Content-based Representations, Indexing and Retrieval of Music

1. Research Team

Project Leaders:	Prof. Shrikanth Narayanan, *Electrical Engineering* Prof. Elaine Chew, *Industrial Systems Engineering* Prof. C-C. Jay Kuo, *Electrical Engineering*
Graduate Students:	Hui-Yun Kao, Ming-Chang Lee, Arpi Mardirossian, Hsuan-Huei Shih, Erdem Unal, Xiaodan Wu
Undergraduate Students:	Ahmet Certen

2. Statement of Project Goals

The proposed research involves development of methods for content-based indexing of music databases using a combination of signal processing and knowledge-based methods, design of statistical algorithms for enabling queries using sung or hummed melodies, and design of robust search techniques for retrieving the queried information especially in the presence of uncertainty. The research approach is user-centric, and can be subdivided into three major topics:

- Representation and Indexing: Algorithms for organizing musical information utilize representations that exploit theories and knowledge about human musical intuition, and music perception and cognition.
- Query Formulation and Interaction Modality: The user interacts with the music database through humming, a natural activity. Algorithms mapping the user's query to music symbols are designed to cater to all levels of musical abilities. Expert knowledge of the desired piece of music is not required.
- Search and Retrieval: Algorithms for retrieval i.e., matching user query against the database, will tolerate and be robust to uncertainties and errors. Such errors occur in the user's query and can also be introduced in the audio signal-to-symbol transduction.

3. Project Role in Support of IMSC Strategic Plan

Content-based multimedia data mining is a rapidly emerging research area. Enabling natural interactions with multimedia information, accommodating a variety of user skills and preferences, is a critical element of such efforts. Music in digital form is an important component of the information explosion that we are witnessing in present times. There is a great need for developing methods and technologies for organizing music databases, and facilitating easy and efficient interactions with digital music information.

4. Discussion of Methodology Used

The proposed architecture comprises a front-end recognizer, that converts the humming signal to note using a statistical pattern recognition approach, which interfaces with a back-end music

database, that is indexed using perceptually viable features. The search process matching a query against indices is formulated as a statistical information retrieval problem. The proposal aims to employ progressively rich indexing representation including repeating patterns for recurring themes, chord, beat and key information. Derivation of such music-theoretic knowledge will benefit from a principled approach of mathematically modeling tonality in music. The statistical framework allows for handling variability and uncertainty in query formulation and retrieval. It also enables providing for quality of solution in the query results.

5. Short Description of Achievements in Previous Years

Development of a statistical pattern recognition based front end for converting humming to notes.
Collect a new humming database for humming recognition.

.

5a. Detail of Accomplishments During the Past Year

Our work on humming recognition last year focused on a new phone level hidden Markov model approach applied to human humming transcription. A music note has two important attributes, i.e. pitch and duration. The proposed system generates multidimensional humming transcriptions, which contain both pitch and duration information. Query by humming provides a natural means for content-based retrieval from music databases, and this research provides a robust front-end for such an application. The segment of a note in the humming waveform is modeled by phone level hidden Markov models (HMM). The duration of the note segment is then labeled by a duration model. The pitch of the note is modeled by a pitch model using a Gaussian mixture model. Preliminary real-time recognition experiments are carried out with models trained by data obtained from eight human objects, and an overall correct recognition rate of around 84% is demonstrated.

Based on previous humming data collection experiences in the past years, a protocol of humming data collection was proposed. Three kinds of humming pieces, which are an ascending-to-descending music scale, ten out of the twenty-two carefully selected music pieces, and three music pieces of free choice, are collected from each participant. The collecting task is still being carried on. The first 60 participants' humming data will be available by this summer.

6. Other Relevant Work Being Conducted and How this Project is Different

The proposed work falls in the general area of content-based indexing of multimedia information [8]. Specifically, we consider progressive incorporation of information obtained from music theory and cognition into designing algorithms for indexing and interactively searching music databases.

Research in Music Retrieval has gained momentum in recent years. The Computer Music Project at Carnegie Mellon University (CMU-url) and the MusicEn Project at the University of Michigan (UMICH-url) are large-scale ongoing projects in this area. Other research centers such as Stanford's Center for Computer-Assisted Research in the Humanities (CCARH-url) also supports research in computer modeling of music.

Our research framework differs from previous and current efforts in this area in that we aim to build a user-centered system that takes a piece of musical information in the form of the user's humming through the conversion-to-note process to retrieval of similar matches from a database that need not be only melodic-based. As such, we will tackle problems of music encoding and retrieval under uncertainty. Both statistical methods and others based on models of tonality will be explored.

7. Plan for Next Year

Plan for next year include completion of humming database creation and organization, statistical analysis of music corpora, and implementation of a music information retrieval system using hummed queries.

8. Expected Milestones and Deliverables

- Algorithms for repeating pattern extraction.
- Initial humming recognition system.
- Refinement of indexing representation by music theoretic models.
- Improved front end.

9. Member Company Benefits

N/A

10. References

[1] USC: "USC Query by Humming project homepage," URL: http://sail.usc.edu/music/.
[2] H.-H. Shih, S.~S. Narayanan, and C.-C.~J. Kuo, "A dictionary approach to repetitive pattern finding in music," in 2001 IEEE International Conference on Multimedia and Expo (ICME2001), August 2001.
[3] H.-H. Shih, S.~S. Narayanan, and C.-C.~J. Kuo, "Music indexing with extracted main melody by using modified lempel-ziv algorithm," International Symposium on the Convergence of IT and Communications (ITCOM 2001), August 2001.
[4] H.-H. Shih, S. S. Narayanan, and C.-C. J. Kuo, "Comparison of dictionary-based approaches to automatic repeating melody extraction," in IS&T/SPIE's Symposium on Electronic Image 2002, 2002.
[5] H.-H. Shih, S. S. Narayanan, and C.-C. J. Kuo, "A statistical approach to humming recognition," In 2002 International Conference on Acoustics Speech and Signal Processing (ICASSP2002), May 2002.
[6] H.-H. Shih, S. S. Narayanan, and C.-C. J. Kuo, "An HMM-based approach to humming transcription," in 2001 IEEE International Conference on Multimedia and Expo (ICME2002), August 2002.
[7] H.-H. Shih, S. S. Narayanan, and C.-C. J. Kuo, "Multidimensional Humming Transcription Using a Statistical Approach for Query by Humming Systems ," In 2003 International Conference on Acoustics Speech and Signal Processing (ICASSP2003), April 2003.

[8] T.~Zhang and C.-C.~J. Kuo, “Content-based Audio Classification and Retrieval for Audiovisual Data Parsing,” Kluwer Academic Publishers, 2001.
[9] CMU: “Music search and retrieval at cmu,”
URL: http://www-2.cs.cmu.edu/afs/cs.cmu.edu/project/music/web/search/.
[10] UMICH: “MUSEN project at the University of Michigan,”
URL: http://musen.engin.umich.edu/.
[11] CCARH: ”CCARH at Stanford,” URL: http://www.ccarh.org.