Query by Humming Project Proposal

DELTA Project
Raymond Kozikowski
2/15/04

Abstract

The purpose of this project is to develop a "Query by Humming" music recall system. The system will catalog songs in a database and develop "keys" to match the inputted query against. The primary purpose of this project in the home is as a luxury item. However, this system will serve as a testbed for voice recognition technology as well as its interaction with cataloging systems so it may be implemented in other areas of the home. This may serve to greatly increase access for disabled people in the home. The database will be built using Berkeley db, the keys will use MIDI music format and the processing will be done using Matlab.

Technical Overview

There are essentially five areas of development for this system.

1. Database- implemented using Berkeley db or another basic databasing program. This facilitates easy programming of a searchable database based upon musical keys. Much of this programming will either be done through the Computer Science department or an intensive study of the programming language is required.
2. Musical Keys- Search items associated with each databased song. Associated with any musical type. There are five sub-steps in this category.
    a. Convert the file to MIDI format
    b. Remove Polyphonic MIDI channel 10. This removes percussion from the MIDI stream because the voice cannot match percussion sounds.
    c. Extract Theme. This basically generalizes the song into a simplified version. Possibly taking the polyphonic MIDI to monophonic.
    d. Convert to Matlab format. Allows for representation of music as numerical array.
    e. Create differential sequence array. Create an array based upon the difference between notes. If the note is higher than the last, a "U" is assigned, if lower a "D" is assigned, and if the same an "S" is assigned.
3. Searching and Error Checking- Find the best way to search the keys and do error checking so that if a note is off, it will not throw off the entire sequence. Error may include adding note, dropping notes or duplicating notes. Once again, the Computer Science Department must be consulted in order to determine possible algorithms to implement. It should also be a system that will work regardless of how the database works and what types of keys are used.
4. Input File- voice rendering of the desired song. This will be treated similarly to the key files.
5. U, D, S coding system- designed to remove many of the undesirable features from music samples and keys. This system removes Key, Note length, and Beat from the sample. This means that any of these factors may be off when inputting a song so long as the relationships between notes remain constant. Disadvantages include

that each note must be clearly defined and while this may be easy for the input files, it will be hard to do for the databased files as the average time between notes is on the order of milliseconds.

Design and Implementation

Present to March 5- background research and fleshing out steps in creating system

March 5 to end of April- Development of keys, research database programming
Summer- test key association with inputs

Fall 2004- build database, searching software, and error checking software. Streamline processing for key files.