Exploring Automatic Chord Recognition Algorithms

Shan Anis
University of Rochester

ABSTRACT

Automatic Chord Recognition

The Task:

- Divide an audio file containing music into frames.
- Assign each frame a chord label according to the analysis of the content.

Advantageous for applications involving Music Information Retrieval e.g. Automatic Music Transcription.

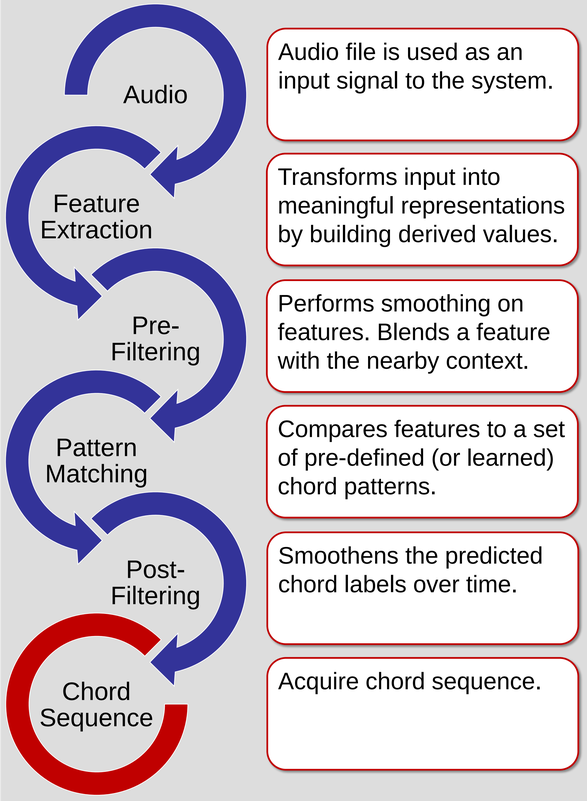
The objective of this project is to analyze current developed Automatic Chord Recognition Systems, performing a systematic evaluation to analyze the different stages of the typical systems.

Our aim is to better understand each stage of the system and propose new directions to potential improvements.

CONTACT

Shan Anis
University of Rochester
Email: sanis@ur.rochester.edu

SYSTEM FRAMEWORK



Chroma Features were used for feature extraction.

Analysis of a musical chord includes the pitch class (Chroma) instead of the absolute pitch (octaves are irrelevant).

Pattern matching approach includes the Hidden Markov Models (HMM) in a supervised training environment.

Chroma vectors were used to train the HMM with one state for each distinguished chord.

SYSTEM IMPLEMENTATION

Chroma Features:

Chroma vector sums up spectral energy into 12 frequency bins corresponding to the 12 pitch classes.

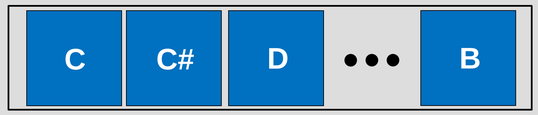
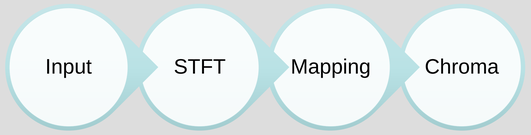


Figure 1: Spectral energy of the 12 pitch classes summed up in 12 bins.

Input is divided into overlapping frames, converted to a short-time Fourier transform (STFT) representation and then mapped to the Chroma features.



Hidden Markov Models (HMM)

In HMMs, the system being modeled is assumed to be a Markov process with hidden states.

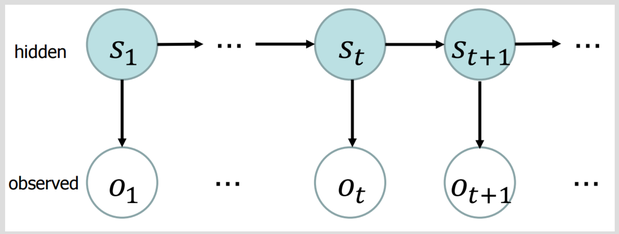


Figure 2: An HMM consists of hidden and observed layers.

An HMM can be represented by its: Initial Probability, Observation Probability and a Transition Probability.

Chords are considered as a Hidden State and Feature Vectors are considered as the Observation.

The Transition Probability is trained from the training data.

DISCUSSION

For pattern matching, template based methods exist which include pre-computed feature templates.

These hand annotated templates provide the required information but creating them is very time consuming.
Therefore we use probabilistic chord models to automate the task.

Furthermore, Chroma features were proven to give the best results as compared to other features.

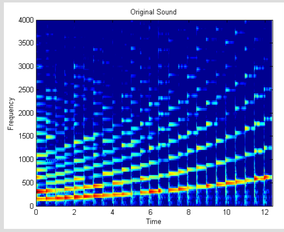


Figure 3: Spectrogram of a recording of a piano playing the chromatic scale (all 12 pitch classes) in ascending order.

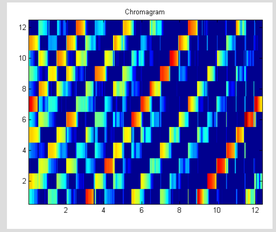


Figure 4: Chromagram of the audio recording in Figure 2.
X-axis represents time.
Y-axis represents pitch class.

EVALUATIONS

Experiments were initialized on 20 Beatles songs from the well-known Beatles data set with ground truth annotations hand-labeled and provided by Christopher Harte.

An overall recognition accuracy of 57% was achieved.

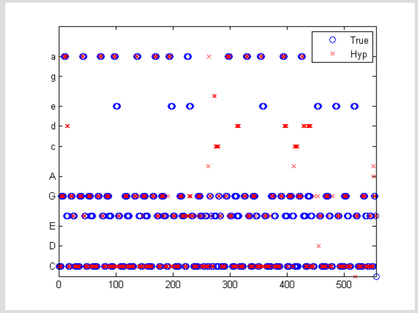


Figure 5: Result obtained using “let it be” as input