Can humans benefit from music IR?

Frans Wiering

v. 0.13, 26-7-2006

ICS Department of Information & Computing Sciences

Universiteit Utrecht

Aim and outline

- brief overview of Music Information Retrieval, in tutorial-like fashion
  - musical content
  - methods and techniques
  - retrieval tasks
- explore (mis)match between research in and use of Music IR
  - is music information?
  - how musical are humans?
- examine some consequences for interaction
 - opportunities for adaptive/multimedial approaches?

What is Music IR?

- Downie 2004:
  - a multidisciplinary research endeavor that strives to develop innovative content-based searching schemes, novel interfaces, and evolving networked delivery mechanisms in an effort to make the world's vast store of music accessible to all
- emerged in 1960s, maturing since late 1990s
  - ISMIR conferences (www.ismir.net)
- Research communities (Futrelle and Downie 2002)
  - Computer science, information retrieval
  - Audio engineering, digital sound processing
  - Musicology, music theory
  - Library science
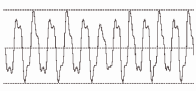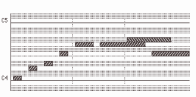 - Cognitive science, psychology, philosophy
  - Law

3

Approaches to Music IR

- Metadata based
  - artist, album, title
  - still the standard solution (e.g. iTunes)
- Content based
  - song text; music treated as blob
  - keyword (e.g. Pandora)
    - description of musical content or musical meaning is notoriously hard
  - internal features (Shazam, Themefinder)
    - frequency, pitch, beat, etc.
- Choice of approach depends on research community
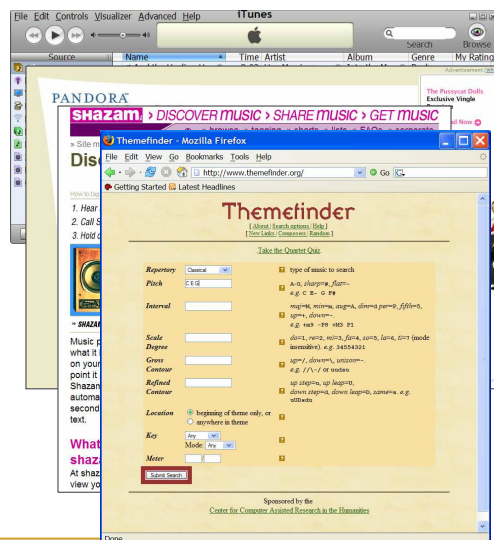
4

Basic representations of musical content

| musical content	example	compare image	compare text	structure	convert to above	convert to below
Digital audio (MP3, Wav)		level 1: primitive features	speech	none	-	hard
Time-stamped events (MIDI)		level 2: objects	text	little	easy	fairly hard (OK job)
Music notation (Finale, Sibelius, MusicXML)		level 2: compound objects	text + markup	much	easy (OK job)	-

5

Audio

- work with raw features
- detect *sound events*
- frequency domain
  - pitch-bands
    - MFCC
    - chroma
  - pitch (F0) estimation
    - solved
  - multiple F0 estimation
    - unsolved
- time domain
  - onset detection
  - rhythm, beat tracking
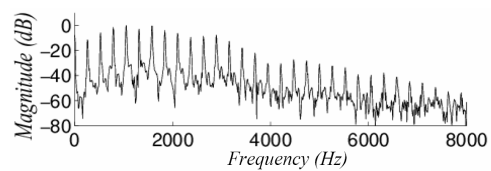
frequency spectrum (Klapuri 2004)
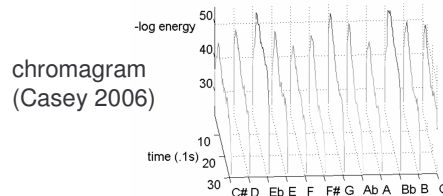
chromagram (Casey 2006)
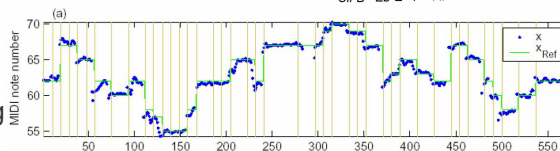
F0-estimation (Klapuri 2004)

6

3

Sound events

- sound events is what we primarily perceive in music
- perceptual properties of sound events
    - pitch:
        - low-high: c. 90 categories
    - duration
        - long-short: multiples of 2 and 3
        - 'quantization' into categories: rhythm
    - loudness
        - soft-loud; continuous
    - timbre, tone quality
- use of discrete pitches and regular pulse are important differences between music and speech (Fitch 2006)
- sound events must be organised to become music
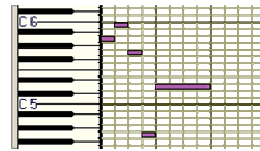
piano roll notation

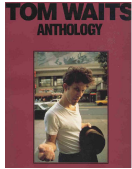www.museumspeelklok.nl

---

Notation

- music notation interprets and structures sound event categories
    - scalar and durational framework
    - successive sounds form *melodies*
    - simultaneous sounds form *chords* or *harmonies*
    - many other features too, e.g. *loudness*
- related to mental processing of music
    - musical meaning resides in groups, not in single notes
- music notation retrieval is not just for professional musicians

pitch

time

melody

Grave

O Tod, o Tod, wie bit - - ter, wie bit - -

f

chord

J. Brahms, Vier ernste Gesänge

Musical similarity

- central issue in music IR
  - connected to musical meaning
- many levels of musical similarity
  - many levels of retrieval
  - different retrieval methods

- similarity examples
  - detailed: performance, arrangement

  - generic: technique, genre, mood

9

Retrieval methods

- Audio data
  - fingerprinting
    - no pitch/rhythm detection
    - exact match only
  - chroma-based matching
    - finds musically similar passages
  - self-organising maps
    - clustering
    - musical genres

- Symbolic data
  - string-based methods
    - usually pitch-only
    - exact, substring, approximate matching methods
  - probablilistic methods
    - Markov chains; similarity derived from transition probabilities
  - set-based methods
    - usually pitch, duration, onset time
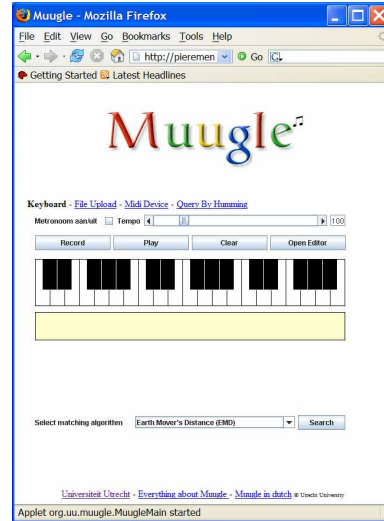   - geometric distance measures such as EMD/PTD, C-Brahms

10

Our demo engine: Muugle

- Musical Utrecht University Global Lookup Engine
- http://give-lab.cs.uu.nl/muugle
- framework
  - interfaces
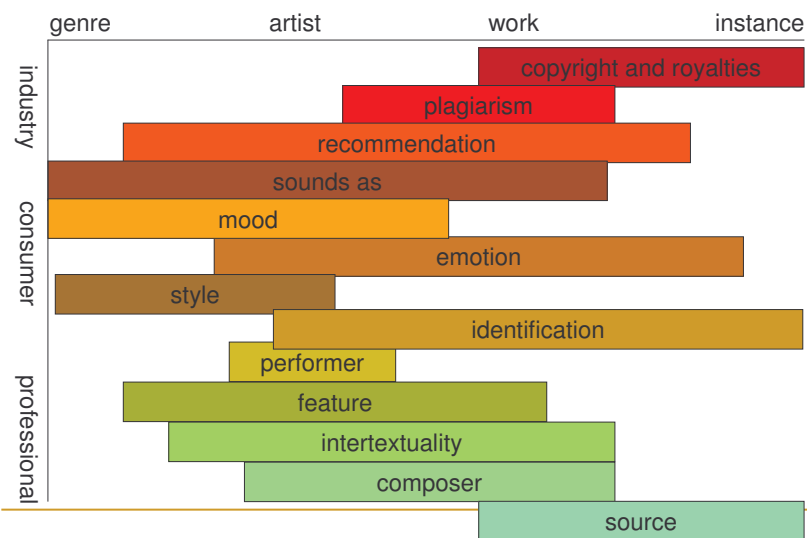 - matching algorithms
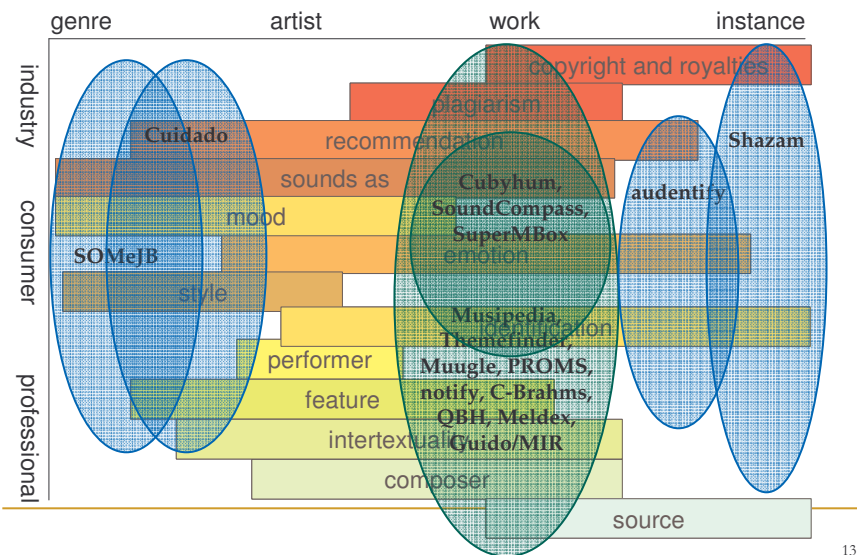 - collections (RISM, Alexa)

Muugle team (12-2005)

11

Retrieval tasks and systems

12

6

Retrieval tasks and systems

Content-based Music IR

- Fragmented
  - representations
  - methods and techniques
  - kinds of musical similarity
  - communities
- (semi) open areas
  - comparison, evaluation (since 2005: MIREX)
  - musical meaning
    - obvious semantic gap
    - what is musical similarity?
  - user studies
  - killer applications
- who wants musical information anyway?

Critique of (naïve) Music Information Retrieval

- It is assumed that music is information,
- that users search that information and
- that they are bad at expressing their information needs.
- This does not take perceptual and cognitive aspects of music into account and
- ignores other user needs (experience, emotion, context),
- which may require novel forms of interaction and
- integration of MIR techniques in more general applications.

15

Is music information?

- Information is best exemplified by 'functional prose' that has 'aboutness'
  - makes sense as a professional view of music
  - information about what? pseudo-information
  - can one separate document and information in music?
  - importance of surprise
- comparison to language
  - Fitch (2006) 9 design principles of music, compared to 13 design principles of language (Hockett 1960)
  - Important differences:
    - interchangeability: language +, music ?/-
    - music lacks features related to semantics
    - music is repeatable, used in performative context
  - music is like language without propositional, combinatorial meaning

16

Meaning of music

- Emotional
  - a-referentially expressive (Fitch 2006)
  - patterns of implication and realisation (Narmour)
  - influence on endorphin production
  - a-referential meaning sticks easy to text and image (films!)
- Social
  - mother-infant bonding
  - synchronises movement, dance
  - supports ritual, work, sports, shopping
  - group identity, coherence
  - entertainment, background
  - provides plausible evolutionary explanation
- many reasons for wanting to have (and search) music
  - more than just *information*
  - professional view obscures this

normal ending

evaded ending

dissonant ending

17

How musical are humans? (Peretz 2006)

- almost everyone has been exposed to music
- very many people have had musical instruction
  - nearly 50% in US and UK
- genuine amusicality is scarce (c. 4%)
  - often genetically determined
- 'nonmusicians'
  - can distinguish subtle style differences
  - are equally good at learning songs as professionals
  - are good at identifying out-of-tune notes
  - recognize patterns of implication and realisation
  - (have excellent beat-tracking abilities)
- Memory test   A   B

18

Case: Query by Humming

- well-studied strategy (since Ghias et al. 1995), many applications
  - user produces melody
  - system tells you what it is
  - Example: Query 🔊 Result 🔊
- users are generally not very good at query formulation
  - 'cannot sing in tune' or
  - 'cannot remember music correctly'
  - (non) interchangeability of music
  - compare face recognition
    - we are very good at recognizing faces
    - but generally not at drawing them

19

Musical needs

- **must a musical query always be a melody?**
  - more 'ecological' querying methods (Lesaffre 2005)
- **does QBH satisfy a widely-felt user need?**
- **user need not an information need but an experience need, composed of**
 - taste
  - mood, emotion
  - expertise
  - function

20

Some requirements for interaction

- evade problems of production
  - employ non-musicians' musical expertise: recognition, beat tapping, movement
- emotion-aware
- context-aware
- solutions to temporal nature and 'presence' of audio
  - common solution: 2D, 3D arrangement
- preserve expressiveness
  - hard if you translate music to space, movement and/or text

Emotion retrieval in music

- Research Micheline Lesaffre (IPEM, Ghent)
  - correlation of structural and affective/emotive descriptors
  - used in recommender system
- Extension (inspired by talk A. Hanjalic)
  - valence/arousal mapping
  - trajectories
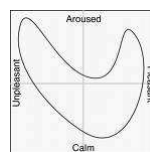
CHEERFUL
- **quick** (0,467**)
- **dynamic** (-0,519**)

SAD
- **soft** (-0,608**)
- **slow** (-0,658**)
- **static** (0 ,558**)
- **void** (-0,604**)
- **flowing** (-0,531**)

CAREFREE
- **slow** (0,298**)
- **dynamic** (-0,318**)

TENDER
- **soft** (-0,830**)
- **slow** (-0,659**)
- **static** (0,527**)
- **void** (-0,726**)
- **flowing** (-0,718**)
- **harmonious** (0,638**)
- **bright** (-0,543**)

AGGRESSIVE
- **hard** (0,731**)
- **quick** (0,513**)
- **dynamic** (-0,410**)
- **compact** (0,639**)
- **stuttering** (0,639**)
- **rough** (-0,612**)
- **dull** (0,510)

Presenting musical output

- General problem: how to create an overview of music
- Islands of music
  - Self-organising map Pampalk (2001)
- playlist generators
  - takes similarity scores as input
  - Van Gulik et al. (2005)
- Musicream
  - musical taps and magnets
  - Goto (2005)
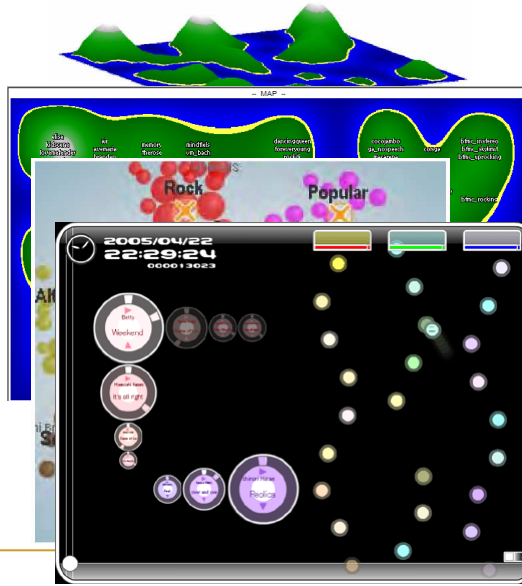
23

How can we improve on this?

- Pieces of music are represented as points
- Can they be given some more 'character'?
- Ideas that recently came up in a brainstorm
  - use facial expression
  - use shape, colour, texture of objects
- Possibilities
  - supports query formulation
  - can be spatially organised
  - makes orientation more easy
  - go listen to an object (or face)

24

Context-awareness

- Experiencing music is not just finding music
  - expect limited use of music search engines
- Cultural heritage applications may involve Music IR tools:
  - exploring musical heritage in context
    - Witchcraft project (Dutch ballads)
  - music as part of cultural context
    - choosing right background music
    - Medieval liturgical practice (idea: Saskia Rolsma)
- Use Music IR to support other types of retrieval
  - use steretyped musical features of Hollywood films
- Great future for Music IR as one of the tools within larger applications

25

The Witchcraft project

- Utrecht University and Meertens Instituut
  - 2006-2010
- Aim: to provide content-based access to field recordings of Dutch ballads
  - collected by Ate Doornbosch, 1957-1994
  - very rich in metadata ('Liederenbank')
- last traces of a song tradition (and a way of living) that has now almost completely disappeared
- Folksong researchers need melody retrieval
  - identify melodies when metadata do not help
  - find similar melodies: trace history, find tune families
- Considerable public interest in folksong
  - very successful website about 'straatliederen'
  - offers lots of context
  - searching could one way of exploring
  - 'under the hood' music IR for directing people to potentially interesting materials

26

Conclusion: can humans benefit from music IR?

- sure, but only if they are music professionals
- for other humans it's better to drop the I-word
- concentrate on needs that better correspond to the experience of music
- interaction part underdeveloped
- put music (I)R under the hood

27