Master's thesis proposal

Nicholas Esterer

Grouping of partial trajectories for additive synthesis:

One model of audio is the sum of sinusoids plus noise model, here referred to as the additive synthesis model. The additive synthesis model is convenient because it is efficient (the waveforms can be synthesized using table look−up) and allows for easy and intuitive manipulation of the sinusoidal components (it is straightforward to change the amplitude and frequency of a sinusoid) and as opposed to STFT synthesis, no phase update schemes are required. The estimation of the sinusoidal trajectories is problematic in that, in their estimation, we are limited to point representations of their values in time, commonly referred to as partials in the context of audio. This means algorithms have to be devised that connect or group these points, ideally taking into consideration the case where an observation point occludes another. Once partial trajectories have been determined, it may be desireable to group these trajectories according to different criteria, e.g., for source separation. For this we consider partial parameters common of those coming from the same source, e.g., their frequency and amplitude modulation, or the fundamental frequency to which they belong.

The computation of partial trajectories involves:

1. Accurate estimation of the partial parameters such as frequency, slope and their modulations.
2. Grouping partials occuring at the same time as coming from the same source.
3. Connecting partials across time as belonging to the same trajectory.

We will implement and compare a variety of modern techniques for (1) and evaluate their contribution to (2) and (3). Specifically, we will investigate how knowledge of the frequency slope and the harmonicity of partials can improve their grouping.

For partial parameters the techniques investigated will be:

1. Basis pursuit[5]
2. Derivative method[4]
3. Reassignment method[3]
4. Parabolic interpolation[2]
5. Adaptive Quasi−harmonic model[1]

For multiple F0 estimation, aiding in the grouping of partials, we will use the technique of Puckette and Doval, with a refinement involving a negated harmonic comb window.[6]

For (2) we will consider K features of each partial and attempt to cluster components based on their R principle components using a clustering technique such as a Gaussian mixture model (GMM).[7][8]

Subsections will be:

1. Principle components analysis description
2. Peak picking
3. Expectation Maximization and its application to GMM

For (3) we will use a technique previously applied to tracking multiple objects across video frames. It has been shown that high quality solutions to this problem can be obtained through linear programming. [9]

References:

- [1] Pantazis, Yannis, Olivier Rosec, and Yannis Stylianou. "Adaptive AM−FM signal decomposition with application to speech analysis." Audio, Speech, and Language Processing, IEEE Transactions on 19, no. 2 (2011): 290−300.

- [2] Serra, Xavier, and Julius Smith. "Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition." Computer Music Journal 14, no. 4 (1990): 12−24.
- [3] Röbel, Axel. "Estimating partial frequency and frequency slope using reassignment operators." In International Computer Music Conference, pp. 122−125. 2002.
- [4] Marchand, Sylvain, and Philippe Depalle. "Generalization of the derivative analysis method to non−stationary sinusoidal modeling." In Digital Audio Effects (DAFx) Conference, pp. 281−288. 2008.
- [5] Chen, Scott Shaobing, David L. Donoho, and Michael A. Saunders. "Atomic decomposition by basis pursuit." SIAM review 43, no. 1 (2001): 129−159.
- [6] Puckette, Miller S., Miller S. Puckette Ucsd, and Theodore Apel. "Real−time audio analysis tools for Pd and MSP." (1998).
- [7] Jolliffe, Ian. Principal component analysis. John Wiley & Sons, Ltd, 2002.
- [8] Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. "Maximum likelihood from incomplete data via the EM algorithm." Journal of the royal statistical society. Series B (methodological) (1977): 1−38.
- [9] Jiang, Hao, Sidney Fels, and James J. Little. "A linear programming approach for multiple object tracking." In Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, pp. 1−8. IEEE, 2007.