

Optimization of K Parameters on KNN in Gamelan Jegog Title Classification Using Time Domain Features

Adis Luh Sankhya Artayani^{a1}, Luh Arida Ayu Rahning Putri^{a2}

^{a1}Informatics Department, Faculty of Math and Science, Udayana University
Bali, Indonesia

¹adisluhsankhya@gmail.com

²rahningputri@unud.ac.id

Abstract

Bali is one of the provinces in Indonesia which has a lot of culture and arts, one of which is the Gamelan Jegog Bali. The technology nowadays can make it easier for humans to search for the title of a song that was previously unknown. This technology can be applied to the unknown title of Gamelan Jegog. The features used in this system are Short Time Energy and Zero Crossing Rate. The feature is extracted from Gamelan Jegog and then used to find the best k parameter from the K-Nearest Neighbor classifier. The results showed that the highest accuracy was 45% when the k parameter is 9. The amount of data used and the classification method used has an effect on the accuracy of this system when compared to similar studies.

Keywords: *Gamelan Jegog, KNN, Parameter K, Short Time Energy, Zero Crossing Rate*

1. Introduction

Bali is one of the provinces in Indonesia which has a lot of culture and arts, such as kidung, gamelan, dance, traditional clothing, and others. Culture and art have been an attachment to social life in Bali for a long time. For example, one of them is the Gamelan Bali. There are many types of Gamelan Bali depending on the musical instrument used, one of which is the Gamelan Jegog Bali.

The technology nowadays can make it easier for humans to search for the title of a song that was previously unknown. Therefore, in this study a system will be made that can find out the title of a Jegog Gamelan using the K-Nearest Neighbor classification. Gamelan Jegog, whose title is not recognized, can be identified because of this system.

In this study, the features used are Short Time Energy and Zero Crossing Rate. This feature is extracted from Gamelan Jegog and then used to find the best k parameter from the K-Nearest Neighbor classifier.

A previous study about music file grouping based on similar characteristics using low-level features, where the time domain features used are short-time energy, and the zero-crossing rate produces the best accuracy of 80.4762% but the classifier used is naïve bayes [1]. Then in another study about classification using K-Nearest Neighbor on musical instruments (without speech) with features used is MFCC, resulted in the highest accuracy of 91.66% [2]. Based on these two studies, this paper will create a system with the K-Nearest Neighbor classifier using the short time energy and zero crossing rate features applied to the gamelan jegog title.

The optimal use of the k parameter will greatly affect the final accuracy results. If the k parameter used is too small, the final accuracy result will be affected by noise, and if the k parameter used is too large, then the final accuracy result will be more blurred because of the boundaries between classes used [3]. For this reason, it is necessary to look for the optimal k parameter for the classification of the Gamelan Jegog Title.

The following structure of this paper is organized as: In Part 2, a brief description of the flow of the system and the methods used are provided. In Part 3, the system results, testing, and

discussion of the best k parameters used for the classification of Gamelan Jegog titles using time-domain features will be explained. In Section 4, contains statements that answer the problems in the previous section and further research work.

2. Research Methods

2.1. Jegog

Jegog is a musical instrument from Jembrana which all instruments are made of bamboo. Jegog was created by an artist named I Wayan Geliguh or Kiyang Geliduh in 1912, who came from Banjar Sebuah, Dangin Tukad Aya Village, Negara District, Jembrana [4].

A set of Jegog instruments has instruments consisting of barangan, kancil/kantil, suir, celuluk/kentung, undir, and jegog. Each Jegog set has eight blades, which are played with two sticks.

2.2. Method

In this paper, the system uses an input in the form of an unknown title audio signal from the gamelan jegog. The audio signal will go through framing and windowing first before feature extract. In the feature extraction process, the audio signal is extracted using several time-domain feature methods, namely short-time energy and zero-crossing rate to produce short-time energy and zero-crossing rate features. These features will then become input for the classification process using the KNN classifier. The classification process is used for making decision to determine the title of the gamelan jegog based on the highest class frequency in k neighbor. The classification process will produce an output in the form of a class which is the title of the gamelan jegog.

a. Framing and Windowing

Framing is used to divide the audio signal into small pieces which simplify the process of sound analysis and calculation. The audio frame used varies between 10 ms - 50 ms, because in that range, the frame is said to contain the characteristics of a stationary signal [5]. Framing causes the audio signal to be discontinuous so that windowing is necessary.

Windowing is the process of converting a signal to a continuous one. Windowing reduces the signal to zero at the start and end of a frame. The type of window used is the Hamming Window. Hamming Window provides greater and better attenuation for outside the bandpass [6]. Hamming Window is defined in equation (1).

$$w[n] = 0.52 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), 0 \leq n \leq N-1 \quad (1)$$

The windowing result is obtained from the n-th window frame ($w[n]$) multiplied by the n-th frame. N is the total frame.

b. Short Time Energy

Short Time Energy is energy from short segments, which is effective as a parameter for both voiceless and silent segments [7]. Short Time Energy is defined in equation (2).

$$E_n = \sum_{m=n-N+1}^n [x(m)w(n-m)]^2 \quad (2)$$

where, n is the window frame variable and N is the window length.

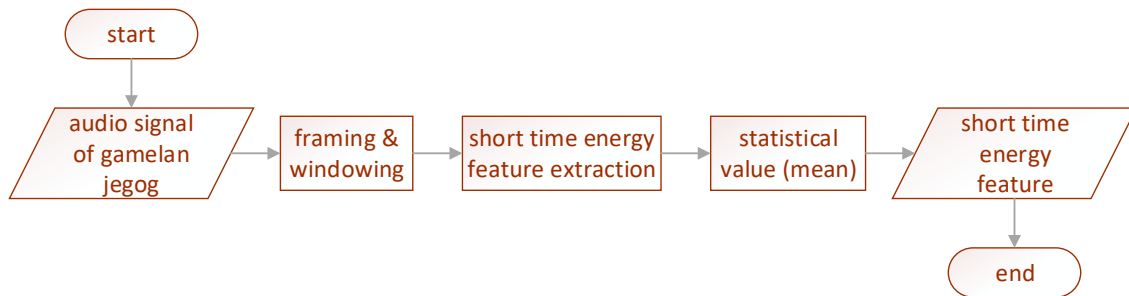


Figure 1. Short Time Energy Feature Extraction

Figure 1 is an illustration of the extraction process of short-time energy. The input is an audio signal from the gamelan jegog in .wav format. The audio signal will go through framing and windowing first. The resulting short-time energy feature will take its statistical value on average.

c. Zero Crossing Rate

Zero Crossing Rate (ZCR) is a measure of times the signal passing through the horizontal axis, which can change from positive to zero then negative or vice versa [8]. ZCR on the audio frame is indicated by the change in the signal signature of one frame. ZCR is defined in equation (3).

$$ZCR = \sum_{m=n-N+1}^n |sgn[x(m)] - sgn[x(m-1) * w(n-m)]| \quad (3)$$

by referring to equation (4).

$$sgn[x(n)] = \begin{cases} 1, & x_i(n) \geq 0 \\ -1, & x_i(n) < 0 \end{cases} \quad (4)$$

where, N is the signal length and x (n) is the sample from frame i (0, 1, ..., N-1).

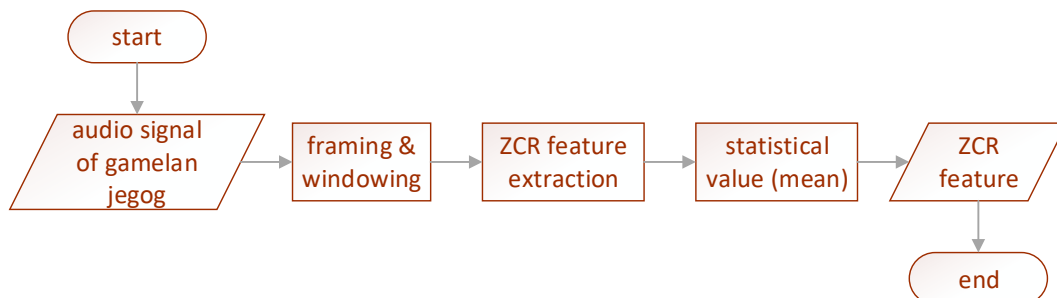


Figure 2. ZCR Feature Extraction

Figure 2 is the extraction process from the zero-crossing rate. The input is an audio signal from the gamelan jegog in .wav format. The audio signal will go through framing and windowing first. The resulting zero-crossing rate feature will take its statistical value on average.

d. Statistical Value

The features produced in a feature extraction process have many frames, so it is necessary to calculate the statistical value for these features into one vector space. The statistical value to be used is the mean (average). The mean (average) is shown in equation (5).

$$Mean(\mu) = \frac{1}{N} \sum_{n=1}^N X_n \quad (5)$$

where, X_n is the result value of the features used in the n-th frame and N is the total frame.

e. K-Nearest Neighbor

K-Nearest Neighbor classifies objects based on training data that has the closest distance to the object. The distance between the testing data and each of training data is calculated by measuring the euclidean distance between them [9]. Euclidean Distance is defined in equation (6).

$$d(a, b) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \tag{6}$$

where, n is the dimension, a is the data training and b is the data testing.

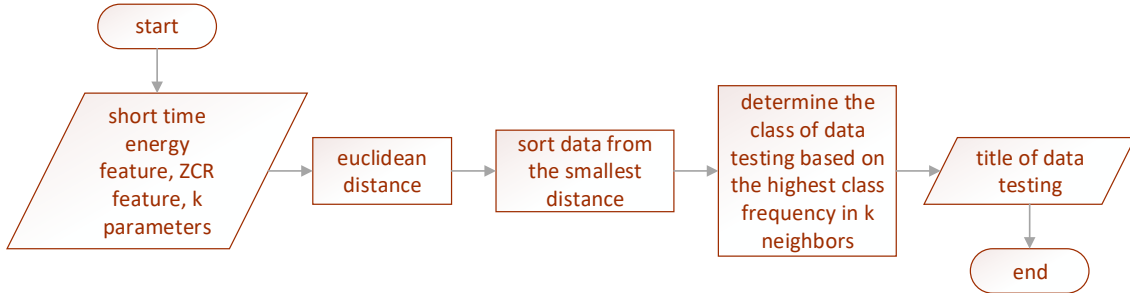


Figure 3. Classification Process

Figure 3 is a classification process with input in the form of a short time energy feature, the ZCR feature, and the k parameter. The classification process is carried out by calculating the distance between the test data and each training data, then sorted from the smallest. The output of this process is the title of the gamelan jegog being tested.

3. Result and Discussion

In this paper, the number of datasets used is 80, which is divided into 4 gamelan jegog titles. In each title, there are 20 pieces of songs with a duration of 10 seconds. The use of datasets for training data and testing data compared 75% and 25%. So, in each gamelan jegog title, there are 15 song pieces for training data and 5 song pieces for test data. The total training data for all titles is 60 data and the total test data for all titles is 20 data. The hop length and frame length used in this system are 256 and 512. Sample rate is 44.1 KHz, so the audio frame used is 11.6 ms.

Table 1. Testing Data Features Extraction

0	1	2	0	1	2
0.045922228	16.58493585	galang-bulan	0.041157238	19.46015542	gegilakan
0.0518779	17.61879352	galang-bulan	0.039001323	20.47368621	gegilakan
0.047968252	9.940023181	galang-bulan	0.039556446	20.99073343	gegilakan
0.051779338	11.94043395	galang-bulan	0.039402371	17.32028088	gegilakan
0.045384099	12.31608909	galang-bulan	0.039017184	17.73432819	gegilakan
0.037324627	41.6061985	gegenderan	0.051860907	6.444998773	jejogedan
0.037884281	29.72585195	gegenderan	0.0419582	4.484859522	jejogedan
0.031364423	27.93020908	gegenderan	0.056667815	3.941764954	jejogedan
0.033589445	27.97838822	gegenderan	0.051841647	5.911505422	jejogedan
0.039094221	37.83240912	gegenderan	0.058853185	12.35995381	jejogedan

Table 2. Training Data Features Extraction

0	1	2	0	1	2
0.031612529	18.27095641	galang-bulan	0.037024407	5.186776372	gegilangan
0.042605088	18.66622817	galang-bulan	0.032231094	9.946087477	gegilangan
0.043574853	16.14020264	galang-bulan	0.037039135	8.393510159	gegilangan
0.046442231	16.23883832	galang-bulan	0.043790105	12.02529955	gegilangan
0.048905162	18.01404933	galang-bulan	0.045165449	20.19197654	gegilangan
0.04797165	18.13142372	galang-bulan	0.043751586	20.79220408	gegilangan
0.046236043	16.76212247	galang-bulan	0.043147749	19.96357066	gegilangan
0.024722212	16.87660543	galang-bulan	0.035090542	8.102800128	gegilangan
0.022659196	17.09745251	galang-bulan	0.034954593	8.019561757	gegilangan
0.025994462	18.47841632	galang-bulan	0.036442095	7.590943033	gegilangan
0.027502356	16.14237583	galang-bulan	0.035990067	6.856960079	gegilangan
0.03027457	15.72707792	galang-bulan	0.037213602	5.75836121	gegilangan
0.030495487	13.06499836	galang-bulan	0.034984049	5.311052744	gegilangan
0.039036443	20.7086726	galang-bulan	0.033062645	6.648407749	gegilangan
0.04461146	18.68439845	galang-bulan	0.033275631	8.647728878	gegilangan
0.029670733	63.53987291	gegenderan	0.040821899	27.03105246	jejogedan
0.028561621	36.64843248	gegenderan	0.053053854	14.01993983	jejogedan
0.034274851	33.77955214	gegenderan	0.050050527	15.51596633	jejogedan
0.036446627	32.54554345	gegenderan	0.059189657	10.22597681	jejogedan
0.032871184	35.05836328	gegenderan	0.0543363	6.111458796	jejogedan
0.028936612	36.17811809	gegenderan	0.053683748	2.774626577	jejogedan
0.034143435	32.63389788	gegenderan	0.05177254	5.265235034	jejogedan
0.03683408	25.18630502	gegenderan	0.046528332	14.23343475	jejogedan
0.040461635	36.62105673	gegenderan	0.048981067	29.17613637	jejogedan
0.034134371	39.89286983	gegenderan	0.049985952	17.65266433	jejogedan
0.033688008	22.48670301	gegenderan	0.056426506	14.55857736	jejogedan
0.033944044	26.49592047	gegenderan	0.041789398	19.93345026	jejogedan
0.03118769	33.54983037	gegenderan	0.052805748	11.41436327	jejogedan
0.036946237	31.31383194	gegenderan	0.046091031	11.86550219	jejogedan
0.026384181	41.9285057	gegenderan	0.04987606	10.54605056	jejogedan

Table 1 and Table 2 show the results of feature extraction on testing data and training data. Column 0 is the extraction result of the zero-crossing rate feature. Column 1 is the result of the extraction of the short-time energy feature. Column 3 is the class of each feature extracted data. The testing scenario will be performed on each parameter k. For each k parameter, all testing data will be tested against the training data to determine the class of each testing data by

calculating the Euclidean distance between the data and each training data using equation (6). The class of each testing data as the output of the classification process determined based on the majority class of its k closest neighbors. After that, the accuracy of all testing data that is predicted to be correct will be calculated for each parameter k using equation (7).

$$Accuracy = \frac{\text{The amount of identified data is correct}}{\text{The total amount of data}} \times 100\% \quad (7)$$

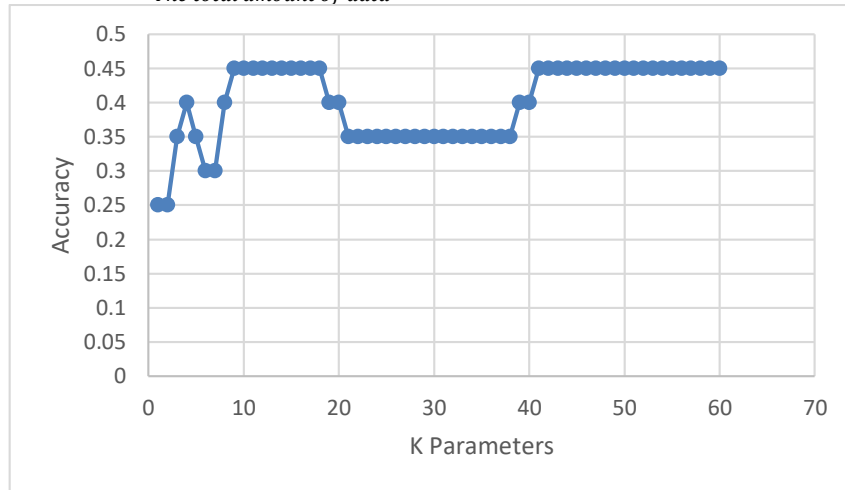


Figure 4. Accuracy of K Parameters

Figure 4 shows the accuracy of each testing process with various k values. The k parameter used is from 1 to the amount of training data, which is 60. If the accuracy obtained from a certain k value is stable, then the testing process should be stopped. However, the testing process is continued until the k parameter reaches the amount of training data to anticipate the possibility of a higher accuracy at other k values once stable accuracy is achieved. From Figure 4, it can be seen that the best accuracy achieved when k=9 with accuracy 0.45 or 45%. The feature extraction mechanism used is same but the accuracy results obtained are very small compared to the previous study [1], under 80.4762%. The difference in the amount of data used and the classification method used has an effect on the accuracy of this system.

4. Conclusion

The results showed that the highest accuracy was 45% when the k parameter is 9. Thus, the optimal k parameter in the gamelan jegog title classification uses zero-crossing rate and short-time energy is 9. Further works are needed to add features or increase the amount of data to get high accuracy.

References

- [1] R. R. Perdana, R. Soelaiman and C. Fatchah, "Implementasi Ekstraksi Fitur untuk Pengelompokan Berkas Musik Berdasarkan Kemiripan Karakteristik Suara," *Jurnal Teknik ITS*, pp. 149-152, 2017.
- [2] M. S. Nagawade and V. R. Ratnaparkhe, "Musical Instrument Identification using MFCC," in *2017 2nd IEEE International Conference On Recent Trends in Electronics Information & Communication Technology (RTEICT)*, 2017.
- [3] M. A. Banjarsari, H. I. Budiman and A. Farmadi, "Penerapan K-Optimal Pada Algoritma Knn untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan IP Sampai Dengan Semester 4," *Kumpulan jurnaL Ilmu Komputer (KLIK)*, pp. 50-64, 2015.

- [4] K. A. P. Negara, G. S. Santyadiputra and I. M. A. Pradnyana, "Film Dokumenter Seni Tabuh Jegog: Sebuah Musik Kegotong-Royongan dari Bali Barat," *KARMAPATI*, pp. 28-39, 2017.
- [5] L. Lu and A. Hanjalic, "Audio Representation," in *Encyclopedia of Database Systems*, New York, 2009, pp. 160-167.
- [6] B. R.G., S. Kopparthi, B. Adapa and B. B.D., "Separation of Voiced and Unvoiced using Zero Crossing Rate and Energy of the Speech Signal," in *American Society for Engineering Education (ASEE) Zone Conference Proceedings*, 2008.
- [7] M. Jalil, F. A. Butt and A. Malik, "Short-Time Energy, Magnitude, Zero Crossing Rate and Autocorrelation Measurement for Discriminating Voiced and Unvoiced segments of Speech Signals," *2013 The International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAECE)*, pp. 208-212, 2013.
- [8] H. Abdulbar, P. P. Adikara and S. Adinugroho, "Klasifikasi Genre Lagu dengan Fitur Akustik Menggunakan Metode K-Nearest Neighbor," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, pp. 8259-8268, 2019.
- [9] P. D. Prasetyo, I. G. P. S. Wijaya and A. Y. Husodo, "KLASIFIKASI GENRE MUSIK MENGGUNAKAN METODE MEL FREQUENCY CEPSTRUM COEFFICIENTS (MFCC) DAN K-NEAREST NEIGHBORS CLASSIFIER," *JTIKA*, Vols. Vol. 1, No. 2, pp. 189-197, 2019.

This page is intentionally left blank