

# Interactivity in audio browsing, both for the user and the researcher

ISMIR Graduate School – October 2004

Nicolas Wack  
nwack@iua.upf.es


Promoter: Prof. Xavier Serra

# Outline

- Personal Background
- Motivation
- Research Proposal
- Conclusion

# Personal Background

- Telecom Engineer School (ENST, Paris)
- Worked on a project called AudioClas

AudioClas :: Prototype The Tape Gallery 

[HOME](#)  
[SIMPLE SEARCH](#)  
[ADVANCED SEARCH](#)  
[SOUND CLASSIFIER](#)  
[CONCEPT BROWSER](#)  
[WORDNET EDITOR](#)  
[CONTACT](#)

**Main Categories**

**TEMPORARY CATEGORY**  
[Aerodrome](#)  
[Amplifiers](#)  
[Car Explosions](#)  
[Gas Explosions](#)  
[Petrol Explosion](#)  
[Stabs](#)  
[Warehouse](#)  
[Explosions](#)

**Musical**  
[Groups / Ensembles](#)  
[Solo / Featured](#)  
[Instruments](#)





















**Sound Effects**  
[Mechanical](#)  
[Natural](#)

**Search**  
 Keywords     
 Disable AudioClas Technology:

in  [\[ show advanced options \]](#)

**Search Results: 128 Sound FX in 26 Categories found** in 10.473 s  
 Searched for Keywords:  
 car [\[ remove \]](#)  
 screeching [\[ remove \]](#)

Sound FXs : 1 - 20 of 128 [\[ Map View \]](#)

#	Title	Duration	
1	 <b>Bucking Bronco Long Slow Screeching Turn ? Bradley Howard</b>	26.67 s	  
2	 <b>Bucking Bronco Short Slow Screeching Turn ? Bradley Howard</b>	19.36 s	  
3	<b>Car Arrives Long Approach On Gravel Sounds Horn Opens Door</b>	16.22 s	  
4	<b>Car Arrives On Gravel Fast With Skid Ver 01</b>	11.47 s	  
5	<b>Car Arrives Skids To Stop Ver 01 ( Tyres Tires )</b>	3.89 s	  
6	<b>Car Arrives Skids To Stop Ver 02 ( Tyres Tires )</b>	2.12 s	  

**Semantically related terms**

=> [car](#) (69)  
 PARTS OF CAR  
 => [automobile horn](#) (1)

=> [car](#)  
 CAR IS PART OF  
 => [train](#) (27)  
 PARTS OF CAR  
 => [suspension](#) (6)

=> [squeak](#) (54)

=> [squawk](#) (1)  
 BROADER TERMS  
 => [shout](#) (15)

=> [screech](#) (89)  
 BROADER TERMS  
 => [noise](#) (3)

=> [scream](#) (18)  
 => [screaming\(a\)](#) (8)

**Co-occurring terms**

=> [metallic](#) (49)

=> [squeak](#) (44)

=> [shout](#) (20)

# Motivation

- The pragmatic point of view:
  - It takes too long to gather data for an experiment
  - Some creativity is lost due to “administrative” tasks...

# Motivation

- 2-level problem:
  1. Features extraction (signal processing)
  2. Classification (statistics and machine learning)
- Independence of both levels
- Can we add more levels?

# Motivation

- Example:



audio file

# Motivation

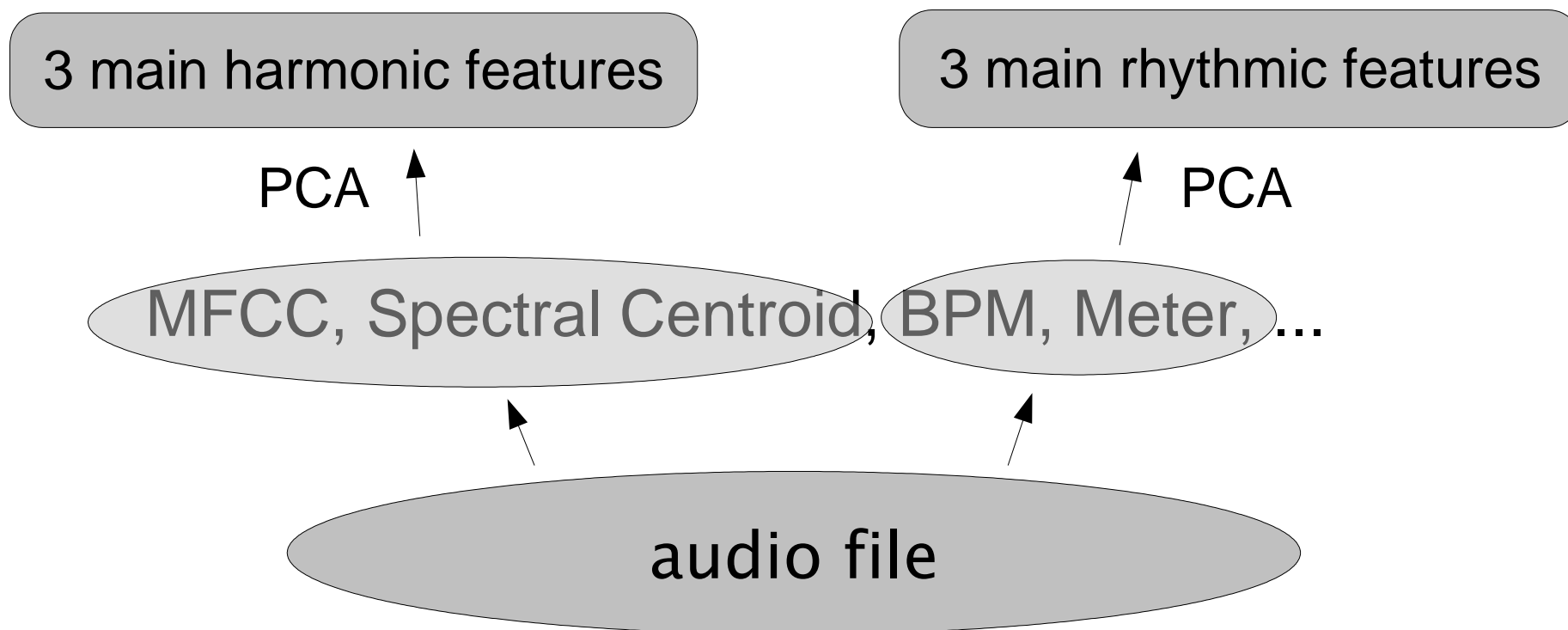
- Example:

MFCC, Spectral Centroid, BPM, Meter, ...



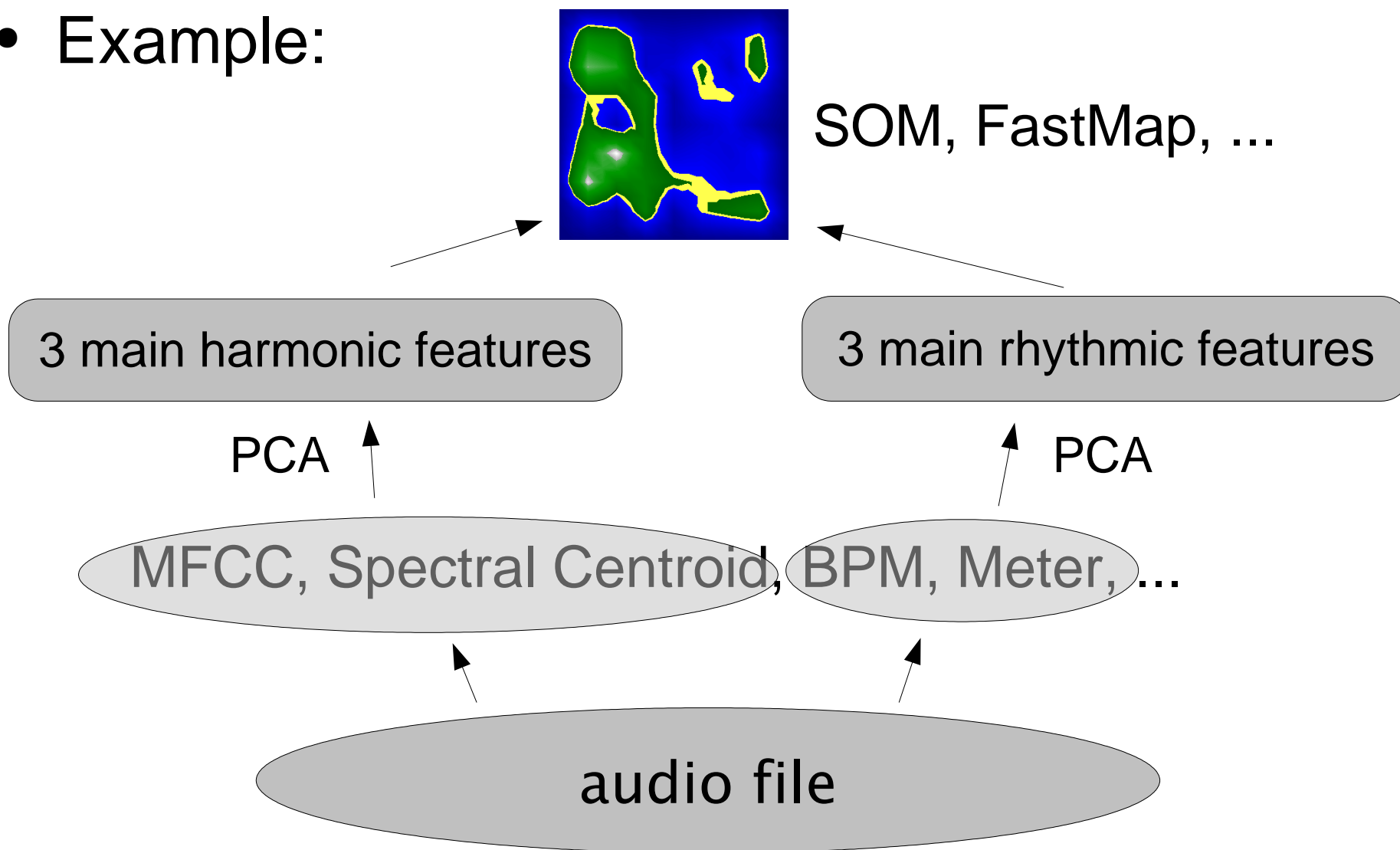
# Motivation

- Example:



# Motivation

- Example:



# Motivation

- There is no ultimate general algorithm to “understand” music
- Generalization: n levels

Musicology, AI

---

???

---

Machine Learning, Clustering

---

Signal Processing

# Research Proposal

- Establish a model of data-exchange between levels
- Define “processing objects” in the OOP sense.

# Research Proposal

- Feedback in decision is important -> allows to model the expectation
- Example: we estimate (using DSP) the pitch of a note with a probability 0.5. Knowing the previous notes, we may obtain better probabilities using melodic rules (using AI)

# Research Proposal

- If time allows:
  - Improve the visualization techniques  
exploration in 3d? new controls to browse the  
audio space? ...
  - Create new musically meaningful descriptors

# Conclusion

Thank You!

Any comments or questions are welcome...