Design of Simulation System based on Digital Processing of Sound to Identify the Genres of Arabic Music

M.E. ElAlami
Computer Science Department
Mansoura University
Egypt

S.M.K. Tobar
Musical Education Department
Mansoura University
Egypt

S.M. Kh
Computer Science Department
Mansoura University
Egypt

Eman A. Esmaeil
Computer Science Department
Mansoura University
Egypt

ABSTRACT

Musical genres identification has attracted considerable attention and interest in many real world applications, and consider an important technology for music information retrieval (MIR). At present, arabic music represents important aspect of arab music heritage in its distinctive and unique nature .This paper presented a system for arabic music analysis, Mel Frequency Cepstral Coefficients (MFCC) and Short Time Energy (STE) are used to extract features for the music signal and (SVM) classifier was technique used for the purpose of classification. Result show that the proposed system is a useful for achieves state-of-the-art performance for arabic musical genres analysis.

Keywords

Support vector machine (SVM) , Mel frequency cepstral coefficients (MFCC) , Music information retrieval (MIR) , Arabic musical genres, Short time energy (STE), Discrete wavelet transform (DWT)

1. INTRODUCTION

Identify genres of arabic music has been an inspiring job in the area of music information retrieval (MIR). through the advent of online music libraries and quick access to music content, people find its increasing to manage the songs that they listen to. One way to categorize and organize songs is based on the genre, which is identified by some characteristics of the music such as rhythmic structure, harmonic content and instrumentation [1]. Music genre identify process mainly contains two steps:

Representation feature and design of classifier. Several works have been done for classifying the music genre. Some of them focus on the representation of feature, and these methods generally aim to make the feature more discriminative [2].

2. RELATED WORK

Classification of music genres is one of the most common MIR tasks and not a new challenge; several others have tried to use algorithms to solve the problem.

Zain and Yue [3], proposed a modern artist classification system , a Convolutional Recurrent Neural Network (CRNN), is applied under a rigorous collection of conditions to the artist20 music artist recognition dataset.

The result reveals that The highest performing model get an average F1 score of 0.937 in three separate experiments, which under identical conditions is a major increase over the corresponding baseline.

Aldona and Bozena [4] , An analysis was raised as to whether separating music tracks in the preprocessing step and expanding the function vector by parameters relevant to particular musical instruments typical of the musical genre enables an accurate automated classification of musical genres in the case of a database comprising thousands of musical extracts and a dozen genres.

The result show that this approach for music genre classification is promising.

Guochao et al [5] , developed a specially designed network for accurately recognizing the music genre. The proposed model aims to take full advantage of low-level information of melspectrogram for making the classification decision.

The result on Various experiments on several benchmark datasets, including GTZAN, Ballroom, and Extended Ballroom, have verified the excellent performances of the proposed neural network.

Snigdha et al [6] , builted multiple classification models and trained them over the Free Music Archive (FMA) dataset. they have compared the performances of all these models and logged their results in terms of prediction accuracies. Few of these models are trained on the mel-spectrograms of the songs along with their audio features, and few others are trained solely on the spectrograms of the songs.

The result show that one of the models, a convolutional neural network, which was given just the spectrograms as the dataset, has given the highest accuracy amongst all other models.

Benjamin and Günther [7], introduced a contribution to the Crowd AI music genre classification challenge "Learning to Recognize Musical Genre.

The results show that best results were obtained by an extreme gradient boosting classifier.

Basilio et al [8], proposed an inquiry into the classification of generated pieces of music, based on the concept of grouping closely related known pieces into separate sets , then automatically creating a new song that is somehow "inspired" from every set.

The result indicates that the results obtained are positive, and indicates the appropriateness of the method used in classification of the music genre.

3. PROPOSED SYSTEM

This paper discusses arabic music genres identification

method which consider an important job in music information retrieval (MIR), The proposed system used arabic database which consist of 35 classes of arabic music genres , the duration of the genre ranges between 5 to 10 second. Identification process of musical genres includes three main steps as shown in Figure.1.
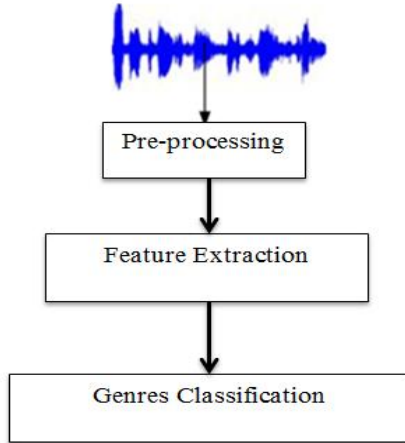

Figure 1 Block diagram of Music Genre identification

3. 1 Pre_processing

Music signal which results from the separation phase contains noise. Pre-emphasis filter is offered to improve the signal to noise ratio (SNR). pre-emphasis filter given by [9]:

$$y_1(t_1) = x_1(t_1) - a * x_1(t_1 - 1) \quad (1)$$

Where $x_1(t_1)$ refers to the original music signal , $y_1(t_1)$ refers to the filtered music signal, and a is set to 0.97.

3.2. Feature Extraction and Framing

In this section, the baseline feature that include discrete wavelet transform (DWT) ,Mel frequency cepstral coefficients (MFCC) and short time energy (STE) have been extracted from the signal.

3.2.1 Discrete wavelet transform (DWT)

For the decomposition of wavelet packet, the wavelet coefficients are subdivided by filtering them into low as well as high frequency. in this context , three levels of wavelet packet sub-bands are utilized for the signal decomposition, then (MFCC) and (STE) are extracted from the music signal.

3.2.2 Mel frequency Cepstral coefficients (MFCC)

MFCC uses a Mel-scale conversion of the fundamental frequency f to approximate the hearing mechanism given as [10]

$$fx^m = 2595\log_{10}\left(\frac{fx}{700} + 1\right) \quad (2)$$

The generalized for extraction of MFCC coefficients can be expressed as:

$$\text{MFCC}_j = \sum_{l=1}^{L} X_l \cos\left[j\left(l - \frac{1}{2}\right)\frac{\pi}{2}\right], j = 1,2,\dots,N1 \quad (3)$$

Where, N1 denotes the number of cepstrum coefficients, and $X_l, l = 1,2,\dots,L$ signifies the Lth filter log-energy output.

3.2.3 Short time energy (STE)

Short time energy (STE) of music signals have been Calculated by the following equation [11].

$$E1_{n1} = \sum_{m1=-\infty}^{\infty}[x(m1)w1(n1 - m1)]^2 \quad (4)$$

where $E1_{n1}$ represents the short-time energy, where as n and m refer to the indexes of frames, $x_1(m_1)$ denotes the discrete time music signal, $w_1(n_1)$ represents the window(applied to music frame).

After feature extraction process first second from the target music signal have been selected, and compared it with all database signals.

After feature extraction process, first second from the target music signal are extracted, and compared it with all database signals. If the similarity is achieved between the first second of the target music signal and any signal of the database with a similarity rate not less than 100%, it is extracted , but the similarity rate has not been achieved , first and second seconds are merged , then compared with all music signal in the database again, and so do until achieve target similarity.

3. 3 Arabic music Genres Classification

When the feature extraction process is done, music signal should be classified, Support Vector Classifier has been used for this task, SVM is an emerging machine learning technology that has already been successfully used for sound classification in general domain. Figure 2 show SVM steps.
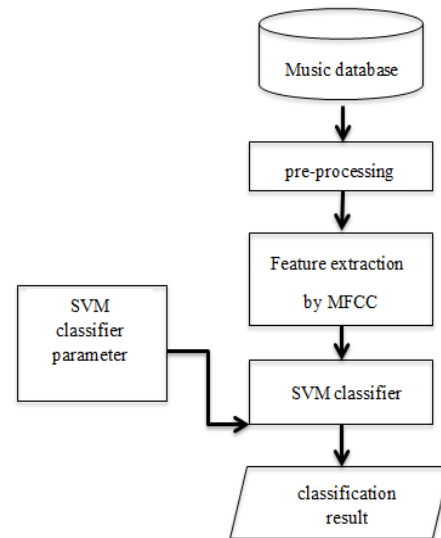

Figure 2: Block diagram of Support Vector classifier

SVM is a supervised technique that uses a designated training data set and is trying to find a decision feature that best classifies the training data. The algorithm's aim is to find a hyper-plane to define decision boundaries separating data points of different groups. It is widely used for problem with pattern recognition and classification. The equation of the hyper plane is given [12]

$$w1^T x1 + b1 \quad (5)$$

Where w is vector of weight, and where b is bias.

Provided the data set labelled for the training $\{x_i, y_i\}_{i=1}^{N}$ with $x_i \in \mathbb{R}^d$ is input vector and $y_i \in \{-1, +1\}$. Where x_i is input vector and y_i Is the appropriate label. SVMs define the d-dimensional input vector x from the input space to the d h - dimensional feature space by non-linear function φ(·):

$\mathbb{R}^d \to \mathbb{R}_h^d$. Hence hyper plane equation becomes [13]

$$w1^T \varphi(x1) + b1 = 0 \qquad (6)$$

With $b1 \epsilon \mathbb{R}$ and w_1 is an unknown vector with the same dimension as $\varphi(x1)$. The resulting of problem with SVM optimisation is given as:

$$\min_{w,\xi,b} j1(w,\xi) = \frac{1}{2} w^T w + c \sum_{i=1}^{n} \xi i \qquad (7)$$

Using Lagrange multipliers, the SVM optimization problem is typically written in dual space by adding restriction in the minimising feature. The dual formulation is shown as [14]

$$\max_{\propto} \sum_{i=1}^{m} \propto_i - \frac{1}{2} \sum_{i,j=1}^{N} \propto_i \propto_j y_i y_j (x_i, x_j) \qquad (8)$$

With the exception of $\propto_i \geq 0$ for all $i = 1, \ldots \ldots m$ and $\sum_{i=1}^{m} \propto_i y_i = 0$.

Thus, the hyper plane can be defined as:

$$f(x) = sgn[\sum_{i=1}^{m} y_i \propto_i (x_i, x) + b] \qquad (9)$$

Multiclass SVM, classifies the specified test sample to 35 classes in this paper. Method "one vs. rest" has been used in the current work where entity belonging to one class is distinguished from another. An optimal hyper plane was calculated to separates each class from the rest of the elements.

A series of binary classifiers are built to get classification of N classes, where each training distinguishes one class from the others. multiclass classification was combined according to the potential output after that sgn function which takes takes a form:

$$:arg \max_{j=1,\ldots M} g^j(x), where \ g^j(x) = \sum_{i=1}^{m} y_i \propto_i^j k(x, x_i) + b^j \qquad (10)$$

And

$$f^j(x) = sng\left(g^j(x)\right) \qquad (11)$$

This has a linear structure with respect to N class and N hyperplanes have computed . In this study 35 classes have been used . Figure 3. showed the main screen for the proposed system.
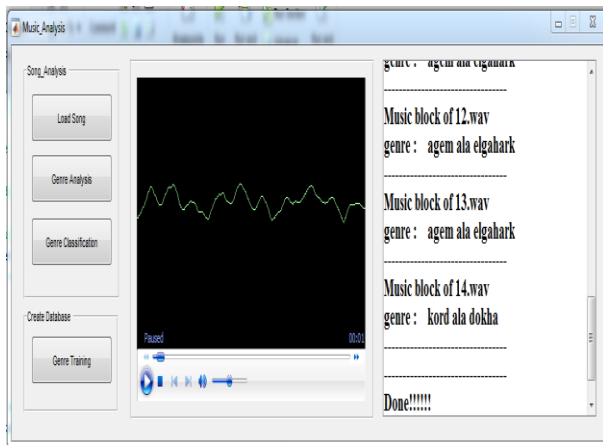

Figure.3 Main screen for the Proposed System

4. EXPERIMENTAL EVALUATION

For testing the performance of arabic music genres classification system , arabic music database which contains 35 classes of arabic musical genres has been used , the duration of the genre range between 5 to 10 second. Figure 4 show experimental result to identify music genres at first 40 second of 3-salamat arabic song by using Short time energy (STE).
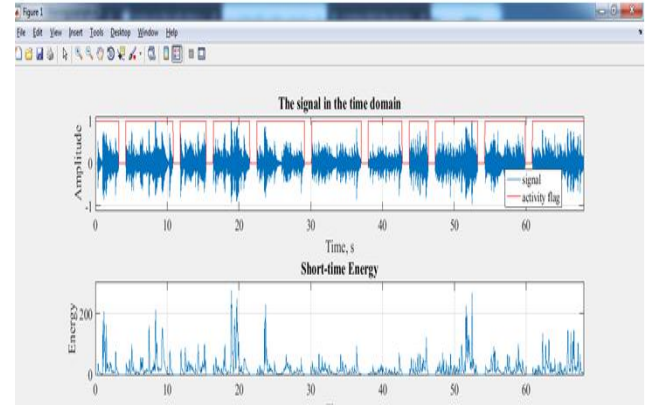

Figure 4: Experimental result to identify music genres at first 40 second of 3-salamat arabic song

5. RESULT

From the experiment result , it showed the best effect of DWT with (MFCC) and (STE) for music signal features extraction, as well MFCC are a good features used for individual tracks . Accuracy classification by using SVM arrive to 98 % by using Recognition as shown through the following equation .

$$Recognition \ Rate = \frac{Number of \ correctly recognized items}{Total \ number of \ items} * 100\%$$

6. CONCLUSION

To process wide music libraries, tools need to extract important details about musical pieces from music signal, among of these important details is music genres analysis. In this paper, a simulation system for arabic music genres analysis has been discussed. Arabic music database which consists of 35 class of music genres have been used. The best effect of (MFCC) and (STE) for music features extraction have been showed with a high accuracy classification by using (SVM).

7. FUTURE WORK

In this paper, a novel supervised method for arabic music analysis has been introduced. The results shown the effective of the proposed system for arabic music analysis.In future work, it will be suggested to focus on melody feature such as structural and temporal feature of frequency .

8. REFERENCES

[1] George Tzanetakis , et al. Musical genre classification of audio signals. (2002) IEEE Transactions on speech and audio processing 10(5):293–302.

[2] Hong Yang , et al. Music Genre Classification using Spectral Analysis Techniques With Hybrid Convolution-Recurrent Neural Network. (2019). International Journal of Innovative Technology and Exploring Engineering, 9(1), pp.149-154.

[3] Nasrullah, Z. and Zhao, Y. (2019). Music Artist Classification with Convolutional Recurrent Neural Networks. 2019 International Joint Conference on Neural Networks (IJCNN).

[4] Rosner, A. and Kostek, B. (2017). Automatic music

genre classification based on musical instrument track separation. Journal of Intelligent Information Systems, 50(2), pp.363-384.

[5] Liu ,c. and Feng ,L (2019) . Bottom-up Broadcast Neural Network For Music Genre Classification. Elsevier 24 Jan (2019).

[6] Chillara ,S. and Haldia ,S (2019). Music Genre Classification using Machine Learning Algorithms: A comparison. International Research Journal of Engineering and Technology (IRJET), Volume: 06 Issue: 05 | May 2019.

[7] Murauer,B .and Specht ,G (2018). Detecting Music Genre Using Extreme Gradient Boosting. Creative Commons Attribution 4.0 International (CC BY 4.0) license.

[8] Goienetxea, I., Martínez-Otzeta, J., Sierra, B. and Mendialdua, I. (2018). Towards the use of similarity distances to music genre classification: A comparative study. PLOS ONE, 13(2), p.e0191417.

[9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems. pages 1097–1105.

[10] Bahoura, M. 2009. Pattern recognition methods applied to respiratory sounds classification to normal and wheeze classes. Computers in Biology and Medicine, 39 (9). 824–843.

[11] Yao, P. and Zhou, K. (2017). Application of short time energy analysis in monitoring the stability of arc sound signal. Measurement, 105, pp.98-105.

[12] Frihia, H. and Bahi, H. (2017). HMM/SVM segmentation and labelling of Arabic speech for speech recognition applications. International Journal of Speech Technology, 20 (3), pp.563-573.

[13] Mittal, T. and Sharma, R. (2016). Integrated search technique for parameter determination of SVM for speech recognition. Journal of Central South University, 23(6), pp.1390-1398.

[14] Londhe, N. and Kshirsagar, G. (2017). Speaker independent isolated words recognition system for Chhattisgarhi dialect. 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS).