

Melody Detection in Polyphonic Audio

Rui Pedro Paiva

October 20, 2007

Centro de Informática e Sistemas da
Universidade de Coimbra

Outline

- ❑ Introduction
- ❑ Overview
- ❑ Multi-Pitch Detection
- ❑ From Pitches to Notes
- ❑ Identification of Melodic Notes
- ❑ Evaluation Procedures
- ❑ Overall Results
- ❑ Conclusions and Future Work

Introduction

□ Motivation and Objectives

- **Goal**

- Extract a symbolic representation of melody from a polyphonic audio musical signal

- **Broad range of applications**

- MIR, music education, plagiarism detection, metadata

- **No general-purpose, robust, accurate solution developed so far**

□ Overall Approach

- **Focus on melody, no matter what other sources are present (figure-ground separation)**

Introduction

□ Melody Definition

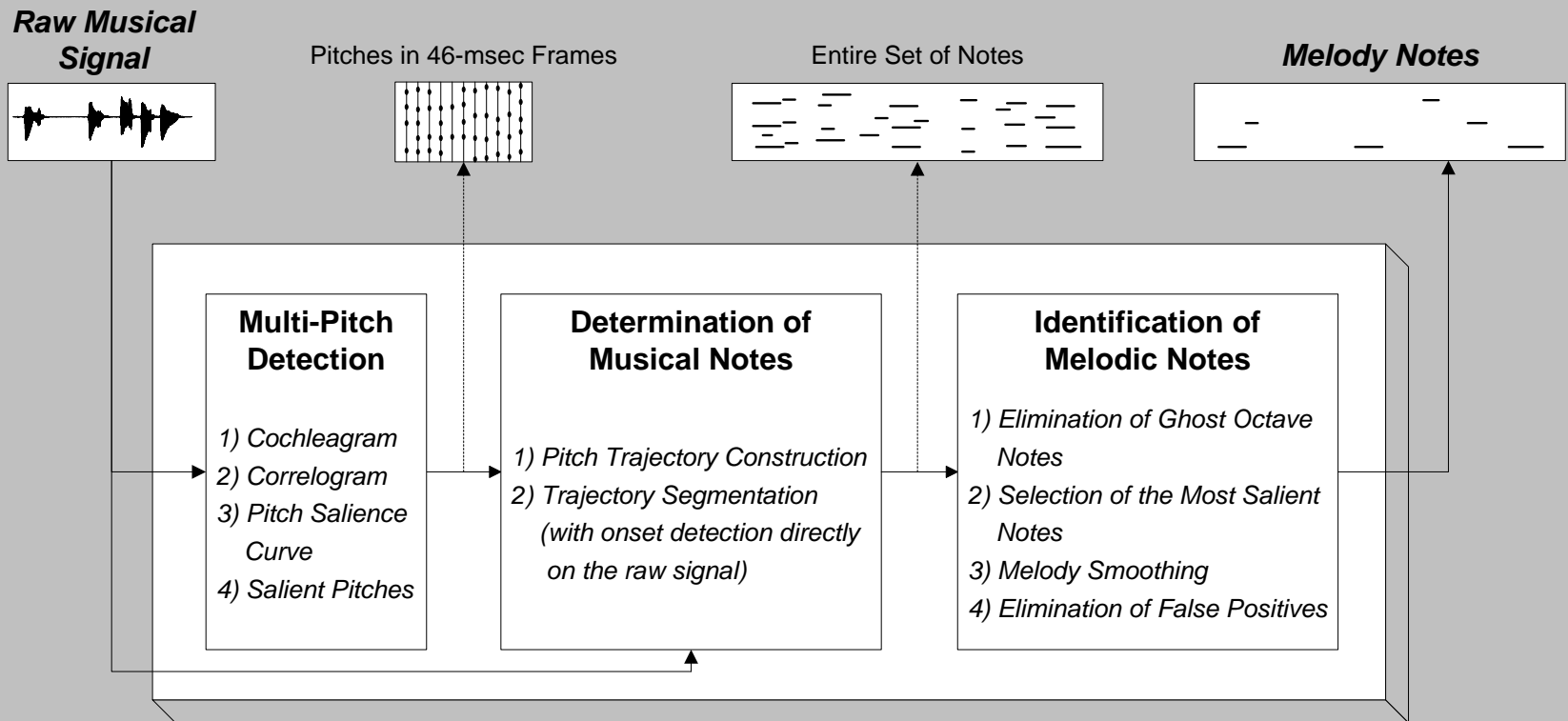
- **Subjective concept**

- Different people may define and perceive melody in different ways
- Proposals addressing cultural, perceptual, emotional or musicological facets of melody

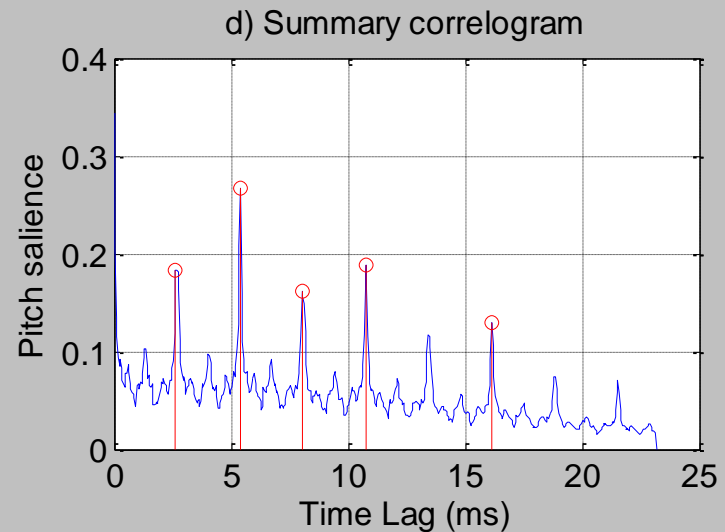
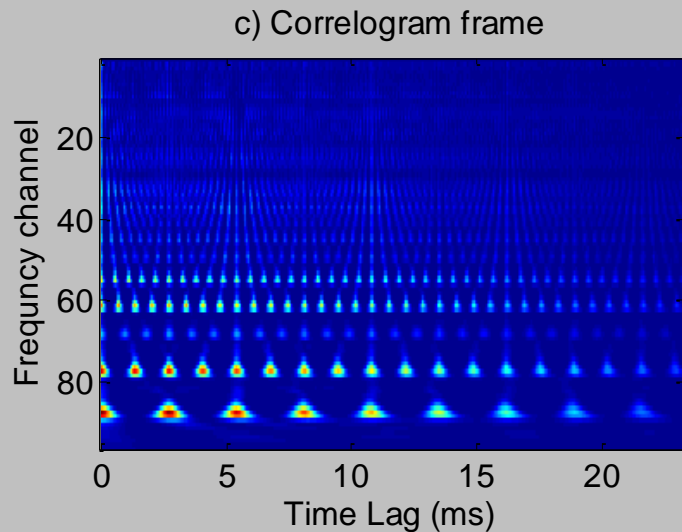
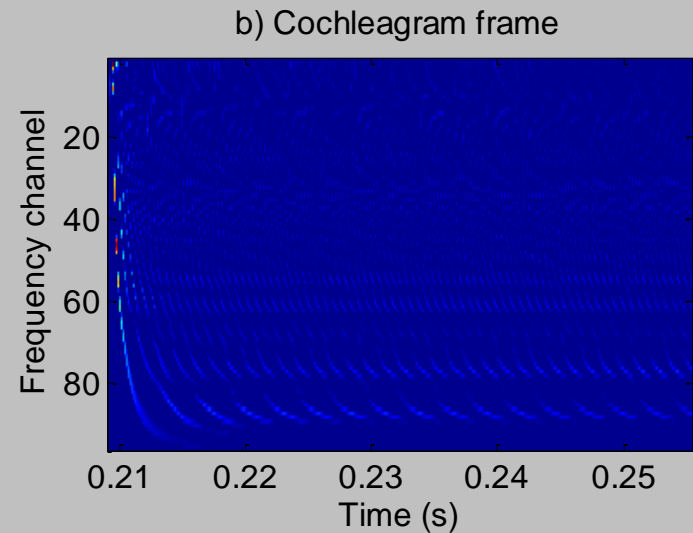
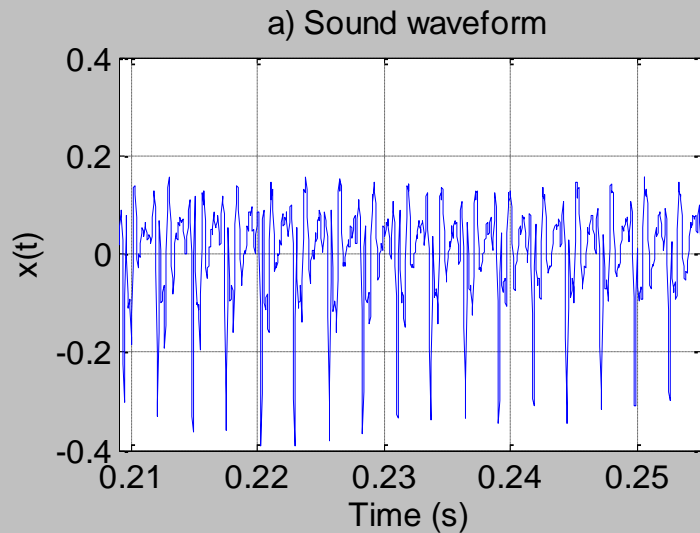
- **Context of our work**

- “Melody is the individual dominant pitched line in a musical ensemble”

Overview



Multi-Pitch Detection



From Pitches to Notes

□ Goal

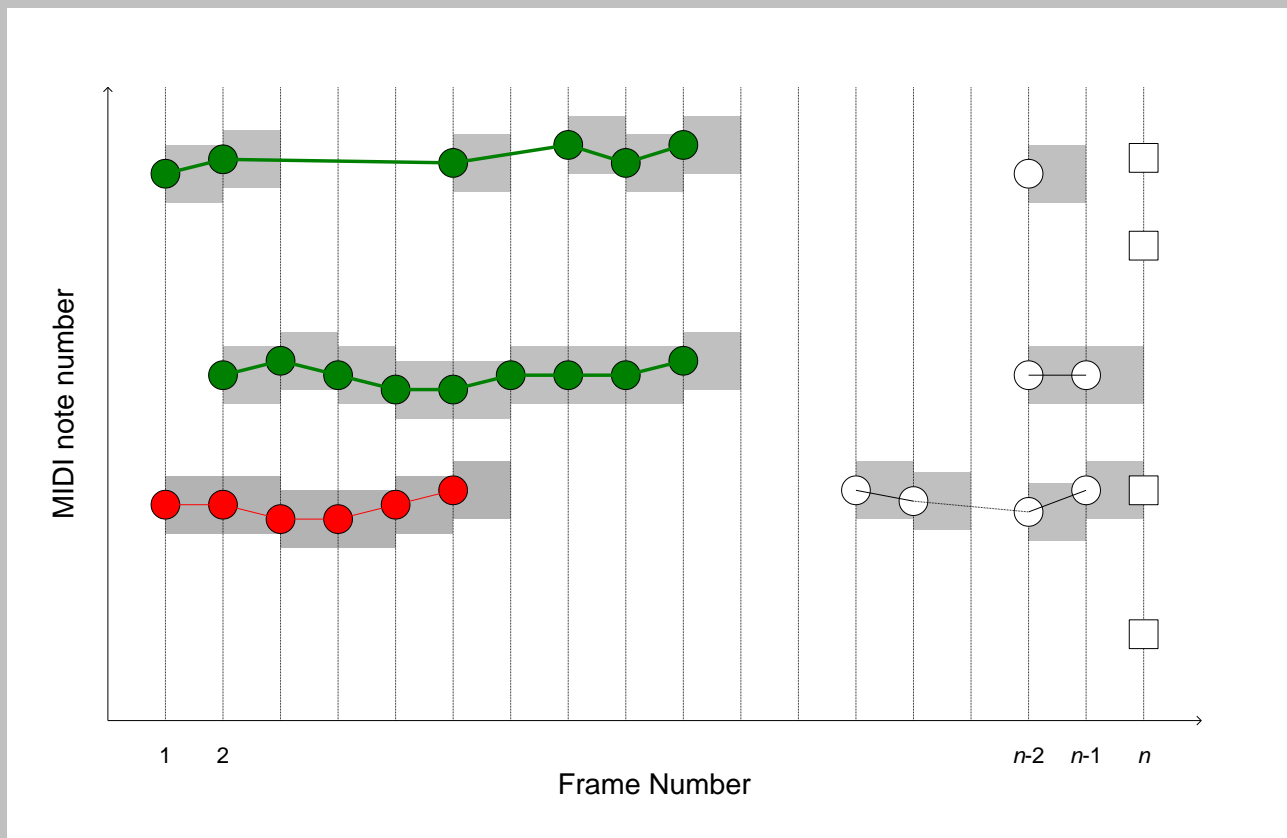
- Quantize temporal sequences of detected pitches into MIDI notes

□ Approach

- Pitch Trajectory Construction [Serra]
- Frequency-Based Track Segmentation
- Salience-Based Track Segmentation

From Pitches to Notes

□ Pitch Trajectory Construction



From Pitches to Notes

□ Frequency-Based Track Segmentation

- Approximation of frequency curves by piecewise-constant functions (PCFs)
 - 1) MIDI quantization

a) Original PCFs



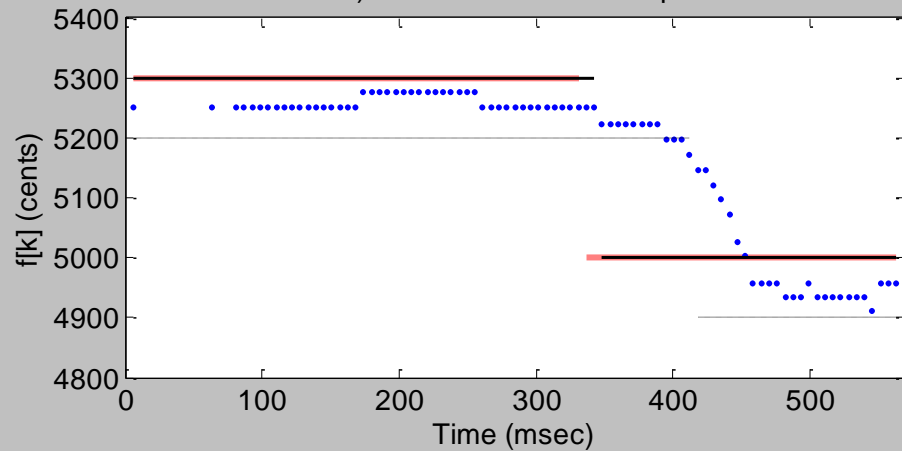
b) Filtered PCFs



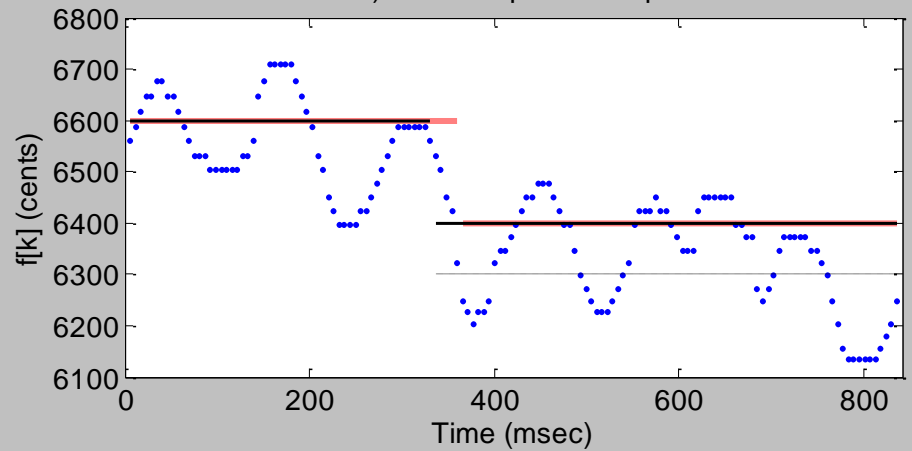
From Pitches to Notes

□ Frequency-Based Track Segmentation

a) Eliades Ochoa's excerpt



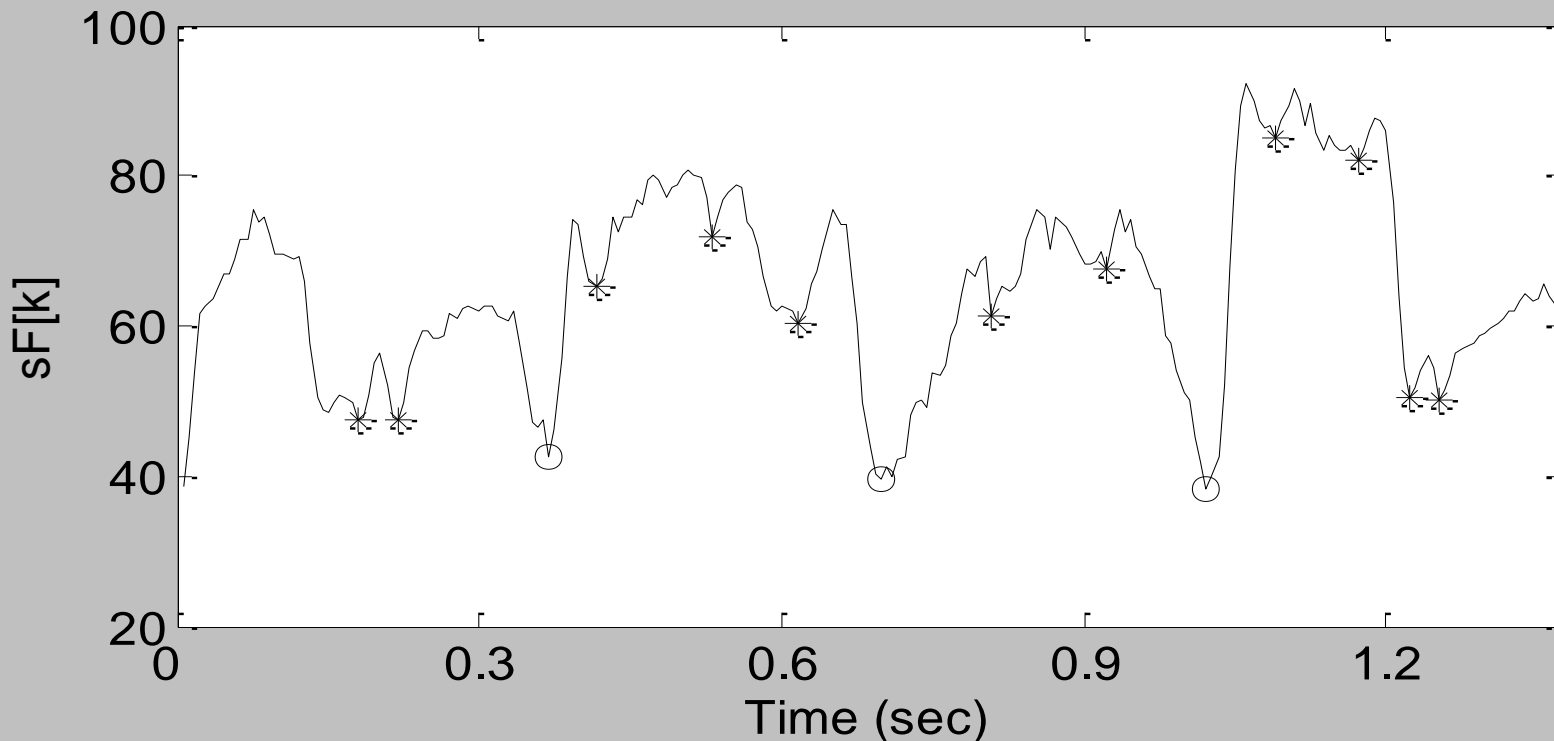
b) Female opera excerpt



From Pitches to Notes

□ Saliency-Based Track Segmentation

- 1) Candidate Segmentation Points



Identification of Melodic Notes

□ Goal

- **Separate the melodic notes from the accompaniment**

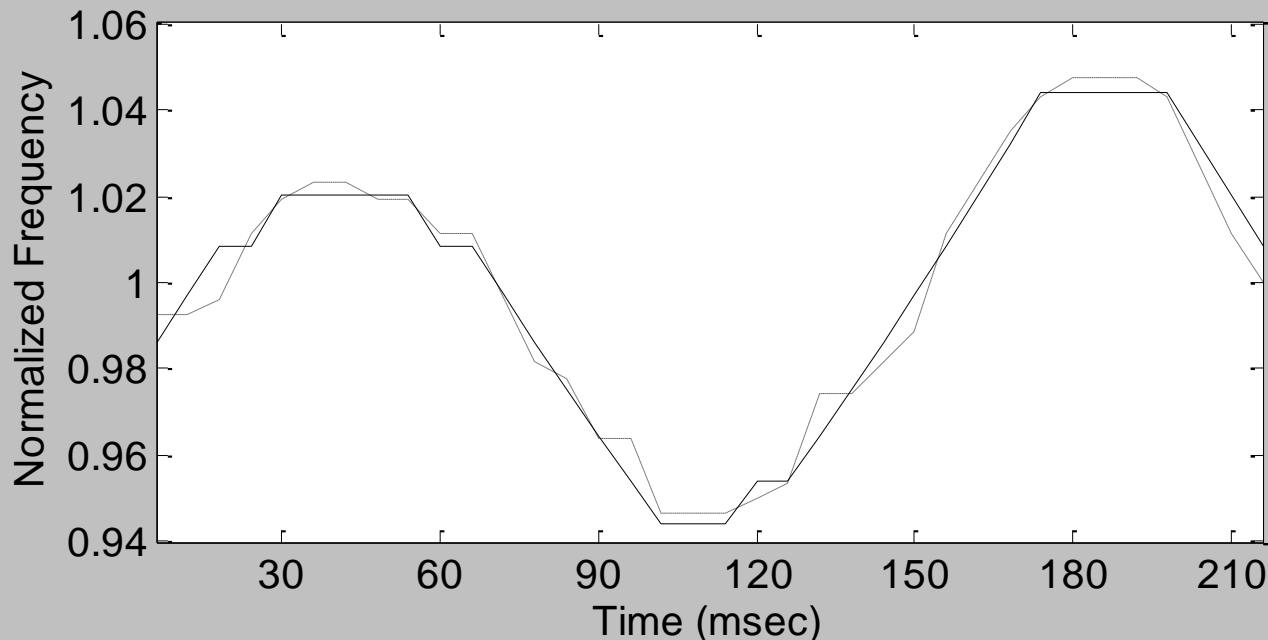
□ Approach

- **Elimination of Ghost Harmonically-Related Notes**
- **Selection of the Most Salient Notes**
- **Melody Smoothing**
- **Elimination of Spurious Accompaniment Notes**
- **Note Clustering**

Identification of Melodic Notes

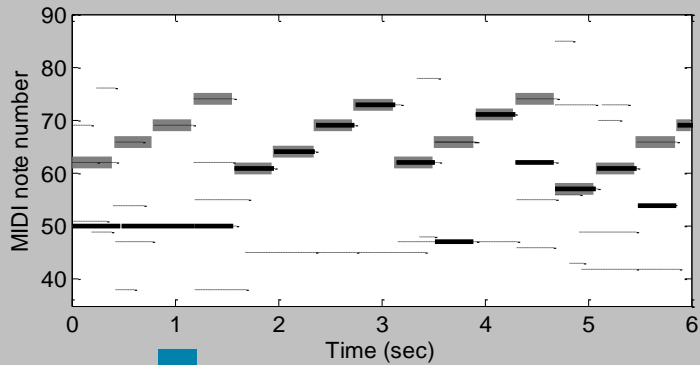
□ Elimination of Ghost Harmonically-Related Notes

- Delete harmonically-related notes that satisfy the “common fate” principle

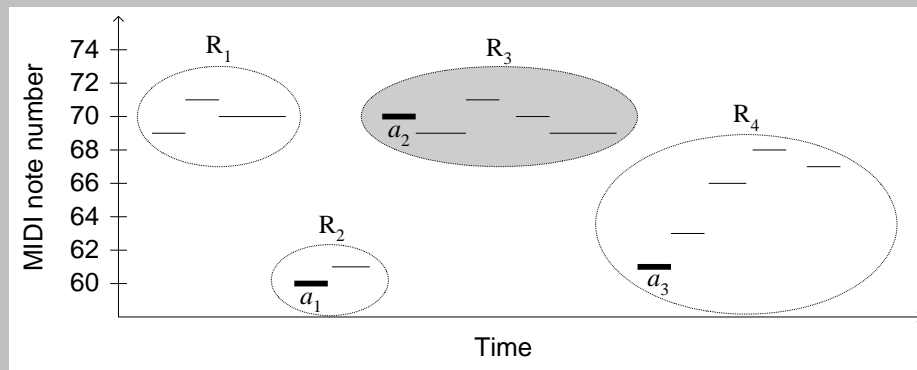
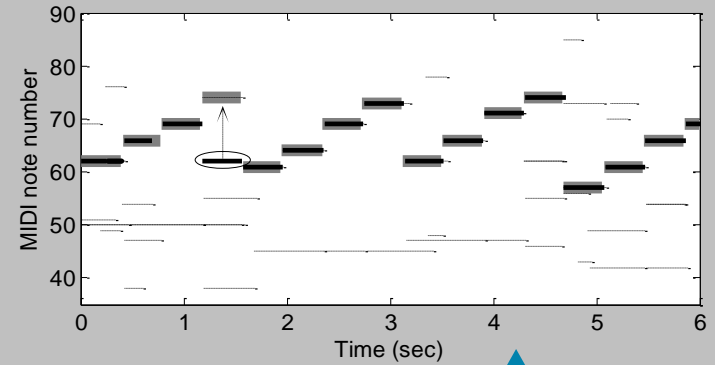


Identification of Melodic Notes

Selection of the Most Salient Notes



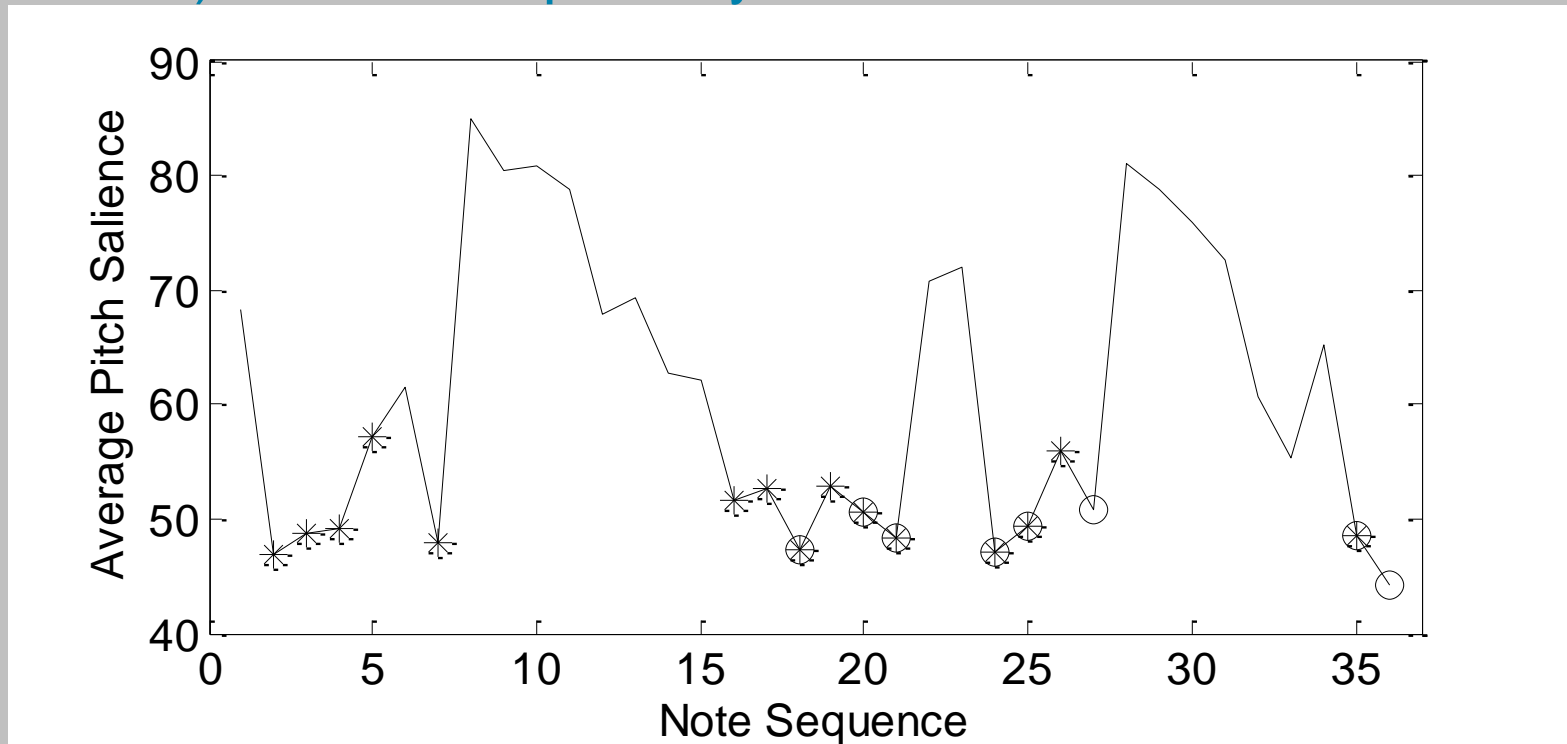
Melody Smoothing



Identification of Melodic Notes

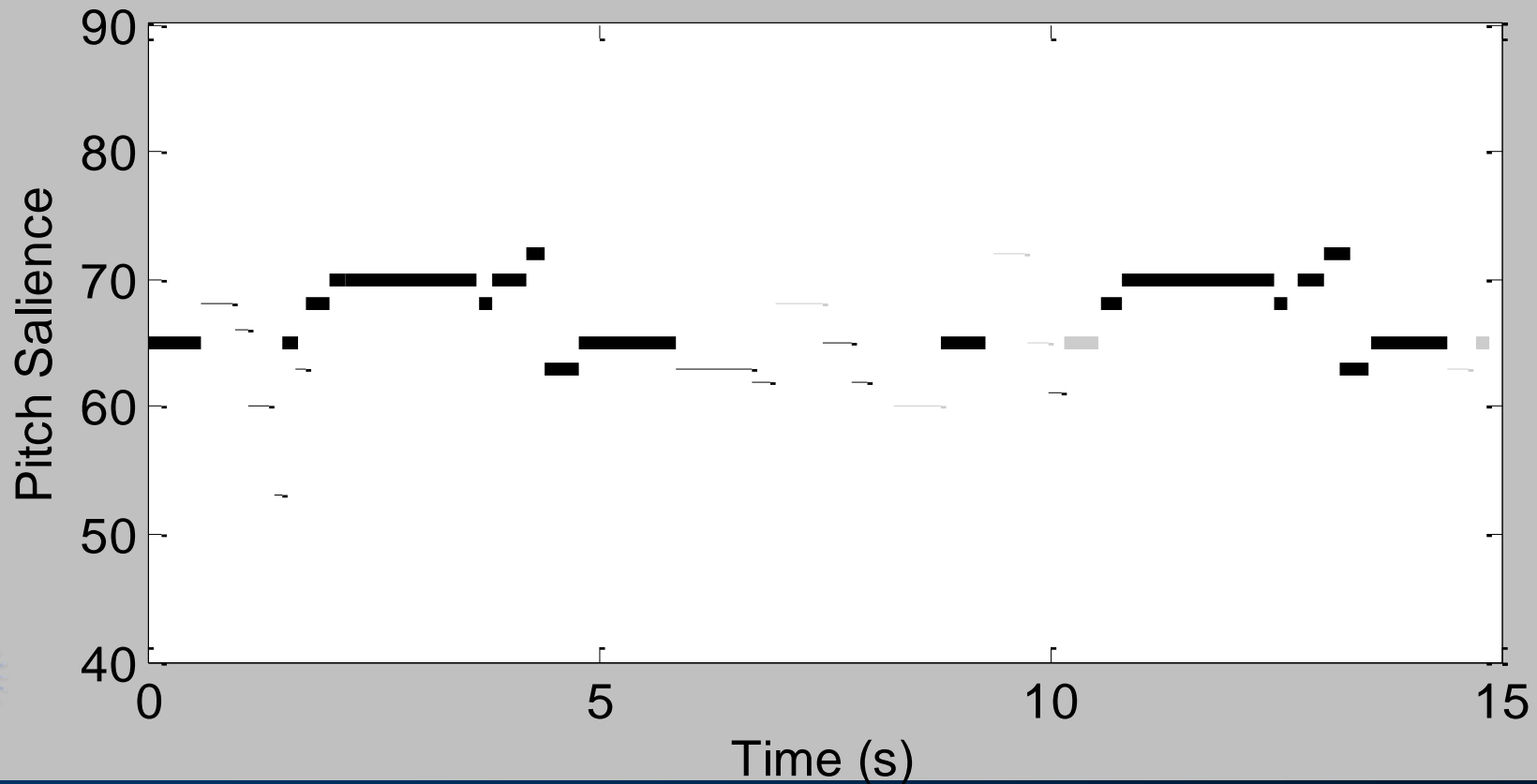
□ Elimination of Spurious Accompaniment Notes

- 1) Delete deep valleys in the salience contour



Identification of Melodic Notes

□ Elimination of Spurious Accompaniment Notes



Identification of Melodic Notes

□ Note Clustering

- **Feature Extraction**

- Spectral shape, harmonicity, attack transient, intensity, pitch-related features

- **Feature Selection and Dimensionality Reduction**

- Forward feature selection
- Principal Component Analysis

- **Clustering**

- Gaussian Mixture Models
 - 2 clusters: melodic and “garbage” cluster

Evaluation Procedures

□ Ground Truth Data

<i>ID</i>	<i>Song Title</i>	<i>Category</i>	<i>Solo Type</i>
1	Pachelbel's "Kanon"	Classical	Instrumental
2	Handel's "Hallelujah"	Choral	Vocal
3	Enya - "Only Time"	New Age	Vocal
4	Dido - "Thank You"	Pop	Vocal
5	Ricky Martin - "Private Emotion"	Pop	Vocal
6	Avril Lavigne - "Complicated"	Pop/Rock	Vocal
7	Claudio Roditi - "Rua Dona Margarida"	Jazz/Easy	Instrumental
8	Mambo Kings - "Bella Maria de Mi Alma"	Bolero	Instrumental
9	Eliades Ochoa - "Chan Chan"	Son	Vocal
10	Juan Luis Guerra - "Palomita Blanca"	Bachata	Vocal
11	Battlefield Band - "Snow on the Hills"	Scottish Folk	Instrumental
12	daisy2	Pop	Vocal
13	daisy3	Pop	Vocal
14	jazz2	Jazz	Instrumental
15	jazz3	Jazz	Instrumental
16	midi1	Pop	Instrumental
17	midi2	Folk	Instrumental
18	opera female 2	Opera	Vocal
19	opera male 3	Opera	Vocal
20	pop1	Pop	Vocal
21	pop4	Pop	Vocal

Evaluation Procedures

□ Evaluation Metrics

- **Pitch Contour Accuracy**

- Melodic Raw Pitch Accuracy (MRPA)
- Overall Raw Pitch Accuracy (ORPA)
- Melodic Raw Note Accuracy (MRNA)
- Overall Raw Pitch Accuracy (ORNA)

- **Note Extraction Accuracy**

- Percentage of correct frames

- **Melody/Accompaniment Discrimination**

- Recall
- Precision

Overall Results

❑ Multi-Pitch Detection

- MRPA = 81.0%

❑ Note Determination

- Frequency-Based Track Segmentation: Re = 72%; Pr = 94.7%
- Salience-Based Track Segmentation: Re = 75%; Pr = 41.2%

❑ Identification of Melodic Notes

- **Elimination of Ghost Notes (EGN)**
 - Eliminated 37.8% of notes (only 0.3% melodic)
- **Overall Melody Extraction**
 - MRNA = 84.4%; ORNA = 77.0%
- **Melody/Accompaniment Discrimination**
 - Recall = 31.0%; Precision = 52.8%
- **MIREX'2004**
 - Train (note accuracy) = 80.1/ 77.1%; Test = 77.4 / 75.1%
 - Train (pitch contour accuracy) = 75.1; 71.1%
- **MIREX'2005**
 - 61.1% overall raw pitch accuracy

Conclusions and Future Work

□ Main Contributions

- Note determination
- Identification of melodic notes in a mixture

□ Conclusions

- Adequate MPD approach in a melody context
- Note determination with satisfactory accuracy
- Identification of melodic notes successful in medium SNR conditions

□ Future Work

- Larger musical corpus
- Reduce computational time for pitch detection
- Melody/Accompaniment discrimination