

# Content-based Representations, Indexing and Retrieval of Music

Hsuan-Huei (Maverick) Shih  
 Prof. Shrikanth Narayanan  
 Prof. C.-C. Jay Kuo  
 Prof. Elaine Chew

## Research Goal

Content-based indexing of music databases using:

- signal processing and knowledge-based methods
- statistical algorithms for query by singing or humming
- robust search techniques for retrieval under uncertainty

The research approach is **user-centric**.

The **THREE MAIN TOPICS** are:

- **Representation and Indexing**
- **Query Formulation and Interaction Modality**
- **Search and Retrieval**

## Role in IMSC

There is a great need for developing methods and technologies for organizing music databases.

Our research agenda will facilitate easy and efficient interactions with digital music information.

### Reasons for pursuing such research at IMSC:

- 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
- Applications to communications and entertainment elements of IMSC's vision



## Accomplishments

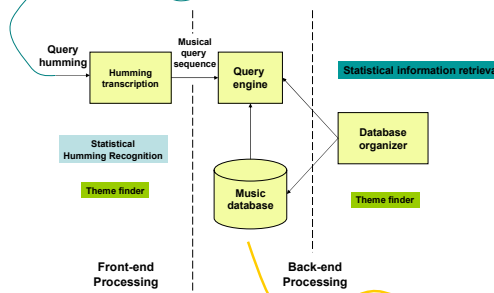
A **System** for Multidimensional Statistical humming recognition

An **NSF-IJR Award (2002)**: A User-centric Content-based approach to Indexing, Query and Retrieval of Music through Signal Processing and Knowledge-based Method  
[www-courses.usc.edu/engr/ise/599musicog](http://sail.usc.edu/music)

A **new course**: ISE599 Engineering Approaches to Music Perception and Cognition  
[www-courses.usc.edu/engr/ise/599musicog](http://www-courses.usc.edu/engr/ise/599musicog)

**Publications**: ICME 2002, ICME 2003, ICASSP 2002, ICASSP 2003  
<http://sail.usc.edu/music/people/publication.html>

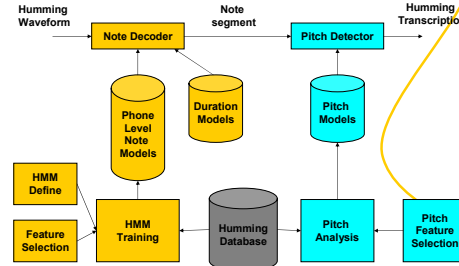
## Research Approach



Multidimensional Statistical humming recognition:

- A humming note has two important attributes which are duration and pitch
- The recognition processing is divided into two stages which are note segmentation stage and pitch detection stage
- Phone-level hidden Markov models with features modeled by Gaussian mixture models which we call Note Models are used at note segmentation stage
- Pitches modeled by Gaussian mixture models which we call Pitch Models are used at pitch detection stage
- During the training phase, note models and pitch models are trained with humming data obtained from real users and currently the database contains 100 user's samples.
- During recognition, the incoming piece of humming waveform is segmented using the trained note models, then the segmented note's pitch is detected using trained pitch models

- Recognition results were improved
- An N-gram based music language model was implemented and applied
- Experimental results showed about 80% accuracy rate on note recognition



## Uniqueness & Related Work

- 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.
- The Computer Music Project at Carnegie Mellon University and the MusicEn Project at the University of Michigan are large-scale ongoing projects in this area. Other research centers such as Stanford's Center for Computer-Assisted Research in the Humanities also supports research in computer modeling of music.

## Five-Year Plan

2002-2004	2004-2006	2006-2008
Algorithms for repeating pattern extraction. Initial humming recognition system.	Refinement of indexing representation by music theoretic models. Improved front end.	Statistical information retrieval using hummed inputs. Usability experiments.

