INESC PORTO

INSTITUTO DE ENGENHARIA DE SISTEMAS E COMPUTADORES DO PORTO

Universidade do Porto
FEUP Faculdade de Engenharia

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA E DO ENSINO SUPERIOR Portugal

ISMIR 2004 Graduate School

07 OCTOBER 2004

Analysis and Recognition of Audio oriented to Music Applications

Luís Gustavo Martins

lmartins@inescporto.pt

http://telecom.inescporto.pt/~lmartins

FEUP / INESC Porto

INESC Porto	Campus da FEUP	R. Dr. Roberto Frias, nº 378	4200-465 Porto PORTUGAL
tel. +351222094000	fax +351222094050	e-mail www@inescporto.pt	internet www.inescporto.pt

Summary

- Short Bio
- R&D Past Experience
- Research Proposal
- Conclusion

Short Bio

- Engineering Academic Degrees:
  - Engineer Degree in Electrical and Computer Engineering – FEUP – 1997
    - Project: **"A Digital 20 Band Equalizer based on a Texas Instruments C31"**
  - MSc. Degree in Electrical and Computer Engineering – FEUP – Jan. 2002
    - MSc. Dissertation: ***"PCM to MIDI Transposition"***
    - Advisor: Prof. Dr. **Aníbal João de Sousa Ferreira** (FEUP / INESC Porto)
    - Reviewers: Prof. Dr. **Xavier Serra** (UPF / MTG) and Prof. Dr. **Francisco José Oliveira Restivo** (FEUP).
  - PhD Student in Electrical and Computer Engineering – FEUP / INESC Porto
    - Started in December 2003 (minimum duration = 3 years)
    - PhD thesis: ***"Analysis and Recognition of Audio oriented to Music Applications"***
    - Advisor: Prof. Dr. **Aníbal João de Sousa Ferreira** (FEUP / INESC Porto)

Short Bio

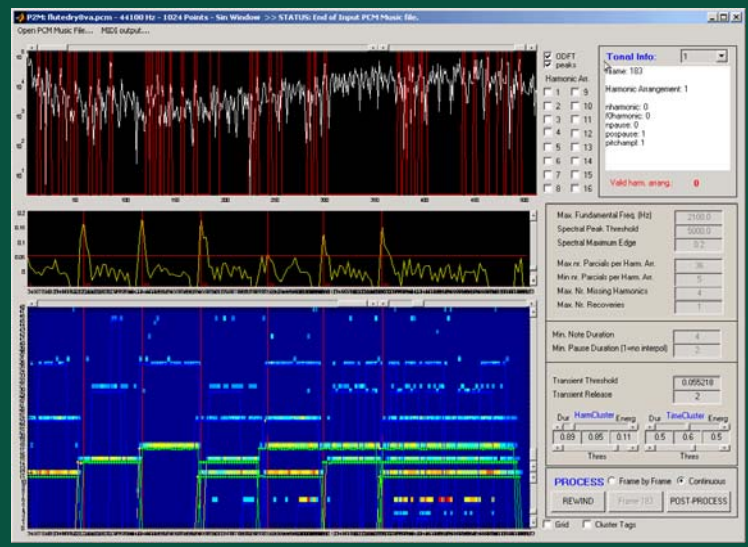
- **Music background**
  - Piano and Music Theory lessons (5th grade) at the Conservatory of Music of Porto – Portugal
- **Research Interests**
  - Acoustics / Psychoacoustics
  - Auditory Scene Analysis
  - Digital Signal Processing
  - Pattern Recognition / Machine Learning
  - Analysis, Processing and Classification of digital audio signals
  - Interactive applications to the musical and sound expressivity

Short Bio

- **Publications**
  - "PCM to MIDI Transposition", 112 AES Conference – Audio Engineering Society – Munich, May 2002.
  - "PCM to MIDI Transposition", MSc thesis, School of Engineering of the University of Porto (FEUP), January 2002.
  - "Implementation of a Real-Time Audio Decoder: ASCdec", 107 AES Conference – Audio Engineering Society – New York, September 1999.
  - "Implementation of a New Method to Digital Audio Equalization", 106 AES Conference – Audio Engineering Society – Munich, May 1999.
  - "20 Band Digital Audio Equalizer, Implementation on an Expand TMS320C31DSK", Technical Report, Texas Instruments "Challenge '97".
- **Participation in Projects**
  - ***VISNET -*** *"NETworked audio Visual media technologies"*
    - FP6-2002-IST-1 European Union Network of Excellence Project
    - audio and multimodal analysis work packages.
  - ***ORBIT*** *- "Object Re-configurable Broadcast Infrastructure Trial"*
    - a contract project between INESC Porto and the BBC
    - development of automatic audio segmentation and classification tools.

R&D Past Experience

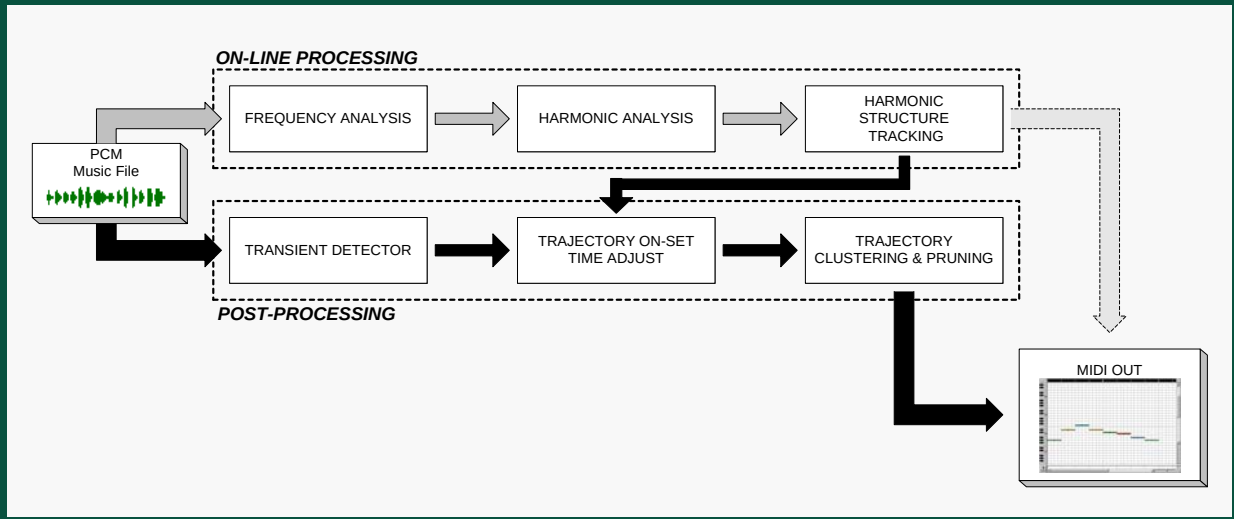
- PCM to MIDI Transcription
  - Extraction of specific features out of a POLYPHONIC musical acoustic signal in order to estimate the musical notes played:
    - Fundamental Freq. → *Pitch*
    - Onsets /Offsets → *Duration*
    - Energy → *Dynamics*



  - Other relevant features to consider
    - Spectral profiles → *Timbre*
    - Beat → *Rhythm/Time Signature*
    - …

R&D Past Experience

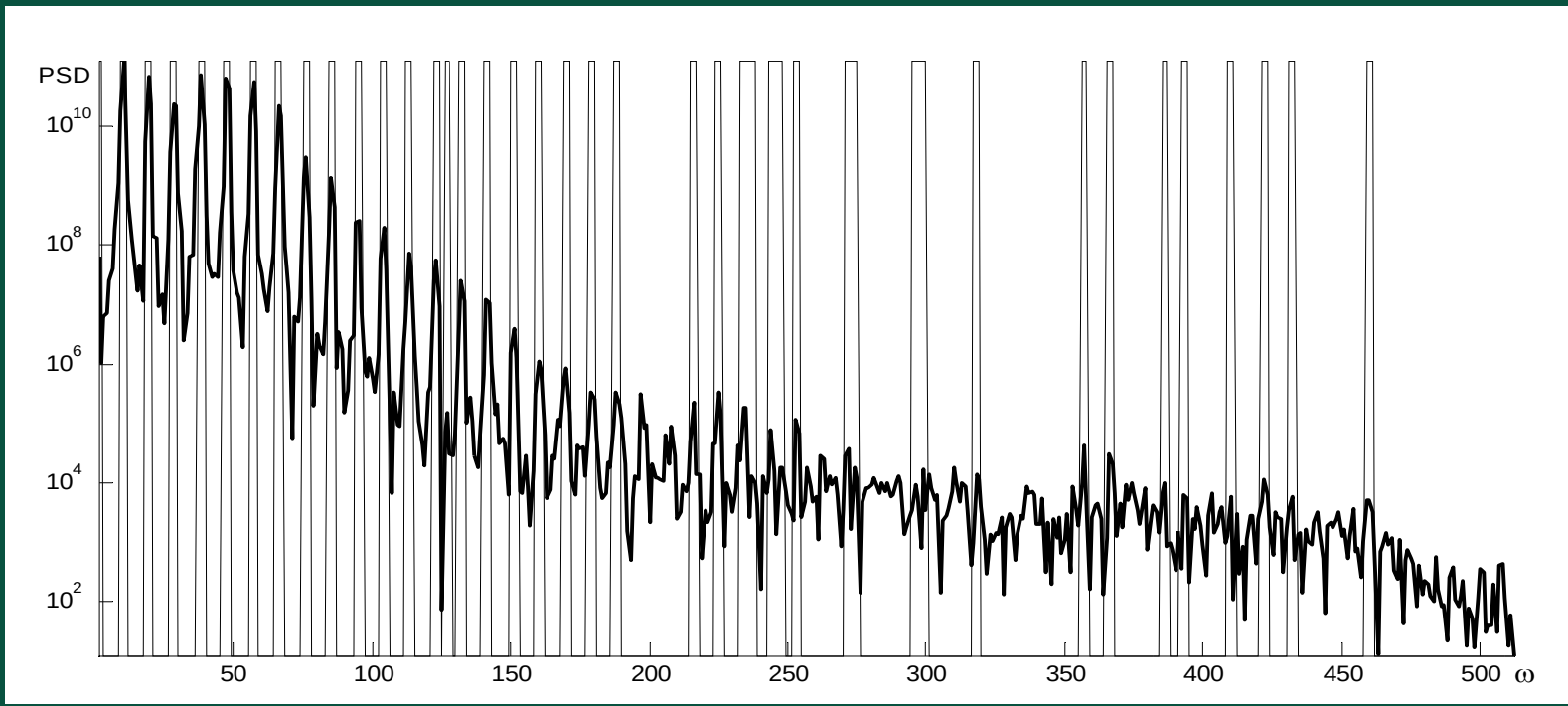
- PCM to MIDI Transcription
  - System Overview



  - Frequency Analysis Front-End
    - Based around a ***50% overlap analysis scheme***
    - Uses an ***N-point sine window and ODFT***
    - Use of ***accurate frequency estimation method*** to increase spectral selectivity without compromise of time resolution

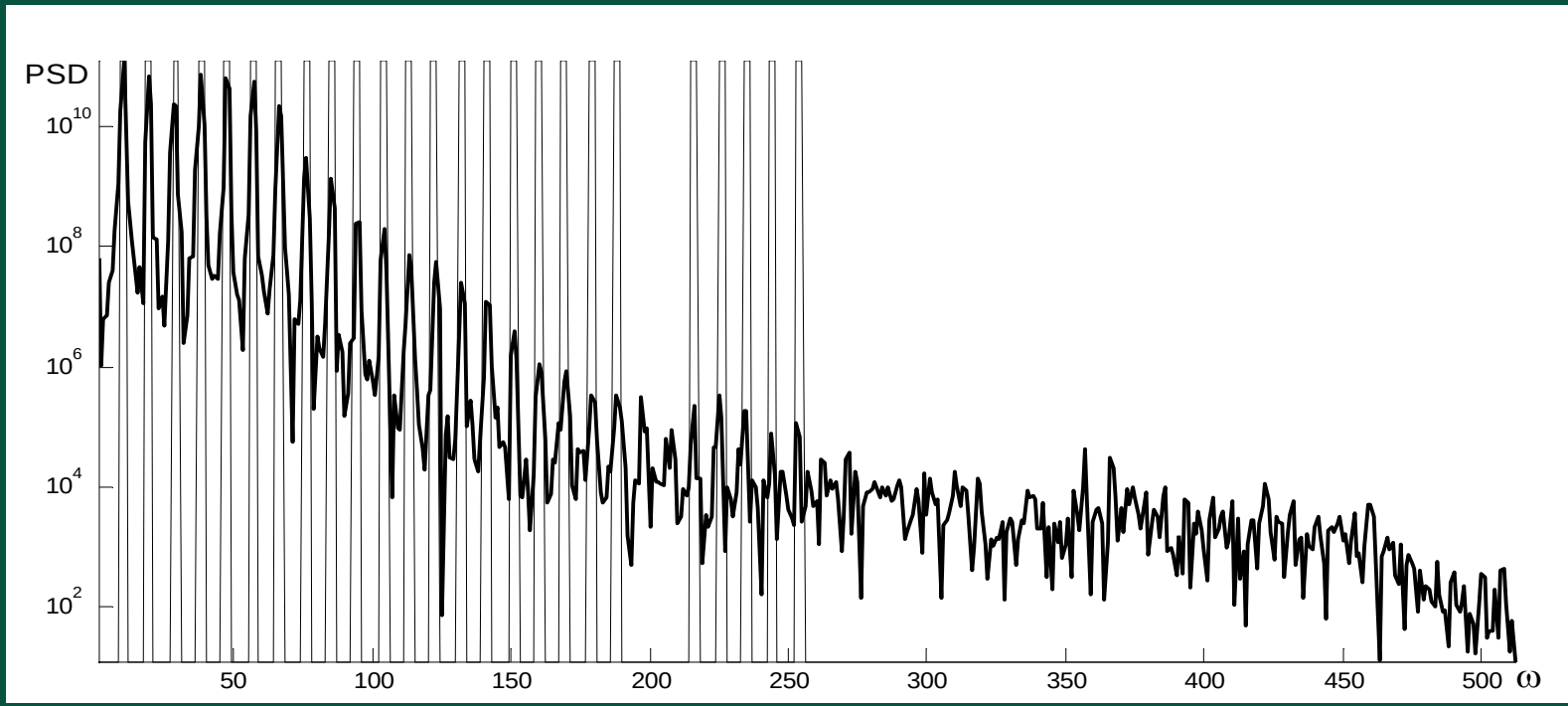
R&D Past Experience

- PCM to MIDI Transcription
  - Harmonic Analysis
    - Peak Detector



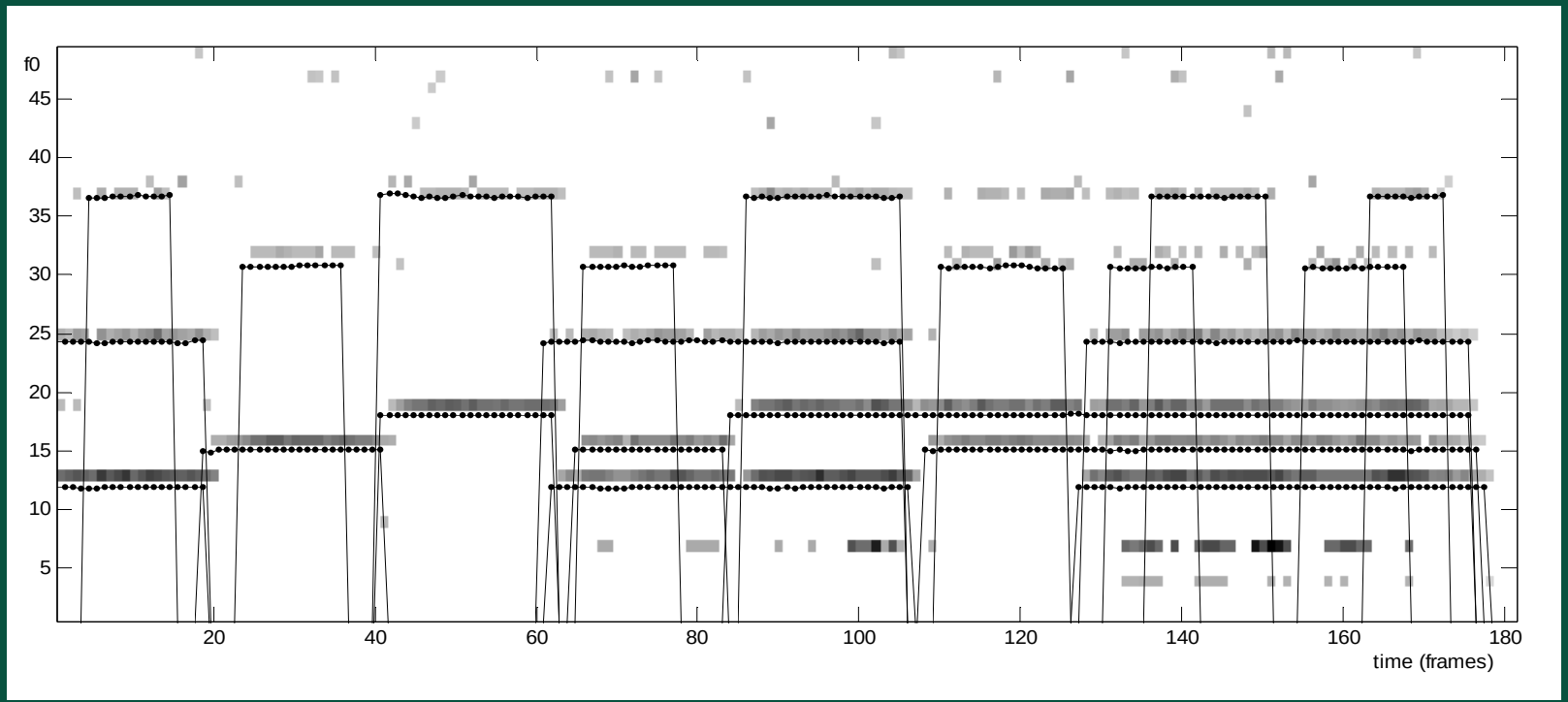
R&D Past Experience

- PCM to MIDI Transcription
  - Harmonic Analysis
    - Harmonic Series Detector → FUNDAMENTAL FREQUENCY



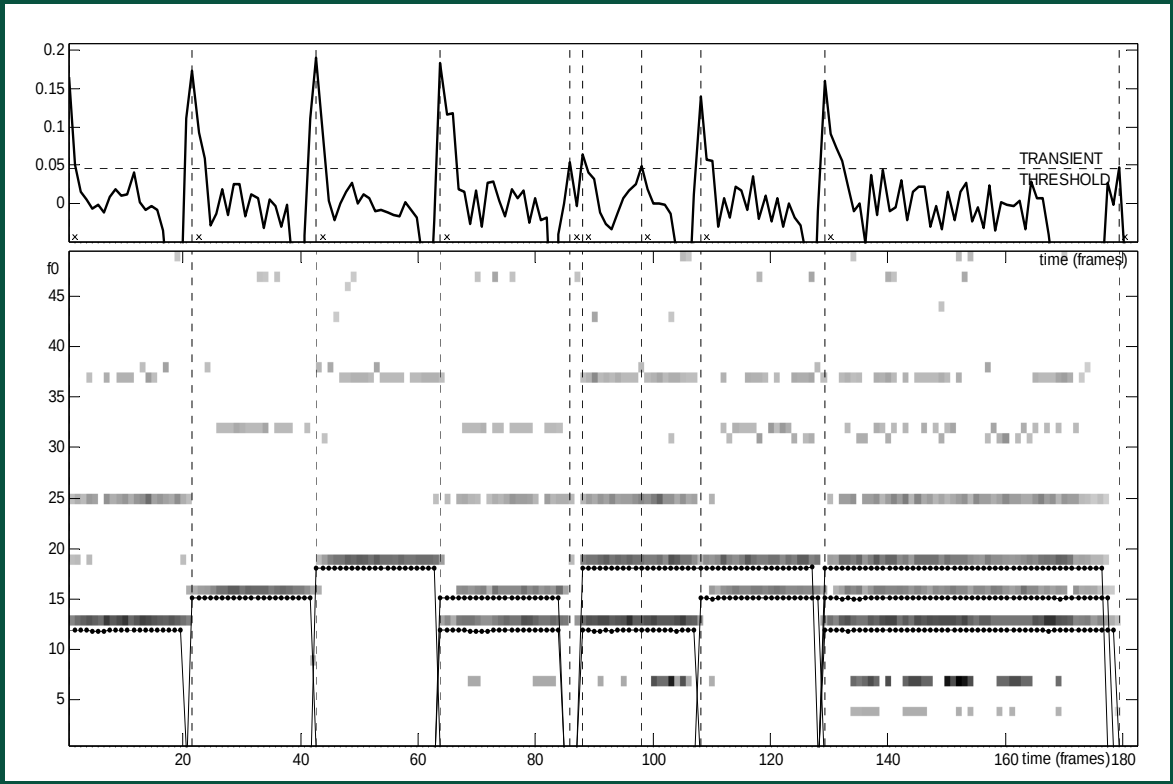
R&D Past Experience

- PCM to MIDI Transcription
  - Harmonic Series Analysis
    - Tracking of *FUNDAMENTAL FREQUENCY* over time → TRAJECTORIES



R&D Past Experience

- PCM to MIDI Transcription
  - Harmonic Series Analysis
    - Selection of best trajectories → estimation of MUSICAL NOTES...

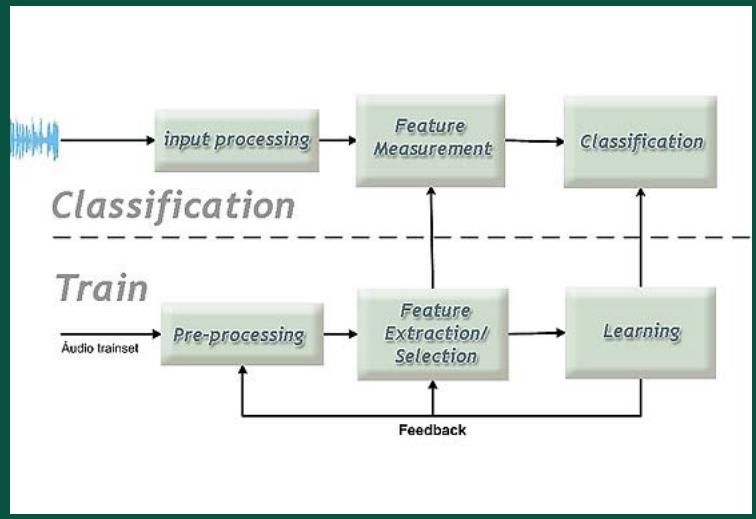


R&D Past Experience

- PCM to MIDI Transcription
  - Some Results and Examples:
    - Simple Piano piece:
      - Original (Audio PCM):
      - Transcribed (MIDI):
    - Taivas:
      - Original (Audio PCM): (from A.Klapuri webpage - http://www.cs.tut.fi/~klap/iiro/dafx2000/)
      - Trancribed (MIDI):
    - Two Singing Voices:
      - Original (Audio PCM):
      - Transcribed (MIDI):

R&D Past Experience

- Audio Classification
  - Time Segmentation and Classification of audio streams



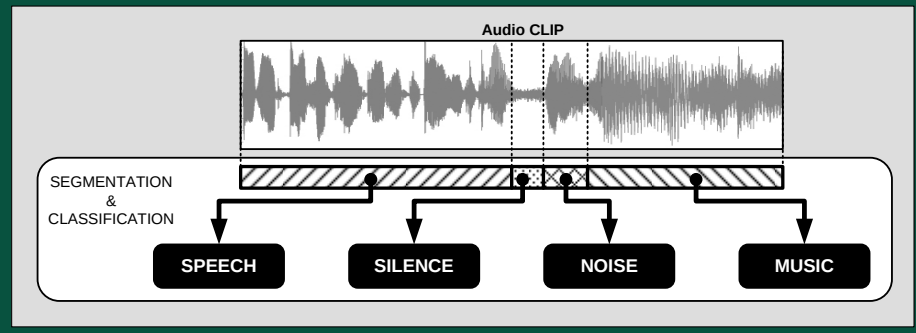
  - Challenges:
    - Genre dependency of features
    - Feature space dimensionality
    - Feature selection
    - Taxonomies
      - user-oriented ***VS*** feature-oriented
      - Flat Classifier ***VS*** Hierarchical Classifier
    - Classify ➔Segment ***VS*** Segment ➔ Classify

R&D Past Experience

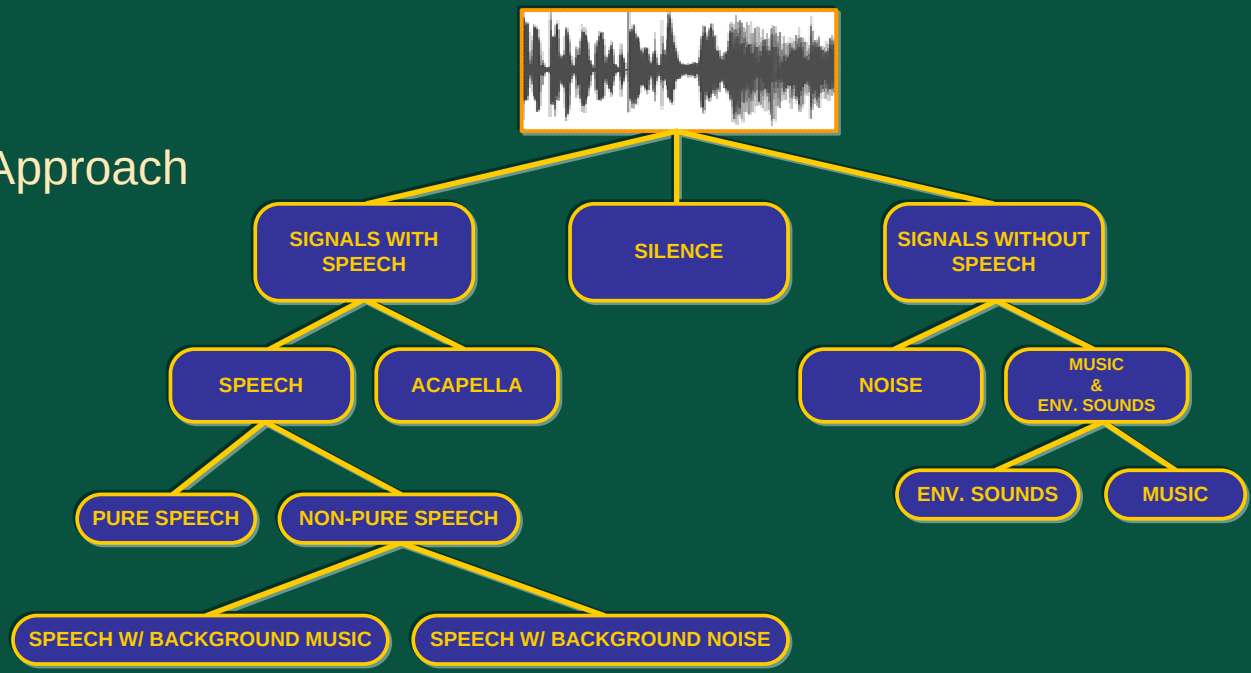
- Audio Classification
  - Feature Extraction
    - Time domain features:
      - Zero-Crossing Rate (ZCR)
      - High Zero-Crossing Rate (HZCR)
      - Energy Ratio
      - …
    - Spectral features:
      - Centroid
      - Roll-Off
      - Flux
      - Flatness
      - MFCC coefficients
      - …
    - LPC analysis
  - Classifiers
    - Statistical classifiers
      - Gaussian
      - GMM
    - Decision Trees
    - Clustering
    - …

R&D Past Experience

- Audio Classification
  - Flat / Direct Approach

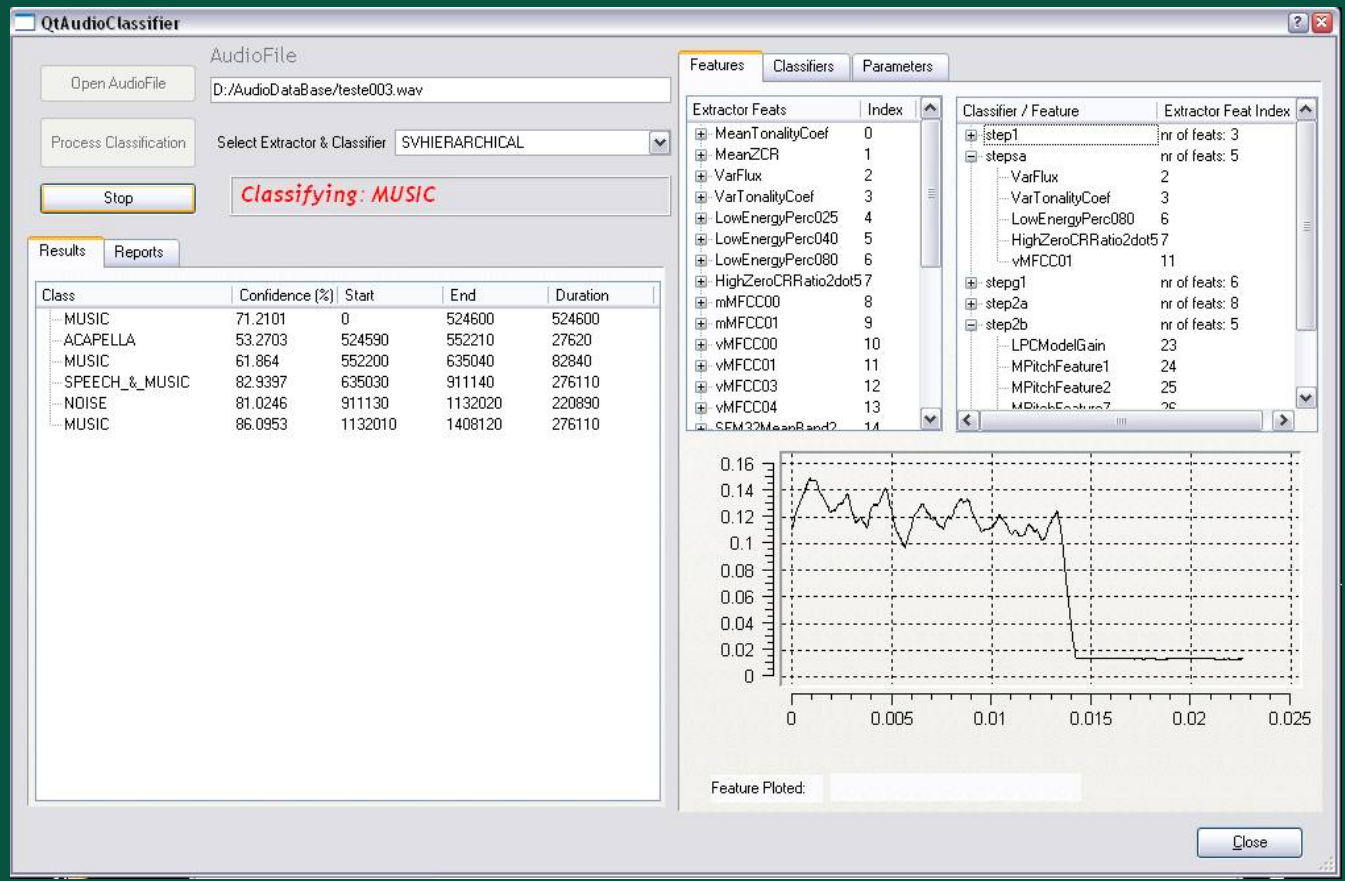


  - Hierachical Approach



R&D Past Experience

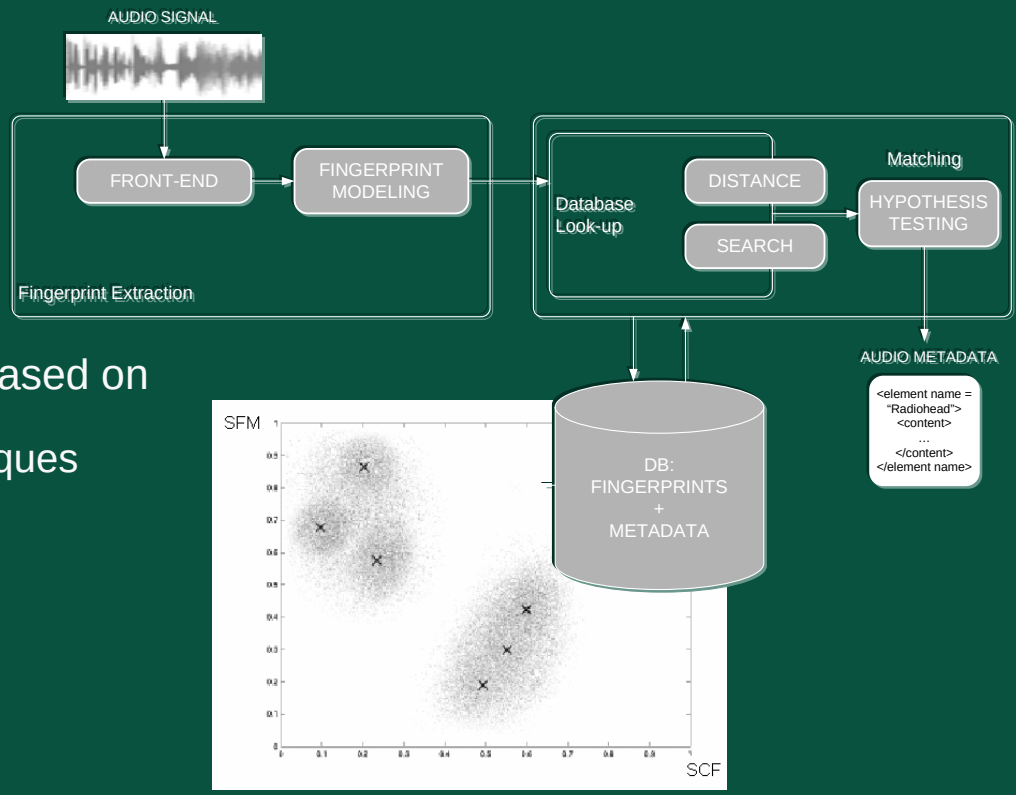
- Audio Classification



DEMO

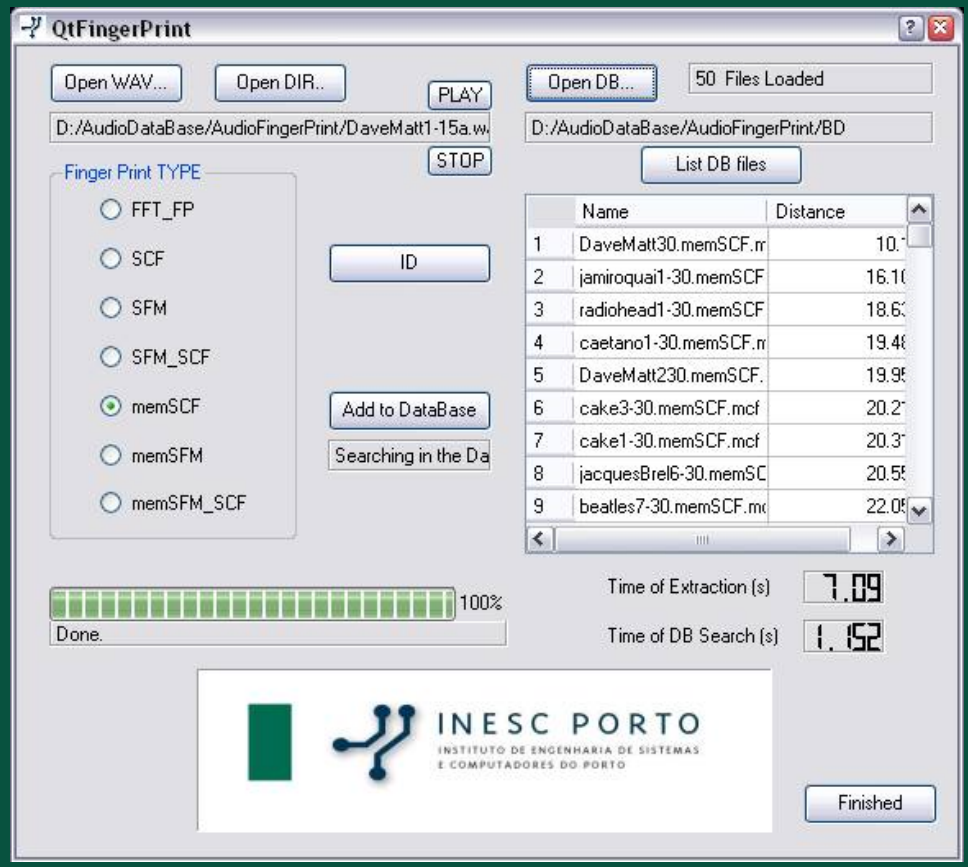
R&D Past Experience

- Audio Fingerprinting
  - Simple experiments based on published solutions
    - Based on perceptual features
      - SFM
      - SCF
      - MFCC
      - Band Energies
    - Perceptual hashing based on
      - Clustering techniques
      - …



R&D Past Experience

- Audio Fingerprinting



DEMO

Research Proposal

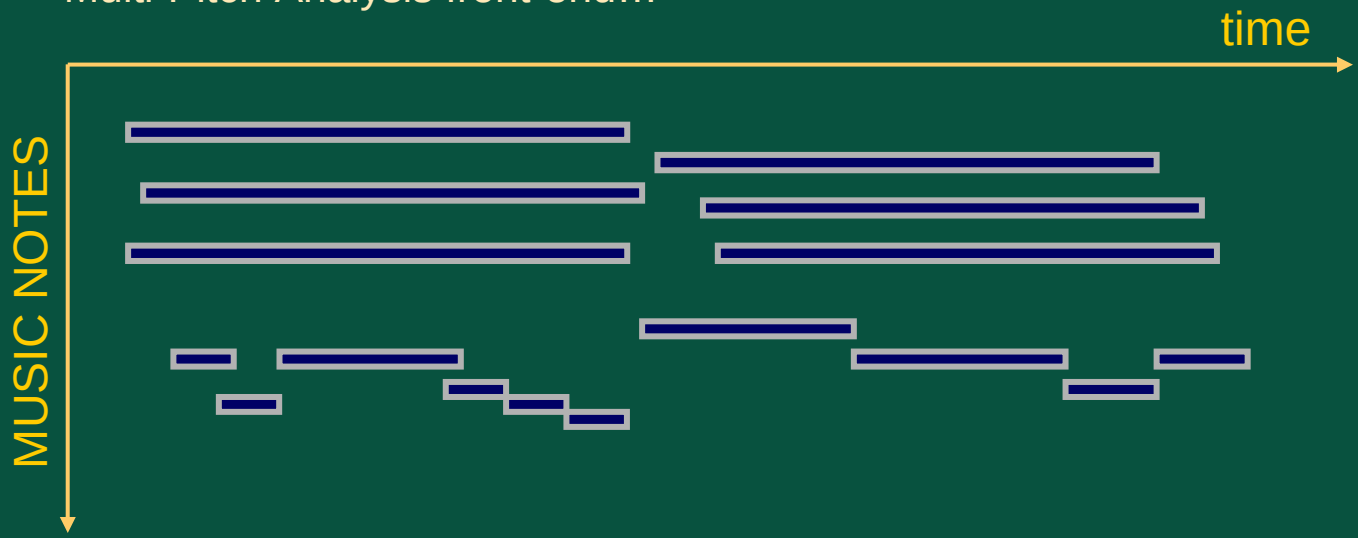
"Analysis and Recognition of Audio oriented to Music Applications"

- Proposal should:
  - Capitalize previous R&D experience in:
    - Multi-pitch estimation
    - Audio Segmentation / Classification
    - *(Audio Fingerprinting)*
  - Address the study and development of new methodologies for:
    - Analysis of audio signals
    - Extraction of **perceptually relevant features** and descriptors
    - Robust classifiers for the automatic classification / description of audio contents

Research Proposal

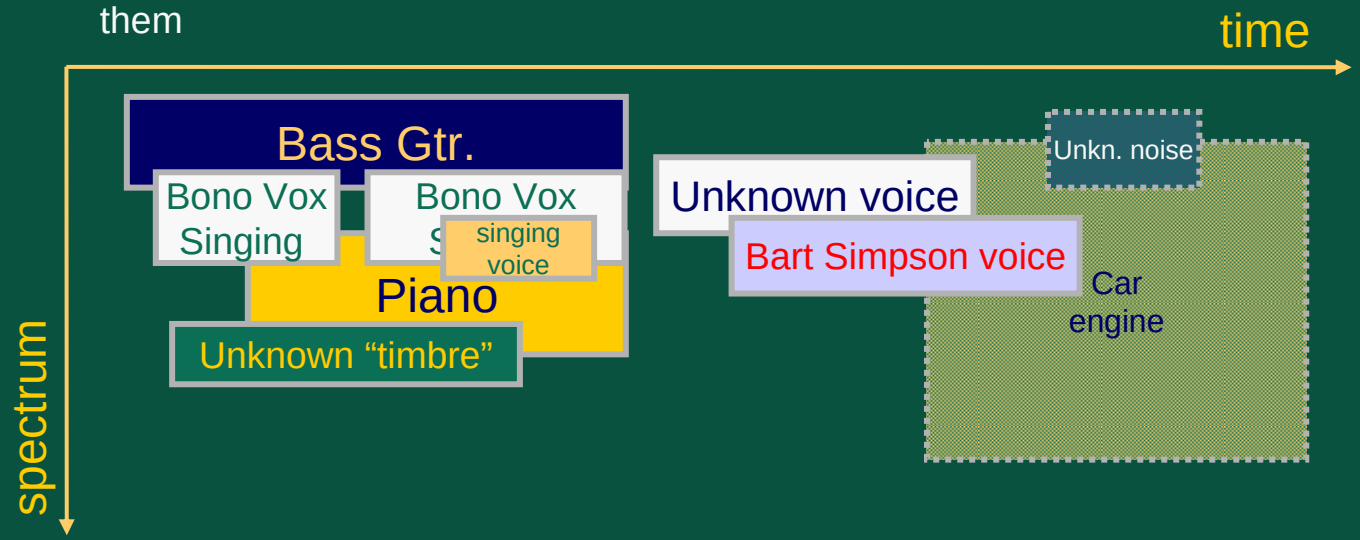
- Combining ideas from…
  - Temporal Audio Segmentation / Classification…

  - Multi-Pitch Analysis front-end…



Research Proposal

- Try to "segment" and recognize distinct and concurrent *Sound Events* in a continuous sound mixture:
  - Similar to the way Human auditory system works
    - Distinct sound sources → different acoustic properties → different perceptual cues
    - Capable to segment sounds from a time/ frequency mixture and classify / recognize them



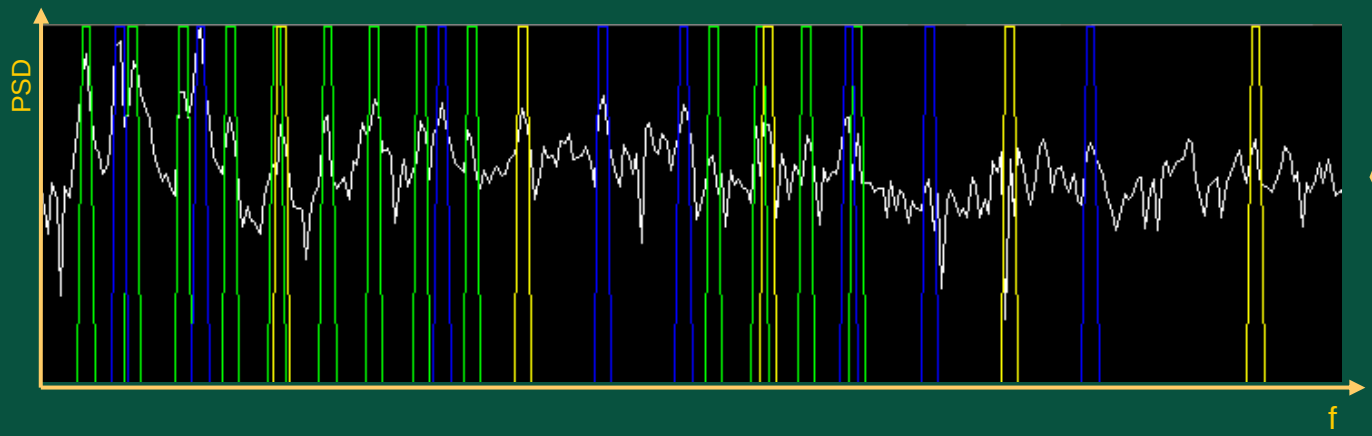
Research Proposal

- Challenges:
  - No general solution for the separation of time and frequency overlapping sounds
  - Same sound source → generates varying sounds over time
  - *Quasi-Harmonic* sounds ***VS*** *Non-Harmonic* sounds
  - **FACT**: Full separation not yet today a reality
- Learn from published work on…
  - Psychoacoustics
  - ASA / CASA
  - Frequency / Harmonic / Multi-Pitch Analysis
  - Sound Separation / Recognition
  - Pattern Recognition / Machine Learning / AI

Research Proposal

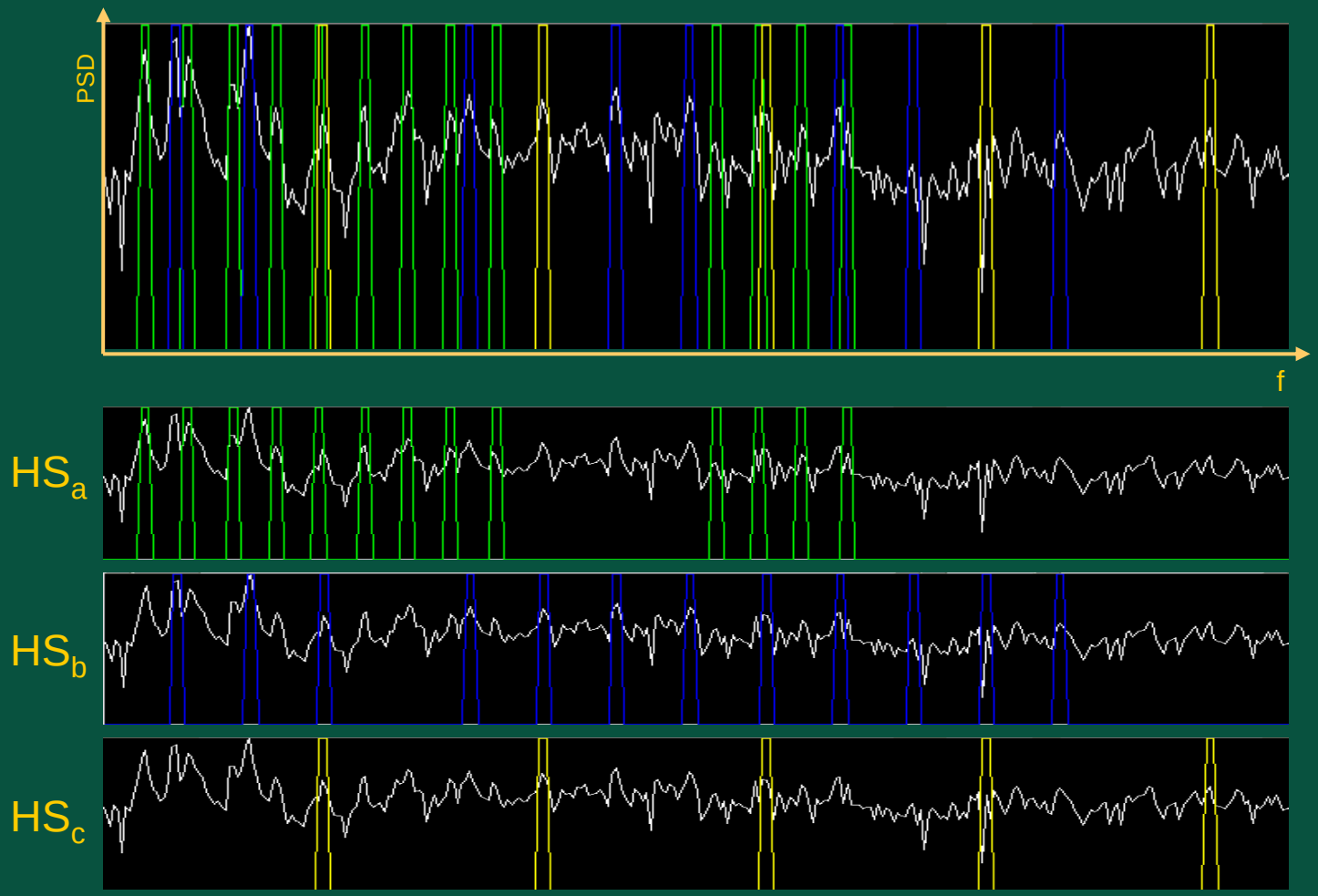
- Proposed Approach for “Segmentation” / Recognition of Harmonic Sounds
  - Accurate identification of *Sinusoidal Components* in the audio signal spectrum

  - Detection of *Harmonic Series* among *Sinusoidal Components*

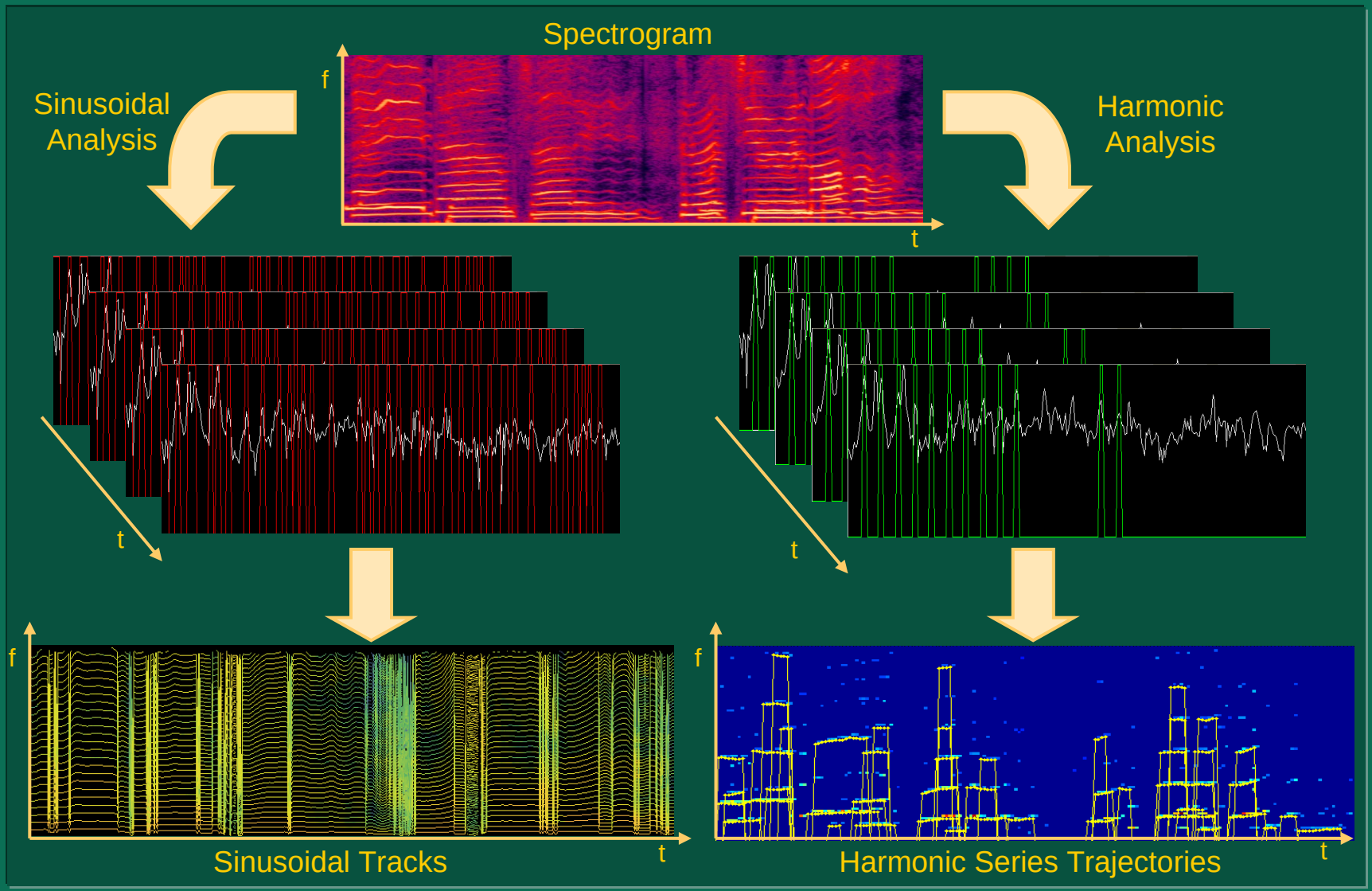


Research Proposal

- Harmonic Series ➔ good evidence of distinct (harmonic) *Sound Events*...

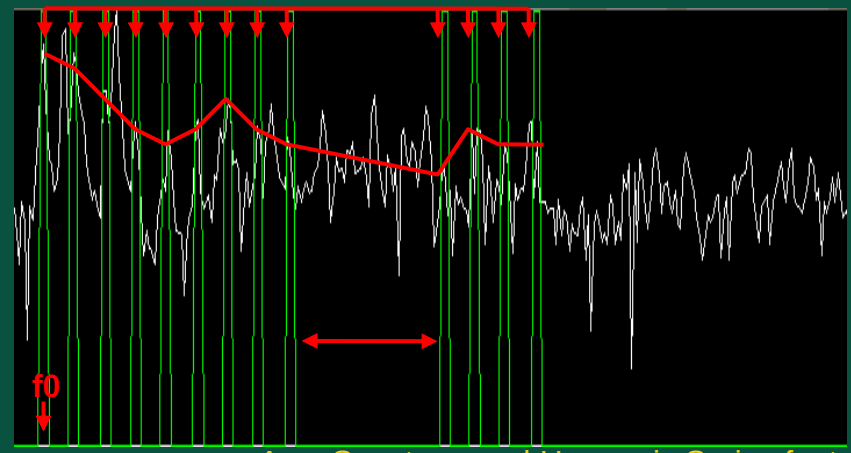


Research Proposal



Research Proposal

- Harmonic Series Trajectories ➔ Sound Event:
  - **Already a good evidence of an harmonic *Sound Event***
  - Possible to extract **relevant features**:
    - Fundamental frequency (fo)
    - Avg. Nr. Of Partials
    - Missing Harmonics
    - Avg. Energy of each partial
    - Energy Decay of each partial
    - Freq. modulation of each partial
    - …

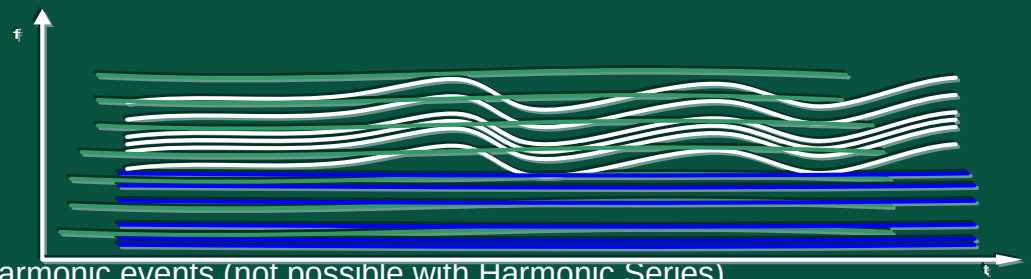


Avg. Spectrum and Harmonic Series feat. over trajectory duration

  - **Source Invariant Features** → RECOGNITION
    - e.g. “Timbre” related features
    - Possible to use Pattern Recognition / Machine Learning methods

Research Proposal

- Sinusoidal Tracks ➔ Sound Event:
  - Sound Events may also be found by grouping sinusoidal tracks:
    - Common fate (ASA):
      - On-set / off-set
      - Modulation
      - …
    - May be used to detect non-harmonic events (not possible with Harmonic Series)
    - Harmonic relation (➔ **Harmonic Series Trajectory!)**
    - Spectral Models (previously learned ”timbre” models)



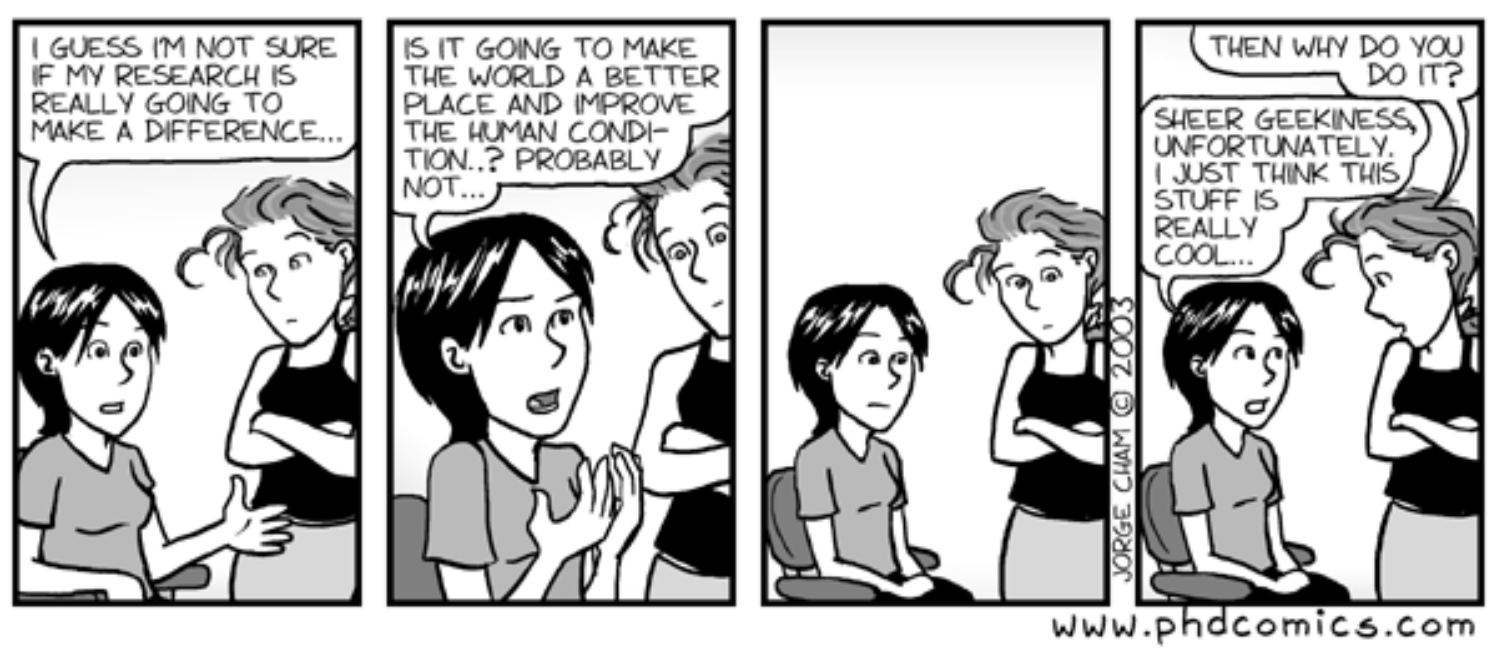
- Difficulties in both Sinusoidal and Harmonic tracking
  - Time / Frequency resolution of freq. analysis front-end
  - Bad modelling capabilities of transient part (i.e. attack) of sounds (!)
  - Shared partials between Sound Events
  - Masked partials
  - Noisy signals
  - …

Research Proposal

- Conclusions
  - Research Proposal tries to **take advantage of previous experiences** in the areas of **Digital Signal Processing** and **Pattern Recognition / Machine Learning, multi-pitch estimation and audio segmentation / classification** (like the “Lazy Classification Algorithms” ;-))
  - Will demand **learning from other research areas** such as **Psychoacoustics, Auditory Scene Analysis**, ... (learning from other areas is supposed to be a good thing, isn’t it? ;-))
  - Try to **derive new perceptually motivated features** for audio and music segmentation/classification
  - Will definitely keep me busy for a long time! ;-)

So, before I really start, please let me know if you think I’m pointing my efforts into a completely wrong direction!!

Thank You!



lmartins@inescporto.pt

http://telecom.inescporto.pt/~lmartins