Single-Frame Discrimination of Non-Stationary Sinusoids

Jeremy Wells & Damian Murphy, Intelligent Systems – Audio Lab,
Department of Electronics, University of York, jjw100@ohm.york.ac.uk

Introduction

The 'sinusoidal-plus-residual' representation of audio for transformation of signals, such as the spectral modelling synthesis (SMS) system of Serra, is one that has many applications in both music analysis and processing. Whilst spectral modelling systems exist that can perform transformations and resynthesis from model data in real-time, the possibility of generating this data in real-time has received little attention. The motivation for the investigation of single-frame sinusoidal discriminators reported here is to explore the possibilities for real-time SMS systems.

Here we present two new methods for discrimination of non-stationary sinusoids within a single analysis frame. Both methods use data from time reassignment of Fourier transform data. These methods are then compared with an adapted version of an existing discrimination method, both in terms of their effectiveness and their computational cost.

Sinusoidal Discrimination Methods

These methods use estimates of *ΔA* and *Δf* obtained using a high-accuracy method, iterative *reassignment distortion analysis* (RDA), previously described by the authors. These estimates are obtained by fitting a second-order polynomial to Fourier time-reassignment data. Using the estimates for *ΔA* and *Δf* the behaviour of other measures obtained from the reassigned Fourier data is investigated to see whether it indicates behaviour expected for a sinusoid with those parameters.

Method 1: Reassignment 'Goodness of Fit'

Here the goodness of the fit of the time reassignment data to the second-order polynomial is tested against that expected. The method compares the variance of the data from the fitted curve and compares it against that expected for a sinusoid with the same *ΔA* and *Δf* estimates. Figure 1 shows how the variance changes with *ΔA* and *Δf*. Although the variance is not a function of *f* or *A* it is affected by the phase, which can be seen in the ridges in this figure.

Method 2: Phase and Amplitude Reassignment Difference

A second means of deriving reassignment measures, by differentiating with respect to amplitude instead of phase, has been proposed by Hainsworth et al. However these measures only agree for stationary sinusoids where non-Gaussian windows are used. Figure 2 shows how the difference in time-reassignment values obtained from each method varies with *ΔA* and *Δf*. Again, whilst this difference is insensitive to changes in *f* and *A*, differences in the phase causes the position of the small ridges to change.

Method 3: Non-stationary Cross Correlation

This is a highly effective but computationally intensive method proposed by Lagrange et al. which is included here for comparison with the two novel methods.

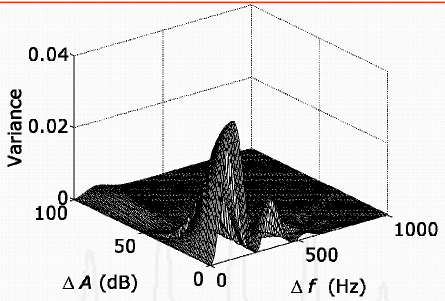

Figure 1: Variance as a function of ΔA and Δf for a single Hann-windowed sinusoid

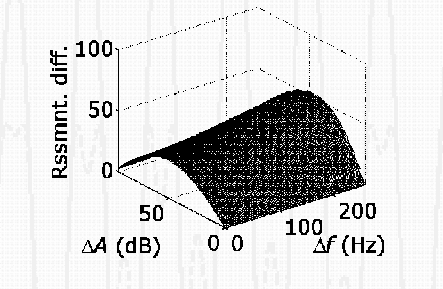

Figure 2: Modulus of difference between amplitude and phase measures of time-reassignment as a function of ΔA and Δf for a single Hann-windowed sinusoid.

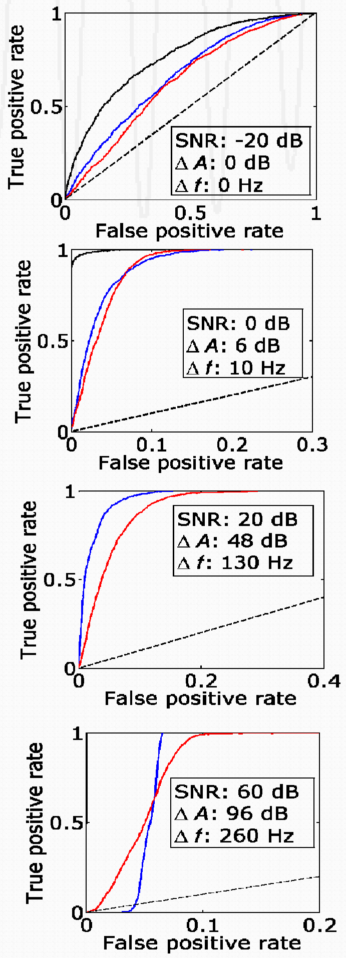

Figure 3: ROC curves for reassignment 'fit' (blue line), reassignment 'difference' (red) and correlation (solid black) discriminators. The dashed line shows the expected performance of a random classifier.

Comparison of methods

The performance of the discriminators, for sinusoids with a range of *ΔA* and *Δf* values combined with noise, are compared here using Receiver Operating Characteristics (ROC) graphs. These graphs offer a straightforward way of evaluating and comparing the performance of binary classifiers. A curve for a thresholding discriminator is produced by varying the threshold across a range of values that will produce a false positive rate (FPR) of 1.0 and a true positive rate (TPR) of 0.0 at one extreme of the range, and a TPR of 1.0 and an FPR of 0.0 at the other. A perfect classifier will produce a line which moves from (0.0, 0.0) to (0.0, 1.0) and from there to (1.0, 1.0). The closer a curve is to this perfect trajectory the better its performance for the signal and conditions for which it is being tested.

For each of the ROC figures presented here 1000 instances of a 1025 sample sinusoid have been generated with randomly chosen parameters. The resultant signal has been combined with white Gaussian noise, whose energy relative to that of the sinusoid is specified for each plot. The steps are as follows:

1. Perform zero-phase, 8x zero-padded FFT on 1025 point Hann-windowed signal.
2. Search the magnitude spectrum for peaks. Discard peaks where the reassigned frequency falls a fixed distance beyond the edges of the bin.
3. For each retained peak find the parameters using the high-accuracy iterative RDA technique.
4. Classify the retained peaks using the discriminator under test.
5. Repeat 4 for the entire range of threshold values and plot FPR and TPR rates as an ROC curve.

The resultant ROC curves are shown in Figure 3. A comparison of the computational cost per candidate peak is given in the following table:

classifier	arithmetic operations	table look-ups
reass. fit	35	4
reass. diff.	15	4
correlation	188 880	0

The correlation discriminator (Method 3) clearly performs best overall and perfectly in the bottom two plots. However its cost is extremely high. The performances of the two other discriminators are much closer to each other. Although inferior in terms of discrimination they are much cheaper to implement. Using the area under the curve as a measure of the overall effectiveness, the 'goodness of fit' measure can be seen to perform better than the 'difference' in all but the last plot in the figure, although its cost is higher.

Conclusions

Two new single-frame sinusoidal discriminators have been described and tested against an existing discriminator. The existing correlation method clearly performs best but is very expensive to implement. The two new discriminators are not nearly as robust but exhibit much better than random performance.

THE UNIVERSITY *of York*