Data & Knowledge Engineering 92 (2014) 60–76

Contents lists available at ScienceDirect

Data & Knowledge Engineering

journal homepage: www.elsevier.com/locate/datak

Editorial

Music genre classification based on local feature selection using a self-adaptive harmony search algorithm

Yin-Fu Huang *, Sheng-Min Lin, Huan-Yu Wu, Yu-Siou Li

Department of Computer Science and Information Engineering, National Yunlin University of Science and Technology, 123 University Road, Section 3, Touliu, Yunlin 640, Taiwan, ROC

ARTICLE INFO

Article history:
Received 1 February 2012
Received in revised form 16 June 2014
Accepted 4 July 2014
Available online 15 July 2014

Keywords:
Classification
Information retrieval
Feature selection
Harmony search algorithm

ABSTRACT

This paper proposes an automatic music genre-classification system based on a local feature-selection strategy by using a self-adaptive harmony search (SAHS) algorithm. First, five acoustic characteristics (i.e., intensity, pitch, timbre, tonality, and rhythm) are extracted to generate an original feature set. A feature-selection model using the SAHS algorithm is then employed for each pair of genres, thereby deriving the corresponding local feature set. Finally, each one-against-one support vector machine (SVM) classifier is fed with the corresponding local feature set, and the majority voting method is used to classify each musical recording. Experiments on the GTZAN dataset were conducted, demonstrating that our method is effective. The results show that the local-selection strategies using wrapper and filter approaches ranked first and third in performance among all relevant methods.

1. Introduction

Recently, the development of digital audio techniques has matured. Through the Internet, increasingly more online music systems, such as 7digital and Amazon MP3, have been developed to provide abundant music services and copyrighted music downloads for consumers. In general, to facilitate music searches, these systems usually categorize music by using several tags such as blues, classical, and country. However, because a vast amount of music is stored in these systems, manually tagging music is time-consuming. Furthermore, most people have different cognitions of music. Therefore, developing an automatic music genre-classification system is necessary for the tagging process to be more effective and become standardized. This paper proposes a music genre-classification system, because genre tags are employed for most musical content descriptions [2].

In other studies, most music genre-classification systems have extracted audio features to achieve satisfactory performance. In general, audio features are long-term features that have been obtained by estimating the overall statistics of short-term features, or are directly estimated for long-term use based on audio-feature descriptions of audio signals. Therefore, conventional bag-of-frames approaches [17,20,26,35] have been proposed and employed to generate feature sets of classification, such as the Marsyas framework [35], which contains several features regarding timbre texture, rhythmic content, and pitch content; the genre collection in this framework is called the GTZAN dataset and is frequently used to compare the effectiveness of different systems [1,3,18,22,29–31]. Moreover, some researchers have focused on proposing refined features, such as modulation spectral analysis [15,22], to improve performance. In addition, the auditory model proposed by [37] maps a musical recording to obtain a 3D representation of its slow spectral and temporal modulations. In our study, we also adopted a bag-of-frames approach and focused on five

* Corresponding author. Tel.: +886 5 5342601x4314; fax: +886 5 5312170.
E-mail address: huangyf@yuntech.edu.tw (Y.-F. Huang).
URL: http://mdb.csie.yuntech.edu.tw (Y.-F. Huang).

http://dx.doi.org/10.1016/j.datak.2014.07.005
0169-023X/

acoustic characteristics (i.e., intensity, pitch, timbre, tonality, and rhythm) to extract useful audio features and form an original feature set for music genre classification.

To obtain sufficient knowledge for classification, generated feature sets commonly contain abundant information, likely with redundant features. To solve this problem, dimensionality reduction techniques are frequently employed and can be classified into two approaches. The first approach is to transform a feature-set matrix from a high-dimensional space to a lower dimensional space through a linear combination of matrix-using techniques, such as principal component analysis (PCA) [16], non-negative multi-linear principal component analysis (NMPCA) [31], non-negative matrix factorization (NMF) [21], and non-negative tensor factorization (NTF) [30]. The second approach is called feature selection, which determines an optimal subset from the original feature set by using search algorithms, such as genetic algorithms (GAs) [19], ant colony optimization (ACO) [7], harmony search (HS) [8], adaptive binary harmony search [36], and support vector machine (SVM) recursive feature elimination (SVM-RFE) [11]. Both approaches can effectively reduce the dimensions of a feature set. In this study, we adopted a feature-selection approach based on the self-adaptive harmony search (SAHS) algorithm [14]; the approach was validated, thus leading to an optimal solution. In our study, we extensively collected useful features and determined which features were more relevant in music genre classification. In general, feature-selection approaches are employed to select a global-feature set for all music genres. However, to achieve global optimization, we adopted a local-selection strategy based on each pair of music genres because it can derive local feature sets that are more relevant when compared with a global-selection strategy. Finally, we verified that using the SAHS achieves higher performance compared to other methods.

For prediction, frequently employed classifiers, such as multi-layer perceptrons (MLPs) [27,32], SVMs [1,15,27,31,32], and linear discriminant analysis (LDA) [15,18,22], have been used to determine optimal linear combinations that discriminate the vector of a feature set for different classes. In this study, the SVM classifier was adopted because, in general, it demonstrates higher performance than other classifiers do [15,31,32] when kernel functions and parameters are appropriately chosen. In this study, we matched each local feature set with an SVM classifier and used the majority voting method to classify each musical recording. Experiments on the GTZAN dataset were conducted, and the results demonstrated that our method is effective. The results showed that the local-selection strategies, which involve adopting two approaches, ranked first and third in performance among all relevant methods.

The remainder of the paper is organized as follows. Section 2 presents the system architecture and briefly describes the music classification procedures. Section 3 introduces five acoustic characteristics and presents the extracted features that form an original feature set. Section 4 describes the SAHS algorithm and correlation measuring method used to derive an optimal feature subset from the original feature set. Section 5 presents the experimental results of differing selection strategies and approaches. Finally, we offer our conclusion in Section 6.

2. System overview

This paper proposes an automatic music genre-classification system that uses suitable features selected using a meta-heuristic optimization algorithm called the SAHS algorithm [14] for classifying music genres. The feature-selection model is applied for each one-against-one classifier to classify ambiguous genres precisely.

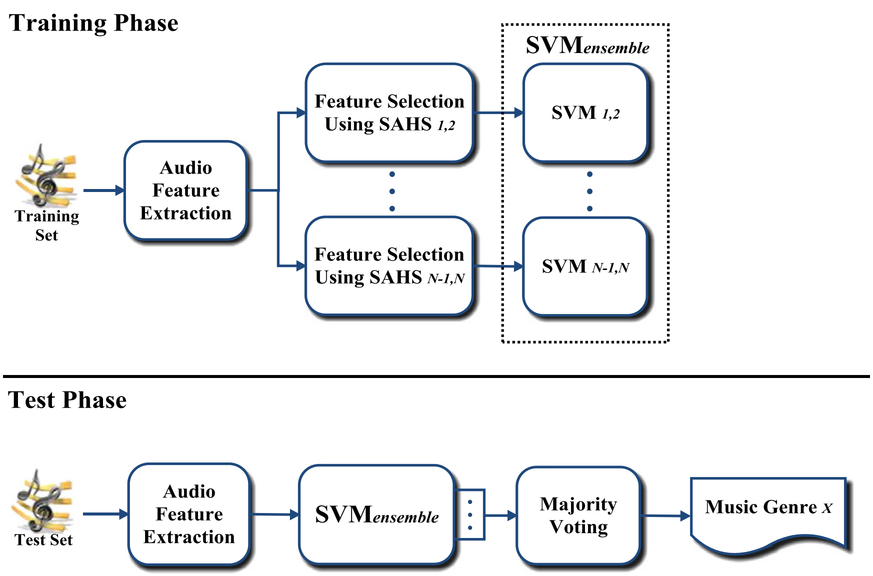

Fig. 1. System architecture.