Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 116 (2017) 174–181

Procedia Computer Science

www.elsevier.com/locate/procedia

2nd International Conference on Computer Science and Computational Intelligence 2017, ICCSCI 2017, 13-14 October 2017, Bali, Indonesia

Indonesian's Traditional Music Clustering Based on Audio Features

Aisha Gemala Jondya[1]*, Bambang Heru Iswanto[2]

[1]*Computer Science Department, School of Computer Science, Bina Nusantara University, Jalan Kebon Jeruk Raya No.27, Jakarta, 11530, Indonesia*

[2]*Universitas Negeri Jakarta, Jalan Rawamangun Muka, Jakarta – 13220, Indonesia*

Abstract

Cluster analysis has been used widely in some applications. In this research, 101 songs from 18 provinces in Indonesia clustered by some set of features extracted directly from the audio data. Before the clustering process, feature selection process with PCA method performed using 60 audio segments from 4 songs to find the optimal set of features which will be used in clustering process. In clustering process, the selected features are extracted from audio signal and clustered by *x*-Means algorithm to find the proper number of cluster. Clustering with this method resulted 4 clusters. The result of this process shows the characteristic of each cluster and some distributions of cultures between areas and provinces. An Agglomerative Hierarchical Clustering method also conducted to compare the result.

Peer-review under responsibility of the scientific committee of the 2nd International Conference on Computer Science and Computational Intelligence 2017.

Keywords: music, clustering, *x*-Means, PCA, AHC

1. Introduction

Indonesia has rich cultures including various style of traditional musics which spread across islands. This traditional musics influenced by some historical background such as multitude of religions and customization of

* Corresponding author. Tel.: +62-812-83600-252
E-mail address: aisha.jondya001@binus.ac.id

1877-0509
Peer-review under responsibility of the scientific committee of the 2nd International Conference on Computer Science and Computational Intelligence 2017.
10.1016/j.procs.2017.10.019

foreign cultures. The aim of this work is to find the features that can differentiate Indonesia's traditional music so that we can explore the similarities between songs from some provinces in Indonesia using the computational approach.

One of the challenges in analyzing music similarity is to find out what it is that allow us to differentiate between music styles which are not directly comparable [1]. Features from each audio data have to be selected and extracted to find the effective set of features. Most of the research in audio analysis conducted using the numerical values of the features that represent audio.

Analyzing the similarity of Indonesian traditional music can be done by using clustering method. Since there is no information about the songs correlation between provinces, clustering results are also used to see the number of clusters. Therefore, x-Means algorithm is used because this method allows the data sets to divided on the optimal number of clusters without prior knowledge [2].

This experiment divided by 2 main process; (1) feature selection and (2) clustering process. This paper is organized as follows: section 2 describe the related work while section 3 described the overall methodology. An overview of audio data explained in section 4. Section 5 described the method and detail of feature selection process and also the validation process while section 6 describe the clustering process using *x*-Means method and compare the result with the AHC method. Discussion and future research are given in section 7.

2. Related Works

For long times, automatic audio music clustering research has been performed for many purposes using manually specified features. A clustering using *k*-Means conducted on classical, rap, metal and Indian's music[3]. Recommendation systems mainly use manual genre-annotations or collaborative filtering, which is time consuming, therefore this study believed that music recommendation systems can be improved a lot by algorithmic music clustering. Instead of using spectro-temporal features, this study clustered music using subjective features provided by Echonest, such as time signature and danceability.

Another clustering was also conducted using manually specified features to compare the culture of 16 Austronesian society [4]. In this study, each song was listened and coded manually based on the scheme developed by the author. Although this two studies used clustering process to help them automatically find similarity between songs, using features that defined manually is not efficient.

Audio analysis process becomes easier and accurate using features that are extracted automatically from audio. The low-level audio features are popular to used in some audio classification research. Ellis et al. classify a set of enviromental sound from a set of audio videos using MFCC features [5]. In this study, human perception of sound textures became principals to construct automatic content clasification.

The low-level audio features has also been used in clustering study. Li et al. studied clustering based on timbral texture features and rhythmic content features extracted automatically from audio for better recommendation system [6]. More over, similar to our study, folk music clustering of four non-European eastern countries, western music and folk music of Cyprus conducted to find the similarity between songs [7]. This study used 25 low-level features and 13 mid-level features to gather information from pitch histogram. Since clustering has no initial knowledge, this study compared clustering in *k*-Means and SOM method where the 2 methods resulted similar result.

3. Methodology

Clustering analysis is one of the technique in machine learning which have the basic purpose to find the informative and useful pattern in a big data [8]. The data mining main process consists of data preparation (data preprocessing), data transformation, data mining, and interpretation of results [9]. Clustering is different from classification where classification's accuracy can be calculated since the labels or prior knowledge of data are available.

In this study to ensure the clustering generates a good result, a feature selection preprocessing is performed by clustering 4 songs segments from 4 different provinces using 11 sets with total 36 recommended features by Giannakopoulos [10]. Eleven sets of time domain and frequency domain features are energy, entropy energy, zero-crossing rate, spectral centroid, spectral entropy, spectral flux, spectral rolloff, Mel-Frequency cepstral coefficient, Harmonic, Chroma Vector and spectral zone. What follows, are definitions of each features.

1. Energy