Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

An Interactive Computational System for the Exploration of Music Voice/Stream Segregation Processes

Andreas Katsiavalos,[1] Emilios Cambouropoulos[2]

School of Music Studies, Aristotle University of Thessaloniki, Greece
[1]akatsiav@mus.auth.gr, [2]emilios@mus.auth.gr

ABSTRACT

Background

In recent years a number of computational models have been proposed that attempt to separate polyphonic music into perceptually pertinent musical voices or, more generally, musical streams. Most models identify monophonic voices although there exist models that allow multi-note chordal streams. Computational models enable the calculation of the closeness of notes in terms of 'belongingness' to the same voice/stream based on a number of auditory streaming principles (Bregman), and, then, decide which notes belong to which voices/streams. The exact way these perceptual principles interact with each other in diverse musical textures has not yet been explored systematically.

Aims

The aim of this research is to explore the way fundamental auditory streaming principles (such as temporal continuity, pitch proximity, tonal fusion, onset synchronisation, pitch co-modulation) interact with each other in order to allow the organisation of musical input into musically 'meaningful' streams. We investigate voice/stream separation, not only in the 'clear' case of pure polyphonic contrapuntal music, but in the general case of any type of musical texture (e.g. homophony, counterpoint, pseudopolyphony, heterophony and all sorts of intermediate situations).

Method

A computational system is developed that accepts as input a musical surface represented as a symbolic note file, and outputs a piano-roll like representation depicting potential voices/streams. The user can change a set of variables that affect the relative prominence of each streaming principle giving, thus, rise to potentially different voice/stream structures. The colour of the connections between notes represents the relative streaming membership strength. For a certain setting of the model's parameters, the algorithm is tested against a small but diverse set of musical excerpts (consisting of contrasting cases of voicing/streaming) for which voices or streams have been manually annotated by a music expert (this set acts as ground truth).

Results

Preliminary qualitative results are encouraging as streaming output is close to the ground truth dataset. However, it is acknowledged that it is difficult to find one stable set of parameters that works equally well in all cases. We are currently working of improving the way streaming membership closeness is calculated. In any case, the interactive application itself is most interesting as a tool for musicians and researchers to explore musical streaming processes.

Conclusions

The proposed musical streaming computational model enables the study of voice/stream separation processes per se, and, at the same time, is a useful tool for the development of more sophisticated applications in Music Informatics (especially MIR) as streaming is a necessary precursor to many interesting musical tasks (e.g. pattern recognition and extraction, melody identification, and so on).

Keywords

Auditory streaming, voice separation, stream separation

Accompanying Files are found in the Proceedings CD-ROM, in directory /Multimedia/258