❐ 1036

Music fingerprinting based on bhattacharya distance for song and cover song recognition

Riyanarto Sarno[1], Dedy Rahman Wijaya[2], Muhammad Nezar Mahardika[3]
[1,2,3]Informatics Department, Institut Teknologi Sepuluh Nopember, Indonesia
[2]School of Applied Science, Telkom University, Indonesia

Article Info

ABSTRACT

People often have trouble recognizing a song especially, if the song is sung by a not original artist which is called cover song. Hence, an identification system might be used to help recognize a song or to detect copyright violation. In this study, we try to recognize a song and a cover song by using the fingerprint of the song represented by features extracted from MPEG-7. The fingerprint of the song is represented by Audio Signature Type. Moreover, the fingerprint of the cover song is represented by Audio Spectrum Flatness and Audio Spectrum Projection. Furthermore, we propose a sliding algorithm and k-Nearest Neighbor (k-NN) with Bhattacharyya distance for song recognition and cover song recognition. The results of this experiment show that the proposed fingerprint technique has an accuracy of 100% for song recognition and an accuracy of 85.3% for cover song recognition.

Corresponding Author:

Dedy Rahman Wijaya,
School of Applied Science,
Telkom University,
Jl. Telekomunikasi Terusan Buah Batu, Bandung 40257, Indonesia.
Email: dedyrw@tass.telkomuniversity.ac.id

1. INTRODUCTION

Music is a popular aspect of human life. Many people make music or listen to it while working, studying, or relaxing. In public places, songs are played to entertain visitors, who are often curious about what music they are listening to and want to know what song they are hearing. In addition, a music recognition system also can be used to detect illegally copied music. One problem that is difficult to handle is detection cover song. Cover songs are sung by different singers. Gender differences in singers, different sound colors, and improvised tones make it harder to detect. Several studies have been done in song recognition and several methods are applied in this field. In the previous studies, song recognition has been done by Multi-band Sum of Spectrogram, with 91% accuracy [1], and the PCA algorithm [2]. Cover song recognition has been done using the raw signal as input [3], [4]. From each of segment of the signal, the chroma will be obtained. This chroma can be used to recognize a cover song with 62% accuracy [5]-[7]. Other experiments used pitch [8]-[10], Information-Theoretic Measures of Similarity [11], music structure segmentation [12] and 2D Fourier Transform [13] to recognize cover songs. In a previous experiment, fingerprinting has been used to identify the title of a song based on the raw signal of the music. In another experiment, the fingerprint was gained and processed in spectrogram form. In our experiment, we used features from MPEG-7. The extraction of music features based on MPEG-7 in the form of a two-dimensional matrix, n x m. MPEG-7 has subband values for each feature. In a previous experiment, song recognition using fingerprinting was done by applying a sliding algorithm on a one-dimensional matrix [14]. This experiment kept the matrix obtained from MPEG-7 extraction. This activity is known as MIR. MIR means processing the music spectrogram to obtain useful information [15]. Song recognition (fingerprinting) and

cover song detection is part of MIR. For example, music recommendation system using collaborative filtering and music genre classification also have been proposed [16], [17]. A fingerprint of song is characteristic for certain types of music. There are many ways to get the fingerprint of a music piece, such as using scikit-learn library or analyzing its spectrogram.

In this paper, we propose an audio fingerprinting technique for song and cover song recognition based on MPEG-7 features. Moreover, MPEG-7 has been reported to successfully detect the mood of music [18], [19] and tempo classification [3]. A cover song means that the singer performs a song originally performed by another artist [20]. Besides the extraction process, the difference between this experiment and previous experiments is in how the fingerprint of the music was obtained. We used MPEG-7 extraction because the result of the extraction process is a number of features that can be useful for obtaining information from a music piece. By MPEG-7 extraction, 17 features can be obtained [21]. However, in this experiment, we used only 3 out these 17 features. For song recognition we used Audio Signature Type, while we used Audio Projection and Audio Spectrum Flatness for cover song recognition. MPEG-7 extraction produces an XML file containing the 17 features from the MPEG-7 DDL scheme. To get the features that exist in the XML file format, XQuery needs to be applied. The selected features are then treated by a sliding algorithm and k-NN algorithm with Bhattacharyya distance.

The remain of this paper is arranged as follows: Section 2 explains the research methods as well as MPEG-7, Bhattacharyya distance, the sliding algorithm, k-NN, discrete wavelet transform, song recognition method, cover song recognition method, the system architecture, and the dataset. Section 3 describes the result that was gotten from the experiment. Section 4 is the conclusion of this paper.

2. RESEARCH METHOD

This section explains the MPEG-7 and required features for this experiment. Furthermore, it also describes the proposed k-NN combined with sliding algorithm, Bhattacharyya distance, discrete wavelet transform, and the dataset.

2.1. Song recognition method

The goal of this method is to recognize the title of a song. The song saved in "wav" format is extracted using MPEG-7 feature extractor. The extraction result is the Audio Signature Type feature. Audio Signature Type from the extraction process is then applied in the sliding algorithm and k-NN using Bhattacharyya distance. In this experiment, k-NN was used because it has been successfully reported for favorable performance in non-stationary signal processing [17], [22], [23]. The details of this process are depicted in Figure 1.
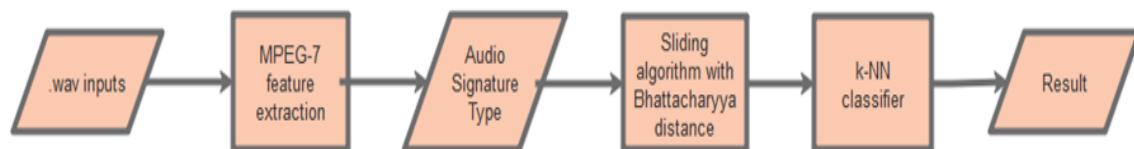

Figure 1. Details of song recognition

2.2. Cover song recognition method

The goal of this method is to identify the title of a cover song. Cover song recognition is an extension of the song recognition method. The song saved in .wav format is extracted using MPEG-7. The difference with song recognition is the number of required features. The extraction process of the cover song recognition method produces Audio Spectrum Flatness and Audio Spectrum Projection. These features are then processed by 2D discrete wavelet transform. The 2D discrete wavelet transform result is applied in the sliding algorithm and k-NN with Bhattacharyya distance. The specifics of this process are shown in Figure 2.
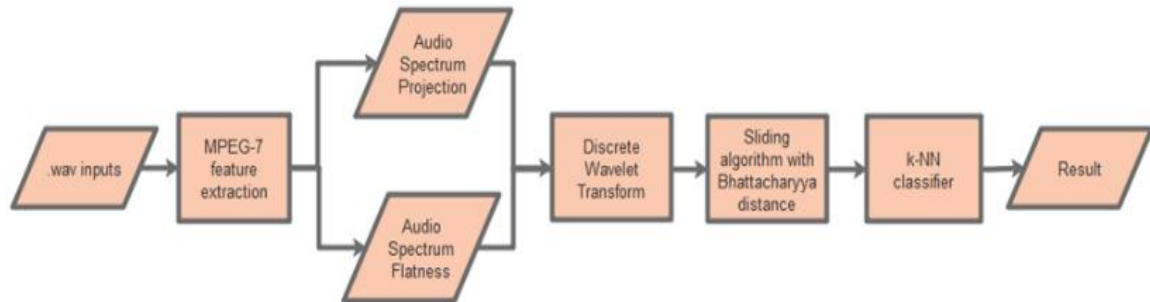
Figure 2. Details of cover song recognition

2.3. Feature extraction

MPEG-7 is a standard description of multimedia content according to the ISO/IEC 15938 standard [24]. The multimedia contents include images, music (sound), and video. However, this study focused only on multimedia content in the form of music. In performing feature extraction, MPEG-7 produces a number of features called low level descriptors. The extracted music features based on MPEG-7 are in the form of metadata which is stored in matrix form. The matrix has a size of n×m. An example of such a matrix can be seen in Figure 3. The m value is called the subband metadata of the music. The subband metadata depend on the features used. The value of n depends on the duration and size of the sound source. Hence, the longer a sound implies the greater value of n obtained by extraction of MPEG-7 features.

$$\begin{bmatrix} Y_{1,1} & Y_{1,2} & \cdots & Y_{1,m-1} & Y_{1,m} \\ Y_{2,1} & Y_{2,2} & \cdots & Y_{2,m-1} & Y_{2,m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ Y_{n-1,1} & Y_{n-1,2} & \cdots & Y_{n-1,m-1} & Y_{n-1,m} \\ Y_{n,1} & Y_{n,2} & \cdots & Y_{n,m-1} & Y_{n,m} \end{bmatrix}$$

Figure 3. Example of an MPEG-7 matrix

The collection of all MPEG-7 features is stored in an XML document with a specific scheme, called the MPEG-7 DDL scheme. In order to get features from the MPEG-7 extraction process, XQuery has to be applied to the XML-document. MPEG-7 has 17 features but we only used three of them in this experiment. For song recognition, we used Audio Signature Type, while for cover song recognition we used Audio Spectrum Projection and Audio Spectrum Flatness. Audio Signature Type represents the 'identity' of an audio signal. Based on MPEG-7 extraction, this feature is the fingerprint of a piece of music. The Audio Signature Type number, m, is 16 based on Figure 1. Audio Signature Type is gained from the audio file saved in "wav" format. The audio files in "wav" format are extracted using MPEG-7 feature extractor. Audio Signature Type is in the form of a matrix. The values in the Audio Signature Type matrix are in the range [-1, 1]. It is this matrix that represents the characteristics of the music and it has to be processed to perform song recognition.

In the cover song recognition experiment we used two features from MPEG-7. Audio Spectrum Projection (ASP) is a feature derived from independent component analysis (ICA) and singular value decomposition (SVD). Audio Spectrum Projection is used to symbolize a low-dimensional feature of the spectrum after projection. Audio Spectrum Projection is a spectrogram that is used for sound classification from various sources. To get the value of Audio Spectrum Projection, several are steps required. Apply (1) to the Audio Spectrum Envelope feature, where l is the index of an AE logarithmic frequency range ($⟦AE⟧\_cb$ (k, l)), and k is the frame index. After that, apply (2) and (3), the goal of which is to normalize each decibel-scale spectral vector with root mean square (RMS), where L is the number of AE spectral coefficients and K is the total number of frames.

$$AE_{cb}(k,l) = 10Log_{10}(AE(k,l)) \tag{1}$$

$$R_l = \sqrt{\sum_{l=1}^{L}(AE_{cb}(k,l))^2}, 1 \leq l \leq L \tag{2}$$

$$X(k,l) = \frac{AE_{cb}(k,l)}{R_l}, 1 \leq k \leq K \tag{3}$$

Audio Spectrum Flatness is the second MPEG-7 feature used in cover song recognition. Audio Spectrum Flatness (ASF) reflects the flatness properties of a signal. The flatness properties are defined by a specified number of frequency bands. ASF shows how the power spectrum of a signal deviates from the frequencies of a flat shape. The term 'flat shape' describes the noise or impulse in a signal. This feature is used to measure the similarity between one audio signal and another. The first step of Audio Spectrum Flatness extraction is calculating the power spectrum of the signal [21]. Then, it is divided into loSign and hiSign.

The value of loSign and hiSign are determined by (4) and (5), respectively, where the value of Y determines the lower band edge. The minimum value for loSign is recommended to be 250 Hz, so n=-8, and D is the desired number of frequency bands. Defining the frequency bands with hiSign and loSign makes Audio Spectrum Flatness too sensitive to the sampling frequency. Hence, it needs to be modified so that all bands slightly overlap each other. The modification of hiSign and loSign can be seen in (6) and (7). Then for each frequency band d, Audio Spectrum Flatness can be calculated with (8).

$$loSign = 2^{\frac{1}{4}Y} \times 1\,kHz \tag{4}$$

$$hiSign = 2^{\frac{1}{4}D} \times loSign \tag{5}$$

$$loF_d = 0.95 \times loSign \times 2^{\frac{1}{4}(d-1)} \tag{6}$$

$$hiF_d = 1.05 \times loSign \times 2^{\frac{1}{4}b} \tag{7}$$

$$ASF = \frac{hiK'_b - loK'_b + 1\sqrt{\prod_{k'=loK'_b}^{hiK'_b} P_g(k')}}{\frac{1}{hiK'_d - loK'_d + 1}\sum_{k'=loK'_d}^{hiK'_d} P_g(k')} \qquad (1 \leq d \leq D). \tag{8}$$

2.4. Bhattacharyya distance

Bhattacharyya distance does not really look for the distance of anything. Unlike Euclidean distance, Bhattacharyya distance is more like finding the similarity between two distributions [25]. The distribution refers to the Bhattacharyya coefficient. This coefficient is usually used to discover the relative closeness of two vectors. The formula of Bhattacharyya distance can be seen in (9), where D_B (r, s) is the distance between r and s distributions or classes.

$$D_B(r,s) = -\ln(BC(r,s)) \tag{9}$$

Where

$$BC(r,s) = \sum_{x \in X} \sqrt{r(x)s(x)} \tag{10}$$

2.5. Discrete wavelet transform (DWT)

Discrete wavelet transform (DWT) is generally used in signal processing. DWT is a method used to reconstruct a signal but still retains the original signal. It was also successfully applied to process signals electronic nose [22], [26], [27]. Signals that wavelet transform has been applied to have a shorter length than the original. There are two filters that can be taken when applying wavelet transform to an audio signal, i.e. low-pass filter and high-pass filter. The low-pass filter yields approximate coefficient which represents the voice of the singer of the song (in this case women and men have almost the same values) and the detail coefficient which represents the sound of the instruments of the song. In this study, the values taken from the DWT were the approximate coefficient only. DWT is highly dependent on the level of decomposition that is used. The higher the level of decomposition indicates the higher the risk that the signal will be damaged. Conversely, the lower the level of decomposition, the more susceptible the signal will be to noise [28]. Level decomposition up to 4 with wavelet family haar were used in this experiment. All levels are gathered by the low-pass filter, so they can be classified. 2-D DWT is applied to each feature, which is still in a matrix form. Figure 4 denotes an example of audio signal that reconstructed by DWT.
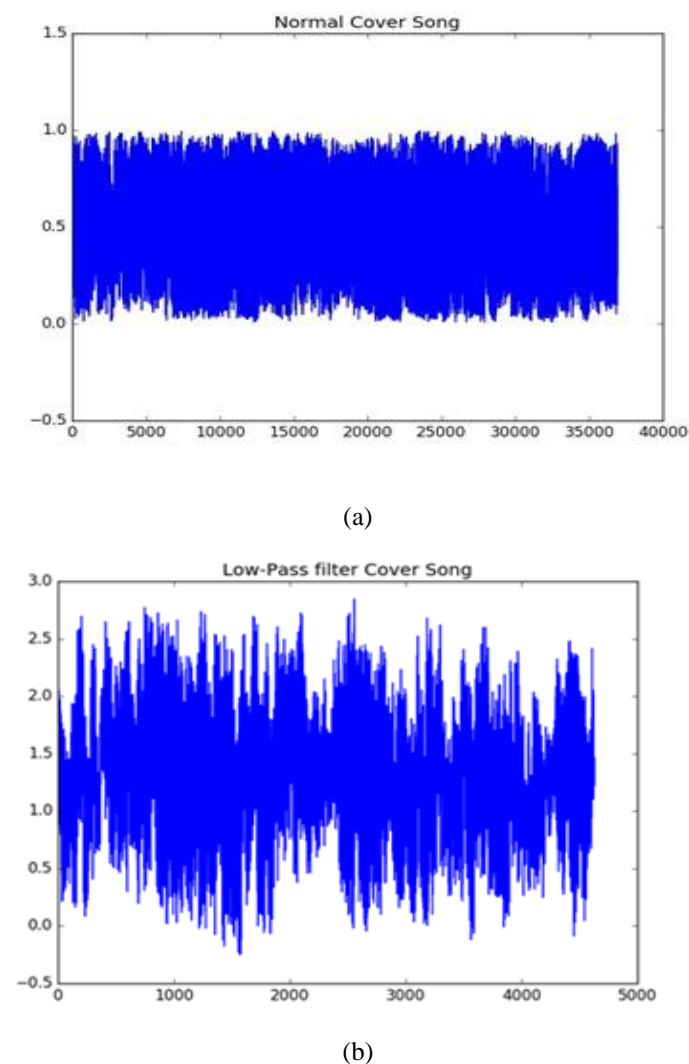
(a)

(b)

Figure 4. The example of audio signal from MPEG-7 extraction, (a) Original signal (b) Reconstructed signal

2.6. k-NN combined with sliding algorithm

As mentioned above, the features of MPEG-7 are in the form of a matrix. Moreover, k-NN is used as a classification method. Common k-NN methods analyze training data and testing data then measure the distance between them using Euclidean distance. However, in this experiment, we used Bhattacharyya distance to measure distance. Both training data and testing data are in the form of a vector, but the feature extraction result of MPEG-7 is in the form of a two-dimensional matrix. In order to keep the matrix form, k-NN is combined with a modified sliding algorithm.

In this experiment, we did not 'slide' each subband (row). Bhattacharyya distance allows us to calculate the average value of each matrix because Bhattacharyya distance measures the distance between vectors, not between points. The average value of each of the matrices is 'slided' and the Bhattacharyya distance is calculated. This method is used both for song recognition and cover song recognition. In this experiment, the k parameter of k-NN is 5. The difference between song recognition and cover song recognition was the type and number of features used.

2.7. System architecture

Figure 5 illustrates the details of the system architecture used in this experiment.
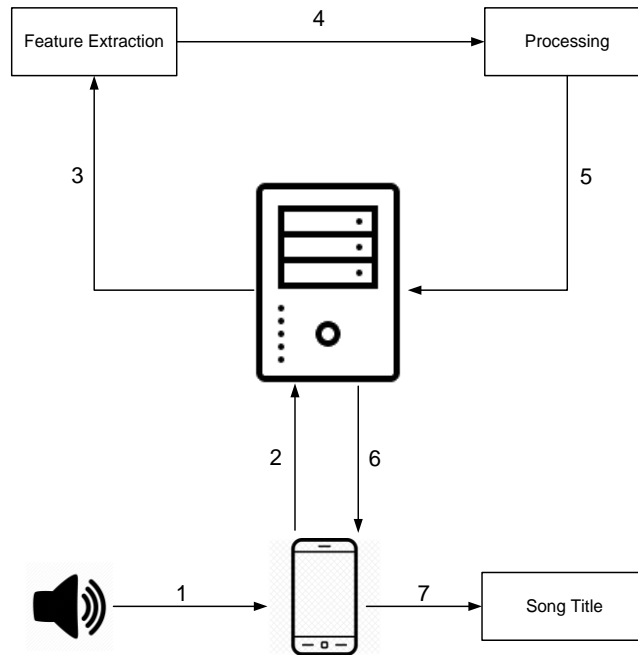
Figure 5. System architecture

The explanation of the process flow is as follows:
1. A song is recorded through a mobile device application. The recorded song should be in "wav" format
2. The recorded song is sent to a server
3. The recorded song is extracted using MPEG-7
4. The result of the extraction process is an XML document that contains several features. Then XQuery is used to obtain the required features
5. The required features are applied in the preprocessing method
6. The result from the preprocessing stage is used according to the method being executed in the classification stage
7. The result from processing will be shown in mobile device application as a result

2.8. Dataset

The audio dataset was obtained from YouTube. The audio dataset has to be in "wav" format. The dataset that was used as training data for song recognition and cover song recognition was one whole song. The training data had duration of 3-4 minutes. After extraction the signal had a length of 300-400 subbands (rows). The dataset for testing data was randomly cut to a length of 30 seconds. After extraction, the signal had a length of 20-30 subbands (rows). The subbands of all training data were then 'slided' along the subbands of the testing data. The result of this 'sliding' is the distance between the testing data and the training data. Then, the shortest distance is searched. The training data and testing data scheme is shown in Figure 6.

For the song recognition experiment, thirty three songs were used to recognize. Each song was randomly selected. Cover song recognition used fifty songs with five unique song titles. For each song there were five cover songs with a male singer and five cover songs with a female singer.
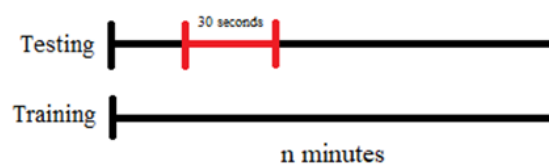

Figure 6. Illustration of the dataset in this experiment

3.   RESULTS AND ANALYSIS
This section discusses the result of experiment including song and cover song recognition.

3.1. Song recognition result
The training data for fingerprint processing consisted of 33 songs. The songs were recorded using a mobile device. The Audio Signature Type feature of each song needs to be saved in a database. In this experiment, we used a whole song as training data. The testing data used were 30 seconds of randomly cut training data. The accuracy of the song recognition method was calculated using (11).

$$ACC = \frac{TRUE}{DATA} \; x \; 100\% \tag{11}$$

Where TRUE is the total of valid predictions and DATA is the total number of testing data. The result of song recognition testing using Bhattacharyya distance is shown in Table 1. The title of the song denotes training data and testing data used in this experiment. In this experiment, all songs were recognized correctly. According to Table 2. and (11), the testing scenario for song recognition using Bhattacharyya distance had a recognition accuracy of 100%.

Table 1. Results of Fingerprint Prediction

No	Title of the Song	Result	No	Title of the Song	Result
1	Middle	√	18	Jera	√
2	Tremor	√	19	Kereta Malam Kop	√
3	Faded	√	20	Kereta Malam Ori	√
4	Turn Up The Speaker	√	21	Kimcil Kepolen	√
5	Saxobeats	√	22	Klepek Klepek	√
6	Livin On A Prayer	√	23	Bizarre Love Triangle	√
7	Careless Whisper	√	24	Asal Kau Bahagia	√
8	It's My Life	√	25	Biar Ku Sendiri	√
9	Thing Will Get Better	√	26	Bagaimana Ku Tahu	√
10	Everybody Knew	√	27	Berdistraksi	√
11	Get Low	√	28	Sayang	√
12	Judi	√	29	Berdua Saja	√
13	Begadang	√	30	Rahasia	√
14	Benang Biru	√	31	How Deep is Your Love	√
15	Goyang Dumang	√	32	Salamku Untuk Kekasihmu yang Baru	√
16	Janji Palsu	√	33	Show Me the Meaning of Being Lonely	√
17	Jaran goyang	√			

Table 2. Confusion Matrix for Fingerprint Testing Scenario

		System Predict	
		True	False
Label	True	33	0
	False	0	0

3.2. Cover song recognition result
To calculate the accuracy of cover song recognition, we used cross validation. The cross validation method used in this experiment was k-fold cross validation. In k-fold cross validation, all entries in the original training data set are used for both training as well as validation. Also, each entry is used for validation just once. K-fold estimates the accuracy of the machine learning model, in this case k-NN, by averaging the accuracies derived from all k cases of cross validation.

In this experiment, the value of k used was 3. Hence, there were ten iterations to calculate the accuracy of our model. The result of cross validation for cover song recognition is shown in Table 3. From the Table 3., we can see that the accuracy of cover song recognition was 85.3%. The standard deviation of cover song recognition was 8.18%, which indicates that the classification was quite stable.

Table 3. Cross Validation Result

K	Accuracy (%)
1	86.6
2	75.5
3	81.1

K	Accuracy (%)
4	81.1
5	97.7
6	75.5
7	90.5
8	96.4
9	90.5
10	78.8
Average	85.3

4.   CONCLUSION

In order to recognize a song, the MPEG-7 feature Audio Signature Type was used as fingerprint or 'identity' of a song. In cover song recognition, the MPEG-7 features Audio Spectrum Projection and Audio Spectrum Flatness features were used. For both song recognition and cover song recognition a k-NN classifier was used. Because the data are originally in matrix form, they were modified by combining them with a sliding algorithm and Bhattacharyya distance. Accuracy for song recognition of this experiment was 100%. After calculating the accuracy using cross validation the recognition accuracy is 85.3%. These results indicate that the proposed method can obtain favorable performance to recognize song and cover song.

REFERENCES

[1]   S. S. Ranjani, K. Karthik, P. K. Bora, and V. Abdulkareem, "Multi-band Sum of Spectrogram Based Audio Fingerprinting of Indian Film Songs for Multi-Lingual Song Retrieval," *2015 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2015*, pp. 1414-1420, 2015.

[2]   I. Mueller and R. Welter, "Robust Audio Fingerprinting for Song Identification," *in 2004 12th European Signal Processing Conference, 2004*, pp. 2095-2098, 2004.

[3]   A. Lazaro, R. Sarno, J. A. R, and M. N. Mahardika, "Music Tempo Classification Using Audio Spectrum Centroid , Audio Spectrum Flatness, and Audio Spectrum Spread based on MPEG-7 Audio Features," in *3rd International Conference on Science in Information Technology (ICSITech)*, pp. 41-46, 2017.

[4]   J.-T. Fang, Y.-R. Chang, and P.-C. Chang, "Deep Learning of Chroma Representation for Cover Song Identification in Compression Domain," *Multidimensional Systems and Signal Processing*, pp. 1-16, Feb. 2017.

[5]   T. M. Chang, E. T. Chen, C. B. Hsieh, and P. C. Chang, "Cover Song Identification with Direct Chroma Feature Extraction From AAC Files," in *2013 IEEE 2nd Global Conference on Consumer Electronics (GCCE)*, pp. 55-56, 2013.

[6]   A. Camarena-Ibarrola, K. Figueroa, and H. Tejeda-Villela, "Entropy Per Chroma for Cover Song Identification," in *2016 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, pp. 1-6, 2016.

[7]   T. Bertin-Mahieux and D. P. W. Ellis, "Large-Scale Cover Song Recognition Using Hashed Chroma Landmarks," in *2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 117-120, 2011.

[8]   N. Chen, J. S. Downie, H. Xiao, and Y. Zhu, "Cochlear Pitch Class Profile for Cover Song Identification," *Applied Acoustics*, vol. 99, pp. 92-96, Dec. 2015.

[9]   N. Chen, W. Li, and H. Xiao, "Fusing Similarity Functions for Cover Song Identification," *Multimedia Tools and Applications*, pp. 1-24, Feb. 2017.

[10] A. Degani, M. Dalai, R. Leonardi, and P. Migliorati, "A Heuristic for Distance Fusion in Cover Song Identification," in *2013 14th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, pp. 1-4, 2013.

[11] P. Foster, S. Dixon, and A. Klapuri, "Identifying Cover Songs Using Information-theoretic Measures of Similarity," *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, vol. 23, no. 6, pp. 993-1005, Jun. 2015.

[12] K. Cai, D. Yang, and X. Chen, "Two-Layer Large-Scale Cover Song Identification System Based on Music Structure Segmentation," in *2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)*, pp. 1-6, 2016.

[13] P. Seetharaman, "Cover Song Identification With 2d Fourier Transform Sequences," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017.

[14] S. D. You, W.-H. Chen, and W.-K. Chen, "Music Identification System Using MPEG-7 Audio Signature Descriptors," *The Scientific World Journal*, vol. 2013, p. e752464, Mar. 2013.

[15] "What is Music Information Retrieval?" [Online]. Available: https://musicinformationretrieval.com/why_mir.html.

[16] M. Sunitha and T. Adilakshmi, "Music Recommendation System with User-Based and Item-Based Collaborative Filtering Technique," *Indonesian Journal of Electrical Engineering and Informatics*, vol. 5, no. 1, pp. 59-68, 2017.

[17] M. Sudarma and I. G. Harsemadi, "Design and Analysis System of KNN and ID3 Algorithm for Music Classification Based on Mood Feature Extraction," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 7, no. 1, pp. 486-495, 2017.

[18] R. Sarno, J. A. Ridoean, D. Sunaryono, and D. R. Wijaya, "Classification of Music Mood Using MPEG-7 Audio Features and SVM with Confidence Interval," *International Journal on Artificial Intelligence Tools*, vol. 27, no. 5, 2018.

[19] J. A. Ridoean, R. Sarno, D. Sunaryo, and D. R. Wijaya, "Music Mood Classification Using Audio Power And Audio Harmonicity Based on MPEG-7 Audio Features And Support Vector Machine," in *Proceeding-2017 3rd International Conference on Science in Information Technology: Theory and Application of IT for Education, Industry and Society in Big Data Era, ICSITech 2017*, vol. 2018, Janua, 2018.

[20] J. Serrà, E. Gómez, P. Herrera, and X. Serra, "Chroma Binary Similarity and Local Alignment Applied to Cover Song Identification," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 16, no. 6, pp. 1138-1151, 2008.

[21] H.-G. Kim, N. Moreau, and T. Sikora, MPEG-7 Audio and Beyond: Audio Content Indexing and Retrieval. John Wiley & Sons, 2005.

[22] D. R. Wijaya, R. Sarno, E. Zulaika, and S. I. Sabila, "Development of Mobile Electronic Nose for Beef Quality Monitoring," in *4th Information Systems International Conference 2017, ISICO 2017, Procedia Computer Science*, vol. 124, no. November, pp. 728-735, 2017.

[23] Hariyanto, R. Sarno, and D. R. Wijaya, "Detection of Diabetes From Gas Analysis of Human Breath Using e-Nose," in *2017 11th International Conference on Information & Communication Technology and System (ICTS)*, vol. 0, pp. 241-246, 2017.

[24] "ISO/IEC 15938-1:2002 - Information technology -- Multimedia content description interface -- Part 1: Systems," ISO.

[25] T. Kailath, "The Divergence and Bhattacharyya Distance Measures in Signal Selection," IEEE Transactions on Communication Technology, vol. 15, no. 1, pp. 52-60, 1967.

[26] D. R. Wijaya, R. Sarno, and E. Zulaika, "Sensor Array Optimization for Mobile Electronic Nose: Wavelet Transform and Filter Based Feature Selection Approach," *International Review on Computers and Software*, vol. 11, no. 8, pp. 659-671, 2016.

[27] D. R. Wijaya, R. Sarno, and E. Zulaika, "Gas Concentration Analysis of Resistive Gas Sensor Array," in *2016 IEEE International Symposium on Electronics and Smart Devices*, pp. 337-342, 2016.

[28] D. R. Wijaya, R. Sarno, and E. Zulaika, "Information Quality Ratio as a Novel Metric for Mother Wavelet Selection," *Chemometrics and Intelligent Laboratory Systems*, vol. 160, no. November 2016, pp. 59-71, 2016.

BIOGRAPHIES OF AUTHORS

Riyanarto Sarno is a Professor, Informatics Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. He received the bachelor's degree in Electrical Engineering from Institut Teknologi Bandung, Bandung, Indonesia in 1987. He received M.Sc and Ph.D in Computer Science from the University of Brunswick Canada in 1988 and 1992, respectively. In 2003 he was promoted to a Full Professor. His teaching and research interest include Internet of Things, Process Aware Information Systems, Intelligent Systems and Smart Grids.

Dedy Rahman Wijaya is currently a lecturer and working towards the Ph.D. degree at Informatics Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. He received the bachelor's degree in Computer Science from Telkom University (former STT Telkom), Bandung, Indonesia in 2006. He received the master's degree in Computer Science from Institut Teknologi Bandung, Bandung, Indonesia in 2010. His main research interests include cyber-physical system, intelligent system, machine learning, and signal processing.

Muhammad Nezar Mahardika is an Bachelor student from the Informatic Department, Institut Teknologi Sepuluh Nopember, Indonesia. His research interest include machine learning, signal processing, and audio analysis.