Audio Fingerprinting

mainly based on

A review of audio fingerprinting (Cano et al. 2005)

presented by

Denis Lebel
Presentation Outline

- Introduction
- Desired Properties
- Usage Modes
- Applications
- Fingerprinting Framework
  - Front-end
  - Fingerprint Models
  - Similarity Measures and Searching Methods
  - Hypothesis Testing
- Conclusion
- References
Introduction

- **Idea**
  - An attempt to mimic human music recognition abilities
- **Audio Fingerprint**
  - Unique identifier of an audio signal
  - Content-based signature that summarizes an audio recording
  - Uses relevant (perceptual) acoustics characteristics of signal
- **Fingerprinting System**
  - Database of known fingerprints
  - Query system

Figure 1: General idea of a fingerprinting system
Desired Properties

- Accuracy
  - Function of correct, missed, and wrong identifications

- Reliability
  - Correct identification method

- Robustness
  - Ability to accurately identify an item (no matter how compressed or distorted it is)

- Granularity
  - Ability to identify a signal from a short excerpt

- Security
  - Vulnerability to cracking
Desired Properties

- **Versatility**
  - Ability to identify a signal regardless of audio format

- **Scalability**
  - Performance with very large databases

- **Complexity**
  - Computational costs of fingerprint extraction, size of fingerprint, search complexity, comparison complexity, etc.

- **Fragility**
  - Integrity verification (detection of changes in content)
Desired Properties

- Properties are interrelated and dependent on system purpose.

- Generally speaking, fingerprint should be:
  - A perceptual digest of the recording
  - Invariant to distortions
  - Compact
  - Easily computable
Usage Modes

- **Identification**
  - Content identification of an audio signal

- **Integrity Verification**
  - Detection of data alteration

Figure 2: Content-based audio identification framework. (Cano et al. 2005)

Figure 3: Integrity verification framework. (Cano et al. 2005)
Usage Modes

- **Watermarking support**
  - Audio fingerprints can be used to derive secrets keys from the audio content

- **Content-based Audio Retrieval and Processing**
  - Extraction of audio features (i.e., low-level and high-level descriptors)
  - Fingerprints can be used to retrieve similar content (i.e., query-by-example scheme)
Applications

- Audio Content Monitoring and Tracking
  - At the distributor end
  - At the transmission channel
  - At the consumer end

- Added-Value Services
  - Content information describing audio excerpt (e.g., tempo)
  - Meta-data describing musical work (e.g., composer, year, ...)
  - Other information (e.g., album cover)

- Integrity Verification Systems
  - Audio fingerprints can be used to ensure user’s audio files have the best quality available
Presentation Outline

- Introduction
- Desired Properties
- Usage Modes
- Applications
- Fingerprinting Framework
  - Front-end
  - Fingerprint Models
  - Similarity Measures and Searching Methods
  - Hypothesis Testing
- Conclusion
- References
Fingerprinting Framework

Figure 4: Content-based audio identification framework. (Cano et al. 2005)
Fingerprinting Framework

- Fingerprint Extraction: Front-End

Figure 5: Fingerprint Extraction Framework. (Cano et al. 2005)
Fingerprinting Framework

- Fingerprint Extraction: Fingerprint Modeling
  - Idea: Reduce redundancies
  - Reduce size of fingerprint
  - Similarity measure and search method depends on the model chosen
  - Several techniques can be used (for a summary: Cano et al. 2005)
Fingerprinting Framework

Figure 4: Content-based audio identification framework. (Cano et al. 2005)
Fingerprinting Framework

- Fingerprint Extraction:
  - Similarity Measures
    - Related to type of model chosen
    - Correlation metric is common
      - Example: Euclidean distance

Figure 6: a) Fingerprint block of original clip
b) fingerprint block of a compressed version.
c) Difference (error) (Haitsma et al. 2002)
Fingerprinting Framework

- **Fingerprint Extraction: Searching Methods**
  - Using brute-force search is inappropriate for large database
  - **Idea**: Optimizing the search

- **Some possible optimizations**
  - Pre-computing distances offline
  - Filtering unlikely candidates with a cheap similarity measure
  - Candidate pruning
  - Others...
Fingerprinting Framework

- Fingerprint Extraction: Hypothesis Testing
  - **Idea**: Whether the query is present in the repository
  - A threshold *must* be used and it depends on:
    - Fingerprint model
    - Similarity of fingerprints in the database
    - Database size
    - Discriminative information of the query
  - The larger the database, the higher the probability of wrong match
    - False Acceptance Rate (FAR)
    - False Rejected Rate (FRR)
Conclusion

- Most existing systems fall more or less into this generic framework
- Large databases still represent a challenge (scalability, complexity, accuracy...)
- P2P systems might be the future (e.g., Music2Share)
References


Links

- http://www.shazam.com/
- http://www.relatable.com/
- http://www.audiblemagic.com/
- http://www.gracenote.com/