turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Linggo Sumarno
Assignment title:	Quick Submit
Submission title:	Chord Recognition using FFT Based…
File name:	T_Based_Segment_Averaging_and…
File size:	621.1K
Page count:	5
Word count:	4,429
Character count:	22,990
Submission date:	13-Nov-2020 11:48AM (UTC+0700)
Submission ID:	1444665717

2020 8th International Conference on Information and Communication Technology (ICoICT)

Chord Recognition using FFT Based Segment Averaging and Subsampling Feature Extraction

Linggo Sumarno
Electrical Engineering Study Program
Sanata Dharma University
Yogyakarta, Indonesia
lingsum@usd.ac.id

***Abstract*—This paper proposes a feature extraction subsystem for a chord recognition system, which gives a fewer number of feature extraction coefficients than the previous ones. The method of the proposed feature extraction is FFT (Fast Fourier Transform) based segment averaging and subsampling. Guitar chords were used in developing the proposed feature extraction. In general, the method of the proposed feature extraction is as follows. Firstly, the input signal is transformed using FFT. Secondly, the left half portion of the transformed signal is processed in succession using SHPS (Simplified Harmonic Product Spectrum), logarithmic scaling, segment averaging, and subsampling. The output of subsampling is the result of the proposed feature extraction. Based on the test results, the proposed feature extraction was quite efficient for use in a chord recognition system. For the recognition rate category above 98%, the chord recognition system only required a number of seven feature extraction coefficients. In addition, for the recognition rate category above 90%, the chord recognition system only required a number of six feature extraction coefficients.**

Keywords*—chord recognition, FFT, feature extraction, segment averaging, subsampling

I. Introduction

A chord will be generated if two or more notes are played at the same time. The naming of a chord is based on the origin of the notes being played. For example, if the origin of the notes are the first, third, and fifth notes on a major scale, then the chord is included in the major chord. There are many kinds of chord variations. This variation depends on the scale and the order of the notes on the scale [1].

Nowadays computers can be developed to be able to recognize the chords that are being played. Based on the previous works on chord recognition, the topic of chord recognition, which is using chroma features, is a popular topic. PCP (Pitch Class Profile) is a feature extraction that uses chroma features. PCP was introduced by Fujishima [2]. PCP produced a number of 12 feature extraction coefficients. Each feature extraction coefficient from PCP represents the power of the fundamental frequencies that exist in a chord. The original PCP from Fujishima, is still popular today. Recent studies still use the original PCP [3] [4].

In addition to the original PCP from Fujishima above, there were also a number of derivatives from the PCP. A number of PCP derivatives include, PCP that used statistical features [5], PCP that used logarithmic compression [6], Enhanced PCP [7], Improved Chromagram [8], CRP (Chroma DCT-Reduced log Pitch) Enhanced PCP [9] , and Improved PCP [10]. A number of PCP derivatives also produced a number of 12 feature extraction coefficients.

Based on a previous work also, there was a topic of a chord recognition whose feature extraction did not use chroma features. Sumarno [11] introduced a chord recognition whose feature extraction used FFT (Fast Fourier Transform)-based segment averaging. This kind of feature extraction gave a number eight feature extraction coefficients. If we look at the number of feature extraction coefficients, we can see that the research of chord recognition for reducing the number of feature extraction coefficients is still open.

This work developed further a chord recognition whose feature extraction used the above the FFT-based segment averaging [11], by adding a process called the subsampling process. The aim of adding this process was to give even a smaller number of the feature extraction coefficients, which was less than eight coefficients. As a note, this work used the guitar chords.

II. Research Methodology

A. Development of the chord recognition system and the feature extraction subsystem

The first step in the research methodology was the development of a chord recognition system, which was focused on the feature extraction subsystem. The block

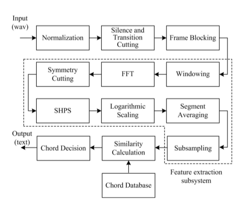

Fig. 1. The block diagram of the developed chord recognition system and its feature extraction subsystem.

978-1-7281-6141-9/2020/$31.00 ©2020

Chord Recognition using FFT Based Segment Averaging and Subsampling Feature Extraction

by Sumarno Linggo

Submission date: 13-Nov-2020 11:48AM (UTC+0700)

Submission ID: 1444665717

File name: T_Based_Segment_Averaging_and_Subsampling_Feature_Extraction.pdf (621.1K)

Word count: 4429

Character count: 22990

Chord Recognition using FFT Based Segment Averaging and Subsampling Feature Extraction

Linggo Sumarno
Electrical Engineering Study Program
Sanata Dharma University
Yogyakarta, Indonesia
lingsum@usd.ac.id

***Abstract*—This paper proposes a feature extraction subsystem for a chord recognition system, which gives a fewer number of feature extraction coefficients than the previous ones. The method of the proposed feature extraction is FFT (Fast Fourier Transform) based segment averaging and subsampling. Guitar chords were used in developing the proposed feature extraction. In general, the method of the proposed feature extraction is as follows. Firstly, the input signal is transformed using FFT. Secondly, the left half portion of the transformed signal is processed in succession using SHPS (Simplified Harmonic Product Spectrum), logarithmic scaling, segment averaging, and subsampling. The output of subsampling is the result of the proposed feature extraction. Based on the test results, the proposed feature extraction was quite efficient for use in a chord recognition system. For the recognition rate category above 98%, the chord recognition system only required a number of seven feature extraction coefficients. In addition, for the recognition rate category above 90%, the chord recognition system only required a number of six feature extraction coefficients.**

Keywords—chord recognition, FFT, feature extraction, segment averaging, subsampling

I. INTRODUCTION

A chord will be generated if two or more notes are played at the same time. The naming of a chord is based on the origin of the notes being played. For example, if the origin of the notes are the first, third, and fifth notes on a major scale, then the chord is included in the major chord. There are many kinds of chord variations. This variation depends on the scale and the order of the notes on the scale [1].

Nowadays computers can be developed to be able to recognize the chords that are being played. Based on the previous works on chord recognition, the topic of chord recognition, which is using chroma features, is a popular topic. PCP (Pitch Class Profile) is a feature extraction that uses chroma features. PCP was introduced by Fujishima [2]. PCP produced a number of 12 feature extraction coefficients. Each feature extraction coefficient from PCP represents the power of the fundamental frequencies that exist in a chord. The original PCP from Fujishima, is still popular today. Recent studies still use the original PCP [3] [4].

In addition to the original PCP from Fujishima above, there were also a number of derivatives from the PCP. A number of PCP derivatives include, PCP that used statistical features [5], PCP that used logarithmic compression [6], Enhanced PCP [7], Improved Chromagram [8], CRP (Chroma DCT-Reduced log Pitch) Enhanced PCP [9] , and Improved PCP [10]. A number of PCP derivatives also produced a number of 12 feature extraction coefficients.

Based on a previous work also, there was a topic of a chord recognition whose feature extraction did not use chroma features. Sumarno [11] introduced a chord recognition whose feature extraction used FFT (Fast Fourier Transform)-based segment averaging. This kind of feature extraction gave a number eight feature extraction coefficients. If we look at the number of feature extraction coefficients, we can see that the research of chord recognition for reducing the number of feature extraction coefficients is still open.

This work developed further a chord recognition whose feature extraction used the above the FFT-based segment averaging [11], by adding a process called the subsampling process. The aim of adding this process was to give even a smaller number of the feature extraction coefficients, which was less than eight coefficients. As a note, this work used the guitar chords.

II. RESEARCH METHODOLOGY

A. Development of the chord recognition system and the feature extraction subsystem

The first step in the research methodology was the development of a chord recognition system, which was focused on the feature extraction subsystem. The block

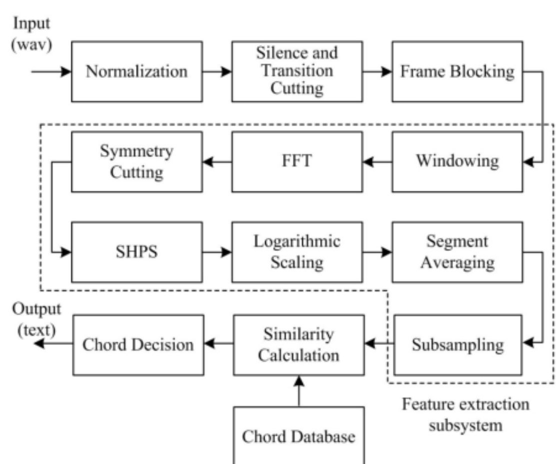

Fig. 1. The block diagram of the developed chord recognition system and its feature extraction subsystem.

diagram of the chord recognition system and its feature extraction subsystem are shown in Fig. 1. As shown in Fig. 1, the input system is a chord signal. This signal is an isolated chord signal recorded in wav format. The system output is a text, which indicates a recognized chord signal.

As a first note, if we look at each process in the feature extraction subsystem, it is the common process. However if we look at as a whole, there is a novelty in the feature extraction subsystem compares with the previous one [11]. As a second note, the implementation of the chord recognition system was carried out using Python software. In a more detail, the input and the function of each block in Fig. 1 are described as follow.

1) Input: The input of the chord recognition system is a chord signal. This chord signal came from the Yamaha CPX 500-II acoustic-electric guitar, as shown in Fig. 2. This chord signal is an isolated chord signal, which recorded in wav format. There are seven chord signals from the major chords of C, D, E, F, G, A, and B [11]. Chord signals recording was carried out using a sampling frequency that met Shannon's sampling theorem [12]:

$$f_s \geq 2\, f_{max} \tag{1}$$

where f_s is the sampling frequency, and the $f_{\max}$ is the highest frequency component of the chord signal to be sampled. This study used a sampling frequency of 5000 Hz. According to the Shannon's sampling theorem above, the sampling frequency had exceeded the highest frequency component of 392 Hz (G4 tone) of the G chord [11].

The duration of the chord signals recording was two seconds. Based on the results of visual evaluations of the amplitude of the chord signal, the choice of two seconds duration was enough to get more than half of the chord signal that already in a steady state condition. As a note, the accurate chord information was available in this steady state condition.

2) Normalization: Normalization is a process for equalizing the maximum value of the sequence of signal data input. In this case, the maximum value is -1 or 1. Normalization is formulated as below.

$$\mathbf{x}_{\text{out}} = \mathbf{x}_{\text{in}} \,/\, \max\left(|\mathbf{x}_{\text{in}}|\right) \tag{2}$$

where $\mathbf{x}_{\text{in}}$ and $\mathbf{x}_{\text{out}}$ are the input and output signal data sequences of the normalization process, respectively. The normalization process is needed because the input signal data sequences have different maximum values.

3) Silence and transition cutting: Silence and transition cutting is a process for removing the silence and transition region of a signal data sequence. This region is on the left side of a signal data sequence. Silence and transition cutting is carried out as follows. Firstly, based on the visual observations, in order to remove the silence region, this work needed the data threshold |0.5|. Starting from the leftmost data of the signal data sequence, if the data was less than |0.5| then the data was removed. Secondly, based on the visual observations also, in order to remove the transition region, this work needed a duration of 200 milliseconds at the leftmost region of the signal data sequence to be removed [11].

Fig. 2. The guitar for this work.

4) Frame blocking: Frame blocking is a process for acquiring a short signal data sequence, which called a signal frame, from a long signal data sequence [13]. Frame blocking is carried out by acquiring a signal frame at the leftmost region of the long signal data sequence. The purpose of using the frame blocking process is to reduce the number of signal data to be processed further. The effect of reducing the number of signal data is the reduction in computational time needed for signal data processing. This work used 256 points of blocking frame length [11].

5) Windowing: In the time domain, windowing is a process for decreasing the discontinuities that appear at the left and the right edges of the signal data sequence [13]. In the frequency domain, this reduction will eliminate the emergence of spectral leakage at the output the FFT process. This work used Hamming window [14] for windowing process. This kind of window has been widely used in the digital signal processing field [15]. The width of the window was the same as the frame blocking length.

6) FFT: FFT is a process for transforming a signal data sequence from the time domain to the frequency domain. This work used FFT radix-2. This kind of FFT has been widely used in the digital signal processing field [15]. The length of the FFT was the same as the frame blocking length. In addition, there were additional calculations of absolute values for the FFT results. This was necessary because the subsequence process, namely SHPS, required positive values.

7) Symmetry cutting: Symmetry cutting is a process for removing the right half portion of the FFT result. As a note, the left and the right half portion of the FFT result show a symmetry property. Therefore, it was sufficient if this work used only the left portion of the FFT result.

8) SHPS: SHPS (Simplified Harmonic Product Spectrum) is a process for reducing the harmonic signal data. The reduction of these harmonic signal data, visually, will show a clearer difference, between a signal data sequence and the other signal data sequences [11]. This work used the SHPS that was introduced by Sumarno [11]. This SHPS is a derivative of HPS (Harmonic Product Spectrum) that was introduced by Noll [16].

9) Logarithmic scaling: Logarithmic scaling is a process for increasing the number of significant local peaks. Based on the previous research [11], in the chord recognition that using segment averaging, if we increase the number of significant local peaks, it can increase the recognition rate. Logarithmic scaling is formulated as below.

$$\mathbf{X}_{out} = \log\left(\alpha \mathbf{X}_{out} + 1\right) \tag{3}$$

where $\mathbf{X}_{in}$ and $\mathbf{X}_{out}$ are respectively the input and the output signal data sequences, from the logarithmic scaling process. The α value is the logarithmic scale factor. Adding the value '1' to the logarithmic scale factor formula is to avoid the zero value logarithm, which will give an infinite value. This work used a logarithmic scale factor of 50 [11].

10) Segment averaging: Segment averaging is a process for obtaining a short signal data sequence from a long signal data sequence. The segment averaging was initially inspired from Setiawan [17]. Next, the segment averaging was developed by Sumarno [18]. Algorithmically, the segment averaging process is shown below.

1. Suppose $\mathbf{Y}_{in}$ is an input signal data sequence in the segment averaging process. $\mathbf{Y}_{in}$ has positive values and has a length N, with $N = 2^q$ for $q \geq 0$ which is a positive integer.
2. Set the segment length L points, with $L = 2^p$ for $0 \leq p \leq q$ which is also a positive integer.
3. Cut $\mathbf{Y}_{in}$ uniformly using the segment length L points. This cutting will give a number of M segments namely $\mathbf{S}_1$, $\mathbf{S}_2$, …, $\mathbf{S}_M$ with

$$M = N / L \tag{4}$$

   As a note, each segment has a data sequence with the length L points.

4. Calculate the average value of the data sequence in each segment, and then arrange it into the following $\mathbf{Y}_{out}$ data sequence.

$$\mathbf{Y}_{out} = \{\, \text{avg}(\mathbf{S}_1), \text{avg}(\mathbf{S}_2), \ldots, \text{avg}(\mathbf{S}_M)\}$$

The $\mathbf{Y}_{out}$ data sequence is the output signal data sequence in the segment averaging process. This $\mathbf{Y}_{out}$ data sequence has the length M points. Based on equation (4), the output of the segment averaging has a length of 2^n points, with $n = q - p$. This work used the segment averaging output length 4, 8, 16, and 32 points.

11) Subsampling: Subsampling is an advanced process of the segment averaging process above, in order to get the signal data sequences that are even shorter. In this work, the output of subsampling process is also called the result of the feature extraction data. The effect of obtaining the shorter feature extraction data is the reduction in storage that needed to store a number of the feature extraction data. Interpolation is a way to do the subsampling process. An interpolation method that has been widely used is spline interpolation [19]. In this work evaluated linear, quadratic, and cubic spline interpolation methods [20]. In addition, this work also evaluated the subsampling output of 1-8 coefficients.

12) Chord database: Chord database is a collection of a number of chords reference feature extraction (C, D, E, F, G, A, and B). The development of the chord database was carried out as follows. The first one was recording a number of 10 training chord samples for each chord (C, D, E, F, G, A, and B). It was assumed that by using a number of 10 training chord samples, all variations of each chord signal from a guitar musical instrument have been obtained. The second one was processing the feature extraction of all the training chord samples by carrying out the normalization process up to subsampling process that are shown in Fig. 1. For each chord (C, D, E, F, G, A, and B), the feature extraction of 10 training chord samples will produce 10 feature extraction data sequence. The third one was carrying out the averaging calculation for each chord as follows.

$$\mathbf{Z}_T = \frac{1}{10}\sum_{i=1}^{10} \mathbf{Y}_{T,i} \tag{5}$$

where T is a chord (C, D, E, F, G, A, or B), $\{\mathbf{Y}_{T,i} \mid 1 \leq i \leq 10\}$ are 10 feature extraction of a T chord, and $\mathbf{Z}_T$ is a reference feature extraction from the average of 10 feature extractions of a T chord. The last one was collecting a number of seven reference feature extractions namely $\mathbf{Z}_C$, $\mathbf{Z}_D$, $\mathbf{Z}_E$, $\mathbf{Z}_F$, $\mathbf{Z}_G$, $\mathbf{Z}_A$, and $\mathbf{Z}_B$ in the chord database.

13) Similarity calculation: Similarity calculation is a process for calculating the similarity values between a chord feature extraction of an input signal, and a number of chords reference feature extraction (C, D, E, F, G, A, and B) stored in a chord database. Thus, at the output of the similarity calculation process, there are a number of seven similarity values. This work used cosine similarity. This kind of similarity has been popularly used [21].

14) Chord decision: Chord decision is a process for determining the output text (C, D, E, F, G, A, or B), which indicates a chord that is recognized. The chord decision process was carried out as follows. The first one was finding the largest similarity value out of a total of seven similarity values, which are the outputs of the similarity calculation process. The second one was determining a recognized chord. A chord (C, D, E, F, G, A, or B) associated with the largest similarity value is determined as a recognized chord. As a note, an output process that is determined based on the largest similarity value indicating that it uses the template matching method [22] [23].

B. Developing Test Chords

The second step in the research methodology is the developing of the test chords, which were used for testing the chord recognition system. This work recorded 20 test chord samples for each chord (C, D, E, F, G, A, and B). Therefore, there were a number of 140 test chord samples.

C. Testing and Recognition Rate Calculation

The final step in the research methodology is testing and calculating the recognition rate. Testing was carried out using 140 test chord samples, for the segment averaging outputs 4, 8, 16, and 32 points, the number of feature extraction coefficients 1-8, and the methods of subsampling namely linear, quadratic, and cubic spline interpolation. The recognition rate calculation is the calculation of the ratio (expressed in percent), between the number of correctly recognized chords, and a number of 140 test chord samples.

III. Results and Discussions

A. Results

The chord recognition system shown in Fig. 1 has been tested for the chords of a guitar shown in Fig. 2. This test was carried out for the segment averaging outputs 4, 8, 16, and 32 points, the number of feature extraction coefficients 1-8, and the methods of subsampling namely linear, quadratic, and cubic spline interpolation. The results are shown in TABLE I. As a note, the number of feature extraction coefficients are correspond with the subsampling outputs.

TABLE I. TEST RESULTS OF THE DEVELOPED CHORD RECOGNITION SYSTEM.
RESULTS SHOWN: RECOGNITION RATE (%)

Subsampling method	Number of feature extraction coefficients (coefficients)							
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Segment averaging output = 4 points								
Linear spline interpolation	14.29	25.71	60.00	68.57	-	-	-	-
Quadratic spline interpolation	14.29	37.14	60.00	68.57	-	-	-	-
Cubic spline interpolation	14.29	37.14	61.43	68.57	-	-	-	-
Segment averaging output = 8 points								
Linear spline interpolation	14.29	45.71	61.43	62.86	88.57	**91.43**	**98.57**	100
Quadratic spline interpolation	14.29	47.14	65.71	68.57	87.14	84.29	90.00	100
Cubic spline interpolation	14.29	51.43	65.71	70.00	87.14	88.57	92.86	100
Segment averaging output = 16 points								
Linear spline interpolation	14.29	51.43	64.29	80.00	87.14	85.71	88.57	98.57
Quadratic spline interpolation	14.29	35.71	67.14	77.14	84.29	82.86	84.29	95.71
Cubic spline interpolation	14.29	25.71	67.14	77.14	85.71	84.29	87.14	97.14
Segment averaging output = 32 points								
Linear spline interpolation	14.29	50.00	60.00	82.86	82.86	91.43	94.29	97.14
Quadratic spline interpolation	14.29	42.86	44.29	71.43	78.57	90.00	92.86	95.71
Cubic spline interpolation	14.29	38.57	34.29	71.43	78.57	88.57	94.29	92.86

B. Discussions

From the point of view of the optimal results, TABLE I indicates that the use of seven and six feature extraction coefficients can give the highest recognition rates of up to 98.57% and 91.43% respectively. These results were achieved by using the segment averaging output eight points, and the subsampling method with linear spline interpolation.

From the point of view of the number of feature extraction coefficients, TABLE I indicates that in general, if the number of feature extraction coefficients is increasing, the recognition rate will increase. This is due, if the number of feature extraction coefficients is increasing, the more dimensions will be used to distinguish one chord pattern from other chord patterns. This one will increase the discrimination level of the feature extraction. Finally, the increase of the discrimination level will lead to an increase of the recognition rate.

From the point of view of segment averaging output, TABLE I indicates that the optimal number of the segment averaging output is eight points. If the number of segment averaging output is less than eight points or even more than 8 points, it will decrease the recognition rate. These cases can be explained as follows. Firstly, if the segment averaging output is less than eight points, it will cause the input of the subsampling process to be rougher. Consequently, this input will cause the output of the subsampling process (which is the result of feature extraction) also to be rougher. This case will decrease the discrimination level of the feature extraction. Secondly, if the segment averaging output is more than eight points, it will cause the input of the subsampling process to be more detail. However, since the subsampling output (which is the result of feature extraction) are less than eight coefficients, it will cause the comparison between the output and the input of the subsampling process will be getting away from the value of 1. This case will also decrease the discrimination level of the feature extraction. Finally, the decrease of the discrimination level will lead to a decrease of the recognition rate.

From the point of view of the subsampling method, TABLE I indicates that for each number of feature extraction coefficient and each number of segment averaging output, out of a total of 28 events, there are a total of 15 events where the use of linear spline interpolation gives the highest recognition rate. This means, in the majority, the use of linear spline interpolation is superior to quadratic and cubic spline interpolation. In addition, this also means that in the majority, the use of 2 evaluation points for each piece of spline interpolation is sufficient to get the highest discrimination level of the feature extraction. As a first note, each piece of linear, quadratic and cubic spline interpolation, requires 2, 3, and 4 evaluation points [20]. As a second note, in line with what has been described above, the highest discrimination level of the feature extraction will give the highest recognition rate.

C. Comparison with the other feature extractions

A comparison of some chord recognition rates for several feature extraction methods is shown in TABLE II. As a first note, TABLE II only compare the number of feature extraction coefficients that have been achieved by the author with previous researchers. As a second note, whatever the test chords used, the previous methods [8] [9] [10] always give a number of 12 feature extraction coefficients.

As seen in TABLE II, the proposed feature extraction method can be considered the most efficient based on the two categories of recognition rate. The first one, for the recognition rate category above 98%, the chord recognition system only requires a number of seven feature extraction coefficients. The second one, for the recognition rate category above 90%, the chord recognition system only requires a number of six feature extraction coefficients.

IV. CONCLUSIONS AND FUTURE WORKS

This work proposes a feature extraction subsystem, which can be used in a chord recognition system. The

TABLE II. THE PERFORMANCE COMPARISON OF SOME FEATURE EXTRACTIONS FOR CHORD RECOGNITION.

Feature Extraction	Number of Feature Extraction Coefficients	Recognition Rate (%)	Test Chords
Improved Chromagram [8]	12	94	1440 test chords from 360 recorded guitar chords
Improved PCP [10]	12	95.83	192 test chords from 192 generated guitar chords
CRP (Chroma DCT-Reduced log Pitch) Enhanced PCP [9]	12	99.96	4608 test chords from 576 generated guitar chords
Segment averaging with SHPS and logarithmic scaling [11]	8	100	140 test chords from 140 recorded guitar chords
Segment averaging and subsampling (this work)	**7**	**98.57**	140 test chords from 140 recorded guitar chords
	6	**91.43**	

feature extraction subsystem is FFT based segment averaging and subsampling.

Based on experiments, for guitar musical instrument, the optimal results can be viewed from the two categories of recognition rate. The first one, for the recognition rate category above 98%, the chord recognition system only requires a number of seven feature extraction coefficients. The second one, for the recognition rate category above 90%, the chord recognition system only requires a number of six feature extraction coefficients. These results were achieved by using the segment averaging output eight points, and the subsampling method with linear spline interpolation.

For further works, it can be explored the use of other subsampling methods besides spline interpolation. In addition, it can also be explored the other variants of the segment averaging feature extraction.

ACKNOWLEDGMENT

This work has been supported by The Institute of Research and Community Services of Sanata Dharma University, Yogyakarta.

REFERENCES

[1] J. Hartquist, Real Time Musical Analysis of Polyphonic Guitar Audio, Master Thesis, California Polytechnic State University, 2012, 125-6.

[2] T. Fujishima, "Realtime chord recognition of musical sound: a system using Common Lisp Music", Proceeding of the International Computer Music Conference (ICMC), Beijing, 1999, pp. 464–467.

[3] K. Muludi, Aristoteles, and A.F.S. Loupatty, "Chord Identification Using Pitch Class Profile Method with Fast Fourier Transform Feature Extraction", International Journal of Computer Science Issues, Vol. 11, Issue 3, No. 1, 2014, pp. 139-144.

[4] P. Gaonkar, S. Varma, and R. Nikhare, "A Survey on Content-Based Audio Retrieval Using Chord Progression", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, No. 1, 2016, pp. 629-636.

[5] M. Muller, F. Kurth, and M. Clausen, "Audio Matching via Chroma-based Statistical Features", Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR), 2005, pp. 288–295.

[6] A.P. Klapuri, A.J. Eronen, and J. Astola, "Analysis of The Meter of Acoustic Musical Signals", IEEE Transactions on Audio, Speech and Language Processing, Vol. 14, No. 1, 2006, pp. 342–355.

[7] K. Lee, "Automatic Chord Recognition from Audio using Enhanced Pitch Class Profiles", Proceedings of the International Computer Music Conference (ICMC), New Orleans, 2006, pp. 306–313.

[8] A.M. Stark, and M.D. Plumbey, "Real-Time Chord Recognition for Live Performance", Proceedings of The International Computer Music Conference, ICMC 09, 2009, pp. 85-88.

[9] P. Rajparkur, B. Girardeau, and T. Migimatsu, "A Supervised Approach To Musical Chord Recognition", Stanford Undergraduate Research Journal, Vol. 15, 2015, pp. 36-40.

[10] K. Ma, Automatic Chord Recognition, Department of Computer Sciences, University of Wisconsin-Madison, http://pages.cs.wisc.edu/~kma/projects.html, 2016.

[11] L. Sumarno, "Chord Recognition using Segment Averaging Feature Extraction with Simplified Harmonic Product Spectrum and Logarithmic Scaling", International Journal of Electrical Engineering Informatics, Vol. 10, No. 4, 2018, pp. 753-764.

[12] L. Tan, and J. Jiang, Digital Signal Processing Fundamentals and Applications 2nd Edition, Oxford: Elsevier Inc., 2013, pp. 15–56.

[13] E. Meseguer, Speech Analysis for Automatic Speech Recognition, MSc Thesis, Norwegian University of Science and Technology (NTNU), Trondheim, 2009, pp. 4-25.

[14] F.J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform", Proceedings of the IEEE, Vol. 66., No. 1, 1978, pp. 51-83.

[15] W.K. Jenkins, "Fourier Methods for Signal Analysis and Processing", The Digital Signal Processing Handbook 2nd Edition: Digital Signal Processing Fundamentals. Madisetti V. K. (ed), Boca Raton: CRC Press, 2010, pp. 1-1 - 1-29.

[16] A.M. Noll, "Pitch Determination of Human Speech by the Harmonic Product Spectrum, the Harmonic Sum Spectrum and a Maximum Likelihood Estimate", Proceedings of the Symposium on Computer Processing in Communications, Vol. 19, Polytechnic Press, Brooklyn, New York, 1970, pp. 779-797.

[17] Y.R. Setiawan, Numbers Speech Recognition using Fast Fourier Transform and Cosine Similarity (in Indonesian), Undergraduate Thesis, Sanata Dharma University, Yogyakarta, 2015, pp. 66-70.

[18] L. Sumarno, "On The Performance of Segment Averaging of Discrete Cosine Transform Coefficients on Musical Instruments Tone Recognition", ARPN Journal of Engineering and Applied Sciences, Vol. 11, No. 9, 2016, pp. 5644-5649.

[19] E. Kreyszig, Advanced Engineering Mathematics 10th Edition, New Jersey: John Wiley & Sons, Inc., 2011, pp. 820-825.

[20] S.C. Chapra, and R.P. Canale, Numerical Methods for Engineers 7th Edition, New York: McGraw-Hill Education, 2015. pp. 511-521.

[21] S. Zhu, J. Wu, H. Xiong, and G. Xia, "Scaling up top-K similarity search", Data and Knowledge Engineering, Vol. 70, 2011, pp. 60-83.

[22] A.K. Jain, R.P.W. Duin, and J. Mao, "Statistical Pattern Recognition: A Review", IEEE Transactions and Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, 2000, pp. 4-37.

[23] S. Theodoridis, and K. Koutroumbas, Pattern Recognition 4th Edition, California: Elsevier Inc., 2009, pp. 481-519.

Chord Recognition using FFT Based Segment Averaging and Subsampling Feature Extraction

ORIGINALITY REPORT

20%	19%	11%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	core.ac.uk Internet Source	5%
2	www.arpnjournals.org Internet Source	3%
3	en.wikipedia.org Internet Source	1%
4	Anssi Klapuri. "Automatic Bass Line Transcription from Streaming Polyphonic Audio", 2007 IEEE International Conference on Acoustics Speech and Signal Processing - ICASSP 07, 04/2007 Publication	1%
5	senatik.stta.ac.id Internet Source	1%
6	www.sersc.org Internet Source	1%
7	commons.erau.edu Internet Source	1%

8	www.scribd.com Internet Source	1%
9	Suwatchai Kamonsantiroj, Lita Wannatrong, Luepol Pipanmaekaporn. "Improving Pitch Class Profile for Musical Chords Recognition Combining Major Chord Filters and Convolution Neural Networks", 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), 2017 Publication	1%
10	repozitorium.omikk.bme.hu Internet Source	1%
11	addi.ehu.es Internet Source	1%
12	www.springerprofessional.de Internet Source	1%
13	risc01.sabanciuniv.edu Internet Source	1%
14	sinta3.ristekdikti.go.id Internet Source	<1%
15	d-nb.info Internet Source	<1%
16	"Sound, Music, and Motion", Springer Science and Business Media LLC, 2014 Publication	<1%

17	ijeei.org Internet Source	<1%
18	epubs.surrey.ac.uk Internet Source	<1%
19	Masaki Otsuka, Tetsuro Kitahara. "An on-line algorithm of guitar performance transcription using non-negative matrix factorization", 2014 12th International Conference on Signal Processing (ICSP), 2014 Publication	<1%
20	www.ijcaonline.org Internet Source	<1%
21	"Intelligent Computing and Communication", Springer Science and Business Media LLC, 2020 Publication	<1%
22	Xin Tu, Guofu Zhu, Jun Hu, Xiaotao Huang, Zhimin Zhou. "A sidelobes/grating lobes suppression method for ultrawindeband ultrasparse array through-the-wall imaging radar", 2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP), 2014 Publication	<1%
23	www.osti.gov Internet Source	<1%

24	publications.aston.ac.uk Internet Source	<1%
25	www.earthdoc.org Internet Source	<1%
26	www.freepatentsonline.com Internet Source	<1%
27	vufind.katalog.k.utb.cz Internet Source	<1%
28	www.yumpu.com Internet Source	<1%
29	www.easychordz.com Internet Source	<1%
30	S.V. Beiden, M.A. Maloof, R.F. Wagner. "A general model for finite-sample effects in training and testing of competing classifiers", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003 Publication	<1%
31	Studies in Classification Data Analysis and Knowledge Organization, 2014. Publication	<1%
32	manualzz.com Internet Source	<1%
33	www.db-thueringen.de Internet Source	<1%

34	Yi Yu, Roger Zimmermann, Ye Wang, Vincent Oria. "Scalable Content-Based Music Retrieval Using Chord Progression Histogram and Tree-Structure LSH", IEEE Transactions on Multimedia, 2013 Publication	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		