Search engines for music and audio

Hugues Vinet, Scientific Director, IRCAM, France

Coordinator, CUIDADO, IST, FP5
Coordinator, SemanticHIFI, IST, FP6

Preliminary remarks

- Specificity of music information: combination of various complementary representation levels:
  - Audio signals,
  - Symbolic information : scores, lyrics, melodies and other musical patterns, piece structure, etc.
  - High-level metadata : editorial information, genres, etc.
  - > classical semanticWeb, ontology-based methods not adapted.
- Design methodology : required combination of the following approaches:
  - Top-down : elicitation of musical knowledge : *musicology,* development of *music cognition*;
  - Bottom-up : state-of-the art in automatic audio/ symbolic indexing; *digital signal processing;*
  - *Machine learning* : relationships between low- and high-level descriptors;
  - *Man-machine interfaces* : interactivity as the main means of convergence between the user's and the machine's knowledge.

SemanticHIFI : main related results

- Inter-document browsing
  - Main assumption : minimal set of pre-defined high-level descriptions
  - Automatically extracted global descriptors: tempo, timbre, voice presence, key, rhythmical structure, etc.
  - Personal classifications with generalization systems.
  - Various complimentary browsing heuristics : textual search, search by similarity, automatic playlist generation through global constraints, query by humming, musical summaries
- Intra-document browsing
  - Temporal segmentation : high-level piece structure (intro, verses, chorus,…), score and lyrics synchronization, etc.
  - Source separation : browsing within the orchestral polyphony, assisted mixing
- Peer-to-peer sharing
  - Publication/ search mechanisms of user-defined categories
  - Identification of related audio data through unique ID encapsuling
  - Adaptation to other users' metadata

Identified research directions

- Music cognition and perception : elicitation and formalization of musical knowledge from a reception viewpoint vs. production
- Indexing, automatic analysis
  - Audio indexing : segmentation, source identification and separation, transcription,…
  - Polyphonic pattern formalization, extraction, search from symbolic information
- HCI : visualization and interactive manipulation of various kinds of musical data.
- Peer-to-peer systems
  - Management of distributed knowledge production and emergence
- Generalization of databases as content-production techniques
  - Orchestration : combinations of individual sound samples in order to produce a sound target by superimposition
  - Concatenative synthesis (speech, music) :
    - Automatic segmentation of music pieces into small, indexed,units
    - Synthesis of a speech/ musical phrase by concatenation of various units according to global properties and concatenation rules

vinet@ircam.fr
http://shf.ircam.fr