Carley Tanoue
3/11/04
ISE 599

Weyde, Tillman. "Integrating Segmentation and Similarity in Melodic Analysis" and "Optimising Parameter Weights in Models for Melodic Segmentation."

In the first paper, an Integrated Segmentation and Similarity Model (ISSM) for melodic analysis is introduced. The ISSM is based on generating and rating interpretations to find the most adequate one using a neuro-fuzzy system to combine knowledge with learning from data. There are two essential aspects of melodic structure: segmentation – structural units such as perceptual groups or melodic motifs, the building blocks of melodic structure – and similarity – which determines motif relations. The ISSM combines segmentation and similarity relations in a model for the recognition of melodic structure with the goal of determining a structural interpretation of a melody by integrating findings from music theory and existing studies with system optimization by experimental data. For the model itself, only a limited context of up to 10 notes is used and some optimizations are employed. GTTM's grouping well-formedness rules also are applied to prevent the generation of implausible interpretations or filter them out before they are rated. The calculations of ratings of interpretation are done by a neural net based on fuzzy rules connected to an adaptive system. ISSM learns from examples of melodic interpretations as an interactive training scheme. From this, generated relative samples are created and by changing the weights in the neural net which corresponds to the fuzzy truth-values of fuzzy rules to create a meaningful interpretation. The most important fact that Weyde conveys poorly is that even the best interpretation found by backpropagation is not guaranteed to be global because finding the guaranteed optimum would require too many calculations and would require a "full Bayesian model." The conclusion that Weyde came to was that an integrated model for segmentation and structural interpretation based on similarity is necessary since they are interrelated.

Weyde's second paper focuses more on the segmentation aspect of his ISSM. He determines that the influential factors for segmentation are: 1) Motif length and duration – the maximum duration of perceptual groups lies in the range of 3 seconds with the lower bound is in the range of 250-500 ms; 2) Regularity – listeners prefer segmentations which are regular in respect to length and duration; 3) Proximity – important for segmentation as temporal distances introduce group boundaries; and, 4) Gestalt and Similarity – the direction of movement which gives a sense of continuity when it is constant and indicates a group boundary when it changes.

Weyde's first experiment used 25 sequences with varying Pitch, Loudness and direction varying between groups of two to groups of three that were played to eight music students who were asked if they preferred a grouping in motifs of two or three notes. The results from this confusing

experiment suggest that a linear model may be inadequate for modeling the combination of segmentation parameters, probably due to high variability and small data sets. Experiment 2 used the ISSM in the interpretation of 20 randomly generated sequences (with the sequences that were subjectively judged hard to interpret musically removed. If he went this far, why not use a default case? Something that is comparable with others?) divided into two sets of ten and then presented to six music students who were asked to describe an adequate segmentation at the lowest grouping level. One set was used to train the adaptive systems (ISSM and a linear model) and the other was used to test them. The results were difficult because the subjects' opinions differed largely and the system could not be trained successfully, an even better reason for using a default repertoire. I also do not understand why there would be errors on a training set unless he retested the training set after he "trained" the system. Also, his "compatible segmentations" section is very convoluted and confusing. He is not clear in describing how it was going to achieve compatibility. His conclusions were that to achieve training results, it was necessary to introduce some tolerance by the definition of compatible segmentations (?) and that the weights of the trained system are partly not in accordance with the literature probably due to the particular set of examples. Right, blame the music that you created.