

Ananda Jacobs

ISE 599

March 11, 2004

Paper review – Miguel Ferrand, Peter Nelson, and Geraint Wiggins

- 1) *Memory and Melodic Density: a model for melody segmentation*
- 2) *Unsupervised learning of melodic segmentation: a memory-based approach*

1) The authors begin with a general concept about segmentation, asserting that the ability to partition a melody allows for structural description of a piece of music. It is, as they word it, a “pre-processing stage for other tasks such as pattern discovery or music search” (1). I found this concept insightful, as it is sometimes difficult to imagine what cognitive or perceptual stage comes before another. Having read the Lartillot paper on musical pattern discovery, it was not hard to conceptualize a system (computer or human) that first segments, then relates patterns.

The motivation for developing a segmentation model is well articulated. Pattern finding algorithms, with their computationally expensive steps, could benefit from a reduction of the initial search space (1).

The Local Boundary Detection Model (LBDM) uses the Gestalt laws of change and proximity to apply parameters to melodies. When a threshold is defined and these Gestalt-based rules are run, a boundary profile is calculated.

It was not clear what the authors meant by “melodic cohesion” in section 3. It is supposed to be proportional to the frequency of occurrence of a given interval, associated with the melody in question. How does this relate to perception, or “perceived changes in melodic density”? At any rate, the authors seem to do a fine job formalizing the notion of melodic density as a weighted sum of the intervallic contributions occurring over a period of time, as limited by a sliding memory window.

The Melodic Density Segmentation Model (MDSM) was compared to the LBDM on accuracy in detecting actual boundaries for patterns. LBDM generated more excessive boundaries, although the two were comparable.

I found it interesting to note that MDSM is robust over tempo changes. This is due in part to the memory window being defined by number of events, e.g. crotchet notes. I agree with the authors when they say that MDSM may be used successfully to reduce search space without eliminating pattern candidates—the LBDM, after all, included more false positives.

2) The second paper provides an extension of the first idea, that of melodic segmentation. The goal is worthwhile: to avoid the use of *a priori* musical knowledge when implementing the models of segmentation.

To represent melodic information, the authors use two criteria: pitch step (PS) and duration ratio (DR). These measure interval distance and the ratio between consecutive events, respectively.

Markov models were used to predict the locations of segment boundaries, as were Mixed-order Markov models. They determine the probabilities of sequences of symbols. (Not clear precisely how it does this, however.) The methodology involved entropy calculations.

I was relieved to see that the authors actually did use real human subjects for comparison purposes of segmentation models. Many of the studies I’ve been reading have only been based on cognitive or perceptual assumptions, without actually using humans in any part of the study. In the current study, listeners pushed buttons to demarcate where they believed the segments were.