REAL-TIME SOUND SIMILARITY RESYNTHESIS BY FEATURES EXTRACTION ANALYSIS

This abstract presents path∼, a PureData external implementing a corpus-based, concatenative analysis-resynthesis engine. Sound is acquired and resynthesised by a similarity match between audio grains, using a set of audio features extracted offline for sound corpora and in real-time for live input.
The basic idea of this work is to segment the audio files, to perform a feature extraction analysis and to construct a k-d tree and a list of k-nearest neighbors for every grain, where every grain is represented as a set of points. In real-time, features are extracted from a live input and the nearest neighbor is found in the tree. Starting from this element and using its k-nn list, a set of grains is available for the resynthesis.
Two types of segmentations are possible; the first one is frame-based, i.e. given a frame, path∼ cuts out the samples every frame and performs the analysis; the second one is onset-based, i.e. an onset detection is made and this fragmentation cares the gestural nature of grains. A multi-descriptors strategy is adopted: the features are mel frequency cepstral coefficients, a spectral centroid and an amplitude descriptor. Using the default parameters, the dimensionality of the feature space, i.e. dimensions that every coordinate has in the feature space, is 16, 14 given by the MFCCs plus one of the spectral centroid and one of the amplitude descriptor. However, a fine tuning is possible: the mel spacing can be adjust to increase or decrease the number of coefficients and can be even deleted from all the calculations. By default, the amplitude descriptor is the amplitude root mean square; it can be changed in the envelope peak.
path∼ has one real-time audio thread and one worker thread, where all calculations are performed. Using an atomic compare-and-swap, it's possible to add or change source audio materials without stopping the resynthesis.
The main focus of this work is the real-time audio mosaicing. Audio mosaicing refers to the process of reconstructing the temporal evolution of a target sound from fragments taken from a source audio materials. Our interest is the application of this technique to the production of mixed contemporary composition. For this reason, we develop several functions and utilities. For example, a vectorized audio outputs control system, that allows to change, during the instantiation of the external, the number of outputs; a script-based preset management gives to the computer music designer the possibility to call different configurations and fix them in real-time thanks to a graphical editor. A debugger helps in finding errors and suggesting the possible solutions. All these tools give a great flexibility especially during rehearsals. Finally, path∼ can load audio files from hard disk and from Pd-s array, so analysis can be done even on on-the-fly recorded materials.
Actually, three new pieces involve path∼ in the electronic part.

1. Link

OSX and Linux binaries and some patch are available from github.