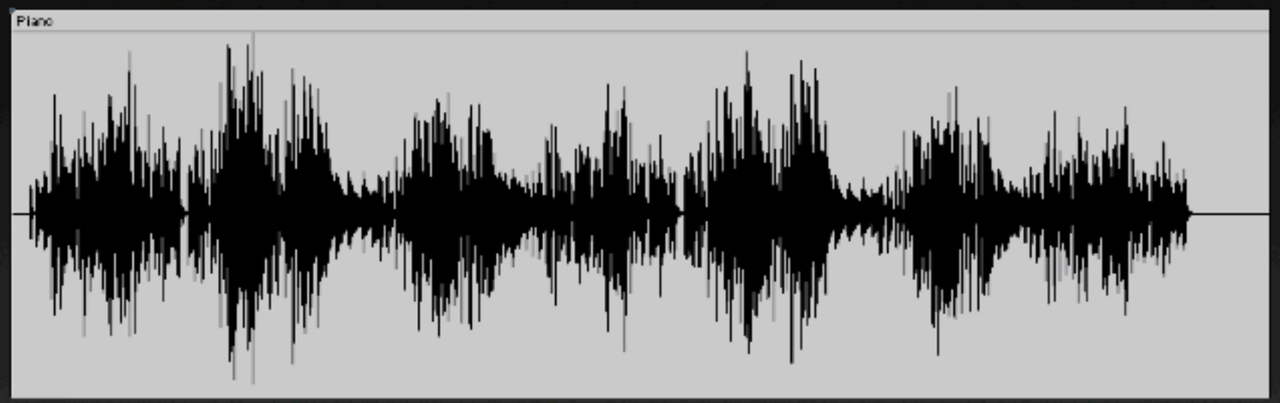
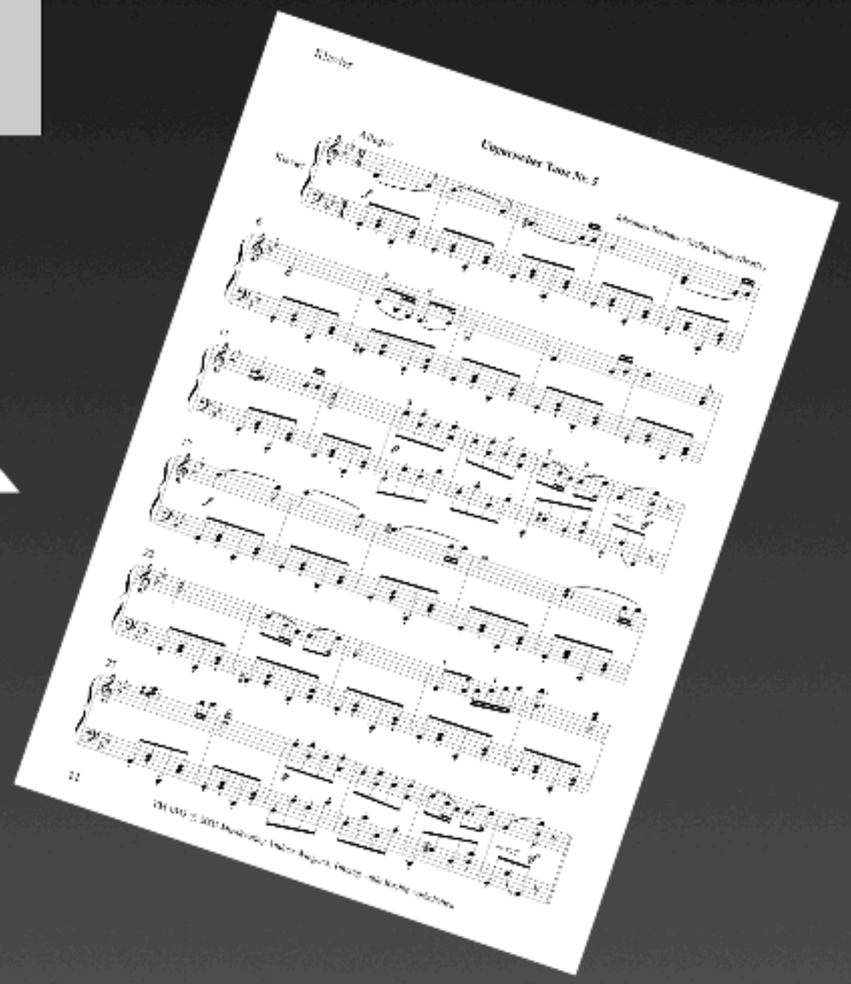
A Musical Approach to Monophonic Audio Transcription and Quantization

Jeff Hentschel

Audio Transcription

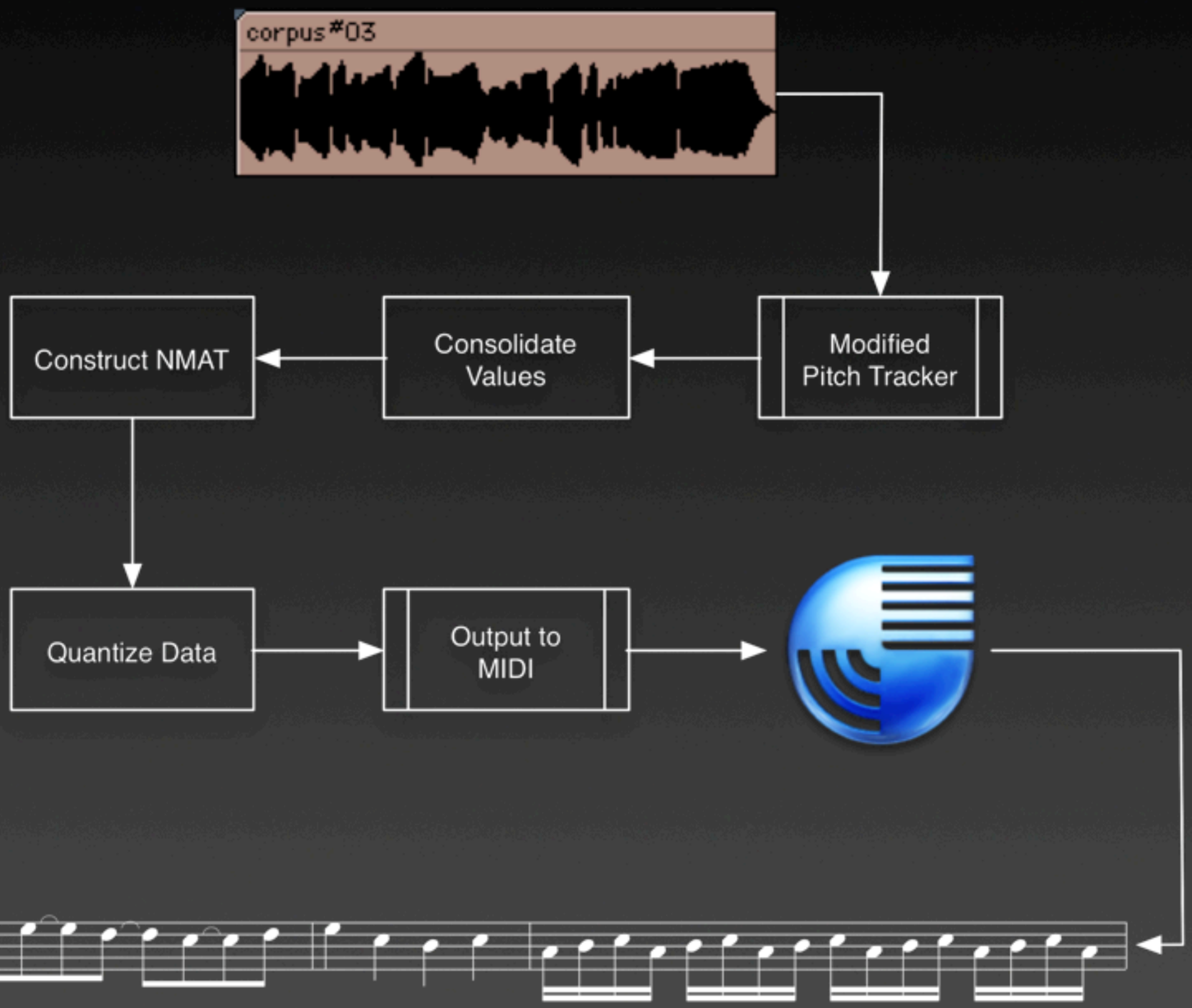




Applications

- Replay improvisation
- Analysis of performances
- For use in searching algorithms

System


corpus#03
Modified Pitch Tracker
Consolidate Values
Construct NMAT
Quantize Data
Output to MIDI

Pitch Tracking

- Based on Boersma's pitch tracking algorithm (praat_pd)
- Modified values for violin
- Frequency range (190-3500)
- Octave cost values
- VoicedUnvoicedCost

Consolidation

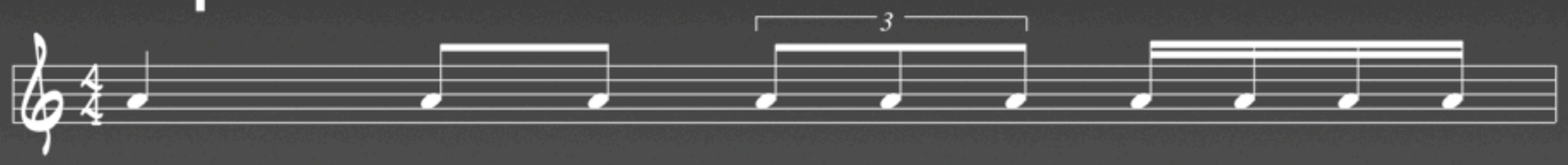
- Takes list of values from pitch tracker, finds notes, note duration, onset time
- Deletes probable extraneous notes (errors in pitch tracker)

$$\frac{60}{BPM} \cdot \frac{1}{tol - (tol \cdot tolVar)}$$

- Round velocities
- Delete end rests

Quantization

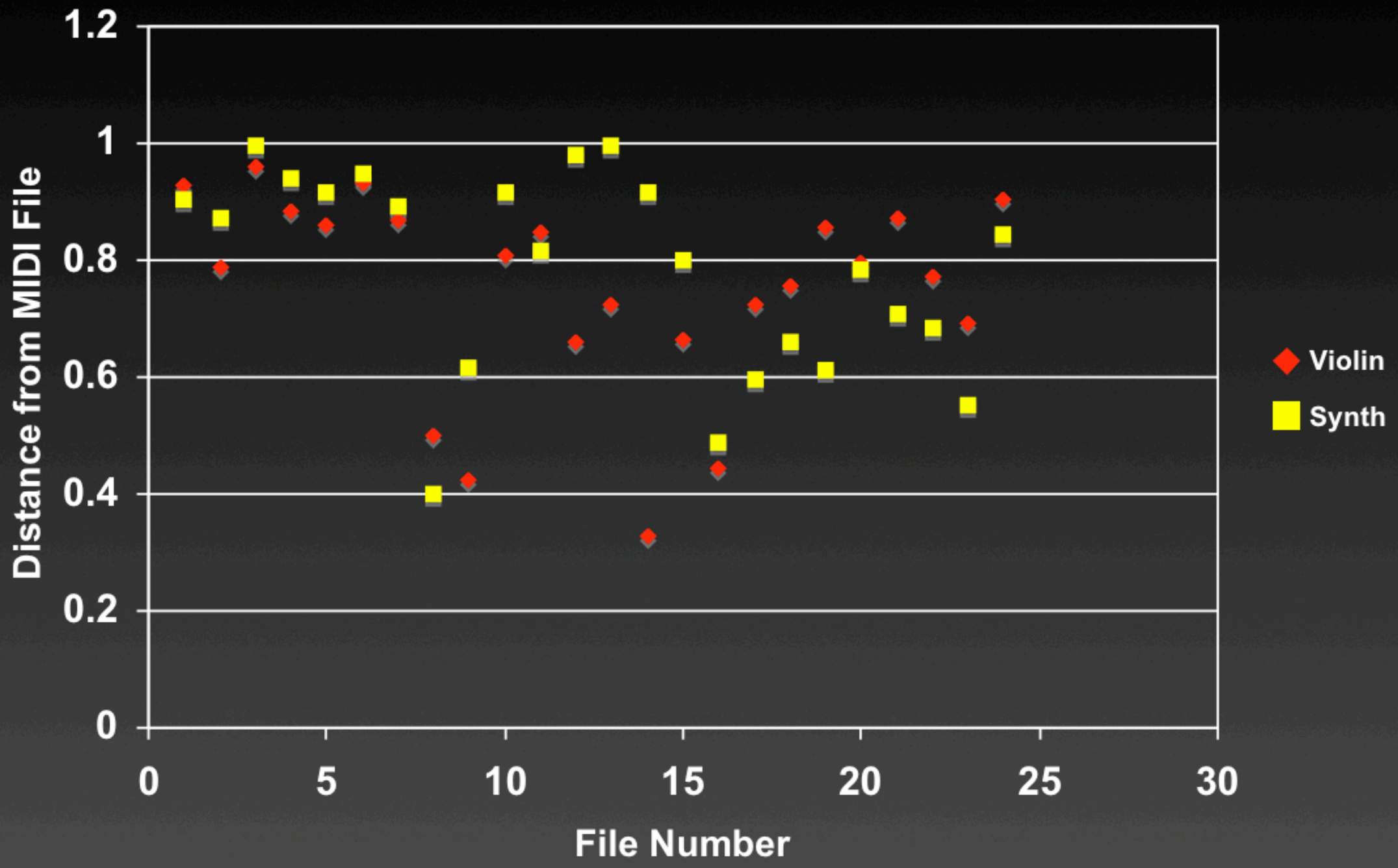
- Takes musical ideas into consideration
- Early note cost
- Long note cost
- Triplets groups of threes
- 4th, 8th, 12th, 16th notes
- Monophonic - 1 note at a time.
- Pickup modifier



Results

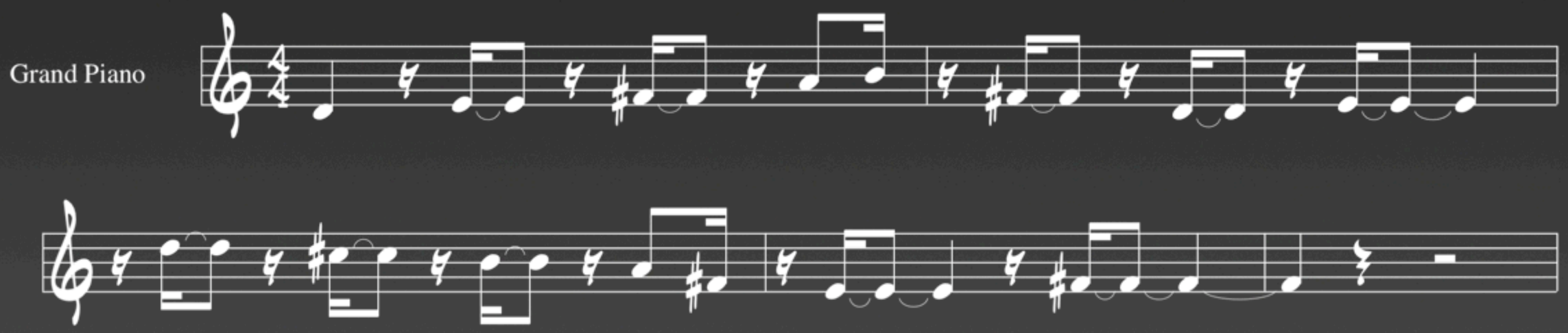
- Overall, fairly good results
- Average accuracy violin: .7497
- Average accuracy synth: .7850
- Range violin: (.3269-.9589)
- Range synth: (.4006-.9959)
- 83% over .60
- 71% over .70

Violin v. Synth Transcription Results



Error Sources

- Repeated note errors
- Octave errors
- Pizzicato (notes too short?)



Future Work

- Onset detection
- Adaptive Oscillators
- Polyphony