ENGAGE

SALERO

Music Technology Group

Gunnar Holmberg
(Head of Project Coordination & Technology Transfer)
Pompeu Fabra University
Barcelona, Spain

http://www.mtg.upf.edu

Music Technology Group - MTG

- Research group focused on audio processing technologies and its music and multimedia applications created in 1994

- Key figures
  - 40+ researchers
  - Research projects since 2001:
    - 12 public funded
    - 15 private funded
  - 40+ publications / year
  - Annual income from research projects: 1,2 M€

MTG activities

- Balance between basic and applied research
- Technology transfer to industry
- Participation in international conferences and events

- MTG offers its experience to work in R&D projects, consultancy services in audio and music fields, and to license our technology

- Sectors that MTG addresses:
  - Professional audio, Musical instruments, Radio and broadcasting, Multimedia database management, Digital music libraries, Electronic music and music creation

Barcelona Music & Audio Technologies

- High-tech solutions for the digital entertainment industry
- “One-stop-shop" in the digital music field
- Spin-off of the Music Technology Group
- 28 researchers and engineers of international reputation with one of the best labs behind

Viver d’empreses de Barcelona Activa

C/ Llacuna, 162
Barcelona, 08018, Spain

Companies & Partners:

INDEX

- Assume that you have a home-entertainment system:

INDEX

- Assume that you have a home-entertainment system:

  - Connected to your TV- and Loudspeaker system

INDEX

- Assume that you have a home-entertainment system:

  - Connected to your TV- and Loudspeaker system
  - Connected to the Home-network/P2P/Internet

INDEX

- Assume that you have a home-entertainment system:

  - Connected to your TV- and Loudspeaker system
  - Connected to the Home-network/P2P/Internet
  - Connected to your portable devises (mobile phone/iPod/mp3-player) etc.

INDEX

- Assume that you have a home-entertainment system:

  - Connected to your TV- and Loudspeaker system
  - Connected to the Home-network/P2P/Internet
  - Connected to your portable devises (mobile phone/iPod/mp3-player) etc.

Seamless access to multiple channels of content:

Download/Upload - Consumer becoming a “Prosumer”

INDEX

- Tired of this?

INDEX

- **Tired of this?**

- What if it was possible for You to influence what was playbacked/broadcasted? A Radio & TV which adopts to Your taste, habits and moods...

INDEX

- Tired of this?

INDEX

- **Tired of this?**

- What if it was possible not only to retrieve - but also to adjust & personalize content?

INDEX

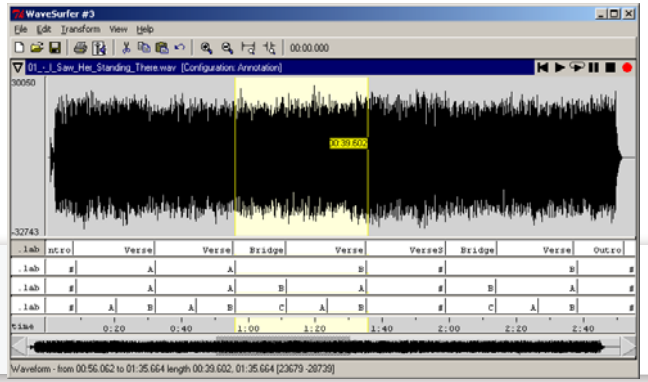
- **Tired of this?**

- What if it was possible not only to retrieve - but also to adjust & personalize content?
  - Time-scaling
  - Add/remove swing
  - Intelligent Skip (across & within songs)

INDEX

- **Tired of this?**

- What if it was possible not only to retrieve - but also to adjust & personalize content?
  - Time-scaling
  - Add/remove swing
  - Intelligent Skip (across & within songs)



INDEX

- **Tired of this?**

- Maybe you want to change your voice?

INDEX

- **Tired of this?**

- Maybe you want to change your voice?
- Or to adjust the volume of the singer?

INDEX

- **Tired of this?**

- Maybe you want to change your voice?
- Or to adjust the volume of the singer?

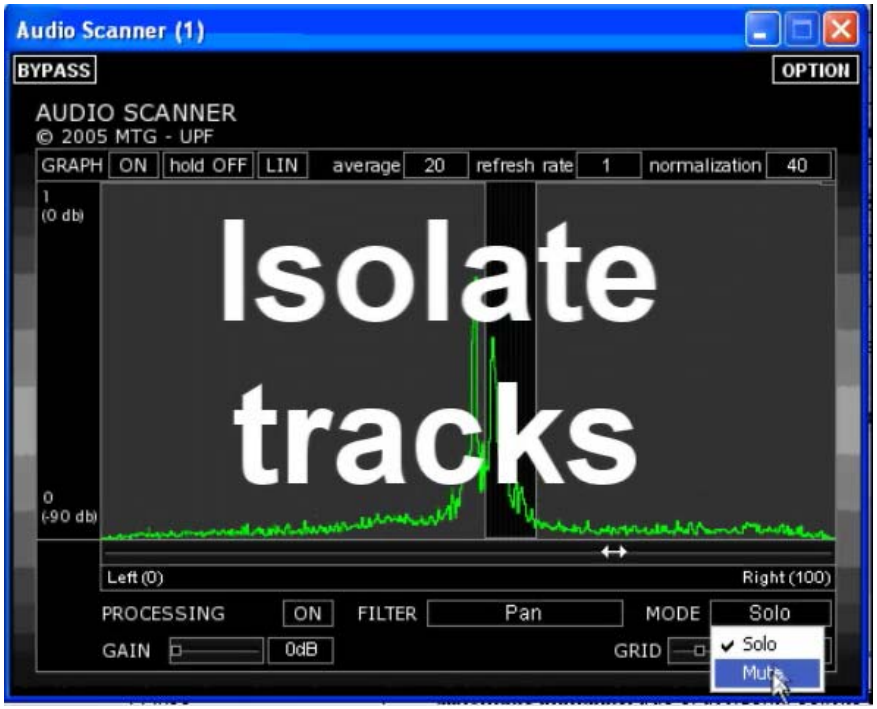
- Or even to change the guitar in a song, as to add more distortion?

INDEX

- How does it work?

INDEX

- How does it work?



INDEX

- MTG and its context
  - UPF, IUA, CIBM, ESMUC ...
- MTG Research topics
  - Voice and Audio Processing
  - Music and Audio Interaction
  - Music and Audio Analysis & Retrieval

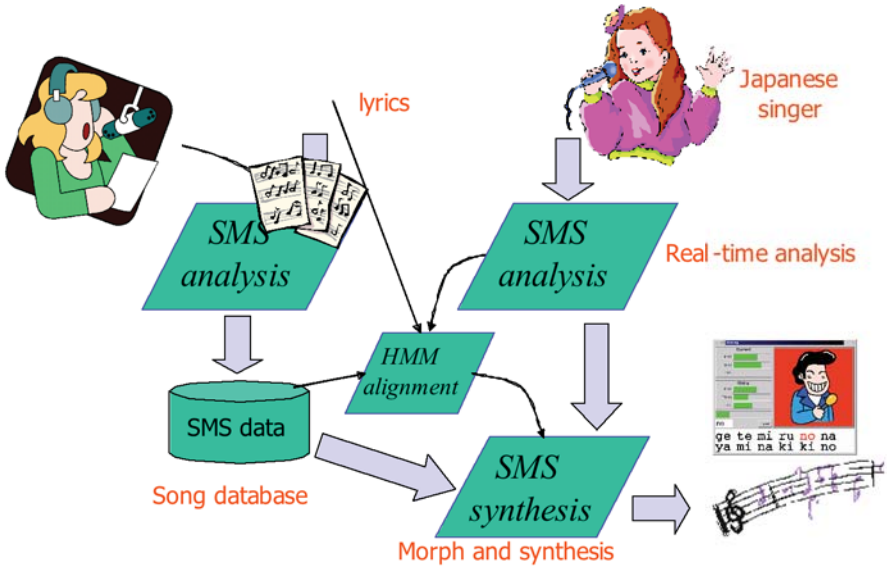
Voice & Audio Processing:

- Real Time Voice Transformation
- Singing Voice Synthesis
- New Instruments

Voice & Audio Processing:

- Real Time Voice Transformation
- Singing Voice Synthesis
- New Instruments

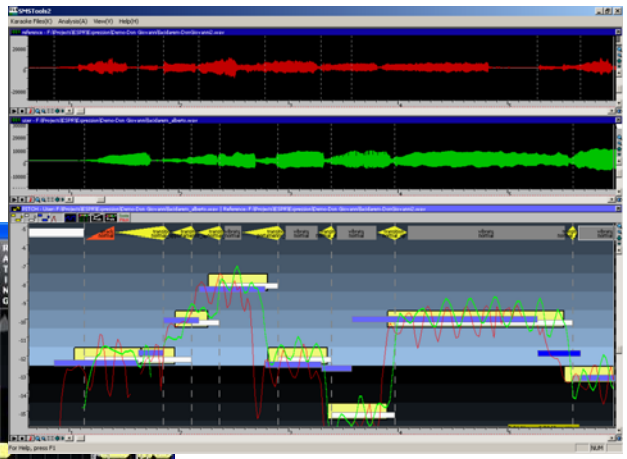
Elvis Project (1997-1999)



YAMAHA

Voice & Audio Processing:

- Rating a karaoke performance as a professional singer would do it.



YAMAHA

Voice & Audio Processing

- Real Time Voice Transformation
- Singing Voice Synthesis
- New Instruments

Voice & Audio Processing

- Real Time Voice Transformation
- Singing Voice Synthesis
- New Instruments

Daisy (1999-2003) turns into the VOCALOID

YAMAHA

Voice & Audio Processing:

- Real Time Voice Transformation
- Singing Voice Synthesis
- New Instruments

Voice & Audio Processing

- Real Time Voice Transformation
- Singing Voice Synthesis
- New Instruments

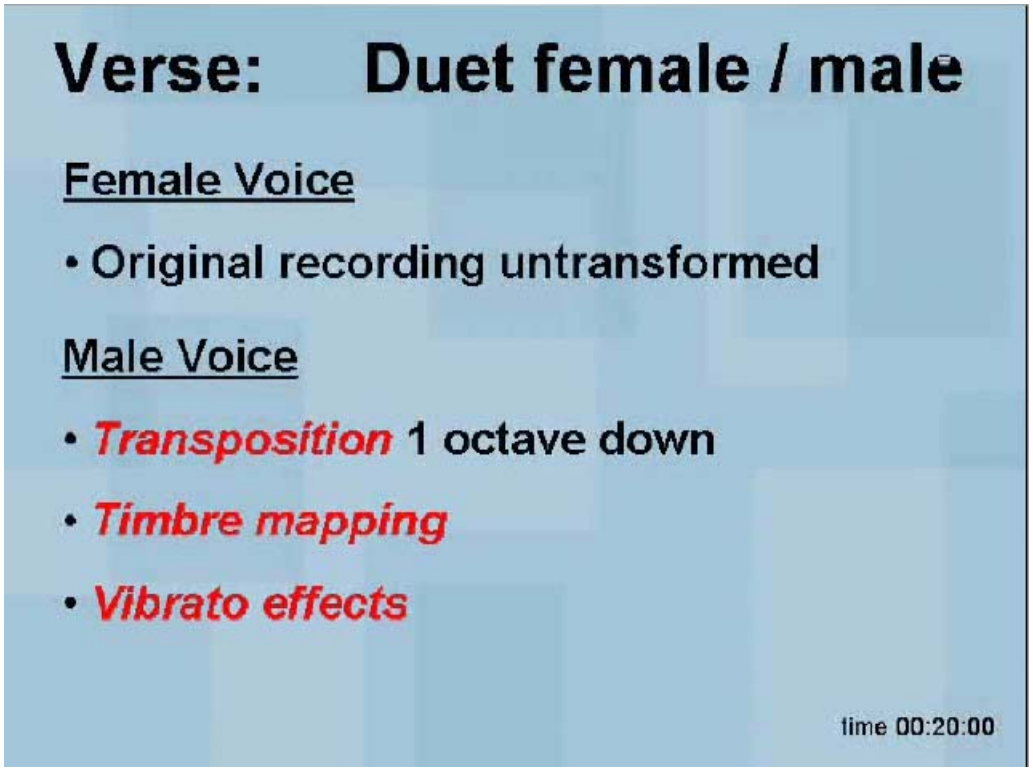
- Wanna to be a bass?

- How about playing the Trumpet?

- Solo to choir
  - Transform a single voice into a full choir

Voice & Audio Processing

- Transforming Female-to-male and Male-to-Female, or singing both at the same time?



INDEX

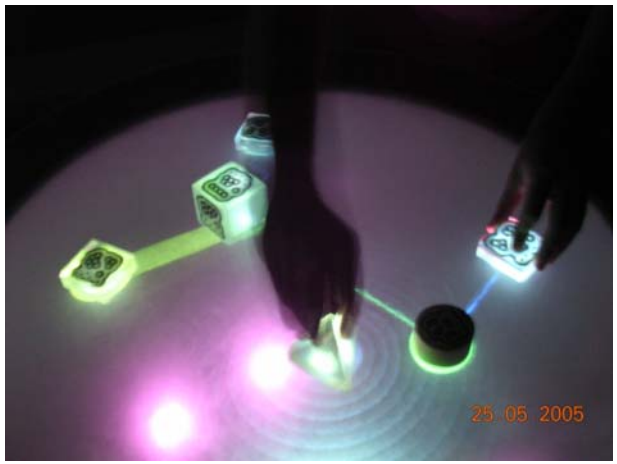
- MTG and its context
  - UPF, IUA, CIBM, ESMUC ...
- MTG Research topics
  - Voice and Audio Processing
  - New Music Interfaces
  - Music and Audio Analysis & Retrieval

New Music Interfaces

- Home DJ & Broadcast Yourself:
  - why should you be forced to invest in expensive hardware (keyboards, mixers etc) for music creation…

New Music Interfaces

- Home DJ & Broadcast Yourself:
  - why should you be forced to invest in expensive hardware (keyboards, mixers etc) for music creation…
  - when it could be enough with just a camera, and some ordinary objects?
  - The future of music creation applications will be in the hands of everyone (from amateur to pro)

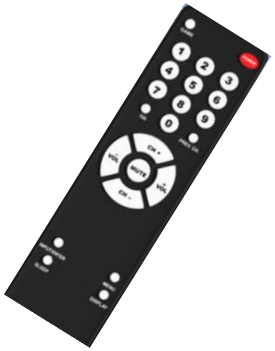


INDEX

- MTG and its context
  - UPF, IUA, CIBM, ESMUC ...
- MTG Research topics
  - Voice and Audio Processing
  - Music and Audio Interaction
  - Music and Audio Analysis & Retrieval

INDEX

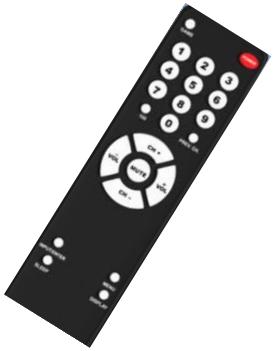
- **A Radio/TV that adjust to Your taste, habits and moods?**



INDEX

- **A Radio/TV that adjust to Your taste, habits and moods?**
- **Requires either (or all)**

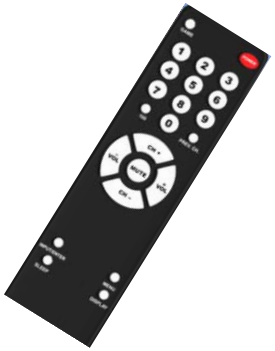
  before fully functional:



INDEX

- **A Radio/TV that adjust to Your taste, habits and moods?**
- **Requires either (or all)**

  before fully functional:

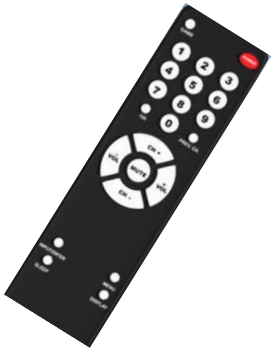


  - **Training (might take weeks/months)**

INDEX

- **A Radio/TV that adjust to Your taste, habits and moods?**
- **Requires either (or all)**

  before fully functional:

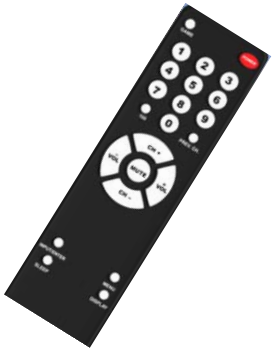


  - **Training (might take weeks/months)**
  - **Massive User-base**

INDEX

- **A Radio/TV that adjust to Your taste, habits and moods?**
- **Requires either (or all)**

  before fully functional:



  - **Training (might take weeks/months)**
  - **Massive User-base**

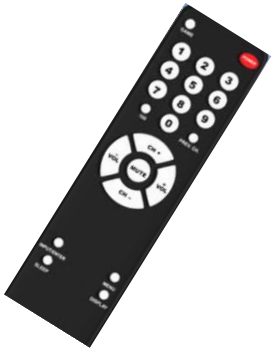
    Cold-start

    Scalability (ex. AMG)

INDEX

- **A Radio/TV that adjust to Your taste, habits and moods?**
- **Requires either (or all)**

  before fully functional:



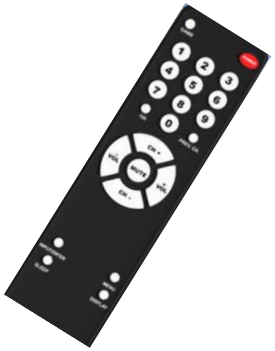
  - **Training (might take weeks/months)**
  - **Massive User-base**

    Cold-start

    Scalability (ex. AMG)
  - **Coverage (ex. Pandora/MusicGenome)**

INDEX

- **A Radio/TV that adjust to Your taste, habits and moods?**
- **Requires either (or all)**
  before fully functional:

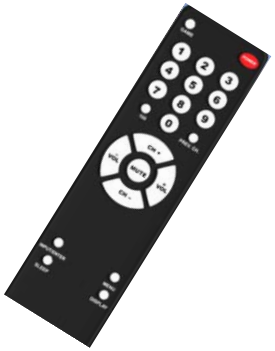


  - **Training (might take weeks/months)**
  - **Massive User-base**
    - **Cold-start**
    - **Scalability (ex. AMG)**
  - **Coverage (ex. Pandora/MusicGenome)**
  - **Novelty (new content)**

INDEX

- **A Radio/TV that adjust to Your taste, habits and moods?**
- **Requires either (or all)**

  before fully functional:

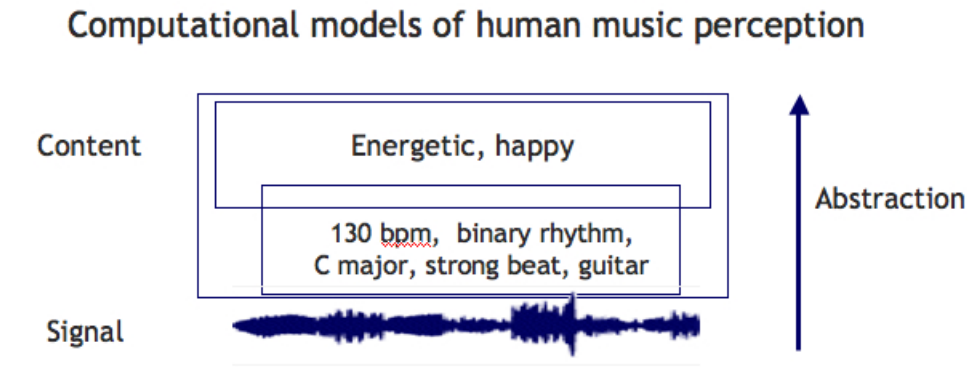


  - **Training (might take weeks/months)**
  - **Massive User-base**
  - **Cold start**
  - **Scalability (ex. AMG)**
  - **Coverage (ex. Pandora/MusicGenome)**
  - **Novelty (new content)**

EXCEPT IF

Music Understanding

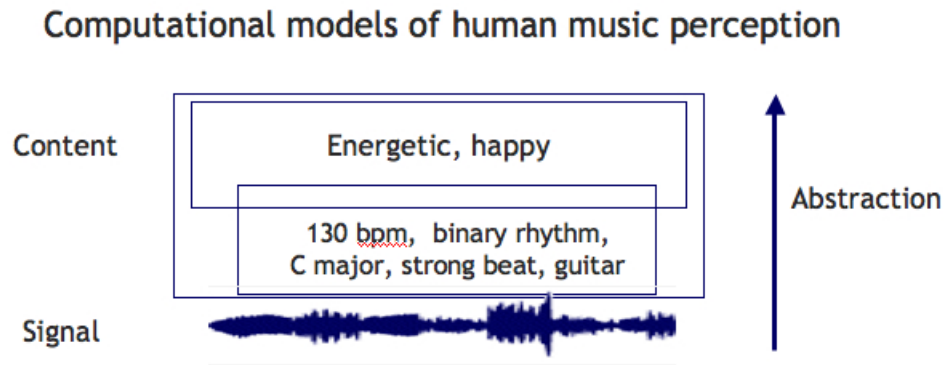
- System Based On Music Understanding:



Disciplines involved: signal processing, musicology, psychoacoustics, artificial intelligence and information retrieval.

Music Understanding

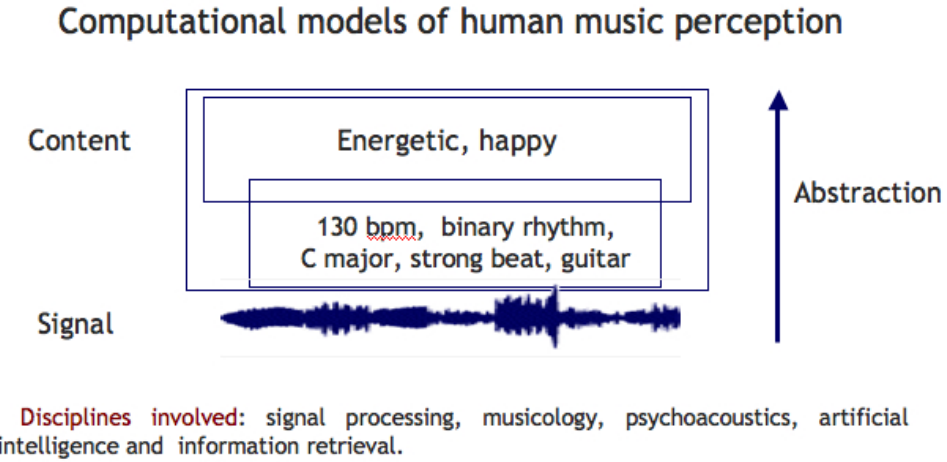
- System Based On Music Understanding:
- Covering Both:



Disciplines involved: signal processing, musicology, psychoacoustics, artificial intelligence and information retrieval.

Music Understanding

- System Based On Music Understanding:
- Covering Both:



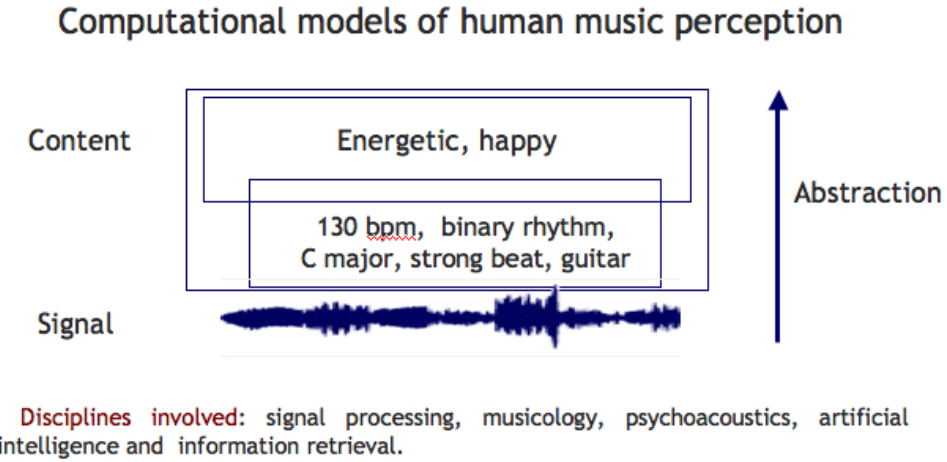
Disciplines involved: signal processing, musicology, psychoacoustics, artificial intelligence and information retrieval.

- human perception & cognition of audiovisual content

AND

Music Understanding

- System Based On Music Understanding:
- Covering Both:



Disciplines involved: signal processing, musicology, psychoacoustics, artificial intelligence and information retrieval.

- human perception & cognition of audiovisual content

AND

- human behavior and interaction surrounding content

Music Understanding
Advantage over competition

- Because of the first (system based on human perception & cognition of content)
- Our system is:

Music Understanding
Advantage over competition

- Because of the first (system based on human perception & cognition of content)
- Our system is:
  - fully functional Day 1 (plug-and-play)
  - large array of devices (mp3, mobile, iPod, desktop/web apps…)
  - connected AND local
  - full granularity (era/genre, artist, album, track… more like this!

Music Understanding
Advantage over competition

- Because of the first (system based on human perception & cognition of content)
- Our system is:
  - fully functional Day 1 (plug-and-play)
  - large array of devices (mp3, mobile, iPod, desktop/web apps…)
  - connected AND local
  - full granularity (era/genre, artist, album, track… more like this!
    But also: Mood, Playlist-type (party, relax…)

Music Understanding
Advantage over competition

- Because of the first (system based on human perception & cognition of content)
- Our system is:
  - fully functional Day 1 (plug-and-play)
  - large array of devices (mp3, mobile, iPod, desktop/web apps…)
  - connected AND local
  - full granularity (era/genre, artist, album, track… more like this!
    But also: Mood, Playlist-type (party, relax…)
  - fully adaptable (more/less danceable, more less
    energetic/calm, faster/slower, higher/lower perc.

Music Understanding
Advantage over competition

- Because of the first (system based on human perception & cognition of content)
- Our system offers:
  - Full costumization:
    - Store Your Own Favorite Settings

Music Understanding
Advantage over competition

- Because of the first (system based on human perception & cognition of content)
- Our system offers:
  - Full costumization:
    - Store Your Own Favorite Settings

Music Understanding
Advantage over competition

- Because of the first (system based on human perception & cognition of content)
- Our system offers:
  - Full costumization:

    Store Your Own Favorite Settings

    … and the system will find those kind even if you change the tracks in your player!

Music Understanding
Advantage over competition

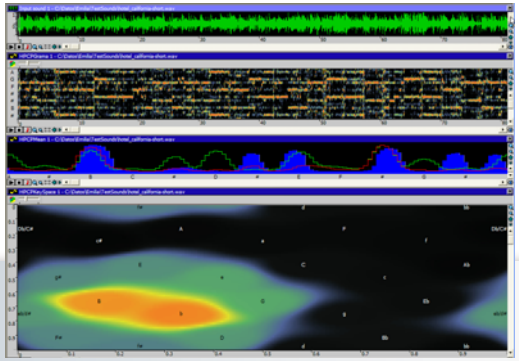
- Because of the first (system based on human perception & cognition of content)
- Our system offers:
  - Full costumization:

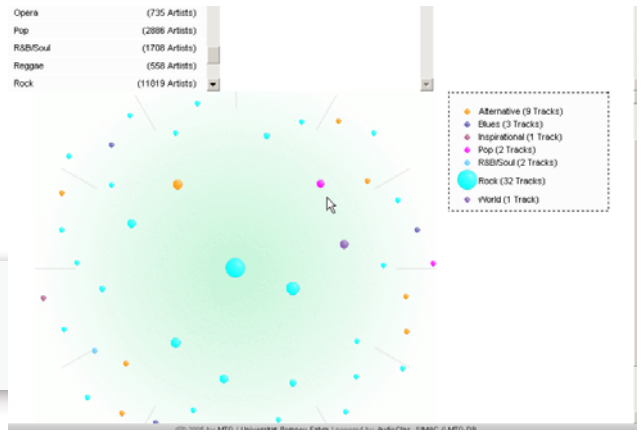
    Store Your Own Favorite Settings

    … and the system will find those kind even if you change the tracks in your player!

    Even the Recommender System itself (down to the very basis for item-selection) can be adjusted by the user:

    … more weight on rhythm, melody, timbre…





Music Understanding
Advantage over competition

- Because of the second (system also monitoring human behavior surrounding the music)
- Our system offers:
  - All Web2.0 functionalities (tagging, RSS, blogs…)
  - Linking any track/artist/album automatically with related content (album-art, artist photo galleries & biographies, news & blogposts, new releases, shopping…)
  - Linking User-to-User & User-to-community
  - Service&Content Discovery (concerts in your town - do you want to buy tickets online?)

BUT: monitoring require metadata
35% of all tracks are unrecognized

CDDB; LAST.FM…

Music Understanding

Audio Fingerprinting

bmat
hear the sound rise

- I like this song! What is it?

- We have developed a system which automatically recognize artist-album- track-name of the songs.

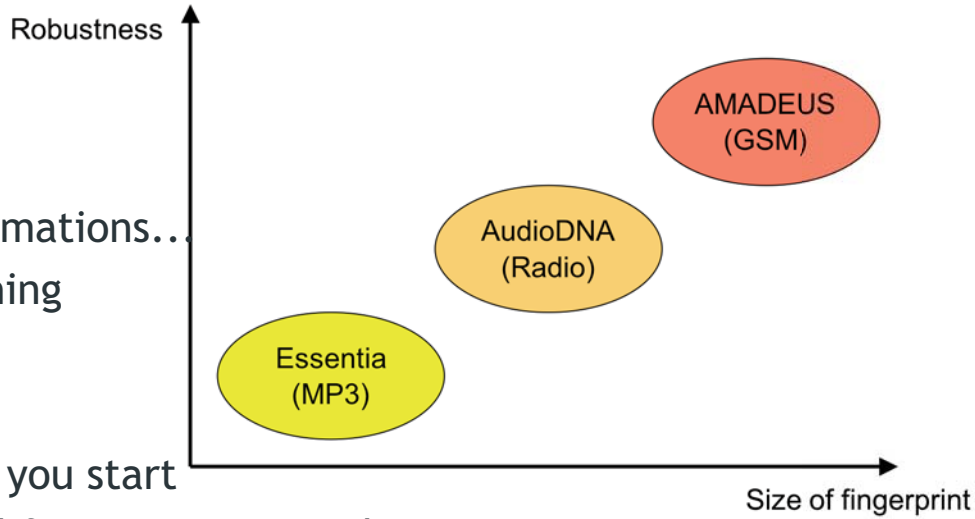
Use-cases:

- Mobile recognition
- Broadcast monitoring, DRM
- Mp3 metadata restoration

Music Understanding
Audio Fingerprinting

3 different systems developed:

RAA (2000-2002) Aida (2002-2004)



- Resistant to noise; background audio; speech over music; mastering transformations...
- Faster than real-time analysis & matching (even with large size databases)
- Works also on-the-fly (streams) - no matter where in the song/stream you start
- Only a few seconds of the song needed for accurate matching

Products & Services:

- Essentia:
  - C/C++/Symbian library for content based feature extraction.
- Mobile Services
  - Automatic Recognition
  - Automatic Playlist Generation
- Content Based Navigation
  - Car Audio Systems
  - MP3-players, iPods, etc.
- Home Entertainment Platform techologies

This presentation has been supported by:

Thank You!

Gunnar Holmberg
gholmberg@iua.upf.edu

Pompeu Fabra University
www.mtg.upf.edu