

DCT based feature extraction and support vector machine classification for musical instruments tone recognition

Linggo Sumarno¹, Rifai Chai²

¹Electrical Engineering Study Program, Sanata Dharma University, Indonesia

²Department of Telecommunications, Electrical, Robotics and Biomedical Engineering, Swinburne University of Technology, Australia

Article Info

Article history:

Received Aug 9, 2020

Revised Mar 2, 2021

Accepted Aug 3, 2021

Keywords:

DCT based feature extraction

Feature extraction

Musical instruments

SVM classification

Tone recognition

ABSTRACT

The conducted research proposes a feature extraction and classification combination method that is used in a tone recognition system for musical instruments. It is expected that by implementing this combination, the tone recognition system will require fewer feature extraction coefficients than those previously investigated. The proposed combination comprises of feature extraction using discrete cosine transform (DCT) and classification using support vector machine (SVM). Bellyra, clarinet, and pianica tones were used in the experiment, with each indicating a tone with one, several, or many major local peaks in the transform domain. Based on the results of the tests, the proposed combination is efficient enough to be used in a tone recognition system for musical instruments. This is indicated in recognizing a tone, it only needs at least eight feature extraction coefficients.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Linggo Sumarno

Electrical Engineering Study Program

Sanata Dharma University

Kampus III Paingan, Maguwoharjo, Depok, Sleman, Yogyakarta, 55282, Indonesia

Email: lingsum@usd.ac.id

1. INTRODUCTION

In a tone recognition system of musical instruments, there are two main parts. The first part takes the characteristics of the tone, or what is called the feature extraction part. The second part classifies the results of the feature extraction part, or what is called the classification part. A transform domain approach could be used to implement the feature extraction part. Discrete fourier transform (DCT) and discrete fourier transform (DFT) are two transformation methods for converting tone signals from the time domain to the transform domain. There are two different ways to feature extraction in the transform domain. The first is feature extraction, which is based on fundamental signals [1]-[6]. The second is feature extraction, which is not based on fundamental signals [7]-[12].

An approach to implementing the classification part in the tone recognition system above is to use a statistical approach. Support vector machine (SVM) is one example of classification that makes use of a statistical approach. SVM is a classification method that originates from statistical learning theory [13], [14]. Initially, SVM was only used for the classification of two classes. In subsequent developments, SVM can be used for multiclass classifications [15]-[21]. The previous tone recognition system research has primarily focused on tones with multiple major (significant) local peaks in the transform domain [9]-[11]. Tone recognition for tones having one, several or many major local peaks in the transform domain has very little to do with it. Previous research [12] proposed combining DFT-based segment averaging for feature extraction and template matching for classification to be used in a tone recognition system. That system could recognize

tones with one, several or many major local peaks in the transform domain. However, to identify a tone, the tone recognition system needs at least 16 feature extraction coefficients. Thus, for the tone recognition system, there is still a chance of obtaining feature extraction coefficients fewer than 16. The advantage of using fewer feature extraction coefficients is that we have fewer data to process.

The conducted research combines feature extraction and classification methods for musical instruments tone recognition. To be more specific, it presents a DCT-based feature extraction and SVM classification for musical instruments tone recognition. As a first note, the feature extraction method does not use fundamental signals in the transform domain. As a second note, the tones used in the conducted research were bellyra, clarinet, and pianica tones, representing a tone with one, several, and many major local peaks in the DCT transformation domain. The DCT transformation domain of the bellyra, clarinet and pianica tones are shown in Figure 1.

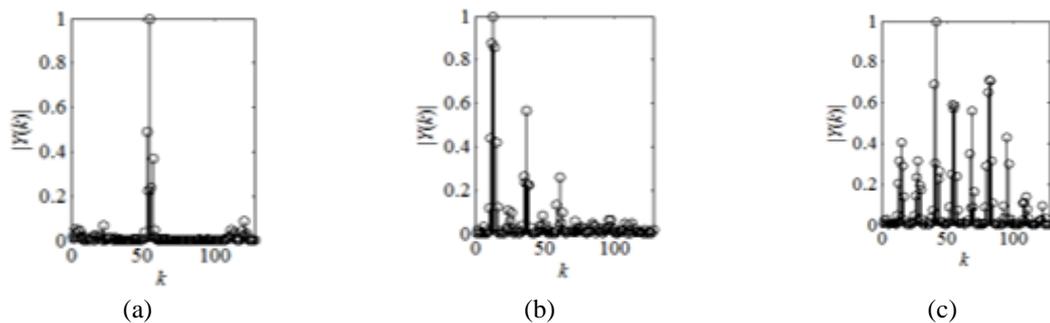


Figure 1. Tone C in the normalized DCT transformation domain representation $|Y(k)|$ with sampling rate 5000 Hz and DCT 128 points, (a) bellyra, (b) clarinet, (c) pianica

2. RESEARCH METHOD

2.1. Materials preparation

A tone signal is used as the tone recognition system's input. The tone signal is an isolated signal that is stored in the waveform audio file format (WAV). This tone signal is acquired from three musical instruments played, namely bellyra, clarinet, and pianica. In this research, we recorded a total of eight tone signals for each musical instrument, namely C, D, E, F, G, A, B, and C'. We recorded the tone signals at a sampling rate of 5000 Hz. Essentially, this sampling rate has met the theorem of Shannon sampling [22] as:

$$f_s \geq 2f_{max} \quad (1)$$

with f_{max} being the highest frequency component of the tone signals and f_s being the sampling rate. The highest frequency components of the tone C' for bellyra, clarinet, and pianica, according to our visual observations using Octave software, were 2097 Hz, 1406 Hz, and 1584 Hz, respectively. Based on our visual observations also, recording a tone for 2 seconds was adequate to acquire a steady-state part of the tone signal. It should be noted that the recorded tone signal can be divided into three parts, namely silence, transition, and steady-state parts. Only in the steady-state part there is accurate tone information.

Three musical instruments were used in this research to acquire the above-mentioned tone signals. They were an Isuzu ZBL-27 bellyra, a Yamaha YCL-255 flute, and a Yamaha P-37D pianica, as shown in Figure 2. The tone signals were captured by using an AKG perception 120 USB microphone.

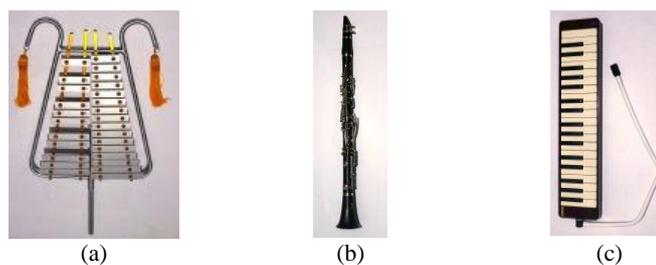


Figure 2. Three musical instruments, (a) bellyra, (b) clarinet and (c) pianica

2.2. System design

The entire tone recognition system in this research is shown in Figure 3 as a block diagram. The system's input is a WAV-formatted tone signal. The system's output is a text, which denotes a recognized tone. We used Octave software to develop the system design. The explanation of all blocks in Figure 3.

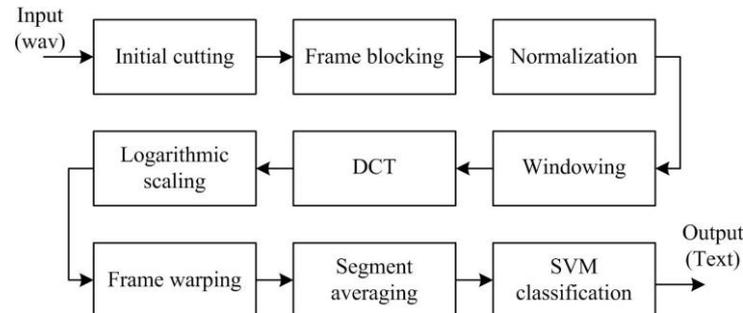


Figure 3. The entire system is shown as a block diagram

2.2.1. Initial cutting

Deleting the silence and transition parts in a tone signal is known as initial cutting. The silence and transition parts need to be deleted because there is no accurate tone information in these parts. The accurate tone information can only be obtained in the steady-state part of the tone signal. Based on our visual observations, the silence part could initially be deleted by making use an amplitude threshold value of $|0.5|$ from the tone signal's highest value. Beginning with the leftmost part of the tone signal, the signal was deleted, if the signal's amplitude was less than $|0.5|$ from the highest value of the tone signal. Following the deletion of the silence part, the transition part was deleted. Based on our observations also, the transition part was deleted for 300 milliseconds from the tone signal's left side. Finally, the steady-state part could be obtained, after the silence and transition parts were deleted.

2.2.2. Frame blocking

Cutting a signal frame from a lengthy signal is known as frame blocking [23]. The objective of this frame blocking is to decrease the amount of signal data. The tone recognition system's computation time will be sped up if the amount of signal data is decreased. The length of frame blocking of 2^n , where n is a positive integer, was evaluated in this research to find the shortest length of frame blocking that gave the highest recognition rate.

2.2.3. Normalization

Setting the highest value of a data signal to 1 or -1 is known as normalization. The objective of this normalization is to decrease the disparity between a data signal's highest value and the others. Normalization is implemented using (2).

$$\mathbf{y}_{out} = \mathbf{y}_{in} / \max(|\mathbf{y}_{in}|) \quad (2)$$

where \mathbf{y}_{in} and \mathbf{y}_{out} are input and output data signal vectors, respectively.

2.2.4. Windowing

Windowing is smoothing discontinuities in a data signal's edges [23]. This discontinuity happens as a result of the data signal being cut in the preceding frame blocking. Discontinuities will give rise extra signals known as harmonic signals visible in the transformed data signal. The visibility of harmonic signals can be reduced by smoothing discontinuities. The Hamming window, which is extensively utilized in signal processing [24], was used in this research. Aside from that, the length of window 2^n was used in this research, where n is a positive integer. This length of window is the same as the above-mentioned the length of frame blocking.

2.2.5. DCT

DCT is converting data signals from the time domain to the transform domain known as the DCT transformation domain. The length of DCT of 2^n was used in this research, where n is a positive integer. The

length of DCT is the same as the length of frame blocking and the length of Hamming window mentioned above. In addition, this research applied the calculation of absolute values of DCT results because the next process, logarithmic scaling, will not allow the calculation of negative values.

2.2.6. Logarithmic scaling

The gap in peak levels in a data signal can be reduced by logarithmic scaling. The logarithmic scaling results indicate an increase in the number of major local peaks. Previous research [11], [12] shows that feature extraction using segment averaging (which is used in this research) produces superior results for data signals with many major local peaks. The following is a mathematical expression of logarithmic scaling.

$$Y_{out}(k) = \log(Y_{in}(k) + 1) \quad (3)$$

where $Y_{out}(k) = \{Y_{out}(0), Y_{out}(1), \dots, Y_{out}(N-1)\}$ and $Y_{in}(k) = \{Y_{in}(0), Y_{in}(1), \dots, Y_{in}(N-1)\}$ are the data signal's output and input, respectively, and N is the length of data signal. The inclusion of value '1' in (3) prevents a logarithmic outcome near to negative infinity if the input data signal has a value close to zero.

2.2.7. Frame warping

Frame warping is reducing the length of a data frame. The results of frame warping show a more dense distribution of data. Basically, this frame warping is carried out by dividing the data frame into two and then combining the two. The algorithm of this frame warping is presented as:

Frame warping algorithm:

1. Consider and input data frame $Y_{in}(k) = \{Y_{in}(0), Y_{in}(1), \dots, Y_{in}(N-1)\}$, where N is an even number that represents the length of the input data frame.
2. Divide $Y_{in}(k)$ into two segments $V_1(k) = \{Y_{in}(0), Y_{in}(1), \dots, Y_{in}((N/2)-1)\}$ and $V_2(k) = \{Y_{in}(N/2), Y_{in}((N/2)+1), \dots, Y_{in}(N-1)\}$.
3. Flip $V_2(k)$ in order to be $V_{2f}(k) = \{Y_{in}(N-1), \dots, Y_{in}((N/2)+1), Y_{in}(N/2)\}$
4. Merge the results of $V_1(k)$ and $V_{2f}(k)$, in order to be an output data frame $Y_{out}(k) = \{Y_{out}(0), Y_{out}(1), \dots, Y_{out}((N/2)-1)\}$, as:

$$Y_{out}(k) = [V_1(k) + V_{2f}(k)]/2 \quad (4)$$

As a note, this frame warping will give an output data frame that is half the length of the input data frame.

2.2.8. Segment averaging

One way to reduce the amount of data in a data frame is to use segment averaging. This research used segment averaging from the previous research [11], [12]. Basically, this segment averaging is carried out by dividing a data frame into a number of data segments and then carry out an averaging operation in each data segment. The algorithm of this segment averaging is presented as:

Segment averaging algorithm:

1. Suppose there is an input data frame $Y_{in}(k) = \{Y_{in}(0), Y_{in}(1), \dots, Y_{in}(M-1)\}$, with the length of frame $M = 2^q$ and $q \geq 0$.
2. Determine the length of a segment L , where $L = 2^p$ for $0 \leq p \leq q$.
3. Cut the input data frame $Y_{in}(k)$ into a number of segments, where the length of each segment is L . The result of dividing the input data frame $Y_{in}(k)$ will give a number of Z of segments as:

$$Z = M/L \quad (5)$$

and also a data sequence $S_v(u) = \{S_v(0), S_v(1), \dots, S_v(L-1)\}$ in each segment.

4. Calculate the output data frame $Y_{out}(v) = \{Y_{out}(0), Y_{out}(1), \dots, Y_{out}(Z-1)\}$ by performing an average operation in each segment as:

$$Y_{out}(v) = \frac{1}{L} \sum_{u=0}^{L-1} S_v(u) \quad (6)$$

The length of segment L was evaluated in this research at 1, 2, 4, ..., $2^{\log_2(M/2)}$ values, where M is the length of the frame resulting from the frame warping. As a note, the (5) can be stated as follows, if N is the length of the input data frame in the frame warping.

$$Z = N/(2L) \quad (7)$$

2.2.9. SVM classification

Classification is determining the pattern class of a data frame. SVM is a method that can be used for this classification. SVM is a linear classification. In training an SVM, the best hyperplane will be explored. This hyperplane separates two data sets, which come from two different pattern classes. Mathematically, this hyperplane is a linear discriminant function.

It is not necessarily that a hyperplane can separate the two data sets from two different pattern classes in the real world. In other words, the two data sets are not linearly separable. Therefore, data transformation needs to be carried out. To carry out this transformation, a function called kernel function [25] can be used. Linear and polynomial kernel functions are two examples of commonly used kernel functions. This research evaluated that two kernel functions.

Initially, SVM was developed for the case of two-class pattern classification. Furthermore, SVM was developed for the case of multiclass pattern classification. This research is a case of multiclass pattern classification. For this multiclass case, one-vs-all (OVA) Tree Multiclass method [20] is used. The selection of this OVA is based on the performance of this OVA, which is comparable to the other methods, particularly [16] and [21]. In this research, we used LibSVM and its default settings [26].

2.3. Feature extraction and SVM training

The SVM shown in Figure 3 needs a number of data in the training process. This data is obtained using the feature extraction proposed in this research, shown in Figure 4. The input is a WAV-formatted tone signal. The output is the feature extraction from the input tone signal. As a note, every block in the proposed feature extraction is the standard one. However, if we look at the complete picture (the series of blocks in Figure 4), we can see that the proposed feature extraction is unique. Thus, we can see a novelty in this research. For each musical instrument (bellyra, clarinet, or pianica) in this research, 10 samples were recorded for each tone signal (C, D, E, F, G, A, B, or C'). So there were a total of 240 tone signals. In addition, the feature extraction of each of the tone signals was processed by the feature extraction shown in Figure 4. SVM training used the results of this feature extraction.

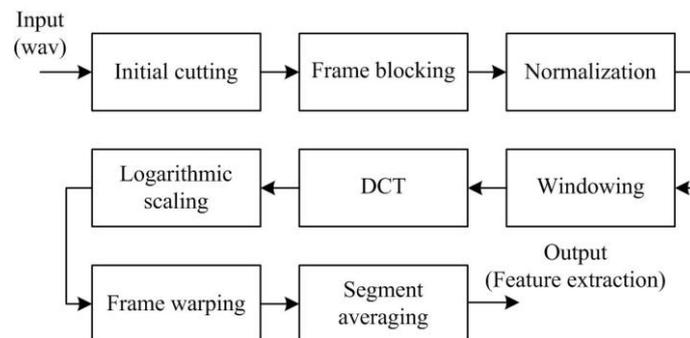


Figure 4. The proposed feature extraction is shown as a block diagram

2.4. Test tones and recognition rate

The number of tone signals used to test the performance of the tone recognition system are the test tones. For each tone signal (C, D, E, F, G, A, B or C') of each musical instrument (bellyra, clarinet, or pianica), a number of 20 samples were recorded for the test tones. So, for each musical instrument there was a total of 160 tone signals. The recognition rate is the magnitude of tone system recognition performance. The following is how the recognition rate is calculated.

$$\text{Recognition Rate} = \frac{\text{Number of correctly recognized tones}}{\text{Test tones}} \times 100 \% \quad (8)$$

3. RESULTS AND ANALYSIS

3.1. Test results

Test results of the developed tone recognition system for using a linear kernel function in SVM classification, for different combinations of the lengths of frame blocking and the number of feature extraction coefficients, are presented in Table 1.

Table 1. Test results for the use of a linear kernel function in SVM classification, for different combinations of the lengths of frame blocking and the number of feature extraction coefficients. results shown: recognition rate (%)

The length of frame blocking (points)	Number of feature extraction coefficients (coefficients)								
	1	2	4	8	16	32	64	128	256
(a) Musical instrument: Bellyra									
64	16.88	27.50	78.75	98.75	100	100	-	-	-
128	13.75	36.88	80.00	100	100	100	100	-	-
256	12.50	37.50	81.88	100	100	100	100	100	-
512	15.63	38.13	76.88	100	100	100	100	100	100
(b) Musical instrument: Clarinet									
64	19.38	27.50	69.38	93.75	100	100	-	-	-
128	22.50	25.63	78.75	97.50	100	100	100	-	-
256	16.88	28.75	65.63	100	100	100	100	100	-
512	12.50	22.40	71.88	98.75	100	100	100	100	100
(c) Musical instrument: Pianica									
64	30.63	51.25	93.75	100	100	100	-	-	-
128	24.38	41.25	91.88	98.75	100	100	100	-	-
256	20.00	35.63	81.25	100	100	100	100	100	-
512	26.25	36.88	72.50	100	100	100	100	100	100

As indicated in Table 1, the recognition rate increases if the number of feature extraction coefficients increases. If the number of feature extraction coefficients increases, it will further increase the dimension of the feature extraction space. The increased dimension of the feature extraction space makes it easier to differentiate between one pattern class and the other pattern classes. The easier differentiation of one pattern class with the other pattern classes will ultimately increase the recognition rate.

3.2. The smallest number of feature extraction coefficients

The goal of this research is to discover the smallest number of feature extraction coefficients that can be used in a tone recognition system. Here, the tone recognition system can recognize the tone with one, several, or many major local peaks in the transform domain. As indicated in Table 1, the use of the smallest number of feature extraction coefficients, i.e. eight coefficients, and the shortest length of frame blocking, i.e. 256 points, can result in the highest recognition rate of up to 100%. In other words, by using at least eight feature extraction coefficients, the tone recognition system can recognize all the tested tones. The tested tones in this case are those with one, many, or many major local peaks in the transform domain.

Furthermore, the number of the smallest (eight) feature extraction coefficients above should be noted. It is linked to the usage of a linear kernel function in SVM classification. For the use of other kernel functions (i.e. polynomial functions) in SVM classification, Table 1 has been reworked for the second and third-order of polynomial functions. The results obtained from the use of both polynomial functions are presented in Table 2.

Table 2. The results of the use of linear kernel and polynomial functions on SVM classification

Kernel function on SVM classification	Number of feature extraction coefficient (coefficients)	Highest recognition rate (%)
Linear	8	100
Second order polynomial	8	98.75
Third order polynomial	16	98.75

The use of linear kernel functions gives the best results, as indicated in Table 2. Only by using eight feature extraction coefficients (the smallest number of coefficients), it can give a recognition rate of up to 100%. This indicates that by using eight feature extraction coefficients, the pattern classes of feature extraction of the tone signals are linearly separable.

3.3. Comparison of some feature extraction and classification combinations

The performance of some feature extraction and classification combinations for musical instrument tone recognition are compared in Table 3. As indicated in Table 3, the feature extraction and classification combination proposed in this research is the most efficient for use in a musical instruments tone recognition system. This is because the tone recognition system needs only eight feature extraction coefficients (the smallest number of coefficients) to recognize the tones with several major local peaks in the transform domain.

Table 3. The best performance comparison of some feature extraction and classification combinations, for the musical instruments tone recognition. Results shown: number of smallest feature extraction coefficients (musical instrument)

Feature extraction/classification	Number of major local peaks in the transform domain			Highest recognition rate (%)
	One	Several	Many	
MFCC/K-NN [9]	-	22 (flute)	22 (cello, piano, trumpet, violin)	91.66
Spectral Features/SVM [10]	-	21 (gamelan)	-	98.93
DCT based segment averaging/template matching [11]	-	16 (soprano recorder)	8 (pianica)	100
DFT based segment averaging/template matching [12]	16 (bellyra)	16 (tenor recorder)	4 (pianica)	100
DCT based segment averaging/SVM (this research)	8 (bellyra)	8 (clarinet)	8 (pianica)	100

4. CONCLUSION

The conducted research proposes a feature extraction and classification combination in a tone recognition system for musical instruments. The purpose of using this combination is to obtain a tone recognition system, which in the recognition process uses the smallest number of feature extraction coefficients. To do this, we combined a DCT based feature extraction and an SVM classification to be used in the tone recognition system. We have discovered from the test results that the proposed feature extraction and classification combination makes the tone recognition system efficient enough. This was because the tone recognition system only needed at least eight feature extraction coefficients to recognize tones with one, several, or many major local peaks in the transform domain. We have also discovered that SVM classification requires only a linear kernel function. This one indicates that the pattern classes from the DCT based segment averaging feature extraction are linearly separable. For further development of this research, we recommend exploring other feature extraction and classification combinations. In this case, these combinations can use different methods for feature extraction (other than DCT based segment averaging) and different methods for classification (other than SVM).

ACKNOWLEDGEMENTS

This research has been supported by the Institute of Research and Community Services of Sanata Dharma University, Yogyakarta, Indonesia.

REFERENCES

- [1] A. Mitre, M. Queiroz, R. R. A. Faria, "Accurate and Efficient Fundamental Frequency Determination from Precise Partial Estimates". *Proceedings of the 4th Audio Engineering Society (AES) Brazil Conference*, pp. 113-118, 2006.
- [2] A. Pertusa and J. M. Inesta, "Multiple fundamental frequency estimation using Gaussian smoothness," 2008 *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 105-108, doi: 10.1109/ICASSP.2008.4517557.
- [3] C. Yeh, A. Roebel and X. Rodet, "Multiple Fundamental Frequency Estimation and Polyphony Inference of Polyphonic Music Signals," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 6, pp. 1116-1126, Aug. 2010, doi: 10.1109/TASL.2009.2030006.
- [4] Z. Duan, B. Pardo and C. Zhang, "Multiple Fundamental Frequency Estimation by Modeling Spectral Peaks and Non-Peak Regions," in *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 8, pp. 2121-2133, Nov. 2010, doi: 10.1109/TASL.2010.2042119.
- [5] S. Gonzalez and M. Brookes, "PEFAC-A Pitch Estimation Algorithm Robust to High Levels of Noise," in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 2, pp. 518-530, Feb. 2014, doi: 10.1109/TASLP.2013.2295918.
- [6] J. K. Nielsen, T. L. Jensen, J. R. Jensen, M. G. Christensen and S. H. Jensen, "Fast and statistically efficient fundamental frequency estimation," 2016 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 86-90, doi: 10.1109/ICASSP.2016.7471642.
- [7] A. Wintarti, D. Juniati and I. N. Wulandari, "Classification of Gamelan Tones Based on Fractal Analysis," *IOP Conference Series: Materials Science and Engineering*, vol. 288, pp. 1-4, 2018, doi: 10.1088/1757-899X/288/1/012022.
- [8] A. I. Tamboli and R. D. Kokate, "An Effective Optimization-Based Neural Network for Musical Note Recognition," *Journal of Intelligent Systems*, vol. 28, issue 1, pp. 173-183, 2019, doi: 10.1515/jisys-2017-0038.
- [9] N. S. Nagawade and V. R. Ratnaparkhe, "Music Instrument Classification using Nontonal MFCC," *Proceeding of the 2nd IEEE International Conference On Recent Trends in Electronics Information & Communication Technology (RTEICT 2017)*, 2017, pp. 2198-2202, doi: 10.1109/RTEICT.2017.8256990

- [10] A. Tjahyanto, Y. K. Suprpto and D. P. Wulandari, "Spectral-based Features Ranking for Gamelan Instruments Identification using Filter Techniques," *TELKOMNIKA Telecommunication, Computing, Electronics and Control*, vol. 11, no. 1, pp. 95-106, 2013, doi: 10.12928/telkomnika.v11i1.895.
- [11] L. Sumarno, "On The Performace of Segment Averaging of Discrete Cosine Transform Coefficients on Musical Instruments Tone Recognition," *ARPN Journal of Engineering and Applied Sciences*, vol. 11, no. 9, pp. 5644-5649, 2016.
- [12] L. Sumarno and Iswanjono, "Feature Extraction of Musical Instrument Tones using FFT and Segment Averaging," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 15, no. 3, pp. 1280-1289, 2017, doi: 10.12928/telkomnika.v15i3.3381
- [13] V. N. Vapnik, "*The Nature of Statistical Learning Theory*," New York: Springer-Verlag, 1995.
- [14] V. N. Vapnik, "*Statistical Learning Theory*," Wiley, 1998.
- [15] S. Abe, "Analysis of multiclass support vector machines," *Proceedings of International Conference on Computational Intelligence for Modelling Control and Automation (CIMCA 2003)*, pp. 385-396, 2003.
- [16] R. M. Rifkin and A. B. R. Klautau, "In Defense of One-vs-All Classification," *The Journal of Machine Learning Research*, vol. 5, pp. 101-141, 2004.
- [17] Yi Liu and Y. F. Zheng, "One-against-all multi-class SVM classification using reliability measures," *Proceedings. 2005 IEEE International Joint Conference on Neural Networks*, 2005., 2005, pp. 849-854 vol. 2, doi: 10.1109/IJCNN.2005.1555963.
- [18] Y. Liu, R. Wang and Y. Zeng, "An Improvement of One-Against-One Method for Multi-Class Support Vector Machine," *2007 International Conference on Machine Learning and Cybernetics*, 2007, pp. 2915-2920, doi: 10.1109/ICMLC.2007.4370646.
- [19] J. Manikandan and B. Venkataramani, "Design of a modified one-against-all SVM classifier," *2009 IEEE International Conference on Systems, Man and Cybernetics*, 2009, pp. 1869-1874, doi: 10.1109/ICSMC.2009.5346200.
- [20] B. Sidaoui and K. Sadouni, "OVA Tree Multiclass Framework for Support Vector Machine," *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS 2013)*, vol. I, pp 36-41, 2013.
- [21] B. Sidaoui and K. Sadouni "Approach Multiclass SVM Utilizing Genetic Algorithms," *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS 2013)*. vol I, pp. 285-290, 2013.
- [22] L. Tan and J. Jiang, "Digital Signal Processing Fundamentals and Applications," 3rd Edition. Elsevier Inc., pp 13-20. 2019.
- [23] N. A. Meseguer, "Speech Analysis for Automatic Speech Recognition," M.S. Thesis, Department of Electronics and Telecommunications, Norwegian University of Science and Technology, Trondheim, 2009, [Online]. Available: <https://core.ac.uk/download/pdf/52128303.pdf>
- [24] W. K. Jenkins, "Fourier Methods for Signal Analysis and Processing," Madisetti VK. Editor. *The Digital Signal Processing Handbook*, Florida: CRC Press, 2010, Ch. 2nd, pp. 1-1-29, 2010.
- [25] B. E. Boser, I. M. Guyon and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," *Proceedings of the Fifth Annual Workshop on Computational Learning Theory (COLT '92)*, pp. 144-152, 1992, doi: 10.1145/130385.130401.
- [26] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp.1-27, 2011, doi: 10.1145/1961189.1961199.

BIOGRAPHIES OF AUTHORS



Linggo Sumarno received doctorate degree from Electrical Engineering Study Program, Gadjah Mada University, Indonesia, in 2010. He is currently a lecturer in Electrical Engineering Study Program, Sanata Dharma University, Indonesia. His research interests include signal processing, image processing, and pattern recognition.



Rifai Chai received the B. Eng. degree from Krida Wacana Christian University, Jakarta, Indonesia in 2000 and a PhD degree in Engineering from the University of Technology, Sydney (UTS) in 2014. From 2000 to 2010 he has worked as a Product Development Engineer; Research and Development Engineer; and Project Engineer with companies in Indonesia and Australia. Currently, he is working as Senior Lecturer with Faculty of Science, Engineering & Technology, Swinburne University of Technology, Melbourne, Australia. His research interests on brain-computer interfaces, biomedical signal processing, biomedical instrumentation and artificial intelligence and pattern recognition. He serves as Associate Editor for Electronics Letters.