

1 Summary of articles

1.1 Folk Music Classification

Chai begins by explaining how it would be useful for computers to be able to classify musical works, and by doing so we could learn how a human can identify a musical work quickly. They then explain the goals of their research being ”: (1) to explore whether there exists significant statistical difference among folk music from different countries based on their melodies; (2) to compare the classification performances using different melody representations; (3) to study how HMMs perform for music classification as a time series analysis problem.”

They used German and Irish folk music pieces for their data. They chose folk songs because of their clear monophonic melody line. They chose German and Irish because of the availability of data.

For the representation of the melody line they used four different methods.

- Absolute Pitch: where they have a value for each note in an octave.
- Absolute Pitch with Duration Representation: Same as absolute pitch but each note is repeated relative to the length of the smallest rhythmical note.
- Interval Representation: The melody is converted to a sequence of intervals.
- Contour Representation: The melody is converted to a sequence of contours, instead of using intervals it just states +/- for small change and ++/- for large change in the melody.

They used 16 different Hidden Markov Models to test how varying structures and number of states for the HMM’s performed in the melody classification. For the HMM’s they used the Baum-Welch re-estimation method for training the markov model for each country. For the identification they used the Viterbi algorithm to decode the sequence and compute its log probabilities.

The data was split 70%/30% training/testing. After training the HMM’s the best representation achieved 63%-77% correctness in identifying the origin of the folk melody.

The results they concluded include that different melodic origins can have varying degrees of similarity making it more difficult to differentiate them. Additionally they discovered that although absolute pitch is a more objective way to represent the melodic line, interval representation performed better, which is consistent with how humans perceive melody lines.

Finally they conclude with stating how HMM’s can be used for melody classification for folk music and how folk music does have significant statistical difference between countries of origin. Ending with stating that melody is an important feature but classification can be greatly improved if other significant features (harmony, instrumentation, performance style etc.) are included in the features used.

1.2 Classification of Musical Genre: A Machine Learning Approach

This paper starts by describing musical genre and how it exists as a terms that define recurrences and similarities between two works of music that the community uses to identify musical events. Thus their analysis is focused on symbolic musical aspects to gain as much information about the dynamic genres, without additional noise. They used six different musical genres with 300 midi songs.

They used 5 different features of the musical work in their research:

- Melodic Intervals
- Instruments
- Instrument Classes and Drum-kits
- Meter/Time Changes
- Note Extension

The experiments performed were all run within The Waikato Environment for Knowledge Analysis. They used 6 different Machine Learning Algorithms for their analysis which include:

- The Naive Bayes
- The Voting Feature Intervals
- J48
- The PART algorithm
- NNge
- RIPPER

They separated the data into training and testing portions with three different proportions of training data 90%, 75%, and 66%. They found that for multi-class categorization the Bayesian classifier performed the best, additionally that classical music is the easiest for their algorithms to recognize followed by Jazz. The binary classifiers outperformed the multi-class in accuracy.

The paper ends with stating that simple musical features can provide enough information to successfully categorize a first level of musical genre but classification of closer genres such as Jazz and Blues would require more complex features.

1.3 Melody Recognition: The Experimental Application of Musical Rules

Cuddy's paper begins by explaining how melodies are recognized even when transposed to different keys despite the fact that the absolute pitches are altered. It then goes on to speculate on the importance of diatonicism, the tones of the diatonic scale, and cadence, a particular order of tones at the end of a melody. The paper then went on to discuss testing of transposed melodies, melodies with errors that fell inside the diatonic scale and errors that fell outside the diatonic scale, in addition to testing melodies in the context of a cadence. Their results showed that when the error in the melody included a non-diatonic tone performance deteriorated, but when inside a diatonic and cadence context performance improved.

1.4 Automatic Identification of Music Works through Audio Matching

This paper begins by explaining a need for automated identification of musical works, then goes on to explain different current types of identification, namely, audio fingerprint, and audio watermarking. It then goes on to begin its discussion on its main work of music identification using audio matching. Their research utilizes Hidden Markov Models, and aims to be able to identify live performances of musical works with diverse instrumentation.

Their approach is based on an *audio to audio* matching process, retrieving all the audio recordings from a database that in some sense represent the same musical content of the query. This is done with the idea that two different performances of the same musical work, even though they could differ greatly, can be generalized to model the features of other performances of the same musical work.

With the goal of analyzing a performance to create a statistical model of the score, they followed the following steps:

1. Segmentation: Extraction of audio sub sequences with a coherent acoustic content
2. Parameter Extraction: Analyzing of the segments to compute a set of acoustic parameters general enough to be matched varying performances of the same musical work
3. Modeling: Automatically building a HMM from the segmentation and parameter extraction to model music production as a stochastic process.

Finally at matching time an unknown recording of a performance is preprocessed to extract the features modeled by the HMMs, and then all the models are ranked based on their probability of creating the same features of the unknown performance.

1.5 Query By Humming – Musical Information Retrieval in an Audio Database

This paper begins by noting how with the growing available databases of music that new forms of querying them are needed and propose humming to be a natural way to query such a database. They then go on to describe how they are using melodic contour to identify each melody line.

They use three main components for their architecture, a pitch-tracking module, a melody database, and a query engine. The process of querying the database would involve: humming into the microphone, which is then sent to a pitch-tracking module to be converted to a contour representation of the melody line, which is then compared against a database of MIDI songs.

The paper then goes on to describe the difficulties encountered when reading in pitch data. After discussing reading in the data, they went on to discuss how they tracked the pitch to get a melodic contour representation of the hummed melody using three different methods:

- Autocorrelation: isolating and tracking the peak energy levels of the signal to measure the pitch.
- Maximum Likelihood: A modification of autocorrelation to increase the accuracy of the pitch and decrease chances of aliasing (Picking an integer multiple of the actual pitch)
- Cepstrum Analysis: The definitive classical method of pitch extraction, which they found didn't work well for tracking humming.

They ended up using a modified form of autocorrelation for the project.

For the query engine they used a fuzzy search string matching algorithm of the contour melody line of the input against the preprocessed contour lines of the songs in the database.

2 References

Mllensiefen, Daniel, and Klaus Frieler. Evaluating Different Approaches to Measuring the Similarity of Melodies. *Studies in Classification, Data Analysis, and Knowledge Organization Data Science and Classification*, pp. 299306., doi:10.1007/3-540-34416-0_32.

Aloupis, Greg, et al. Algorithms for Computing Geometric Measures of Melodic Similarity. *Computer Music Journal*, vol. 30, no. 3, 2006, pp. 6776., doi:10.1162/comj.2006.30.3.67.

Chai, Wei, and Barry Vercoe. "Folk music classification using hidden Markov models." *Proceedings of international conference on artificial intelligence*. Vol. 6. No. 6.4. sn, 2001.

Basili, Roberto, Alfredo Serafini, and Armando Stellato. "Classification of musical genre: a machine learning approach." ISMIR. 2004.

Cuddy, Lola L., Annabel J. Cohen, and Janet Miller. "Melody recognition: The experimental application of musical rules." *Canadian Journal of Experimental Psychology* 33 (1979): 148.

Miotto, Riccardo, and Nicola Orio. "Automatic identification of music works through audio matching." *International Conference on Theory and Practice of Digital Libraries*. Springer, Berlin, Heidelberg, 2007.

Ghias, Asif, et al. "Query by humming: musical information retrieval in an audio database." *Proceedings of the third ACM international conference on Multimedia*. ACM, 1995.