Real-Time Sound Similarity Reconstruction by Features Extraction Analysis

Marco Matteo Markidis José Miguel Fernández

@ Conservatorio di Trento

June 19, 2016

path∼

This presentation describes a real-time audio analysis and resynthesis environment written for compositional research and for the development of music pieces for acoustic instruments and electronics. This work, called path∼, is implemented as an external for Pure Data framework and it consists in an object that can be integrated in a Pd general real-time patch and used in combination with its internal and external objects. This environment is based on an ahead-of-time construction of a database and the related adjacency lists through a features extraction analysis of chosen samples, on a real-time features extraction and concatenative synthesis upon a research of similarity nearest neighbors on the inner database.

In Max/MSP:

In Pd:

In Max/MSP:

- CataRT [1]

In Pd:

In Max/MSP:

- CataRT [1]
- FTM plus MuBu libraries [2]

In Pd:

In Max/MSP:

- CataRT [1]
- FTM plus MuBu libraries [2]

In Pd:

- timbreID [3]

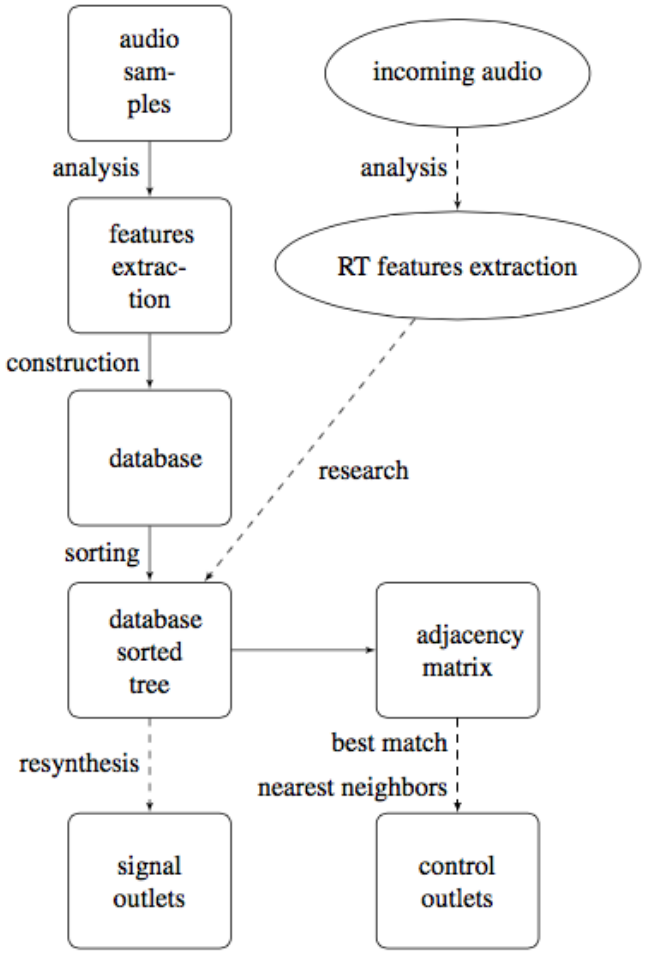

Figure: path∼ computation structure. Ovals and dashed lines refer to real-time computations, rectangles and solid lines to ahead-of-time computations.

Feature Extractions

- Mel Frequency Cepstrum is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency:

$$MFCC_i = \sum_{k=0}^{N-1} X_k cos[i(k + \frac{1}{2})\frac{\pi}{N}] \ . \qquad (1)$$

Feature Extractions

- Mel Frequency Cepstrum is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency:

$$MFCC_i = \sum_{k=0}^{N-1} X_k cos[i(k + \frac{1}{2})\frac{\pi}{N}] \ . \tag{1}$$

- Spectral centroid is defined as the center of mass of magnitude spectrum:

$$SC = \frac{\sum_{k=0}^{N/2} f(k)|X(k)|}{\sum_{k=0}^{N/2} |X(k)|} \ . \tag{2}$$

Feature Extractions

- Mel Frequency Cepstrum is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency:

$$MFCC_i = \sum_{k=0}^{N-1} X_k cos[i(k + \frac{1}{2})\frac{\pi}{N}] \ . \tag{1}$$

- Spectral centroid is defined as the center of mass of magnitude spectrum:

$$SC = \frac{\sum_{k=0}^{N/2} f(k)|X(k)|}{\sum_{k=0}^{N/2} |X(k)|} \ . \tag{2}$$

- Loudness is the root mean square amplitude.

Creation arguments

Name	Type	Default Value
k	int	64
mel spacing	int	250
window size	int	1024
out channels	int	2
virtual channels	int	64
random selection	bool	1
init file name	string	path∼.txt

Methods

Analysis

- analyze (samples)
- add (array)
- bang ()
- threshold (float)
- click ()
- weight (list)
- preset (int/string)

Resynthesis

- hopsize (int)[int,list]
- amp (float)[float,list]
- concatenate (int)
- nosound (bool)
- norandom (bool)
- envelope (int)

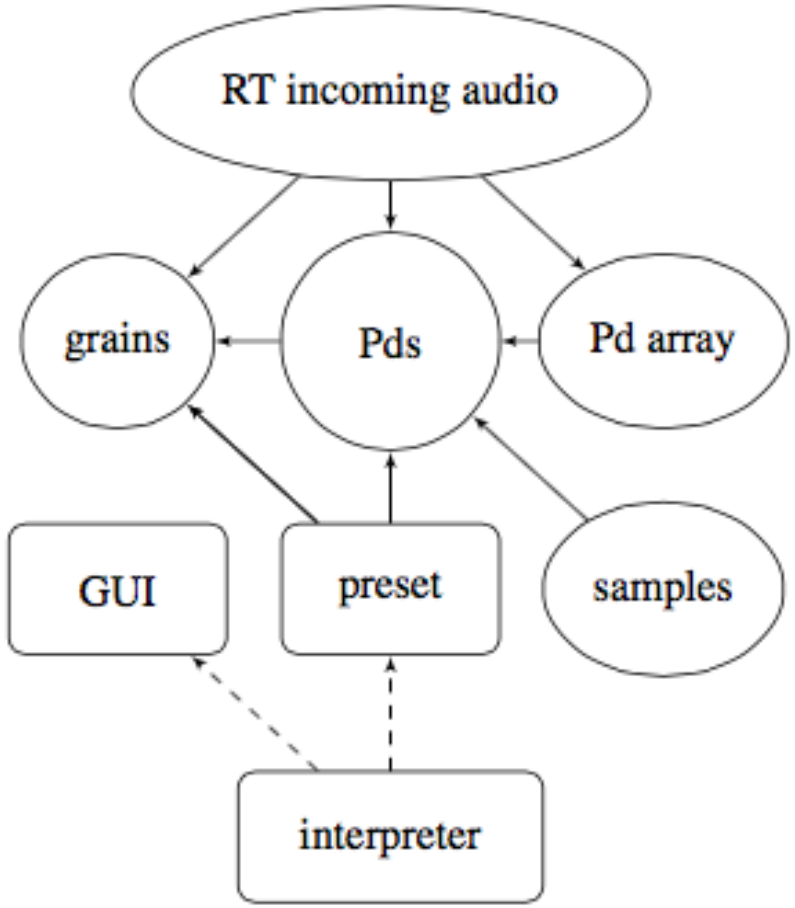

Figure: path∼ architecture. Rectangles and dashed lines are creation-time elements while ovals and solid lines stand for changeable elements.

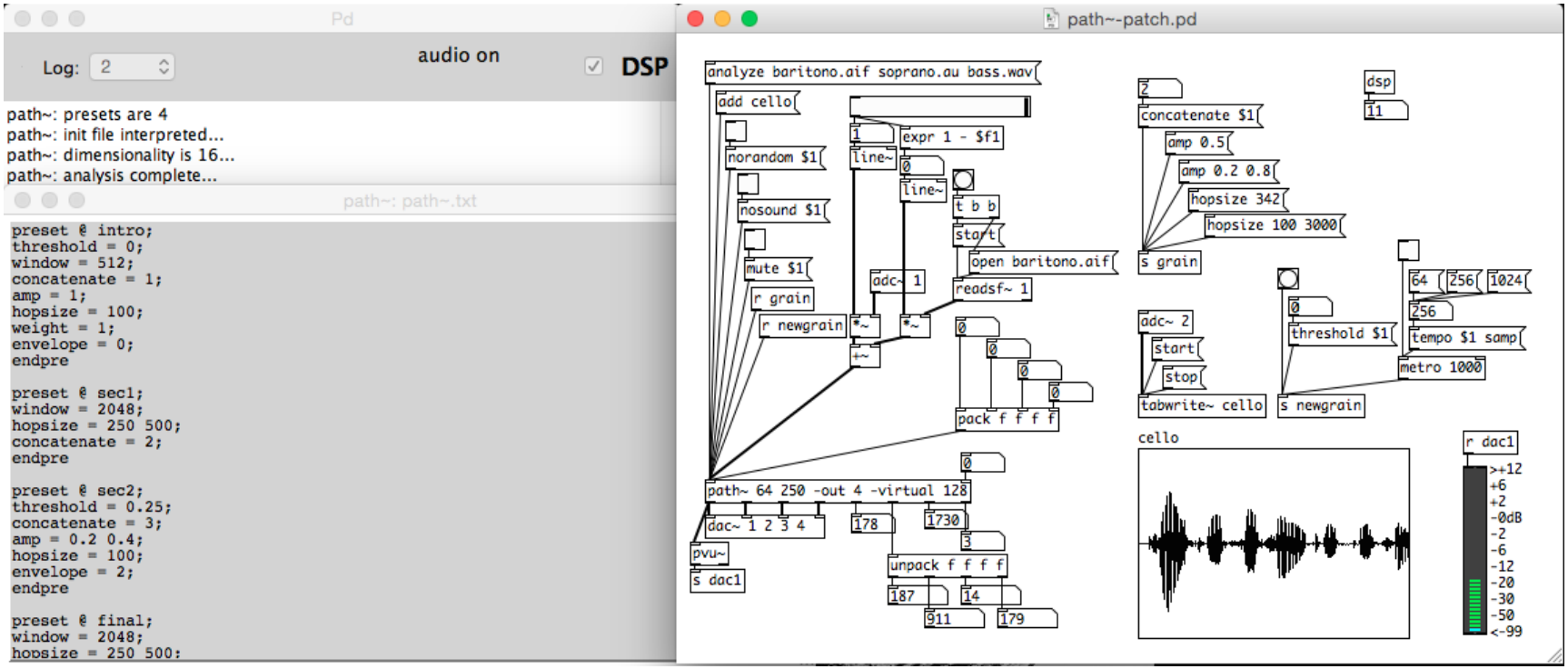

Figure: path~-patch.pd plus GUI within presets.

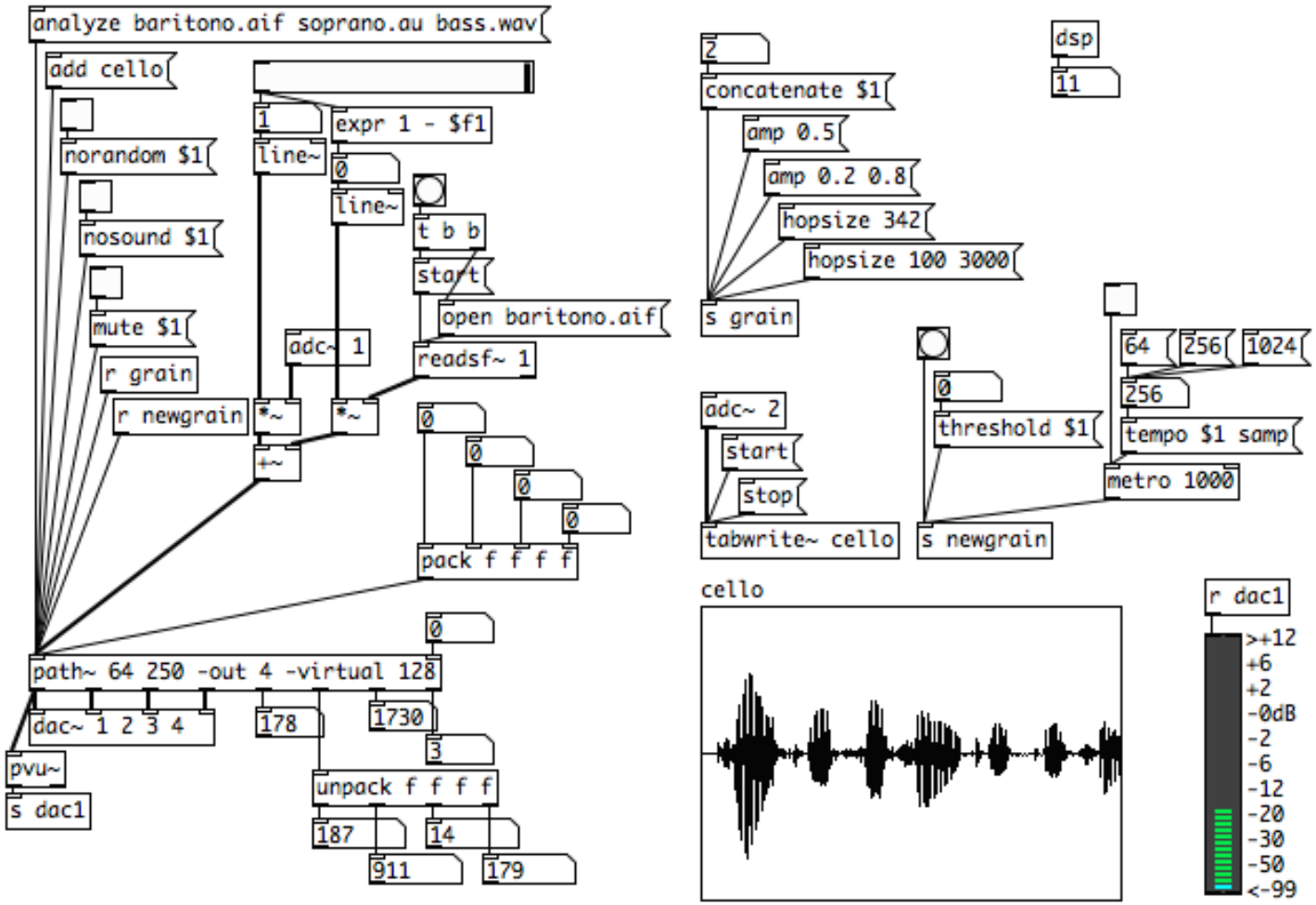

Figure: Detail of path~-patch.pd

- Introduction of new features and their dynamic choice

- Introduction of new features and their dynamic choice
- Implementation of new GUI functionalities

- Introduction of new features and their dynamic choice
- Implementation of new GUI functionalities
- Polyphonic input

- Introduction of new features and their dynamic choice
- Implementation of new GUI functionalities
- Polyphonic input
- Cluster analysis

- Introduction of new features and their dynamic choice
- Implementation of new GUI functionalities
- Polyphonic input
- Cluster analysis
- Better multi-thread support: lock-free queue, thread pool server, thread safe analysis...


D. Schwarz, G. Beller, B. Verbrugghe, and S. Britton.
Real-time corpus-based concatenative synthesis with Catart.
Proceedings of the International Conference on Digital Audio Effects (DAFx), 2006.


N. Schnell, M. A. S. Cifuentes, and J.-P. Lambert.
First steps in relaxed real-time typo-morphological audio analysis/synthesis.
Proceedings of an Inter- national Conference on Sound and Music Computing (SMC), 2010.


W. Brent.
Cepstral analysis tools for percussive timbre identification.
Proceedings of International Pd Conference (PdCon09), 2009.