

Music Genre Classification Using Feature Subset Search

Jihae Yoon, Hyunki Lim, and Dae-Won Kim

Abstract—With the growing number of digital music, the automatic genre recognition problem has been receiving the spotlight in music retrieval information field. A large number of musical acoustic features are reported to degrade the genre classification performance and lead to heavy computational cost. In this paper, we propose a new method for selecting genre-discriminative feature subset from a large number of musical features. We show that the proposed method is able to improve the genre recognition accuracy compared to the traditional selection method.

Index Terms—Genre classification, feature selection, mutual information, incremental search.

I. INTRODUCTION

For a long time, musical genre has been one of the most common description of music contents [1]. Generally human expertise manually annotates genre after listening the music [2]. These procedures require effort and time consuming. However, the suitability of genre annotated by human expertise depends on expertise's musical knowledge and previous experience [3]. With the explosive number of published digital music, an automatic genre recognition has been receiving the spotlight in music retrieval information (MIR) field [4].

Over the past year, many researchers have been employed for designing a learning algorithm that is able to identify the relationship between musical acoustic features and genre. For extracting musical features, many music analysis methods such as MIRtoolbox, JAudio and Marsyas are widely used [5], [6]. According to various options chosen by users, over the nine hundreds of musical features can be extracted from given music. However, all of features are not considered equally important in genre recognition. In some cases, irrelevant features degrade genre recognition performance and also lead to computational inefficiency. One straightforward solution is to select informative feature subsets from the large number of features.

Our goal is to identify a compact feature set that consists of genre discriminative features. The feature subset is found by incrementally searching informative features that take both of two notions, relevance and redundancy, into account. While the learning time is decreased by reduced number of features, the classification accuracy is improved than the case of all musical feature being involved in learning. From various

tests, the present work to music genre classifications shows its potential and direction for improvement.

II. RELATED WORKS

In recent approaches to music genre recognition studies, there are two types of methods; the informative musical feature extraction and the classifier for constructing genre recognition model [7]. P. R. Lisboa de Almeida *et al.* proposed a dynamic ensemble approach to music genre recognition. They used a pool of classifiers consisted of weak classifier SVMs [8]. R. Popovici *et al.* proposed a method for genre classification using a self-organizing map (SOM) [9]. The method identifies the musical similarity based on pitch and timbre features. A. Anglade *et al.* used the harmony rules that are automatically induced from the music [10]. Low-level features and high-level harmony features were combined to improve genre classification performance. Their final results show that the harmony-based rules contain the useful information for genre recognition. S. Doraisamy *et al.* used simple feature selection methods for improving genre classification performance [11]. They used correlation-based feature selection (CFS) and chi-square feature evaluation for selecting the informative musical features for Malay traditional genres. Despite its simplicity, their experimental results show that the genre recognition performance can be improved by informative features.

Of the well-known feature selection methods in machine learning community, the min-redundancy and max-relevance (mRMR) is the most widely used method [12]. The proposed method extends mRMR with conditional mutual information for music genre classification. The mRMR incrementally selects a sub of features that minimize the mean of mutual information among a pair of features selected in a subset (i.e., redundancy) and maximize the mean of mutual information among features and label (i.e., relevance). However, the time complexity of mRMR is $O(N^2)$ as increasing the number of selected features. The efficiency is regarded as its weakness. In this study, the proposed method is designed to improve the time efficiency and the selection performance is reliably maintained.

III. PROPOSED METHOD

Given a music data set, N -dimensional feature vector $F = \{f_1, \dots, f_N\}$ with a target class variable C . Feature selection aims to identifying the optimal subset S consisted with n features, especially $n \ll N$. With the optimal subset, classification performance can be improved. In general, the dependency with a feature and class should be maximized. On the other hand, the dependency between a feature and a feature should be minimized.

To this end, the mRMR algorithm optimizes the following

Manuscript received October 30, 2015; revised February 8, 2016. This research was supported by Ministry of Culture, Sports, and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2015.

The authors are with the Chung-Ang University, Seoul, Korea (e-mail: jihae.cau@gmail.com, hyunki05@gmail.com, dwkim@cau.ac.kr).

topologically maximum similarity in the Euclidian space. In this test, we set this k parameter to 1, which means that we only consider the most similar sample to classify a genre. Because each feature vector has diverse range values, a feature scaling is required. To adjust feature scales, we used min-max and mean-standard deviation normalizations. We compared the performance of both of the two normalized data sets. To obtain statistically reliable classification results, we conducted a hold-out cross-validation for our experiments. Our dataset is randomly divided into two parts; 80% of the songs are used for learning and the rest 20% of the songs are used for testing the recognition performance. Our experiment was repeated 30 times. We assessed the genre recognition performance with an accuracy measure. Table II represents the detailed information about our experimental setup.

TABLE I: CHARACTERISTICS OF COLLECTED GENRE DATA

Feature extractor	MIRtoolbox	JAudio
# feature	369	916
# sample	155 songs	
# class	5 genres	
class	Genre	# songs
	Rap	30
	Electronic	32
	Jazz	34
	Metal	29
	Classic	30

TABLE II: EXPERIMENTAL SETUP

Classifier	naïve Bayes	k -nearest neighbor
Data type	continuous	
Discretization	mean-standard	-
Normalization	-	min-max, mean-standard
Input parameter	-	$k = 1$
Iteration no.	30	
Validation	20% hold-out cross-validation	

TABLE III: COMPARISON OF THE PROPOSED METHOD AND CONVENTIONAL METHODS

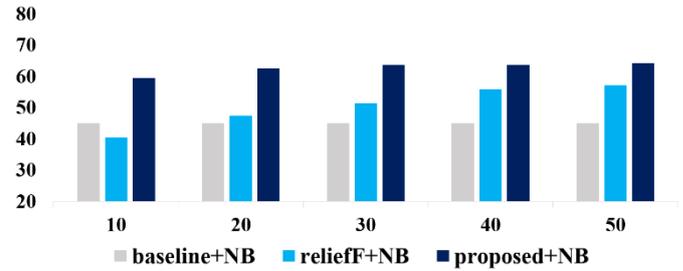
	reliefF	mRMR	Proposed
MIRtool	0.21	17.23	10.02
Jaudio	0.21	40.99	28.31

C. Comparison of Classification Performance

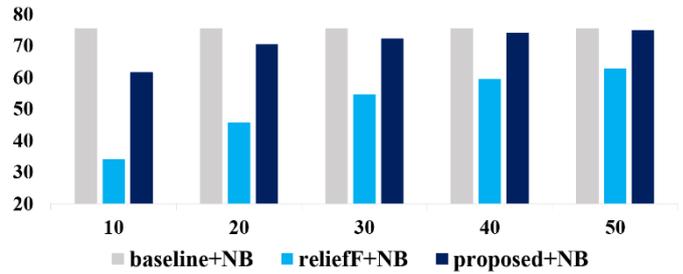
We first investigated the efficiencies of the two feature selection methods. We compared the execution time in seconds for selecting 30 features. Table III indicates that the time cost comparison of conventional methods. The experimental results indicate that our proposed method consumes lower execution time than mRMR. Although the reliefF method finds feature subset faster than other methods, the genre classification performance shows that the feature set selected by reliefF is not useful. Figs. 1-3 show the genre classification results for different classifiers and feature extractors, respectively. The x-axis represents the number of features (10, 20, 30, 40, 50) and the y-axis represents the classification accuracy.

Fig. 1 shows the genre classification accuracy using NB classifier. The base line (gray line) in Fig. 1(a) indicates the result of using all of the features for learning. From the results, it is observed that some irrelevant features are included in a large number of musical features. These irrelevant features degrade the genre recognition performance. Our proposed method is able to filter those irrelevant features and selects

useful musical features that in much informative to genre classification. We see that the proposed method improved the classification performance from 45% (baseline) to 64%. Even though the reliefF+NB showed worse accuracy than the proposed method did, it yielded better results than the baseline in most cases. Fig. 1(b) shows the classification result for a dataset using MIRtoolbox features. Although our result reports the lower recognition performance than the base line that uses 369 full features, we found that our method gives a similar classification accuracy when the size of feature subset reaches 50.

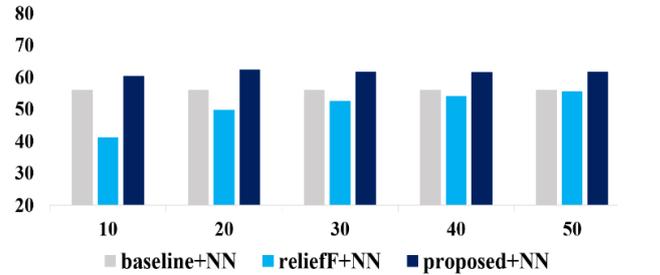


(a) Accuracy for a dataset with JAudio features

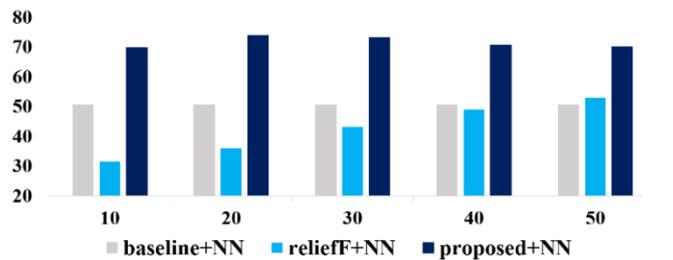


(b) Accuracy for a dataset with MIRtoolbox features

Fig. 1. Comparison of classification accuracy for baseline+NB, reliefF+NB, and the proposed+NB for two data sets.



(a) Accuracy for a dataset with JAudio features

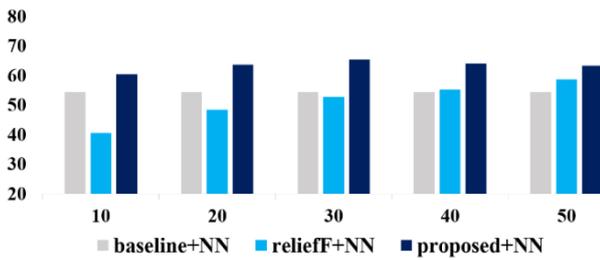


(b) Accuracy for a dataset with MIRtoolbox features

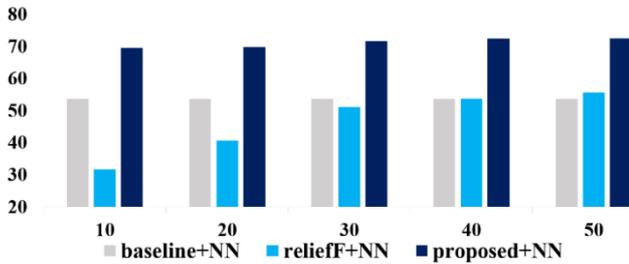
Fig. 2. Comparison of classification accuracy for baseline+NN, reliefF+NN, and the proposed+NN for two data sets. (min-max normalized).

Fig. 2 shows the genre classification accuracy using NN classifier for two normalized data sets. Our result shows the improvement of classification performance, regardless of feature sets. Especially on MIRtoolbox feature set, the proposed method achieved over 20% improvement against

the baseline+NN. The accuracy of the reliefF+NN method is often lower than that of baseline+NN, whereas the proposed method showed reliable improvement in all cases.



(a) Accuracy for a dataset with JAudio features



(b) Accuracy for a dataset with MIRtoolbox features

Fig. 3. Comparison of classification accuracy for baseline+NN, reliefF+NN, and the proposed+NN for two data sets (mean-standard normalized).

	R	E	J	M	C	
R	21	14	8	4	1	R
E	2	8	2	1	1	E
J	4	4	16	0	1	J
M	2	2	1	22	4	M
C	1	4	7	2	23	C

(a) confusion matrix of ReliefF+NN

	R	E	J	M	C	
R	22	7	1	1	0	R
E	4	16	2	0	0	E
J	3	4	25	2	2	J
M	1	2	2	26	0	M
C	0	3	4	0	28	C

(b) confusion matrix of the proposed method+NN

Fig. 4. Confusion matrix of reliefF+NN and the proposed+NN using MIRtoolbox features.

Fig. 3 shows the genre classification accuracy using NN classifier for two mean-standard normalized data sets. The genre classification performance is found to be improved when the proposed feature selection method is combined with NN classifier. We observe that the proposed method provides reliable accuracies regardless of normalization schemes. It is noteworthy that the highest performance is yielded when combined with mean-standard deviation

normalization. Fig. 4 shows the confusion matrices of genre recognition for the reliefF+NN and the proposed+NN; the features are extracted by MIRtoolbox. Due to the space limitation, each genre is represented with a capital; R(rap), E(electronic), J(jazz), M(metal) and C(classic). The rows and columns of each matrix represent actual genre of input songs and their predicted genre that is classified by NN, respectively. Fig. 4(a) is the confusion matrix of the reliefF+NN method. Fig. 4(b) shows the confusion matrix of the proposed+NN. We can see from the confusion matrices that the proposed method is more effective for the classification of electronic and jazz genres than the reliefF+NN method. Compared with the reliefF+NN, genre is represented with a capital; R(rap), E(electronic), J(jazz), M(metal) and C(classic). The rows and columns of each matrix represent actual genre of input songs and their predicted genre that is classified by NN, respectively. Fig. 4(a) is the confusion matrix of the reliefF+NN method. Fig. 4(b) shows the confusion matrix of the proposed+NN. We can see from the confusion matrices that the proposed method is more effective for the classification of electronic and jazz genres than the reliefF+NN method. Compared with the reliefF+NN, the accuracy of the proposed method is improved from 58.06% to 75.48%.

V. CONCLUSION

In this paper, we presented a feature selection algorithm for improving the genre classification. Through the experimental results, we note that our method is able to identify the genre-discriminative musical features. In comparison of other traditional method, the proposed method gives an effective performance for further improvement. We also found that with the only 50 musical features, a classification algorithm is able to learn the inherent rules between musical features and genres considerably well.

REFERENCES

- [1] X. Chen and P. J. Ramadge, "Music genre classification using multiscale scattering and sparse representations," in *Proc. the 47th Annual Conference on Information Sciences and Systems*, MD, USA, March 20-22, 2013, pp. 1-6.
- [2] C. H. Lee, J. L. Shih, K. M. Yu, and H. S. Lin, "Automatic musical genre classification based on modulation spectral analysis of spectral and cepstral features," *IEEE Transaction on Multimedia*, vol. 11, no. 4, pp. 670-682, 2009.
- [3] J. Valverde-Robaza, A. Soriano, L. Berton, F. de Oliveira, M. Cristina, and A. De Andrade Lopes, "Music genre classification using traditional and relational approaches," in *Proc. the Brazilian Conference on Intelligent Systems*, São Carlos-SP, Brazil, October 18-23, 2014, pp. 259-264.
- [4] T. George, E. Georg, and C. Perry, "Automatic musical genre classification of audio signals," presented at the 2nd International Symposium on Music Information Retrieval, Indiana, USA, October 15-17, 2001.
- [5] Y. M. D. Chaturanga and K. L. Jayaratne, "Automatic music genre classification of audio signals with machine learning approach," *Journal on Computing*, vol. 3, no. 2, 2013.
- [6] S. Poria, A. Gelbukh, A. Hussain, S. Bandyopadhyay, and N. Howard, "Music genre classification: a semi-supervised approach," *Pattern Recognition*, Springer Berlin Heidelberg, pp. 254-263, 2013.
- [7] R. Popovici and R. Andonie, "Music genre classification with self-organizing maps and edit distance," presented at International Joint Conference on Neural Network, Killarney, Ireland, July 11-16, 2015.
- [8] P. R. Lisboa de Almeida, A. de Souza Britto, E. J. da Silva Junior, L. E. Soares de Oliveira, T. Montes Celinski, and A. L. Koerich, "Music

genre classification using dynamic selection of ensemble of classifiers,” presented at IEEE International Conference on Systems, Man and Cybernetics, Seoul, Korea, October 14-17, 2012.

- [9] R. Popovici and R. Andonie, “Music genre classification with self-organizing maps and edit distance,” in *Proc. International Joint Conference Neural Networks*, Killarney, Ireland, July 12-17, 2015, pp. 1-7.
- [10] A. Anglade, E. Benetos, M. Mauch, and S. Dixon, “Improving music genre classification using automatically induced harmony rules,” *Journal of New Music Research*, vol. 39, no. 4, pp. 349-361, 2010.
- [11] S. Doraisamy, S. Golzari, N. Mohd, and N. I Udzi, “A study on feature selection and classification techniques for automatic genre classification of traditional malay music,” in *Proc. the ISMIR*, Pennsylvania, USA, September 14-18, 2008, pp. 331-336.
- [12] H. Peng, F. Long, and D. Chris, “Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226-1238, 2005.
- [13] J. Lee and D-W Kim, “Efficient multivariate feature filter using conditional mutual information,” *Electronics Letters*, vol. 48, no. 3, pp. 161-162, 2012.
- [14] Y. C. Lin, Y. H. Yang, H. H. Chen, I. B. Liao, and Y. C. Ho, “Exploiting genre for music emotion classification,” in *Proc. IEEE International Conference on Multimedia and Expo*, NY, USA, June 28-July 3, 2008, pp. 618-621.



Jihae Yoon is currently a master student of the School of Computer Science and Engineering in Chung-Ang Univ. in Seoul, Korea. She started her position from 2015.



Hyunki Lim is currently a Ph.D. candidate at the School of Computer Science and Engineering, Chung-Ang Univ. in Seoul, Korea, in which he joined in 2012. Prior to Ph.D. candidate, he also did his M.S. and the B.S. at Chung-Ang Univ. He is currently interesting in machine learning, specifically mathematical optimization.



Dae-Won Kim is currently a professor in the School of Computer Science and Engineering, Chung-Ang Univ. in Seoul, Korea. Prior to coming to CAU, he did his postdoc, Ph.D., M.S. at KAIST, and the B.S. at Kyungpook Nat'l. Univ., Korea. His research interest includes advanced data mining algorithms with innovative applications to bioinformatics, music emotion recognition, educational data mining, affective computing, and robot interaction.