to achieve better downbeats prediction. (iv) The author proposes the NPOS model, which is the normalization form of POS model. It solves the problem of POS that the fastest clock will always be considered. More specifically, this model has two assumptions. First, positive evidences contain more information contributed to downbeat detection and prediction. Second, subjects tend to tap at a relatively static frequency regardless of the pattern presentation. Note that the original POS model is fundamentally flawed. It always picks the fastest clock as the downbeat. NPOS can be regarded as a refinement of POS that eliminates the effect of frequency of clocks.

The rules used in NPOS model as follows. First, an isolated onset is accented. Second, the second of two adjacent onsets is accented. Third, the first and last of three or more adjacent onsets are accented. I would be quite interested in the reason of these rules being picked. Seems it would not be only three rules for the use of prediction. But there is no formal proof for how many rules being used is able to ensure the completeness of the algorithm (maybe it is NP hard). For the clock setting part, only clocks of less than half the period are generated. And clocks with incompatible frequency of beats are discarded.

Scores are being calculated from the onsets information. In NPOS, only positive evidences are considered, with the normalization of clocks. Because the human rate of beat detection might be slightly different from the assumed clock rate, a modified model of NPOS, TNPOS, with the consideration of tactus, is also proposed.

Three data sets are tested on NPOS with comparison of NEG, POS, and HYBRID model. The first two sets are for beat assignment while the last one is for patten memorization and reproduction. In the first two data sets, NPOS shows the most accuracy. And for the last data set, NPOS achieves good performance with fewest categories. I wonder why HYBRID method is not the best way. Both rest and onsets contribute the information about downbeat at different level. Though HYBRID has more parameters to configure, but it should has a better performance. Maybe it is still because the parameters used in the tests are not optimal for HYBRID. I would be interested for a normalized HYBRID model, and investigate the performance about that model.

Another point is I am not quite pleased with the reference format used in this paper. A common way to reference is to use the number according to the reference section. The author always uses the descriptive phrase for reference, which is very confusing in my reading.