SIPL SIGNAL & IMAGE PROCESSING LAB
Department of Electrical Engineering, Technion

memories

Automatic Transcription of Polyphonic Piano Music

Gur Harary and Yoav Shechtman

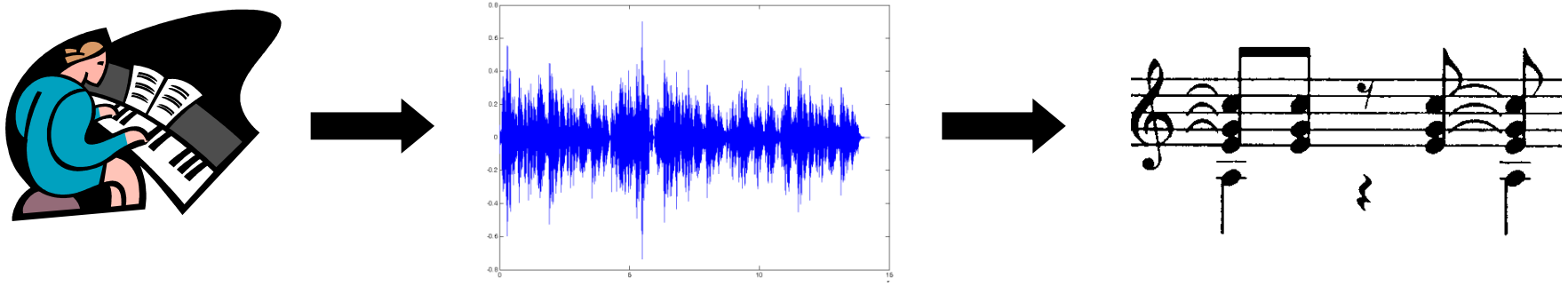
Supervisor – Zvika Ben-Haim

Winter 2006/7

August 2007

The Problem

- Given a polyphonic piano musical piece (wav file) – return the score (midi file).



Musical Background

- Note's pitch: A vibration with **fundamental frequency** f_0.
- Several **harmonics.** Ideally:

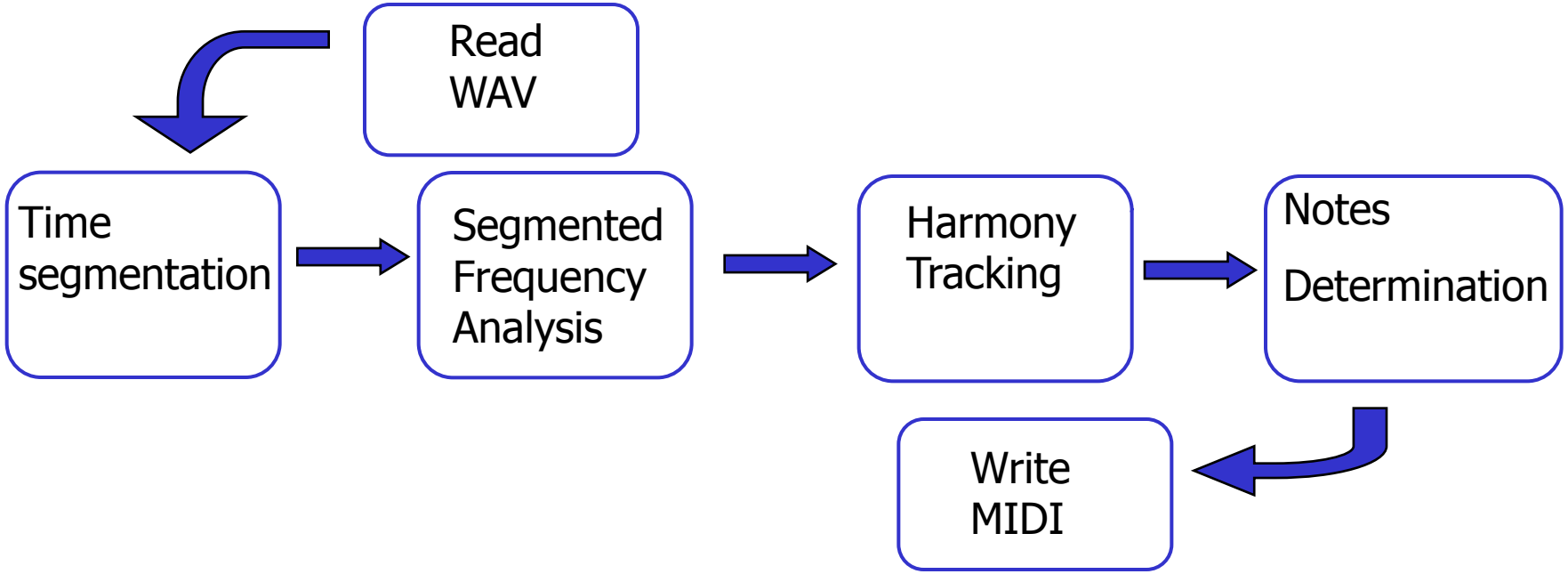
$$f_k = k \cdot f_0 \qquad k = 2, 3, 4...$$

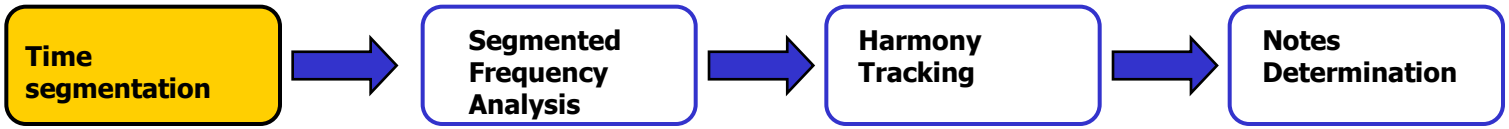
- Harmonic's amplitude distribution gives each instrument unique sound.
- Monophonic = One note maximum at a time
- Polyphonic = Possibly more than one note at a time

The Proposed Solution

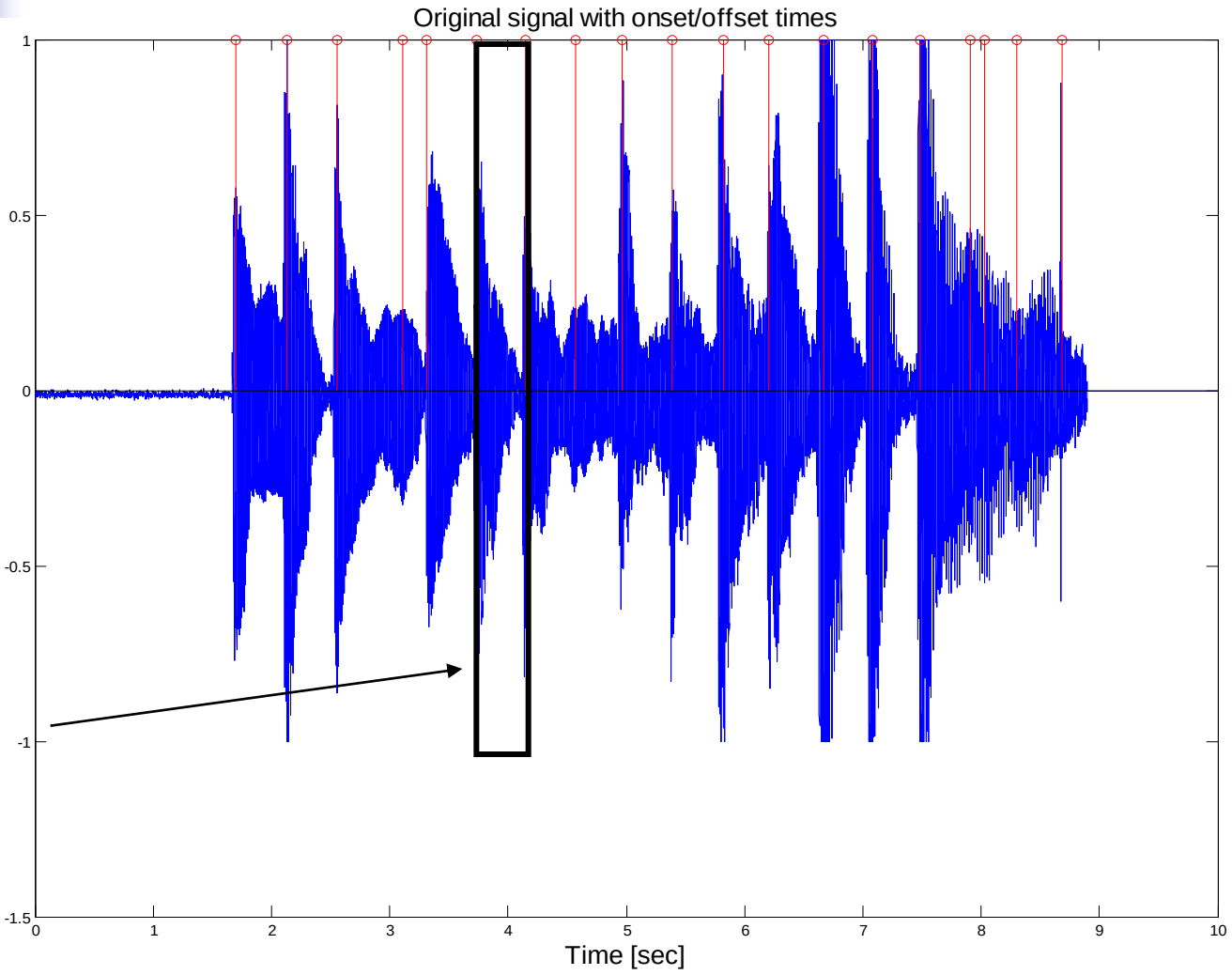
- Implementation based on :

A. Kobzantsev . Automatic Transcription of Polyphonic Music. Master Thesis . Technion , Electrical Engineering , February 2004. Supervised by Dr. Chazan Dan and Prof. Zeevi Yehoshua



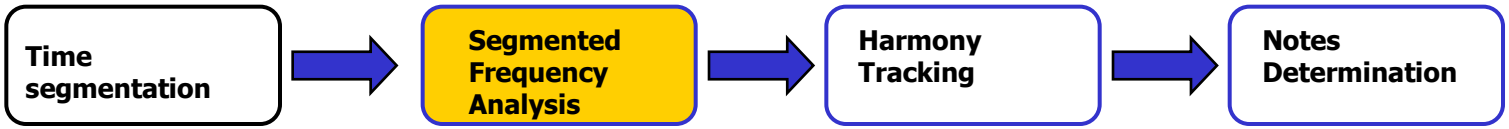


Final Time Segmentation



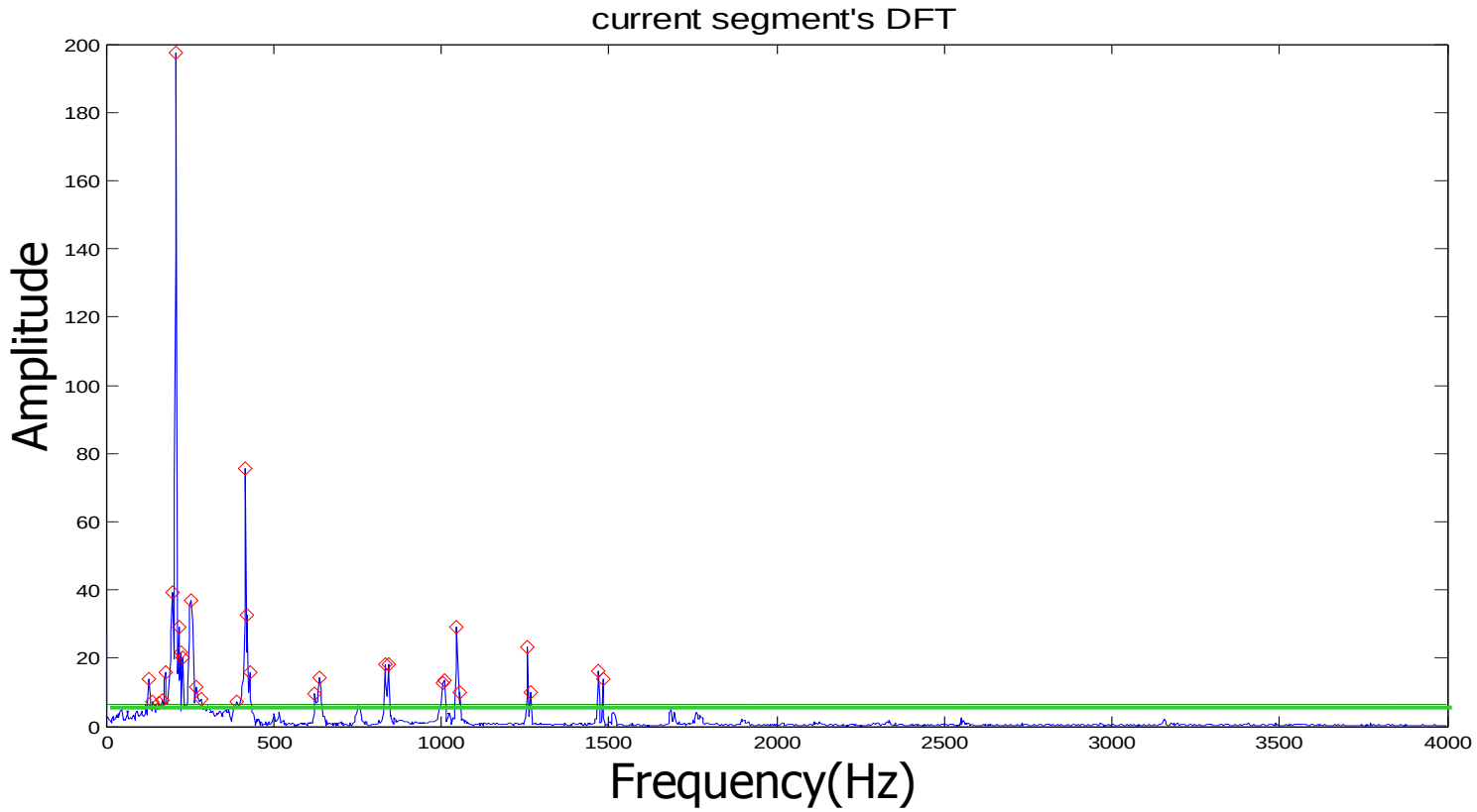
NEXT STEP:

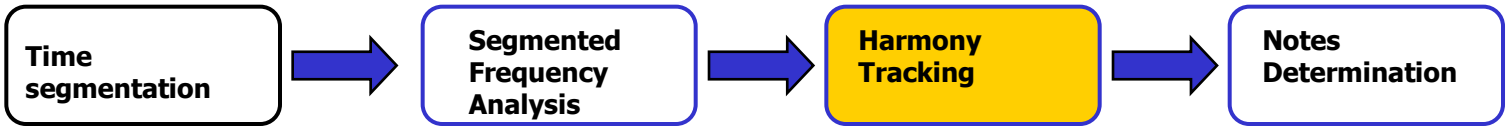
Analyze each segment



Example

- For segment no. 6:





The Inherent Polyphonic Problem

- Problem: One fundamental frequency could be another's harmonic. A common situation in western music, because – it sounds good!

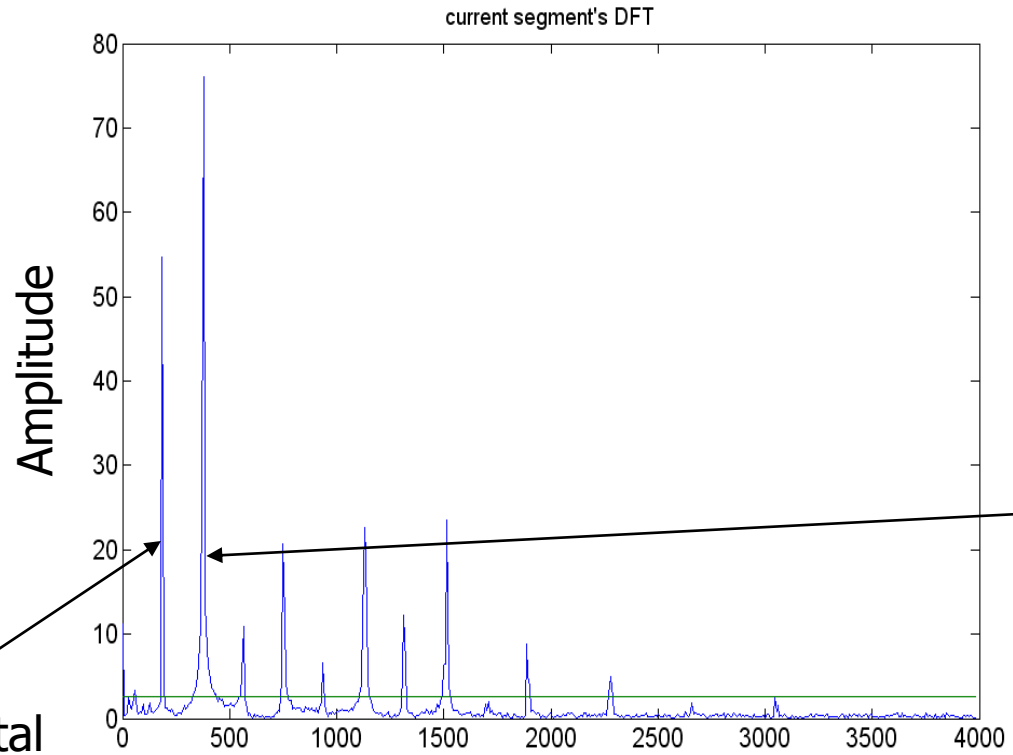
C2 Harmonics: 261.6 , 523.2, 784.8, 1046.5…

C3 Harmonics: 523.2, 1046.5,1569.6…

G4 Harmonics: 784, 1568, 2352…

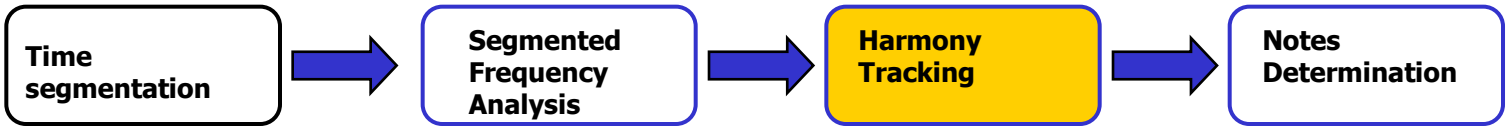
The Polyphonic Problem

- Example: G2+G3 played together



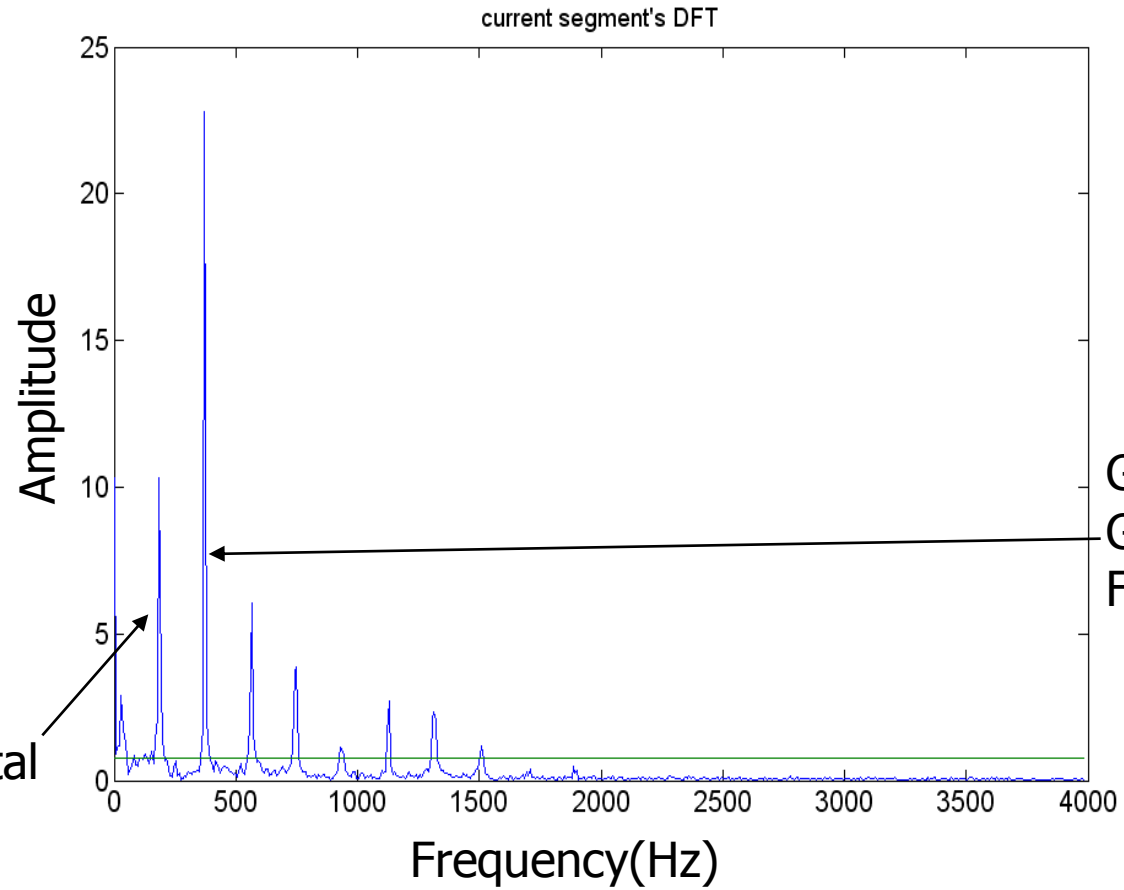
G2 Harmonic – AND G3 Fundamental Frequency

G2 Fundamental Frequency



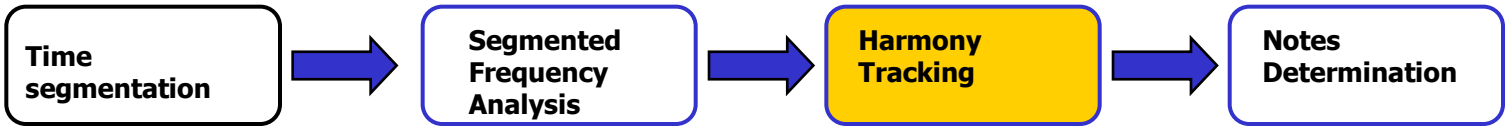
The Polyphonic Problem

- Example: G2 played alone



G2 Harmonic – AND G3 Fundamental Frequency

G2 Fundamental Frequency

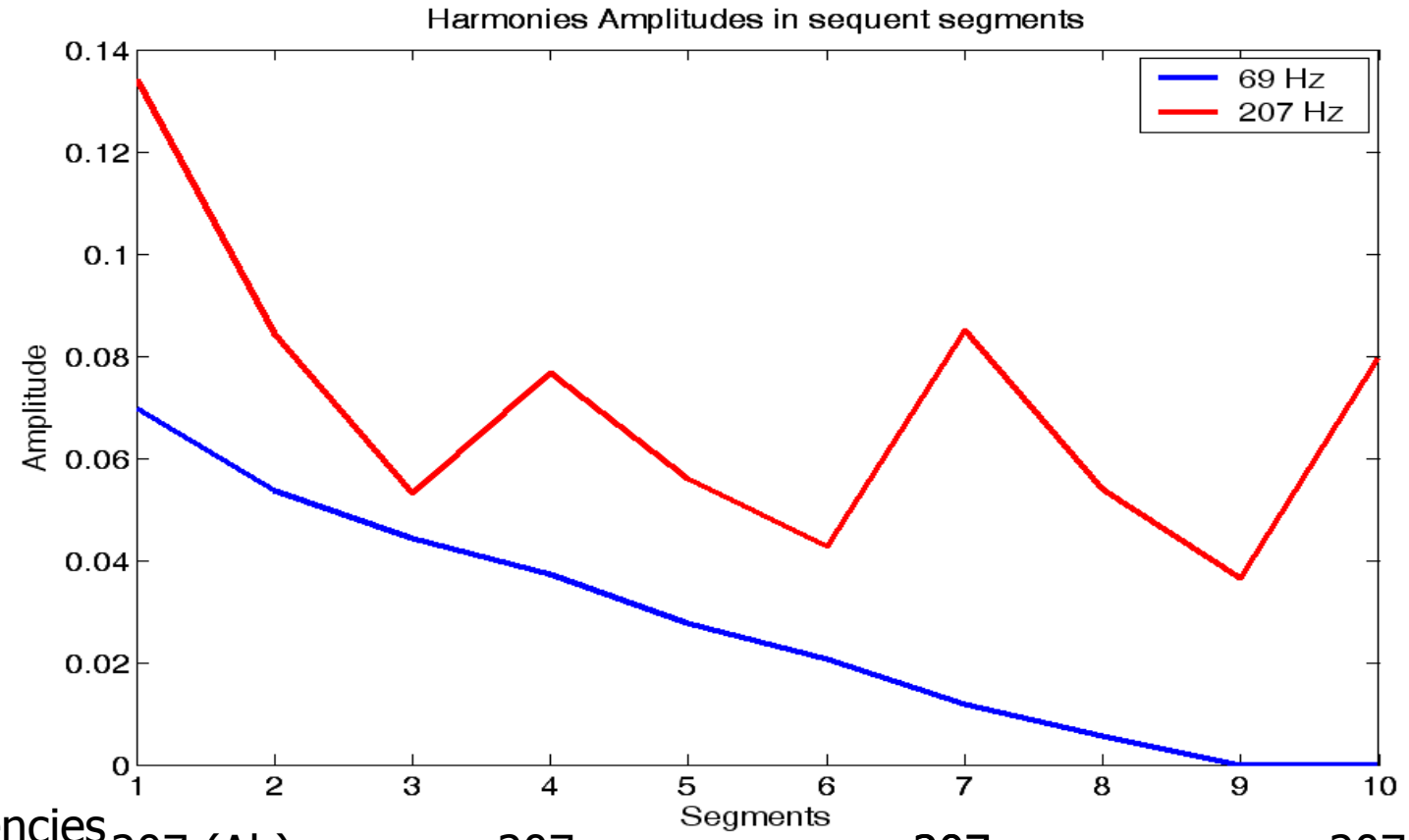


Harmony Tracking

The Basic Idea:

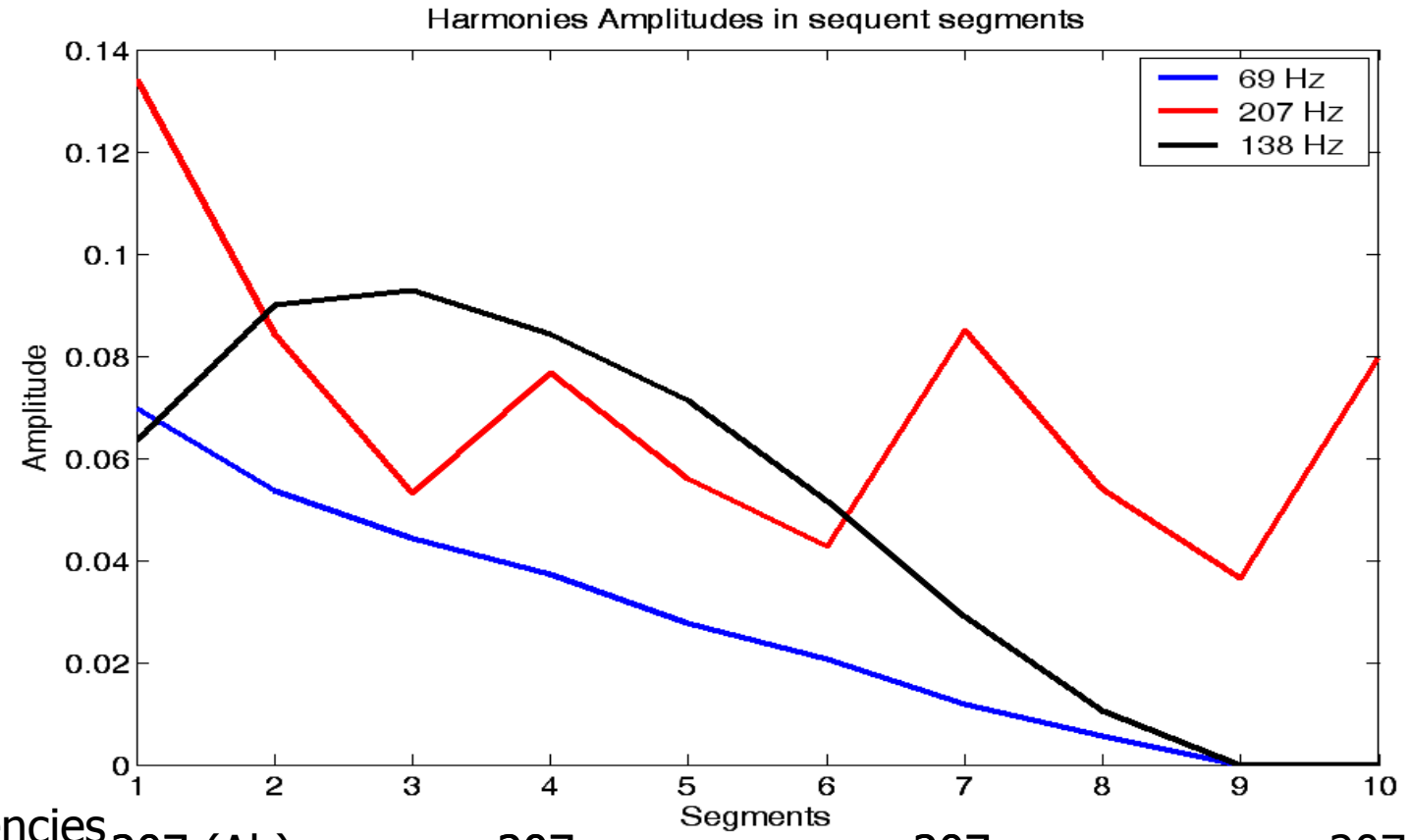
- Divide all frequency peaks in segment into 2 groups – increasing peaks (gained amplitude from last segment) and non-increasing peaks.
- Reason: A newly played note will have increasing peaks. A continuing note will have non-increasing peaks.

Harmony Behavior

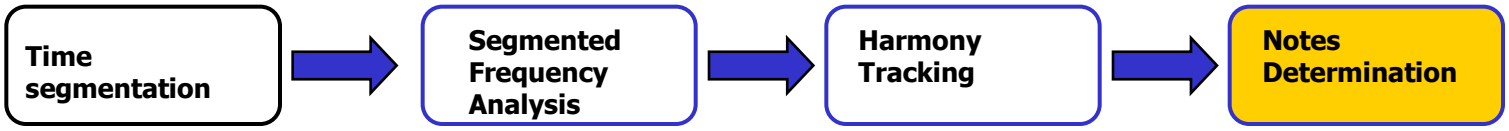


Basic frequencies actually played:				
	207 (Ab)	207	207	207
	69 (Db)	(Ab)	(Ab)	(Ab)

Harmony Bad Behavior



Basic frequencies actually played:				
	207 (Ab)	207	207	207
	69 (Db)	(Ab)	(Ab)	(Ab)



Notes Determination

- Newly played notes: Find the lowest grown frequency with over 2 harmonics – it's a new note.
- Lingering notes: If a previously played note still contains harmonics (above threshold) in this segment – it's still playing.

Comparison to market

Comparison with 2 market programs

- Intelliscore: By Innovative Music Systems Inc.

- Amazing MIDI: By Arakisoftware

Comparison to market

The Entertainer – Scott Joplin

- Original wav: 
- Intelliscore: 
- Amazing MIDI: 
- Our MIDI: 

Comparison to market

Fur Elise – Beethoven

- Original wav: 
- Intelliscore: 
- Amazing MIDI: 
- Our MIDI: 

Results

Musical Piece	Intelliscore	Amazing Midi	Gur and Yoav
Moonlight (15 sec)	84% Recognition of original notes 37% False notes		82% Recognition of original notes 18% False notes
Moonlight (30 sec)		94% Recognition of original notes 69% False notes	76% Recognition of original notes 21% False notes
Fur Elise (15 sec)	76% Recognition of original notes 26% False notes		85% Recognition of original notes 19% False notes
Fur Elise (30 sec)		95% Recognition of original notes 54% False notes	75% Recognition of original notes 22% False notes

False notes = Reconstructed notes that are not in the original piece

Summary and Conclusions

- Transcription of Polyphonic music is a difficult problem with no complete solution today. Partial solution achieved – e.g. lingering notes.
- Algorithm is not instrument dependant
- False octave detection – remains a problem. Future directions – focus on the attack of the notes.