Singer Identification using Autocorrelation Method

Sharmila Biswas, Sandeep Singh Solanki

Abstract: songs are the compositions embedding voice and different instrument’s sound. Different human emotions can be created by playing the appropriate song. Autocorrelation algorithm is used here to find out singer identification. In the first experiment three singers with three hindi songs (vocal) are taken as data set. Then autocorrelation is proposed on concerning a total of three singers. Using bartlett test we have found the most significant autocorrelation values of those songs of three singers. In second experiment three singers with one hindi song (vocal) are taken as data set. Here rms is used as musical features. Then autocorrelation is proposed on concerning those three singers. Using bartlett test we have found the insignificant autocorrelation values of the song of three singers. The first experiment is used to identify the singers for each song. Here three singers identify their own identification test giving most significant values of their songs. The second experiment gives the insignificant value. The insignificant values of musical features of three singers does not give the singer’s identification test.

keywords: tempo, rms, autocorrelation, song.

I. INTRODUCTION

Autocorrelation[1-2] is used here to identify the singers. It measures the relationship between a variable’s current value and its past values. Tempo [3-5] and its application on autocorrelation is proposed, concerning the three singer’s songs. For experiment purpose Musical feature or Audio feature[6-9] like tempo and rms are used as a dataset. In the first experiment three singers with three Hindi songs (vocal) without accompanying instruments each are taken as data set. Then autocorrelation is proposed on tempo concerning the three singer’s voice. Autocorrelation defines to the correlation between members of a series of numbers arranged in time. Here the members are tempos which are taken 5 sec time duration. Bartlett test [10-11] is applied on these autocorrelation values of the singers for significance test. Using the Bartlett test we have found the most significant values of those songs of three singers. Three singers identify their own identification test giving most significant values of their songs. Another experiment we have found the insignificant values of musical feature of these three singers. So it does not identify the singer’s own song. Autocorrelation is often used in signal processing and time domain signals. Using autocorrelation speech and music results are given [12]. Different musical instruments like tabla, harmonium, guitar, and flute are recognized by autocorrelation process [13].

Autocorrelation is also used as a feature for classification of musical instruments [8]. The previous work of singer identification researchers used different classifiers like Random Forest Classifier, Decision Tree Classifier, K-Nearest Neighbour Classifier, Naive Bayes Classifier, MLP Classifier etc. In our paper a new technique of singer identification used based on autocorrelation algorithm. The significant results for all singer’s identify the singers successfully. In this paper Section 2 describes methodology, Section 3 results and discussion and Section 4 conclusion.

II. METHODOLOGY

The block diagrammatic representation of the proposed singer identification process is presented in Figure 1. Step 1 songs excerpts are taken. Step 2 features are extracted. Step 3 Autocorrelation algorithm is used on the extracted features. Using this algorithm the autocorrelation function (acf) is found out. Step 4 Bartlett Test is used on this (acf). If the acf value is greater than critical value, the singer is identified with significance.

Fig 1. Block Diagram of singer identification

A. Database collection- In this research, we have recorded three singer’s three Hindi songs (vocal) without accompanying instruments. For recording purpose we used behringer mic, behringer mixer, creative 5.1 sound card and sonic foundry sound forge 7.0 software. In experiment 1 three singers with three Hindi songs (vocal) without accompanying instruments each are taken as data set. These songs are clipped in .WAV format. These songs are clipped 10 divisions with 5 sec duration. The clips are taken as 0-5sec, 5-10sec, 10-15 sec, 15-20 sec, 20-25sec, 25-30sec, 30-35sec, 35-40 sec, 40-45sec and 45-50 sec time duration. In these experiments we used 10 tempo and rms features as input dataset. The input audio files have various attributes such as file type (wav), sampling rate (44.1k), audio type (mono) etc. The clipping process is done by the MATLAB14 and MIRTOOLBOX1.5.

Revised Manuscript Received on October 20, 2020.

* Correspondence Author
SharmilaBiswas PhD student in the Electronics and Communication Engineering Department, Birla Institute of Technology, Deemed University.
Dr.Sandeep Singh Solanki Professor Electronics and Communication Engineering Department, Birla Institute of Technology, Deemed University.
B. Feature Extraction

In this research, we have found out Tempo and Rms for experiment1 and experiment 2.

Tempo is measured by the speed of a musical piece. It is defined the number of beats per minute. Tempo is evaluated by detecting periodicities from the onset detection curve. Once the onset detection curve is formed by detection function, a determination of periodicities in the curve gives us an estimate of the tempo.

Root Mean Square (RMS) energy denotes the square root of the mean square of the amplitude values of the audio signal

\[ x_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2} \]  

(1)

where \( x_i \) denotes the magnitude of the \( i^{th} \) sample and the total number of samples is \( n \).

C. Autocorrelation

The autocorrelation function (ACF) at \( lag \ k \) is denoted by \( \rho_k \) of a stationary stochastic process.

ACF is defined as \( \rho_k = \gamma_k / \gamma_0 \)  

(2)

where \( \gamma_k = \text{cov}(x_{i+k}, x_i) \) for any \( i \) and \( \gamma_0 \) is the variance of the stochastic process.

D. The Bartlett test

Bartlett test is the significance test of the autocorrelation function values greater than critical values are rejected at the 95% level. Significance measurement for all singer’s songs gives most significant values. Same singer’s features gives significant values for every three songs. Three singers identify their own identification test giving most significant values of their songs. Fig 3, gives the graphical representation of significant values of three singer’s different three Hindi songs (vocal) without accompanying instruments.

| Table 1 | First singer’s 1st song’s autocorrelation using tempo |
|------------------------|------------------------|
| Tempo      | lag   | acf   |
| 80.81429   | lag1  | -0.62371 |
| 76.03075   | lag2  | 0.050851 |
| 166.0723   | lag3  | 0.40284 |
| 59.78754   | lag4  | -0.74896 |
| 86.87201   | lag5  | 0.705083 |
| 167.7571   |       |       |
| 45.32916   |       |       |
| 189.4219   |       |       |
| 80.17233   |       |       |
| 75.33856   |       |       |

| Table 2 | Second singer’s 2nd song’s autocorrelation using tempo |
|------------------------|------------------------|
| Tempo | lag | acf |
| 65.45388 | lag1 | -0.32137 |
| 108 | lag2 | -0.4174 |
| 155.5944 | lag3 | 0.79353 |
| 77.335 | lag4 | -0.23434 |
| 123.6755 | lag5 | -0.33731 |
| 150.8989 |       |       |
| 102.2857 |       |       |
| 123.089 |       |       |
| 119.3986 |       |       |
| 133.5894 |       |       |
Table 3: First singer’s 3rd song’s autocorrelation using tempo

| Tempo   | lag 1 | lag 2 | lag 3 | lag 4 | lag 5 |
|---------|------|------|------|------|------|
| 73.69403| -0.09252 | -0.4009 | -0.28048 | -0.0493 | -0.80003 |
| 161.744 | 111.3461 | 75.43913 | 76.9693 | 99.77903 | 76.05898 | 110.6176 | 109.0149 | 119.8549 |

Table 4: Second singer’s 1st song’s autocorrelation using tempo

| Tempo   | lag 1 | lag 2 | lag 3 | lag 4 | lag 5 |
|---------|------|------|------|------|------|
| 78.68252 | -0.10915 | -0.69351 | 0.392267 | 0.591371 | -0.60736 |
| 84.4606 | 110.8491 | 110.1496 | 66.03057 | 111.7673 | 165.7399 | 84.38965 | 82.63125 | 118.8151 |

Table 5: Second singer’s 2nd song’s autocorrelation using tempo

| Tempo   | lag 1 | lag 2 | lag 3 | lag 4 | lag 5 |
|---------|------|------|------|------|------|
| 138.8445 | 0.325111 | -0.11794 | -0.56994 | -0.71024 | -0.49476 |
| 183.8151 | 97.15035 | 107.0583 | 69.42679 | 52.56768 | 72.3562 | 107.0178 | 177.9451 | 106.4288 |

Table 6: Second singer’s 3rd song’s autocorrelation using tempo

| Tempo   | lag 1 | lag 2 | lag 3 | lag 4 | lag 5 |
|---------|------|------|------|------|------|
| 158.0637 | 0.837289 | 0.559859 | 0.335426 | 0.054805 | -0.63029 |
| 168.7915 | 162.7669 | 172.5092 | 175.9941 | 96.84827 | 56.05359 | 64.73223 | 64.51314 | 68.66323 |

Table 7: Third singer’s 1st song’s autocorrelation using tempo

| Tempo   | lag 1 | lag 2 | lag 3 | lag 4 | lag 5 |
|---------|------|------|------|------|------|
| 129.9735 | -0.15889 | 0.025811 | 0.00502 | -0.82717 | 0.057235 |
| 180.5715 | 144.9542 | 68.5158 | 150.5224 | 67.23679 | 119.9162 | 148.8025 | 134.7729 | 171.6447 |

Table 8: Third singer’s 2nd song’s autocorrelation using tempo

| Tempo   | lag 1 | lag 2 | lag 3 | lag 4 | lag 5 |
|---------|------|------|------|------|------|
| 70.77796 | -0.2734 | 0.098204 | 0.116549 | 0.611106 | -0.75607 |
| 70.83192 | 158.8872 | 57.6677 | 81.94267 | 82.90668 | 161.5633 | 53.3687 | 172.5636 | 159.9205 |

In experiment 2, we find out autocorrelation function on rms concerning the three singer’s Hindi songs (vocal) without accompanying instruments each using the equation (1). First singer’s song’s autocorrelation function are \((0.170632, -0.07472, -0.245, -0.341)\) at lag1, lag2, lag3, lag4 and lag5. Second singer’s song’s autocorrelation function are \((-0.20939, 0.114115, -0.05261, -0.4455, 0.46502)\) at lag1, lag2, lag3, lag4 and lag5. Third singer’s song’s autocorrelation function are \((0.588672, 0.101042, -0.34188, 0.481768)\) at lag1, lag2, lag3, lag4 and lag5. Table 10 to Table 12 shows the different autocorrelation function at different lag values. Using the Bartlett test equation (2) we find the critical values. The critical values are 0.6198. The sample size is \(T=10\).
We compare the value of the sample coefficient with the critical values. The first singer’s first song, second singer’s first song and third singer’s first song’s autocorrelation function values are not greater than critical values. Since the values do not fall outside the bands, the null hypothesis is not rejected at the 95% level. These three acf values are insignificant. So different features do not give significance values.

**Table 10 First singer’s 1st song’s autocorrelation using rms**

| rms     | lag  | acf     |
|---------|------|---------|
| 0.02118 | lag1 | 0.170632|
| 0.017478| lag2  | -0.07472|
| 0.028429| lag3  | -0.245  |
| 0.019021| lag4  | -0.34188|
| 0.013902| lag5  | 0.481768|
| 0.068257|       |         |
| 0.044609|       |         |
| 0.045443|       |         |
| 0.023467|       |         |
| 0.020736|       |         |

**Table 11 Second singer’s 1st song’s autocorrelation using rms**

| rms     | lag  | acf     |
|---------|------|---------|
| 0.008089| lag1 | -0.20939|
| 0.107087| lag2  | 0.114115|
| 0.054396| lag3  | -0.05261|
| 0.041515| lag4  | -0.4455 |
| 0.117884| lag5  | 0.46502 |
| 0.045762|       |         |
| 0.19527 |       |         |
| 0.116155|       |         |
| 0.075295|       |         |
| 0.05627 |       |         |

**Table 12 Third singer’s 1st song’s autocorrelation using rms**

| rms     | lag  | acf     |
|---------|------|---------|
| 0.014589| lag1 | 0.588672|
| 0.01474 | lag2  | 0.101042|
| 0.016473| lag3  | -0.13514|
| 0.016249| lag4  | -0.53226|
| 0.021794| lag5  | -0.42252|
| 0.021889|       |         |
| 0.01878 |       |         |
| 0.018129|       |         |
| 0.019465|       |         |
| 0.018707|       |         |

**IV. CONCLUSION**

In this paper, we correctly identified the autocorrelation of similar features among three singers. The autocorrelation is used on the similar features related to tempo. To find the significance between similar musical features, Bartlett test is used. The first experiment considers the autocorrelation in tempo features concerning the three singer’s rendition of songs without accompanying instruments. Using these autocorrelation values, we calculate the sample coefficient values of these songs.
So we can say all significant values of three songs determine that the singer is the same. Three songs of each singer give significant values. In this way we can identify the three singers individually. This is an identification process of a singer. The second experiment gives the insignificant values of the three singer’s songs. Another feature is not used to identify the singer’s individuality. So it does not identify the singers.

REFERENCES

1. Sato,S.I, Alejandro Bidondo, A. “Synthesis of Music Signals by using Autocorrelation Function.” 6th international symposium on temporal design, joint with 26th annual meeting of Taiwan Institute of Acoustics(TIA)/TAIPEI, 16-17 November (2013)

2. Eck,D. “Meter and autocorrelation.” In 10th Rhythm Perception and Production Workshop (RPPW ’05), Blitzen, Belgium, pp.1-25 (2005)

3. Alonso,M., David,B., Richard,G. “Tempo and beat estimation of musical signals.” In Proc. International Conference on Music Information Retrieval. Barcelona: Audiovisual Institute, Pompeu Fabra University, pp. 158–163 (2004)

4. Gouyon,F.; Dixon,S. “A review of automatic rhythm description systems.” Computer Music Journal, 29(1), pp.1-23 (2005)

5. Peeters, G. “Time variable tempo detection and beat marking.” In Proc. Int. Comput. Music Conf. (ICMC), Barcelona, Spain (2005)

6. Mahto, K., Hota, A., Solanki, S.S., Chakraborty, S. “A study on artist similarity using projection pursuit and mfcc:identification of gharana from ghara performance.” In: International Conference on Computing for Sustainable Global Development, IEEE, New Delhi, India, pp. 647-653 (2014)

7. Datta, A.K., Solanki, S.S., Sengupta, R., Chakraborty, S., Mahto, K., Patranabis, A. “Music information retrieval.” in (1st edn), Signal Analysis of Hindustani Classical Music, Springer Singapore, pp.17-33 (2017)

8. Chandwadkar, D. M., Sutaone, M. “Selecting Proper Features and Classifiers for Accurate Identification of Musical Instruments.” International Journal of Machine Learning and Computing, Vol. 3, No. 2, pp.172-175 (2013)

9. Deshmukh, S.H., Bhurud, S.G. “North Indian Classical Music’s Singer Identification by Timbre Recognition using MIR Toolbox.” International Journal of Computer Applications, Volume 91 – No.4, pp.1-5 (2014)

10. Levich, R.M., Rizzo, R.C. “Alternative Tests for Time Series Dependence Based on Autocorrelation Coefficients.” Technical report, Stern School of Business, New York University, pp.1-23 (1998)

11. Kokoszka, P.S., Politis, D.N. “Nonlinearity of arch and stochastic volatility models and bartlett’s formula.” Probability and mathematical statistics, Vol. 31, Fasc. 1, pp.47–59 (2011)

12. Banchhor, S.K., Sahu, O.P., Prabhakar. “A Speech/Music Discriminator based on Frequency energy, Spectrogram and Autocorrelation.” International Journal of Soft Computing and Engineering (ISJCE) ISSN: 2231-2307, Volume-2, Issue-1, pp. 480-483 (2012)

13. Banchhor, S.K., Khan, A. “Musical Instrument Recognition using Spectrogram and Autocorrelation.” International Journal of Soft Computing and Engineering (ISJCE) ISSN: 2231-2307, Volume-2, Issue-1, pp.1-4 (2012)

AUTHORS PROFILE

Sharmila Biswas PhD student in the Electronics and Communication Engineering Department, Birla Institute of Technology, Deemed University. Her research interests are music signal processing and signal processing.

Dr. Sandeep Singh Solanki Professor in the Electronics and Communication Engineering Department, Birla Institute of Technology, Deemed University. His research interests are music and speech signal processing and automation.