

**EE 473**  
**DIGITAL SIGNAL PROCESSING**

**Term Project**

**CONTENT-BASED CLASSIFICATION  
OF MUSICAL INSTRUMENT SOUNDS  
USING GAUSSIAN METHODS**

*Project Members*

Murat Aksoy  
Hasan Ayaz  
Ender Konukoglu

## ***Problem Statement***

The objective of this project is to classify musical instrument sounds using their cepstral coefficients. The selection is made using the pre-obtained sample set in a database consisting of several notes from various musical instruments.

## ***Introduction***

Classification of audio signals according to their content has been a major concern in recent years. There have been many studies on audio content analysis, using different features and different methods. It is a well known fact that audio signals are baseband, one-dimensional signals. General audio consists of a wide range of sound phenomena such as music, sound effects, environmental sounds, speech and nonspeech signals.

In this project we are mainly concerned with the classification of audio signals that are sampled and recorded from different kinds of musical instruments. The classification of musical instrument sounds, at first step, requires the extraction of certain features related to the input sound sample, which may include root-mean-square amplitude envelope, constant Q transform frequency spectrum, Multidimensional Analysis Scaling trajectories, cepstral coefficients, spectral centroid and presence of vibrato[1].

There are two main approaches to this problem of content based classification based on previous extracted features : The first which uses deterministic methods and the one that utilizes probabilistic techniques.

There are many research efforts, high accuracy audio classification is only achieved for the simple cases such as speech/music discrimination. Pfeiffer et al , presented a theoretic framework and application of automatic audio content analysis using some perceptual features. On the other hand, Saunders presented a speech/music classifier based on simple features such as zero crossing rate and short time energy for radio broadcast.

Scheirer et al conducted many experiments with different classification models including GMM (Gaussian Mixture Model), BP-ANN (Back Propagation Artificial Neural Network) and KNN (K-Nearest Neighbour). Many other works have been done to enhance audio classification algorithms such as pre-classification of audio recordings into speech, silence, laughter and non-speech sounds, in order to segment discussion recordings in meetings. The usage of taxonomic structures also helps to enhance classification performance. In the work by Zhang and Kuo, pitch tracking methods are introduced to discriminate audio recordings into more classes, such as songs, speeches over music, with a heuristic -based model. Accuracy of above 90% is reported. Srinivasan et al, try to detect and classify audio that consists of mixed classes, such as combinations of speech and music together with background sound. The accuracy of classification is over 80% [2]

## *Methodology*

The project makes use of the probabilistic methods and is planned to put forward algorithms to classify musical instruments via their cepstral characteristics. As the first step, the cepstral coefficients from the pre-obtained database are to be obtained by first taking FFT of the signal, then taking the logarithm, and re-taking the inverse FFT. A mean and covariance matrix will then be extracted from this information and fit into a Gaussian function, which will be the main tool to be used in the identification of the input signal.

In order to compare the cepstral characteristics, the cepstral coefficients from the input sample are also extracted and this value is put into the Gaussian distribution stated above, and decision-making is accomplished using Bayesian analysis. The Bayesian decision is made by determining the class whose Gaussian distribution results in the highest probability by this input pattern vector.

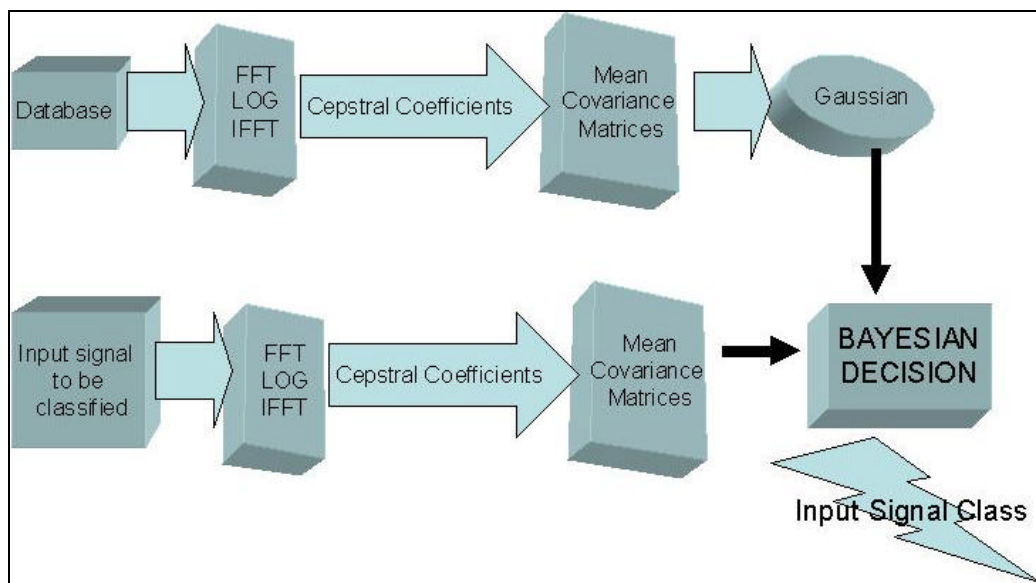


Figure 1. Flow Chart of the classification algorithm

## ***Time Schedule***

Database searching and determination of the instruments to be worked on	7 days
Creating the cepstral coefficients matrix for each sample in the database and fitting into appropriate Gaussian functions	7 days
Implementation of Bayesian Decision algorithm	15 days
Testing the performance of the algorithm by using pre-recorded database samples	10 days
Enhancements on the overall algorithm: ? Extra features for detection (pitch, brightness, etc.) ? Classification of polyphonic sounds	10 days (optional)

## **References**

- [1] Multi-feature Musical Instrument Sound Classifier w/user determined generalisation performance, Ian Kaminskyj, Electrical & Computer Systems Engineering, Monash University
- [2] A Robust Audio Classification and Segmentation Method  
Lie Lu, Hao Jiang and Hong-Jiang Zhang  
Microsoft research, China
- [3] Indexing Audio Databases with Musical Information; Alicja A. Wierzchowska, Polish-Japanese Institute of Information Technology; Zbigniew W. Ras, University of North Carolina, Computer Science Dept., Polish Academy of Sciences, Inst. of Comp. Science,
- [4] Sound Identification and Direction Detection in MATLAB for Surveillance Applications, M. Cowling, R. Sitte, Griffith University Faculty of Engineering and Information Technology, Queensland, Australia
- [5] Automatic Sound Classification Inspired by Auditory Scene Analysis  
Silvia Allegro, Michael Büchler, Stefan Launer Signal Processing Department, Phonak AG, Switzerland + Department of Otorhinolaryngology, University Hospital Zurich, Switzerland
- [6] Julius T. Tou, Rafael C. Gonzalez, Pattern Recognition Principles, Addison-Wesley Publishing Company, 1974
- [7] Richard O. Duda, Peter E. Hart, Pattern Classification and Scene Analysis, John Wiley & Sons Inc, 1973