GENRE CLASSIFICATION

SEFIK EMRE ESKIMEZ

ECE 477 – COMPUTER AUDITION FINAL PROJECT

ABSTRACT

In this project, the problem of automatic genre classification is studied and a new genre classification system is developed. The main goal is to achieve classification of a large music database depending on their music genre with high accuracy. The following genres are used in this project: Jazz, Metal, Hip-Hop, Classical and Disco. The results are presented.

INTRODUCTION

- Music genres are human made labels, basically divides music into categories.
- Categories mainly depend on the instruments that are used in the music piece, harmonics, and rhythmic content of the piece.
- The most popular music genres are rock, pop, metal, classic, jazz, blues and electronic.
- The current media on the Internet are manually labeled (manual labor).
- With increasing number of music content every day, automatic genre classification becomes non-trivial task.

FEATURE VECTOR

Feature vector is formed from various musical descriptors, which is believed mostly represents genre characteristics of musical piece.

Timbre

- First 5 Mel-Frequency Cepstrum Coefficients (MFCCs) are good enough to represent timbre.
- Spectral Centroid can be defined as center of gravity of STFT.
- Spectral Rolloff is the frequency point, which %85 of the magnitudes are distributed below that point.
- Spectral flux indicates the spectral change.
- Zero-cross rate determines the noisiness of the signal.
- Low-energy feature is the percentage of energies whose RMS values are lesser than the average RMS values of musical piece's frames.

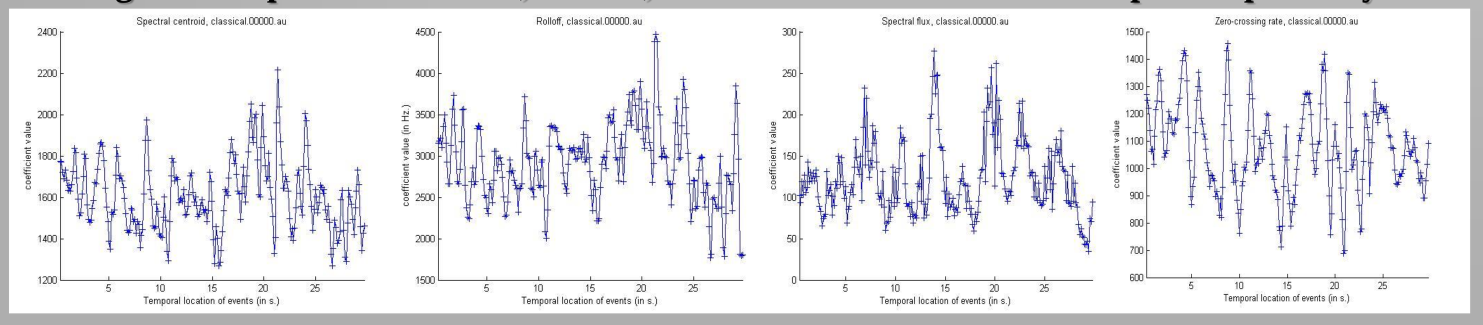
Beat and Tempo

- Means of variances of Beat spectrum is used.
- Dominant tempo for the whole musical piece is extracted and used.

Pitch

- Dominant pitch for the whole musical piece is extracted and used.

Figure 1 – Spectral Centroid, Rolloff, Flux & Zero-Cross Rate Examples respectively.



CLASSIFICATION

Musical Analysis, Retrieval and Synthesis for Audio Signals (MARSYAS) genre collection database is used in this project [3, 17]. Feature vectors for all training and test data are extracted. 20 songs from each genre class are picked as test data and rest of them are treated as train data.KNN classifier, with Euclidian distance measure, does classification. Accuracy is calculated according to ground truth data.

Table 1 shows 3-nearest neighborhood musical genre classification results. Table 2 shows 1-nearest neighbor results. Selecting k = 1 yields better result than 3-nearest neighbor result.

Table 1 - Accuracy values for KNN classification results for first set (3-Nearest Neighbors)

Accuracy	
Total	43.80%
Jazz	33.33%
Metal	42.85%
Hip-Hop	52.38%
Classical	57.14%
Dance	33.33%

Table 2 - Accuracy values for KNN classification results for second set (3-Nearest Neighbor)

Accuracy	
Total	%51.42
Metal	%57.14
Country	%47.61
Pop	%66.66
Classical	%57.14
Reggae	%28.57

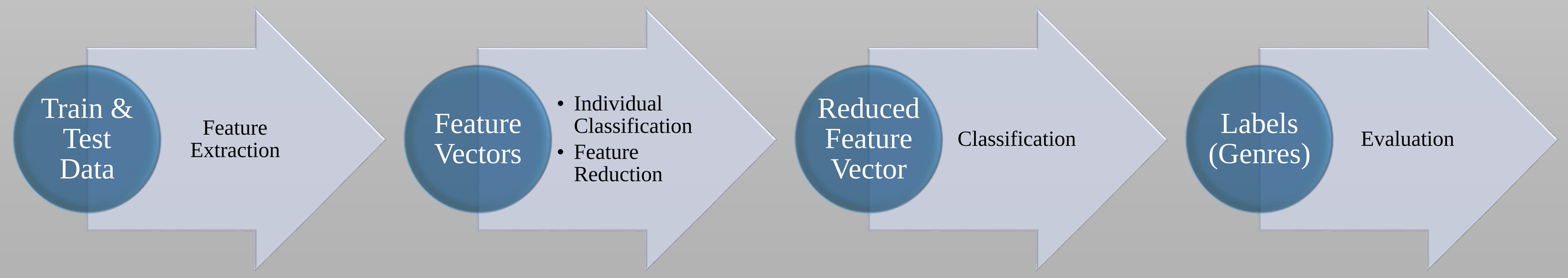
FEATURE IMPACTS ON CLASSIFICATION RESULTS

To measure impact of a particular feature on classification, only one feature is used for classification and total accuracy results are stored. Table 3 shows resulting accuracies for corresponding feature.

Table 3 - Feature's success measure on genre classification.

Feature	1st set Accuracy	2nd set Accuracy	Feature	1st set Accuracy	2nd set Accuracy
Mean of MFCC(2)	30.47%	27.61%	Variance of Beat Spectrum	40.00%	35.23%
Mean of MFCC(3)	33.33%	38.09%	Sum of Beat Spectrum	32.38%	22.85%
Mean of MFCC(4)	26.66%	23.80%	RMS value	35.23%	41.90%
Mean of MFCC(5)	32.38%	34.28%	Mean of Spectral Centroids	44.76%	34.28%
Mean of MFCC(6)	34.28%	33.33%	Variance of Spectral Centroids	28.57%	39.04%
Variance of MFCC(2)	40.00%	40.95%	Mean of Spectral Rolloff	42.85%	49.52%
Variance of MFCC(3)	40.00%	36.19%	Variance of Spectral Rolloff	37.14%	38.09%
Variance of MFCC(4)	34.28%	31.42%	Mean of Spectral Flux	44.76%	46.66%
Variance of MFCC(5)	31.42%	33.33%	Variance of Spectral Flux	46.66%	48.%
Variance of MFCC(6)	38.09%	34.28%	Mean of Zero-crossing rate	39.04%	33.33%
First Dominant Tempo	35.23%	23.80%	Variance of Zero-crossing rate	29.52%	33.33%
Second Dominant Tempo	37.14%	27.61%	Low Energy	23.80%	25.71%
Mean of Beat Spectrum	32.38%	22.85%	Dominant Pitch	35.23%	34.28%

Figure 2 Flowchart of Genre Classification System



FEATURE REDUCTION

Some features are not contributing to classification; instead they are reducing the accuracy of results. Therefore a threshold of %38 is applied to accuracy results of features.

Table 4 - Accuracy values for KNN classification results for first set

	Accuracy	Precision
Total	62.85%	
Jazz	52.38%	57.89%
Metal	47.61%	58.82%
Hip-Hop	61.90%	59.09%
Classical	80.95%	80.95%
Dance	71.42%	57.69%

Table 5 - Accuracy values for KNN classification results for second set

	Accuracy	Precision
Total	60.00%	
Jazz	57.14%	52.17%
Metal	52.38%	50.00%
Hip-Hop	61.90%	65.00%
Classical	85.71%	85.71%
Dance	42.85%	47.36%

EVALUATION

- The most important features for the genre classification are listed.
- Variance of second, third and sixth MFCCs, Beat Spectrum and Spectral Flux
- Means of Spectral Centroid, Spectral Rolloff, Spectral Flux and Zero-crossing rate.
- Those features have above %38 accuracy and they are found by experimenting.
- Compared to Table 1, accuracy is increased by %43, after feature reduction.
- Precision for each class is above %55.

CONCLUSION & FUTURE WORK

- A system is developed to classify musical pieces into genres (labels).
- Additional feature reduction step is implemented to increase accuracy and precision of genre classifier.
- It is found that some of the proposed features are not as effective as other features and lowers the overall genre classification quality.
- Additional features can be tested to see their impact on genre classification.
- Different classifiers can be tested and feature reduction step can be arranged according to new classifiers. Classifier's impact on genre classification can be explored.