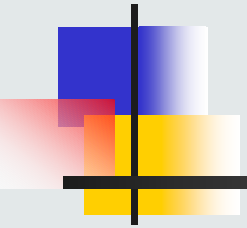
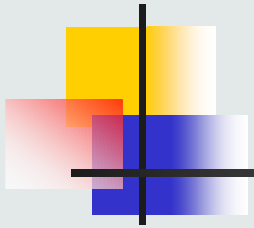


Automatic Transcription of Monophonic Audio Signals

Narciso Trevilatto Junior

Jayme Garcia Arnal Barbedo

Amauri Lopes



Context

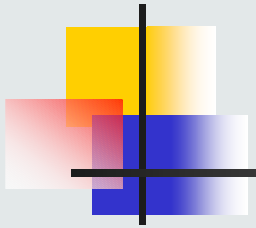
- Potentially Usefull for Musicians and other Professionals of Music
- Good Results for Monophonic Signals
- Treating Complex Signals is Still a Problem
- This Work: First Step of More Sophisticated Techniques

Fundamental Frequency Estimation

- Autocorrelation Method

$$f_0 = F_s / n_d$$

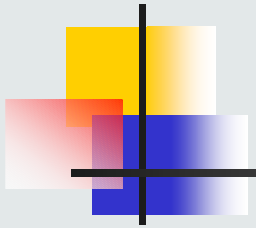
- *f0* Tracking
  - Window size of 50 ms
  - Hop size of 25 ms
  - Detects frequencies above 40 Hz



f0 Extraction

- Time Expansion to Eliminate Harmonics
- Peak Selection

Frequency and Duration Estimation

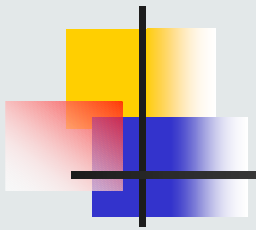


- MIDI number extraction
- Rounding of MIDI numbers
- Determination of temporal bounds of the notes

Results

Sound Source	**Number of Notes**	**Correct Detect**	**False Detect**	**Index *I***
Strings	507	484	45	0.87
Wind	1805	1712	93	0.90
Speech	492	463	69	0.80
Total	2804	2659	224	0.87

$$I = (CorrectNotes - FalseNotes) / TotalNotes$$



Conclusions and Future Work

- Good results for simple audio excerpts
- Do not take into account effects like *vibrato* and *glissando*
- Future improvements
  - Use of improved techniques for harmonic rejection
  - Incorporation of logics based on musical theory
  - Extension to complex signals