Comparative study of pitch detection algorithm to detect traditional Balinese music tones with various raw materials

I P B W Brata¹ and I D M B A Darmawan²

¹Informatics Department, Faculty of Mathematics and Natural Sciences, Universitas Udayana, Kampus Bukit Jimbaran, Badung, Indonesia
²dewabaya@unud.ac.id

Abstract. The characteristics of a tone are formed from the fundamental frequency values that make up the tone. A musical instrument's tone has different fundamental values in building the timbre of the instrument's tone. This study compares four PDA (Pitch Detection Algorithm) methods and tries to find the best algorithm for detecting Balinese instruments' tone with different raw materials, namely leather, metal, and bamboo. The PDA method analyzed was the Autocorrelation, ZCR, HPS, and FFT methods. This study also explores the performance of each PDA (Pitch Detection Algorithm) approach in checking the pitch of traditional Balinese musical instruments and focuses on which of the four PDA (Pitch Detection Algorithm) methods detect Balinese musical instruments by comparing the results of the pitch analysis of each line with the original frequency. This study used 193 tone data consisting of Kendang to represent skin, Gangsa instruments to represent metal, Rindik, and Balinese flutes to represent bamboo material. Based on the test results, the PDA (Pitch Detection Algorithm) method that produces the closest value of the four musical instruments is HPS (Harmonic Product Spectrum) and FFT (Fast Fourier Transform) with an average difference of 4.81 Hz and 4.85 Hz.

1. Introduction
Pitch is the perception of sound level, which is closely related to the fundamental frequency and is a feature that can represent a tone's characteristics. The tones on musical instruments usually have different values, identities, and rhythm. Pitch is a qualitative measurement of how humans perceive sound about the signal's frequency, which describes the height of the perceived sound. Increasing the frequency of a sound will also make the tone of the sound higher. The tone frequency \( f \) is a pure sine wave corresponding to the sound's perceived tone. The fundamental frequency \( F_0 \) is the inverse of the length of the tone period, \( P_0 \), which represents the smallest repetition unit of a signal.

There are several PDA (Pitch Detection Algorithm) methods in detecting the pitch [2] comparing the Autocorrelation algorithm and ZCR (Zero Crossing Rate). Both algorithms are suitable for detecting the frequency of traditional Balinese musical instruments (gender wayang). Another study comparing several pitch tracker methods for real-time guitar effects [3] concluded the YIN method the algorithm overall performed best with mean errors below 0.01 Hz and latency of 27.4 ms and better than FFT, PRAAT, and LTP. Our interest is also obtained from this research because the FFT approach has problems with the stepped C major scale, so we want to test FFT in the case of percussive musical instruments. [4] was comparing Autocorrelation and AMDF in detecting the pitch of the voice with Thai and did not find the significant effects of preprocessing on AMDF. The analysis of the HPS (Harmonic Product Spectrum) method was also carried out [5] by comparing the algorithm with the ZCR (Zero Crossing Rate) in the case of Kidung Bali. It found that the HPS had a small
standard deviation value but with more considerable processing time. Pitch detection is essential in a wide variety of sound processing cases. However, many studies have discussed tone detection algorithms, but these studies usually only compare them with the same musical instrument. If we know which algorithm is best, using one algorithm will make a computer work more straightforward, such as making music transcriptions in a Balinese cultural performance with various musical instruments.

This study will compare several PDA (Pitch Detection Algorithm) methods to detect various musical instruments. The musical instruments studied were the Kendang, Gangsa, Rindik, and Balinese flute. This research will focus on how the four algorithms in detecting the four Balinese musical instrument's fundamental frequencies are used by analyzing each instrument's original frequency using the Adobe Audition 2020 application, then making them a reference. The difference between the results of Adobe analysis with each algorithm will determine the performance of the algorithm. This algorithm's performance is very diverse, mostly if we use different musical instruments, and this research wants to find out from the three raw materials whether each algorithm has its advantages.

2. Method

This study will compare the Autocorrelation algorithm, ZCR (Zero Crossing Rate), HPS (Harmonic Product Spectrum), and FFT (Fast Fourier Transform) in detecting various types of musical instruments. These musical instruments are Kendang (leather), Gangsa (metal), Rindik, and Balinese Suling (bamboo), each made of different raw materials. The musical instrument's sound signal will undergo pitch extraction according to each algorithm's calculation with the Scipy [6] and Numpy [7] library in python programming.

2.1. Fast Fourier Transform

Fast Fourier Transform (FFT) is a method used to transform a signal from the time domain to the frequency domain. As a pitch detection algorithm [8], FFT has proven to be accurate and robust in practice on natural sounds such as voice and classic musical instruments. The concept of Fourier transformation uses to convert the signal from time-domain data to frequency space [1]. Continuous Fourier transforms defined as follows equation (1).

\[ X(f) = \int_{-\infty}^{\infty} x(t) e^{-2\pi i f t} dt. \]  

(1)

2.2. Autocorrelation function

Autocorrelation (ACF) [3] is an algorithm for estimating a sound signal's fundamental frequency (f0). This method works in the time domain based on the clipping center method and the cross-product of a signal with itself. There are several usages of the ACF outside the field of audio processing. Still, for our purposes, when regarding pitch, it can be seen as a way of searching for fundamental frequency hidden within the signal. ACF method defined as in equation (2).

\[ r_x(\tau) = \sum_{n=-\infty}^{\infty} x_n x_{n+\tau}. \]  

(2)

2.3. Zero-Crossing Rate

Zero crossing rate (ZCR) measures the number of times the signal in the time domain changes its sign. In this case, what is meant is an algebraic sign different from each sample in the previous piece. The ZCR describes the amount of high-frequency energy in the signal called the fundamental frequency (f0) even though calculated in the time domain. The ZCR proved to be highly discriminatory for the percussion-shaped class of instruments [9]. The ZCR of the x(n) time-domain signal defines by equation (3).

\[ ZCR(n) = \frac{1}{2N} \sum_{i=1}^{N} | \text{sign} [ x(n+i) ] | - | \text{sign}[x(n+i-1)] |. \]

(3)
2.4. HPS (Harmonic Product Spectrum)
Harmonic Product Spectrum (HPS) algorithm is described in [10], a tone detection algorithm that works in the frequency domain. FFT is used to represent the signal in spectral form. This method assumes the analyzed signal is in the way of a harmonic part. This method will calculate the spectral energy of the harmonic positions for each candidate $F_0$. The fundamental frequency ($f_0$) was obtained from the most considerable spectral energy.

![Image of HPS algorithm](Signal, Spectrum, Fundamental Frequency)

Figure 1. Harmonic Product Spectrum
Source: [9,10]

3. Result and Discussion
This study used the recording of the Tascam dr-44wl, which amounted to 193 sound data for the four musical instruments. The data is the result of recording all the notes played by the four instruments. Of course, each tool has a different number of tones. The drum instrument has six notes with a total dataset of 30, a ten-note gangsa instrument with a total dataset of 50 data, a thirteen note index instrument with a total dataset of 65 data. The latter is a ten-tone Balinese flute with 48 data points - each dataset recorded with a single channel (mono) 44100 Hz sampling frequency. The data length varies depending on the instrument. The first test is to find each instrument tone's original frequency using the Adobe Audition 2020 application to analyze each note's frequency on a musical instrument. Adobe audition is a software that can calculate the pitch of the voice with a frequency analysis feature. According to the filter effect, Adobe Audition uses a frequency analysis panel to identify problematic frequency bands. This research uses python programming with Numpy and Scipy python libraries to calculate each pitch detection method. This Python library has credible functions in calculating Autocorrelation, Zero Crossing Rate, Fast Fourier Transform, and Harmonic Product Spectrum. The signal file in the form of .wav becomes the system input, and then the library uses to perform calculations.

3.1. Original Frequency Analysis Results
Table 1 shows the results of tone analysis with Adobe Audition on each musical instrument. The research results will be the primary reference for the original frequency of the musical instrument tone. Where each musical instrument represents the respective material (leather, metal, and bamboo).

| Table 1. Instrument frequency analysis |
|---------------------------------------|
| Instrument | Type | Tone | average frequency |
| Kendang    |      |      |                  |
| Right hand | Cung |       | 161.39 Hz        |
| Right hand | De   |       | 98.02 Hz         |
| Right hand | Tek  |       | 27.25 Hz         |
| Left hand  | Plak |       | 25.55 Hz         |
| Left hand  | Pung |       | 224.03 Hz        |
| Left hand  | Teng |       | 231.99 Hz        |
| Gangsa     |      |      |                  |
| 1           | Ndong|       | 292.97 Hz        |
| 2           | Ndeng|       | 324.96 Hz        |
| 3           | Ndung|       | 409.58 Hz        |
3.2. Comparison of all Pitch Detection Algorithm

3.2.1. Result of Kendang instrument. Table 2 shows each pitch detection algorithm's analysis results on the Kendang instrument, while table 3 shows the difference between the original frequency and each Pitch Detection Algorithm.

Table 2. Comparison of the pitch detection algorithm with the original frequency on the Kendang Instrument

| Type   | Tone | Real frequency | FFT  | ZCR  | ACF  | HPS  |
|--------|------|----------------|------|------|------|------|
| Right  | Cung | 161.39         | 164.99| 173.01| 166.75| 164.99|
| Right  | De   | 98.02          | 86.11 | 114.04| 91.96 | 86.10 |
| Right  | Tek  | 27.25          | 1.62  | 133.45| 214.85| 1.62  |
| Left   | Plak | 25.55          | 2.48  | 151.99| 148.18| 1.46  |
| Left   | Pung | 224.03         | 225.65| 220.75| 108.85| 225.66|
| Left   | Teng | 231.99         | 227.96| 240.76| 205.05| 227.93|

Table 3. The difference between each pitch detection algorithm and the original frequency on the Kendang instrument

| Type   | Tone | FFT  | ZCR  | ACF  | HPS  |
|--------|------|------|------|------|------|
| Right  | Cung | -    | -    | -    | -    |
| Right  | De   | -    | -    | -    | -    |
| Right  | Tek  | -    | -    | -    | -    |
| Left   | Plak | -    | -    | -    | -    |
| Left   | Pung | -    | -    | -    | -    |
| Left   | Teng | -    | -    | -    | -    |
Table 3 shows the difference and average of the original frequency with the frequency of each PDA (Pitch Detection Algorithm). The FFT (Fast Fourier Transform) algorithm and the HPS (Harmonic Product Spectrum) algorithm have the smallest difference for detecting frequencies than the other two algorithms. However, if observed, the four methods cannot correctly see the Kendang instrument's two tones, especially the Tek-right tone and the left-Plak tone. This error can be affected by playing the Kendang using the palms and making a short and complex sound that makes it very difficult to detect the Kendang instrument frequency accurately.

3.2.2. Result of the Gangsa instrument. Table 4 shows that each Pitch Detection Algorithm analysis results on the Gangsa instrument, while table 5 shows the difference between the original frequency and each Pitch Detection Algorithm.

### Table 4. Comparison of the pitch detection algorithm with the original frequency on the Gangsa Instrument

| Type | Tone  | Real frequency | Pitch Detection Algorithm |
|------|-------|----------------|----------------------------|
|      |       |                | FFT | ZCR | ACF | HPS |
| 1    | Ndong | 292.97         | 298.65 | 369.03 | 264.82 | 298.65 |
| 2    | Ndeng | 324.96         | 325.74 | 509.55 | 200.72 | 325.73 |
| 3    | Ndung | 409.58         | 409.74 | 492.73 | 414.44 | 409.74 |
| 4    | Ndang | 449.23         | 448.05 | 581.93 | 453.83 | 448.05 |
| 5    | Nding | 550.15         | 555.29 | 635.33 | 533.64 | 555.29 |
| 6    | Ndong | 593.09         | 598.23 | 709.25 | 591.06 | 598.23 |
| 7    | Ndeng | 662.78         | 660.81 | 779.72 | 580.30 | 660.81 |
| 8    | Ndung | 806.46         | 811.54 | 1009.57 | 429.17 | 811.54 |
| 9    | Ndang | 865.46         | 861.38 | 911.63 | 266.86 | 861.38 |
| 10   | Nding | 1099.86        | 1096.49 | 1124.35 | 218.69 | 1096.49 |

### Table 5. The difference between each pitch detection algorithm and the original frequency on the Gangsa instrument

| Type | Tone  | FFT  | ZCR  | ACF  | HPS  |
|------|-------|------|------|------|------|
| 1    | Ndong | 5.68 | 76.06 | 28.15 | 5.68 |
| 2    | Ndeng | 0.78 | 184.59 | 124.24 | 0.77 |
| 3    | Ndung | 0.16 | 83.15 | 4.86  | 0.16 |
| 4    | Ndang | 1.18 | 132.7 | 4.6   | 1.18 |
| 5    | Nding | 5.14 | 85.18 | 16.51 | 5.14 |
| 6    | Ndong | 5.14 | 116.16 | 2.03  | 5.14 |
| 7    | Ndeng | 1.97 | 116.94 | 82.48 | 1.97 |
| 8    | Ndung | 5.08 | 203.11 | 377.29 | 5.08 |
| 9    | Ndang | 4.08 | 46.17  | 598.6 | 4.08 |
| 10   | Nding | 3.37 | 24.49 | 881.17 | 3.37 |

| Average | 3.258 | 106.855 | 211.993 | 3.257 |
Table 5 shows the difference and average original frequency with the frequency of each PDA (Pitch Detection Algorithm). FFT (Fast Fourier Transform) and HPS (Harmonic Product Spectrum) still hold the smallest difference as in the drum instrument with an average difference of 3.258 Hz and 3.257 Hz to detect the tone frequency of the Gangsa instrument. However, it should be noted that Autocorrelation is also good at detecting this metal instrument. The Autocorrelation can still detect the Gangsa instrument’s pitch from the first to the sixth tone, but after the eighth note, which is a high note, the Autocorrelation is not able to correctly detect metal musical instruments.

3.2.3. Result of Rindik instrument.
Table 6 shows the analysis results of each Pitch Detection Algorithm on the Rindik instrument, while table 7 shows the difference between the original frequency and each Pitch Detection Algorithm.

### Table 6. Comparison of the pitch detection algorithm with the original frequency on the Rindik Instrument

| Type  | Tone | Real frequency | FFT | ZCR | ACF | HPS |
|-------|------|----------------|-----|-----|-----|-----|
| 1     | Ndong | 158.32         | 159.63 | 858.97 | 159.41 | 159.65 |
| 2     | Ndeng | 189.88         | 188.45 | 320.30 | 186.11 | 188.45 |
| 3     | Ndung | 222.82         | 216.47 | 347.83 | 215.34 | 216.48 |
| 4     | Ndang | 240.57         | 244.33 | 479.90 | 238.46 | 244.31 |
| 5     | Nding | 274.00         | 273.38 | 463.56 | 273.28 | 273.38 |
| 6     | Ndong | 324.95         | 326.05 | 471.56 | 326.34 | 326.07 |
| 7     | Ndeng | 364.60         | 364.20 | 555.70 | 363.90 | 364.21 |
| 8     | Ndung | 444.89         | 441.01 | 680.00 | 437.79 | 441.00 |
| 9     | Ndang | 502.29         | 504.11 | 720.83 | 499.27 | 504.12 |
| 10    | Nding | 546.77         | 551.27 | 805.21 | 550.68 | 551.27 |
| 11    | Ndong | 663.60         | 659.94 | 951.24 | 657.49 | 659.92 |
| 12    | Ndeng | 747.18         | 743.41 | 916.81 | 730.56 | 743.37 |
| 13    | Ndung | 884.42         | 880.38 | 1256.08 | 674.86 | 880.33 |

### Table 7. The difference between each pitch detection algorithm and the original frequency on the Rindik instrument

| Type  | Tone | FFT | ZCR | ACF | HPS |
|-------|------|-----|-----|-----|-----|
| 1     | Ndong | 1.31 | 700.65 | 1.09 | 1.33 |
| 2     | Ndeng | 1.43 | 130.42 | 3.77 | 1.43 |
| 3     | Ndung | 6.35 | 125.01 | 7.48 | 6.34 |
| 4     | Ndang | 3.76 | 239.33 | 2.11 | 3.74 |
| 5     | Nding | 0.62 | 189.56 | 0.72 | 0.62 |
| 6     | Ndong | 1.1 | 146.61 | 1.39 | 1.12 |
| 7     | Ndeng | 0.4 | 191.1 | 0.7 | 0.39 |
| 8     | Nding | 3.88 | 235.11 | 7.1 | 3.89 |
| 9     | Ndang | 1.82 | 218.54 | 3.02 | 1.83 |
| 10    | Nding | 4.5 | 258.44 | 3.91 | 4.5 |
| 11    | Ndong | 3.66 | 287.64 | 6.11 | 3.68 |
| 12    | Ndeng | 3.77 | 169.63 | 376.62 | 3.81 |
| 13    | Ndung | 4.04 | 371.66 | 209.56 | 4.09 |

The Rindik instrument also shows the same thing as the other two musical instruments. Table 7 shows that the FFT and HPS algorithms still show the smallest difference, with an average value of 2.81 Hz and 2.82 Hz. However, even though the average ZCR algorithm is smaller than the Autocorrelation
algorithm, on the detection of tones of types 1 to 10, ACF has a small difference closer to the original sound compared to ZCR, only when the 11 and 12 tone types this algorithm has a very drastic difference with the original frequency.

3.2.4. Result of Balinese flute instrument. Table 8 shows the analysis results of each Pitch Detection Algorithm on the Balinese flute instrument, while table 9 shows the difference between the original frequency and each Pitch Detection Algorithm.

Table 8. Comparison of the pitch detection algorithm with the original frequency on the Balinese flute Instrument

| Type  | Tone | Real frequency   | FFT     | ZCR     | ACF     | HPS     |
|-------|------|------------------|---------|---------|---------|---------|
| Low   | Dang | 434.15           | 431.33  | 997.39  | 431.57  | 431.32  |
|       | Deng | 323.30           | 323.69  | 333.35  | 323.93  | 323.69  |
|       | Dung | 402.78           | 399.89  | 1016.00 | 399.94  | 399.95  |
|       | Dang | 878.74           | 875.51  | 872.45  | 876.45  | 875.51  |
|       | Deng | 640.82           | 640.68  | 672.57  | 640.72  | 640.68  |
| Medium| Ding | 539.83           | 540.13  | 1357.79 | 540.22  | 540.13  |
|       | Dong | 584.13           | 584.09  | 694.68  | 584.42  | 584.09  |
|       | Dung | 806.31           | 810.37  | 809.32  | 567.45  | 810.36  |
| High  | Ding | 1091.16          | 1090.32 | 937.28  | 636.93  | 1090.32 |
|       | dong | 1355.01          | 1354.46 | 1215.20 | 677.77  | 1354.46 |

Table 9. The difference between each pitch detection algorithm and the original frequency on the Balinese flute instrument

| Type  | Tone | FFT     | ZCR     | ACF     | HPS     |
|-------|------|---------|---------|---------|---------|
| Low   | Dang | 2.82    | 563.24  | 2.58    | 2.83    |
|       | Deng | 0.39    | 10.05   | 0.63    | 0.39    |
|       | Dung | 2.89    | 613.22  | 2.84    | 2.83    |
|       | Dang | 3.23    | 6.29    | 2.29    | 3.23    |
|       | Deng | 0.14    | 31.75   | 0.1     | 0.14    |
| Medium| Ding | 0.3     | 817.96  | 0.39    | 0.3     |
|       | Dong | 0.04    | 110.55  | 0.29    | 0.04    |
|       | Dung | 4.06    | 3.01    | 238.86  | 4.05    |
|       | Ding | 0.84    | 153.88  | 454.23  | 0.84    |
| High  | Ding | 0.55    | 139.81  | 677.24  | 0.55    |
|       | dong | 0.05    | 139.81  | 677.24  | 0.55    |

Table 9 also shows that the FFT and HPS algorithms are stable in detecting the Balinese flute instrument's frequency. The autocorrelation algorithm is also suitable for calculating the frequency of the Balinese flute. When measuring the high notes, the ACF cannot calculate the correct frequency for this instrument. In contrast to the ZCR, which has the highest average frequency, it cannot correctly detect Balinese flute instruments.

Table 10. Average of all algorithm

| Algorithm | Average  |
|-----------|----------|
| FFT       | 4.81144872 |
| ZCR       | 162.068295 |
| ACF       | 118.800173 |
| HPS       | 4.85678205 |

7
Table 10 shows the average difference of all Balinese musical instruments studied. FFT and HPS are reliable algorithms in calculating Balinese musical instrument's fundamental frequency, even with low signal conditions.

4. Conclusion

Autocorrelation is not suitable for metal-based instruments such as the Gangsa, while the Zero Crossing Rate scores low on average for bamboo-based instruments. The best algorithms for pitch detection are obtained by HPS and FFT, with the similarity between the mean frequency differences from the analysis results. These two methods are very well used to measure each instrument's frequency with an absolute difference of 4,811 Hz and 4,856 Hz.

References

[1] McLeod P 2008 Fast, Accurate Pitch Detection Tools for Music Analysis (University of Otago)
[2] Darmawan I D M B A 2015 Perbandingan Metode Zcr Dan Autocorrelation Untuk Menghitung Frekuensi Pada Gambelan Gender Wayang J. Ilm. Ilmu Komput. Univ. Udayana 8 1–6
[3] Von Dem Knesebeck A and Zölzer U 2010 Comparison of pitch trackers for real-time guitar effects 13th International Conference on Digital Audio Effects, DAFx 2010 Proceedings (DAFx-10) pp 1–4
[4] Tan L and Karnjanadecha M 2003 Pitch Detection Algorithm: Autocorrelation Method and Amdf 3Rd Int. Symp. Commun. Inf. Technol. 551–556
[5] Darmawan I D M B A 2017 Implementasi Real Time Pitch Detection Untuk Mendeteksi Nada Kidung Bali Dengan Python Semin. Nas. Sains dan Teknol. IV 2017 29–36
[6] Virtanen P, Gommers R, Oliphant T E, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt S J, Brett M, Wilson J, Millman K J, Mayorov N, Nelson A R J, Jones E, Kern R, Larson E, Carey C J, Polat İ, Peng Y, Moore E W, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero E A, Harris C R, Archibald A M, Ribeiro A H, Pedregosa F, van Mulbregt P, Vijaykumar A, Bardelli A Pietro, Rothenberg A, Hilboll A, Kloeckner A, Scopatz A, Lee A, Rokem A, Woods C N, Fulton C, Masson C, Häggström C, Fitzgerald C, Nicholson D A, Hagen D R, Pasechnik D V., Olivetti E, Martin E, Wieser E, Silva F, Lenders F, Wilhelm F, Young G, Price G A, Ingold G L, Allen G E, Lee G R, Audren H, Probst I, Dietrich J P, Silterra J, Webber J T, Slavíč J, Nothman J, Buchner J, Kulick J, Schönhäger J L, de Miranda Cardoso J V, Reimer J, Harrington J, Rodriguez J L C, Nunez-Iglesias J, Kuczynski J, Trizt K, Thoma M, Newville M, Kümmerer M, Bolingbroke M, Tartre M, Pak M, Smith N J, Nowaczyk N, Shebanov N, Pavlyk O, Brodtkor P A, Lee P, McGibbon R T, Feldbauer R, Lewis S, Tygier S, Sievert S, Vigna S, Peterson S, More S, et al 2020 SciPy 1.0: fundamental algorithms for scientific computing in Python Nat. Methods 17 261–72
[7] Harris C R, Millman K J, van der Walt S J, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith N J, Kern R, Picus M, Hoyer S, van Kerkwijk M H, Brett M, Haldane A, del Rio J F, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C and Oliphant T E 2020 Array programming with NumPy Nature 585 357–62
[8] Marchand S, Efficient A and Using P A 2001 An Efficient Pitch-Tracking Algorithm Using a Combination of Fourier Transforms Sylvain Marchand To cite this version : Proceedings of the Digital Audio Effects (DAFx01) Conference pp 170–174
[9] Klapuri A and Davy M 2007 Signal processing methods for music transcription (Springer Science & Business Media)
[10] Schroeder M R 1968 Period Histogram and Product Spectrum: New Methods for Fundamental-Frequency Measurement J. Acoust. Soc. Am. 43 829–34