Carley Tanoue
3-4-04
ISE 599

Kilian, Jurgen and Holger H. Hoos. "Voice Separation – A Local Optimisation Approach."

In this paper, Kilian and Hoos propose a new voice separation algorithm based on a stochastic local search method with a focus on finding a range of voice separations that can be seen as "reasonable solutions in the context of different types of score notation," not to find "the correct" voice separation. Kilian and Hoos claim that their approach is different from the traditional split point separation and the rule based approaches.

Their algorithm is based on the idea of slices which is calculated using note overlapping. The number of different voice separations for a single slice depends on the number of notes in the slice and on the maximum number of voices in the desired output of the separation algorithm meaning that the user can interactively tune the parameters voice separation process in order to obtain a desired output. The cost function, used for assessing and optimizing the quality of a voice separation, is a weighted sum of terms that penalize individual, undesirable features that include: 1) Pitch Distance Penalty, which increases with the interval size between two succeeding notes in a voice with the first note of a voice having a fixed penalty for starting a new voice; 2) Gap Distance Penalty, which is imposed if adding a note from the current slice to a voice introduces a rest and increases with the duration of the rest; 3) Chord Distance Penalty, which increases with the range of a chord, with the differences in durations of its notes, and with the distance between the respective onset times: Also included in this penalty is the Range Penalty – the pitch difference between the lowest and the highest note in the chord, the Duration Penalty – the relation between the shortest and longest note in the chord, and the Onset Time Penalty – the ratio of the difference between the onset times of the latest and the earliest note in the chord over the duration of the longest note in the chord; and 4) Overlap Distance Penalty, which increases with the amount of overlap of overlapping notes that are assigned to the same voice without combining them into a chord. With the cost function in place, a stochastic local search approach is implemented for finding a cost-optimized voice separation for each slice which is based on a randomized greedy choice method.

A couple of thoughts about this paper. In the implementation section, the authors noted that small overlaps, defined as having a time difference between the offset of the earlier note and the onset time of the later note being small compared to the durations of the two notes, are eliminated by shortening the duration of the earlier note. This could possibly lead to a major problem especially if it was implemented in the GUIDO NoteServer as projected as a future project. It could explain the "choppiness" of a midi-to-gmn conversion. Similarly, the authors noted that the onset time correction, resulting from minor differences in the onset times of two notes that result from imprecise playing and is therefore replaced with their average if the durational overlap of the two notes is large compared to the onset time distance and if the onset time distance is very small, forces their algorithm to combine the respective notes into

chords if they get assigned to the same voice and hence facilitates the recognition of chords. This could possibly lead to another major problem if implemented in the GUIDO NoteServer. It could explain the "elongation" of a midi-to-gmn conversion.

Also of worthy of commenting is the fact that this paper included many "opinionated" statements in which the authors declared the superiority of their algorithm to any other while creating excuses for poor performance with comments such as "We could not improve this result by changing parameter settings. It seems that when only considering the notes, without any knowledge about the composer and style of the piece, there is no reason why another separation should be preferred." Other offending comments included, "requires only minimal manual modifications in order to obtain a satisfying result," "our algorithm achieves good results," "Empirical tests (not reported here) suggested …", and "for which we obtained good empirical results." If they are allowed to use "good" as a clarifier, then I am good as well.