The ISMIR Cloud: A Decade of ISMIR Conferences at Your Fingertips

Maarten Grachten, Markus Schedl, Tim Pohle & Gerhard Widmer
music@jku.at | http://www.cp.jku.at

Introduction

- Many available ISMIR publications nowadays
- Typical access to publications: by meta data or literal text search
- Our proposal: access by latent semantic structure of ISMIR corpus. Aims:
  - Facilitate concept oriented access to cumulative ISMIR proceedings
  - Provide an overview of significant ISMIR topics (and their evolution over the years)

Method

Latent semantic indexing:

- Build term list: noun phrases from publications, via part-of-speech tagging
- Filter term list: remove if corpus frequency < web frequency + ε
- Build term-document matrix $V$: term-frequencies per document

Nonnegative matrix factorization (NMF):

$V \approx W \cdot H$

$W_{ij}$: activation of term $i$ for concept $j$

$H_{jk}$: activation of document $k$ for concept $j$

$V$, $W$, and $H$ provide basis for analysis and interactive applications of the ISMIR publication corpus

Evolution of ISMIR topics over time

- 22 concepts extracted from ISMIR corpus using NMF
- Let $K$ be all documents for year $y$, then $\sum_{k \in K} H_{jk}$ is the total activation of topic $j$ in $y$
- This can reveal trends in popularity of ISMIR topics
  - e.g. Music-recommendation and annotation-based processing are booming; Less attention for QBH and drum-transcription since 2005

Web-Application: The ISMIR Cloud Browser

- Public access: http://www.cp.jku.at/projects/ISMIR-cloud
- Search ISMIR publications for (conjunctions of) terms
- Shows related terms and documents for query terms
- Shows which related terms occur in which document
- Links to online pdf documents

Implementation:

- Term cloud representation via PCA on submatrix of $V$
- Term size computed as concept co-activation ($W$) with query
- Related documents: the $k$ documents from corpus $D$ with highest ‘concept cosine similarity’ to query $Q$:
  $$\arg \max_{x \in D} T_{lq} \cos(W_{q} \cdot H_{d})$$

ISMIR 2009, 10th International Society on Music Information Retrieval Conference, 26th – 30th October 2009, Kobe, Japan