The Europeana Sounds Music Information Retrieval Pilot

Alexander Schindler
alexander.schindler@ait.ac.at

Sergiu Gordea
sergiu.gordea@ait.ac.at

Harry van Biessum
hvbiessum@beeldengeluid.nl

Europeana Sounds - Motivation and Goals

- Emphasize on Europe's cultural audio heritage
- Aggregating content provided by 20 partner institutions
  - Digital libraries and audio archives
- Made publicly accessible through the Europeana portal

Music Information Retrieval Pilot

- Process the huge corpora of European cultural heritage
- Evaluating the applicability of technologies deriving from the MIR domain
- Implements a query-by-example functionality
  - Using audio-content based similarity search
- Preceded by an elaborated evaluation of the Europeana Sounds collection

Europeana Sounds Data

- 400,615 Meta-data descriptions
  - 28 Different languages
  - Varying granularities of quantity and quality
- 389,120 Referenced audio tracks
- 312,096 Processable audio files
- Distribution of Content
  - Music mostly
  - Spoken Word (e.g. interviews, old dialects)
  - Recorded radio broadcasts
  - Environmental Sounds

Audio-Content Descriptors

- Evaluated Feature Collection
  - Mel Frequency Cepstrum Coefficients (MFCC)
  - Statistical Spectrum Descriptors (SSD)
  - Rhythm Patterns (RP)
  - Chroma
  - Root Mean Square (RMSE)
  - Spectral Centroid (SPEC CENT)
  - Tempo
  - TONNETZ
  - Zero Crossing Rate (ZCR)
- Feature-Set Composition
  - 5 Acoustic categories:
    - Timbre, Rhythm, Harmony, Loudness, Noise
  - Empirically evaluated feature weights

Category	Feature	Description	f. W.	c. W.
Timbre	MFCC	Timbre description	23%	
	SSD	General spectral description	8%	
	SPEC CENT	Pitch description	8%	39%
Rhythm	RP	Rhythmic description	18%	
	BPM	Tempo	7%	25%
Harmony	CHROMA	Harmonic Scale	12%	
	TONNETZ	Traditional harmonic description	12%	24%
Loudness	RMSE	Loudness description	9%	9%
Noise Behaviour	ZCR	Noisiness description	3%	3%

Table 1: Overview of the audio-content *descriptors*, their corresponding acoustic *categories*, their assigned feature weight *(f. W.)* as well as the cumulative category weight *(c. W.)*.

Implementation

- Similarity Search
  - Nearest Neighbor search using Canberra Distance
  - Late fusion to combine feature spaces
  - Feature weighting to balance influence of single features
- Supported Use-Cases
  - Term-based queries on meta-data
  - Query by Example
  - Use external audio file to query for similar content in Europeana

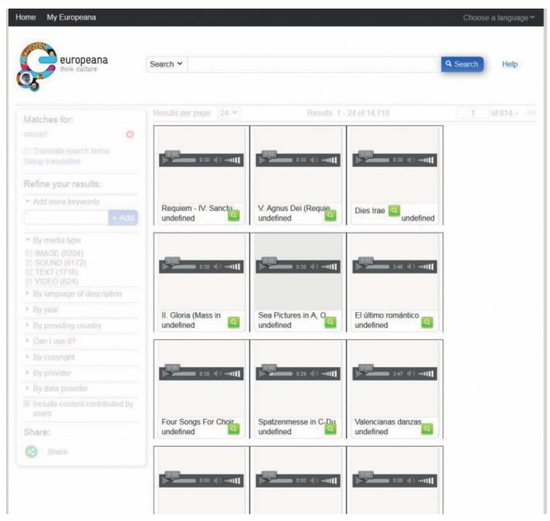

Figure 1: User interface of the Music Information Retrieval Pilot showing a result list of audio items which sound similar to the top-left entry.

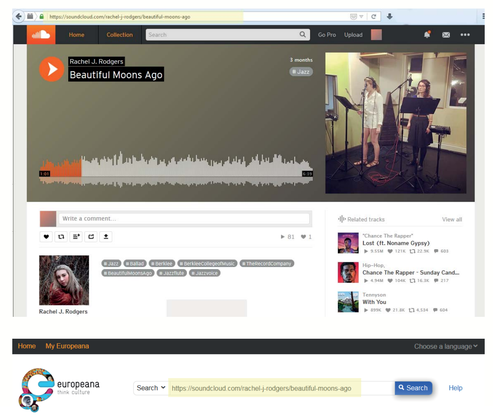

Figure 2: Use Soundcloud content to search for similar sounding audio tracks within the Europeana collection. Copy the URL of the Soundcloud entry and paste it to the search form of the MIR pilot.

Automatic Evaluation

- Semantic descriptive categories derived from meta-data
- Different granularities
  - *Classic quartet allegro major* (Genre / Tempo, Rhythm / Tuning)
  - *Jazz / Smooth Jazz* (Genre / Sub-Genre)
  - *Piano Concerto* (Instrumentation)
- Evaluated at different cut-off Points

Query	#	1	2	3	5	10
Jazz	31801	38,0	35,0	31,4	31,7	28,6
Smooth Jazz	2419	49,1	45,9	43,8	25,8	20,8
Ragtime	57	24,6	15,8	12,3	7,3	3,6
Classical	28569	44,3	42,1	40,5	38,3	35,1
Classic G major	304	17,1	14,8	14,0	12,6	9,3
Classic q.a.m.	191	9,4	6,3	7,3	8,1	5,6
Piano Concerto	510	38,6	32,0	28,0	23,9	17,6
Requiem	463	32,6	26,9	22,0	16,2	10,7
Opera	8278	26,8	24,7	22,7	21,1	18,9
Operette	1081	27,7	22,9	20,8	17,3	14,6
Flamenco	1827	40,7	33,0	29,2	24,3	18,2
Flamenco Guitar	287	22,3	17,1	15,3	13,5	10,0
Tarantella	152	33,6	28,0	22,4	16,1	8,5
Tango	3716	30,2	24,9	22,3	19,5	16,0
Animal Sounds	1097	89,7	87,0	85,1	82,8	78,7
Animal Sounds Crickets	113	59,3	55,3	56,6	53,0	48,1
Interview	484	77,5	74,3	72,0	68,6	60,8

Table 2: Precision values for the computational evaluation at cut-off points 1,2,3,5,10. Abbreviations: #: number of class items; *Classic q.a.m.*: Classic quartet allegro major.

User Evaluation

- 13 participants / 90min sessions
  - 3 different groups:
    music lovers, hobby musicians, music professionals
- Perception of the calculated similarities
  - overall similarity
  - specific characteristics (e.g. tempo, rhythm, harmonic, timbre)
- Low variance between groups
- Agreement upon timbre, instrumentation and harmony

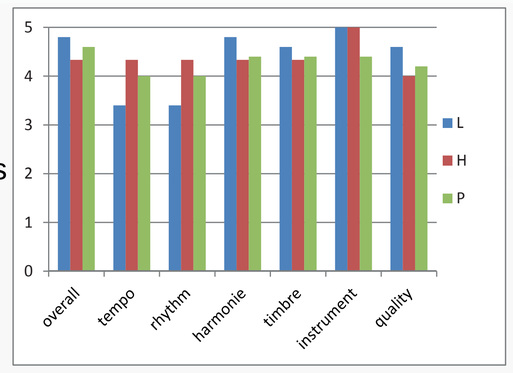

Figure 1: Quantitative evaluation of the similarity perception over different music properties, tailored by the user groups music lovers (L), hobby musicians (H) and music professionals (P).