Journal of Engineering Science and Technology
Vol. 19, No. 2 (2024) 684 - 704

QUANTITATIVE ASSESSMENT OF SOUND CHARACTERISTICS USING THE AFFINITY (A), MEAN AFFINITY (MA), BRIGHTNESS (S), MEAN CONTRAST (MC), HARMONICITY (H) AND MONOTONY (M) OF GAMELAN TIMBRE

SININ HAMDAN[1,*], AALIYAWANI EZZERIN SININ[2], KHAIRUL ANWAR MOHAMAD SAID[1], JOHARI ABDULLAH[3], AHMAD FAUDZI MUSIB[4], HARINI SOSIATI[5]

[1]Faculty of Engineering, Universiti Malaysia Sarawak, 94300, Kota Samarahan, Sarawak, Malaysia
[2]Department of Science and Technology, Faculty of Humanities, Management and Science Universiti Putra Malaysia Bintulu Campus, 97008 Bintulu, Sarawak, Malaysia.
[3]Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, 94300, Kota Samarahan, Sarawak, Malaysia
[4]Faculty of Human Ecology, Universiti Putra Malaysia, 43400, Serdang, Selangor Darul Ehsan, Malaysia
[5]Department of Mechanical Engineering, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Yogyakarta 55183, Indonesia
*Corresponding Author: hsinin@unimas.my

Abstract

This study investigates the sound characteristics of Peking Gamelan. The characteristics of Peking Gamelan sound is defined by focusing on acoustics related to the description, composition, and distribution of harmonic and secondary frequencies (high-pitched tones) in a sound. These descriptors are calculated from the Fast FourierTransform (FFT) spectra using the Python Programming Language. Frequency spectrum characteristics are obtained for each sound from the Peking Gamelan using PicoScope oscilloscopes to investigate the fundamental and overtone frequencies. With this FFT spectrum, we can determine the amplitude a_i for each frequency f_i. The FFT spectrum is essentially a discrete collection of N frequencies (f_i) and N amplitudes (a_i). This paper suggests sound characteristic descriptors that allow for the extraction of sound features from FFT. Sound characteristics are related to the fundamental frequency f_o and secondary frequency f_i (which exist in FFT). The centroid $\bar{f}$ indicates the presence of frequencies other than f_o with amplitude distribution in such a way that $\bar{f}$ has a magnitude greater than f_o. $\bar{f}$ does not represent the harmonic nature of sound because it is not correlated with the natural scale series of any musical instrument. The acoustic descriptor used in this work is Affinity (A), Mean Affinity (MA), Brightness (S) Or Sharpness, Mean Contrast (MC), Harmonicity (H) And Monotony (M). Peking 1 has the highest affinity (minimum A value). Peking 5 has the brightest sound (maximum S value) and the most harmonic sound (minimum H value). Peking 6 exhibits the greatest increase in amplitude with the largest increase in the M value. Peking 1 shows the greatest decrease in amplitude with the largest decrease in the M value. The MA value of Peking 5 indicates dense secondary sound close to $\bar{f}$ and has the maximum MC value because its secondary frequencies are very small. The coefficients used can discriminate the differences and similarities in sound characteristics. Mean Affinity MA or Mean Contrast MC allows musical instruments to be uniquely identified. The coefficients MA, MC, H, and M are sufficient to describe the distribution of harmonics in the FFT.

Keywords: , Dimensionless coefficients, Fast Fourier transform (FFT), FFT-based sound characterization, Peking Gamelan analysis.

1. Introduction

Sound characteristics are common attributes that discriminate between different sounds (even if they have the same frequency, intensity, and duration) [1]. Sound characteristics in musical instruments are also referred to as timbre. Quantifying sound characteristics as measurable magnitudes remains an open research topic [2, 3].

Two complementary approaches to analysing sound characteristics are (i) aspects related to the psycho-physical perception of sound by listeners who can differentiate and recognize sound sources and (ii) focusing on acoustics related to the description, composition, and distribution of harmonic and secondary frequencies (high-pitched tones) in a sound. This paper uses the second approach i.e. focusing on the acoustic related to the distribution of harmonic and secondary frequencies in Peking gamelan.

This following the original idea of physicist Georg Simon Ohm (1789-1854), who stated that differences in sound characteristics arise from the presence of harmonics and their relative amplitudes [4]. The presence of harmonics and their relative amplitudes help characterise the sound from Peking gamelan. Advances in digital technology for sound recording and reproduction involve collecting amplitude and frequency information from spectral decomposition using FFT. All variations in FFT for digital sound are free from environmental and stimulus response variations received by the listener. An important question in this research is how to measure or express the variations in sound characteristics of musical instruments and the intensity levels of Fourier spectra. The variations in sound characteristics of musical instruments and the intensity levels of Fourier spectra yield the distribution of harmonic and secondary frequencies in Peking gamelan.

What elements of sound characteristic variations can be extracted from FFT-based analysis of audio recordings? To address this question, sound characteristics will be analysed from audio recordings of a set of Peking Gamelan instruments, evaluating changes in the dimensionless coefficients of sound characteristics proposed by Gonzales and Prati [5, 6]. Peking Gamelan is chosen as the subject of study because the sound characteristics is very smooth and can be measure or express in intensity levels of Fourier spectra.

Six functions are required to describe the monophonic musical sound characteristics in terms of normalized amplitude and frequency. The measurement of f_o against the average frequency $\bar{f}$ is called Affinity (A). The measurement of frequency dispersion against $\bar{f}$ is called Mean Affinity (MA). The measurement of amplitude a_o against the collection of amplitudes ai is called Brightness (S) or Sharpness. The measurement of the average amplitude of the pulse collection is called Mean Contrast (MC). The descriptor indicating the approximation of secondary pulses to the integer multiples of f_o is called Harmonicity (H). The descriptor for envelope through the average envelope in the pulse collection is called Monotony (M). These A, MA, S, MC, H and M will differentiate the sound characteristics of each Peking gamelan.

The study intends to close a gap in the literature by concentrating on Peking Gamelan, a traditional music genre. Although gamelan's culture and instruments have been studied, we have yet to look at how to tune these instruments precisely. This research examines the harmonic features of the Peking set, employing novel techniques to identify and interpret the distinct tones. In contrast to previous research, which frequently adopted a social science methodology, this study integrates scientific and cultural viewpoints. Because of the metals utilized, traditional Gamelan instruments produce unique tones. However, the explanations for these sounds are still unknown.

Using technology and cultural customs, the study presents a novel method for identifying tunings in Peking gamelan sets. By eliminating the constraints of conventional tuning, this innovative approach seeks to comprehend sound qualities fully. It can be challenging to replicate gamelan music on Western instruments like the piano because of its distinctive tunings. The study debunks myths about Peking gamelan tunings in Western music and provides a fresh approach to understanding and retrieving them. We gain important insights regarding the differences across instruments by concentrating on the Peking set. As a result, this study encourages conversation and increases understanding of the significance of creating instruments to recognize sound features in Peking gamelan sets. For example, it offers scholars, cultural enthusiasts, and musician valuable insights. The study also suggests that artificial intelligence can be utilized to preserve and enhance Peking gamelan, which would significantly advance traditional music fusion with contemporary technology.

The formation of a finite set of frequencies requires a fundamental frequency (f_o), and others may be concordant (harmonic frequencies) or non-concordant (secondary frequencies fi or high-pitched frequencies). The number of harmonics and f_i, as well as the relative intensity of the sound, determine the specificity of sound characteristics that characterize each musical instrument [7, 8].

This paper proposes an acoustically motivated sound characteristic descriptor. This descriptor allows for computer extraction of sound characteristic information using FFT from music recordings. Dimensionless coefficients from sound characteristics indicate descriptors in the spectrum that characterize the evaluated Peking Gamelan.

Sound characteristic assessment enables us to create precise musical parameters, as well as identify, classify musical instruments, and assess the quality of sound recordings. The acoustic descriptor is motivated by the acoustics of music using spectra from FFT. Digitizing sound using computers through FFT represents a significant advancement in (i) methods for retrieving music information [9, 10], (ii) recognizing and identifying musical instruments [11, 12], (iii) characterizing audio music recordings.

To characterize musical sound, attributes of (i) frequency, (ii) intensity, (iii) duration, and (iv) sound characteristics are essential. The first three attributes are directly measurable quantities. The fourth attribute, sound characteristics, is multi-dimensional and is an attribute that allows one to distinguish between sounds that have the same frequency, intensity, and duration. Sound characteristics allow us to distinguish between sounds from different instruments even if they have the same musical notes (i.e., the same intensity and duration).

Specific musical sound characteristics with f_o in the musical scale are related to the attack, sustain, and decay due to the presence of harmonics [13, 14]. Sound is essentially composed of a set of waves characterized by magnitude, amplitude, and frequency. At any given moment, sound is determined by amplitude and frequency.

As a result, the Fourier spectrum for monophonic audio signals is essentially finite, a collection of pairs of numbers related to amplitude and frequency components. The discrete distribution of these number pairs can be described in terms of (i) maximum frequency and amplitude, (ii) f_i following f_o for the considered sound, and (iii) statistical measures of average amplitude and middle frequency.

Quantitative assessment of sound characteristics involves considering all three aspects: (i) fundamental frequency (f_o) and fundamental amplitude (a_o), (ii) measures of f_i or harmonic frequencies, (iii) relative intensity (measured relative to a_o), and (iv) the presence of frequency sets (rhythmic, monotonic, harmonic series, and so on). Two magnitudes (amplitude and frequency) are used for these three aspects, i.e., musical sound, value distribution shape, and minimum value for secondary sound components.

The Affinity (A) measures f_o against the average frequency $\bar{f}$ Mean Affinity (MA) measurement of frequency dispersion against $\bar{f}$, measurement of amplitude a_o against the collection of amplitudes ai is called Brightness (S) or Sharpness, measurement of the average amplitude of the pulse collection is called Mean Contrast (MC). descriptor indicating the approximation of secondary pulses to the integer multiples of f_o is called Harmonicity (H), descriptor for envelope through the average envelope in the pulse collection is called Monotony (M).

Musical frequencies form a finite and countable discrete set, comprising 12 distinct values (C, C#, D, D#, E, F, F#, G, G#, A, A#, and B) in each musical octave (for a total of 96 f_o in 8 octaves audible from 20Hz to 20KHz). Sound characteristics can be characterized by a limited set of dimensionless coefficients, which are quantities related to frequency and amplitude in the Fourier spectra of audio recordings. Motivated by the acoustics of music, these dimensionless coefficients serve as tone descriptors and can be explained by functions with discrete distributions of normalized frequencies and amplitudes. When the amplitudes from FFT spectra are normalized (using the ratio of amplitude a_i for each f_i to the maximum amplitude in each spectrum), we can compare normalized amplitudes. These can be aggregated into descriptors of f_o (from a musical scale consisting of 96 f_o values). The FFT values are essentially a discrete collection of pairs of different amplitudes and frequencies and can be summarized with dimensionless parameters. Six functions are required to describe the monophonic musical sound characteristics in terms of normalized amplitude and frequency as follows:

i. The measurement of f_o against the average frequency $\bar{f}$ is called Affinity (A).

ii. The measurement of frequency dispersion against $\bar{f}$ is called Mean Affinity (MA).

iii. The measurement of amplitude a_o against the collection of amplitudes ai is called Brightness (S) or Sharpness.

iv. The measurement of the average amplitude of the pulse collection is called Mean Contrast (MC).

v. The descriptor indicating the approximation of secondary pulses to the integer multiples of f_o is called Harmonicity (H).

vi. The descriptor for envelope through the average envelope in the pulse collection is called Monotony (M).

Table 1 provides equations for sound characteristics resulting from 3 aspects of 2 variables in FFT, namely amplitude and frequency [6].

Table 1. Sound characteristics based on 2 variables in FFT, namely amplitude and frequency.

Description	Formula
Centroid, $\bar{f}$	$\bar{f} \equiv \frac{\sum_{i=1}^{N} a_i f_i}{\sum_{i=1}^{N} a_i}$
Affinity, A	$A = \frac{1 \sum_{i=1}^{N} a_i f_i}{f_0 \sum_{i=1}^{N} a_i} = \frac{f}{f_0}$
Sharpness, S	$S = \frac{a_0}{\sum_{i=1}^{N} a_i}$
Harmonicity, H	$H = \sum_{j=1}^{N} \left(\frac{f_j}{f_0} - \left\lvert \frac{f_j}{f_0} \right\rvert \right)$
Monotony, M	$M = \frac{f_0}{N} \sum_{j=1}^{N} \left(\frac{a_{j+1} - a_j}{f_{j+1} - f_j} \right)$
Mean affinity, MA	$MA = \frac{\sum_{i=1}^{N} \lvert f_i - \bar{f} \rvert}{N f_0}$
Mean contrast, MC	$MC = \frac{1}{N} \sum_{j=1}^{N} \lvert a_0 - a_j \rvert$

1.1. Fundamental frequency descriptor

In acoustic signal analysis, centroid $\bar{f}$ is commonly used to describe sound characteristics. Normalized amplitude ai is obtained by dividing the value of amplitude at f_i by the fundamental frequency amplitude a_o. $\bar{f}$ indicates the presence of frequencies other than f_o with amplitude distribution in such a way that $\bar{f}$ has a magnitude greater than f_o. $\bar{f}$ does not represent the harmonic nature of sound because it is not correlated with the natural scale series of any musical instrument. f_o only coincides with $\bar{f}$ from the distribution if harmonic and secondary frequencies do not exist. This is impossible because $\bar{f}$ for each resonator in any musical instrument is subject to some distortion, superposition, and patterns generated by the geometry of the musical instrument.

The distortion between f_o and $\bar{f}$ is significant from the perspective of music acoustics. The quantity assessed through dimensionless coefficients is called Affinity A, which explains the spectrum distance, i.e., how far f_o is from $\bar{f}$. If f_o and $\bar{f}$ are close, then the sound is considered more affinity with $\bar{f}$, and A has a value close to 1. The variation of amplitude a_o against amplitude a_i in the spectrum is assessed using the Sharpness coefficient (S). S is a measure of amplitude or relative intensity. Relative intensity is perceived from a_o amplitudes in the a_i distribution. The larger S for musical sounds, the easier it is to interpret. In the ideal case, S=1, meaning pure sound has an outstanding maximum without any f_i. In actual conditions, S is always less than 1.

1.2. Descriptor for frequency distribution

To characterize the distribution of f_i with respect to f_o in the FFT spectrum, statistical descriptors of f_i are used to assess kurtosis (the peak sharpness of the frequency distribution curve), uniformity (if it exists), and statistical parameters. In a spectrum, the relationship between f_i (whether harmonic or not) and f_o always exists. If f_i approaches integer multiples of f_o, the sound is considered more harmonic compared to f_i that is very different from f_o. To describe this property, the Harmonicity coefficient (H) is proposed. The H function assesses the extent to which f_1, f_2, f_3 …f_j are harmonics. Any f_j is a harmonic of f_o if the ratio between them is an integer. If all f_j are harmonics of f_o, then H=0. Whenever there is one or more f_j that are not harmonics, H is non-zero and always increases.

Another remarkable aspect of the maximum distribution in FFT is the variability of amplitude with respect to frequency. Sound characteristic coefficients called monotony (M) assess the uniformity of the f_i distribution (whether harmonic or not). After f_o, the next and subsequent maxima may have increasing or decreasing amplitudes, meaning they may increase in amplitude (increased monotony) or decrease in amplitude (decreased monotony). Monotony M indicates whether harmonics appear in consecutive increments (positive M), or harmonics appear in consecutive decrements (negative M) after f_o. Monotony M indicates whether harmonics are present in most consecutive increments (positive M) or in most consecutive decrements (negative M) after f_o.

1.3. Statistical distribution

Another aspect perceived in FFT is f_i, which covers different frequency ranges, sometimes very close (or clustered), while other times they cover widely separated (or scattered) frequency ranges in a larger space. When f_i is very close to each other (or close to f_o), the sound appears denser (or thicker). Conversely, spacing in the frequency domain provides clarity (or transparency) to f_o. This descriptor measures the frequency distribution and assesses the spacing of f_i in relation to $\bar{f}$. Clarity or transparency is a sound characteristic coefficient defined by Mean Affinity (MA). The MA coefficient assesses the density of frequency distribution relative to the minimum value (including f_o). A low MA refers to a dense distribution of secondary frequencies that are close to $\bar{f}$. Mean Contrast (MC) is a coefficient that measures the amplitude of f_j, namely a_j relative to the amplitude a_o.

This paper introduces a valid characteristic descriptor that uses Fast Fourier Transform (FFT) analysis. The goal is to capture and measure the distinct sound features of Peking Gamelan. Within the suggested description, the study aims to pinpoint and quantify sound properties, such as frequency, intensity, duration, and spectral features. Additionally, the study investigates how sound digitization via FFT advances the fields of music information retrieval, musical instrument recognition, and audio recording characterization within the Peking Gamelan set.

The research aims to confirm if dimensionless coefficients obtained from acoustic properties serve as descriptors for the assessed Peking Gamelan spectrum. It also evaluates the descriptor's ability to categorize instruments, set exact musical settings, and assess the quality of sound recordings. To statistically evaluate critical characteristics such as relative intensity, fundamental frequency, harmonic frequencies, and the existence of frequency sets, the study looks into several research issues. The goal is to differentiate between sounds made by various instruments of the Gamelan. The approach seeks insight into Peking Gamelan sounds' attack, sustain, and decay characteristics, especially regarding harmonics and other distinctive melodic sound elements. In general, these research questions direct the investigation toward revealing the potential and uses of the suggested acoustic descriptor in the intricate examination of Peking Gamelan.

The paper explores recent advances in digital technology for sound recording and reproduction, emphasizing spectrum decomposition using FFT to obtain amplitude and frequency information. This method eliminates the listener's perception of ambient and stimulus-response fluctuations from FFT variations for digital sound. Determining ways to quantify or communicate changes in the Fourier spectra's intensity levels and the sound properties of musical instruments is a major issue driving this research. The primary concern is which components of fluctuations in sound characteristics can be retrieved by means of FFT-based analysis of recorded audio. This investigation of digital technology and FFT provides an essential basis for comprehending and measuring the subtleties in sound properties relevant to the study.

2. Methodology

Figure 1 displays a set of 6 Peking gamelan used in this work. Figure 2 illustrates the scheme for audio recordings of the Peking Gamelan [15]. The cast bronze Peking 1, 2, 3, 5, 6 and 1' was chosen from a range of Malay gamelan ensemble. The acoustic spectra of the measured sets of just-tuned cast bronze Peking which were made in Indonesia was captured using PicoScope oscilloscopes to investigate the fundamental and overtone frequencies. Excitation was done by beating the Peking with a mallet by an expert Peking player. The microphone was held above the top surface along the axis of symmetry of the Peking at a distance of about 20 cm.

The frequency reading was verified by recording a sound of 1KHz from a signal generator. The microphone is a flat response microphone capable of capturing only 20Hz-20kHz. The setup to capture the sound is based on Owsinski (2009). The arrangement of microphone and apparatus for the measurement are shown in Fig. 2. The microphone was placed right above the bar. The PicoScope computer software (Pico Technology, 3000 series, Eaton Socon, UK) was used to view and analyse the time signals from PicoScope oscilloscopes (Pico Technology, 3000 series, Eaton Socon, UK) and data loggers for real time signal acquisition. PicoScope software enables analysis using FFT, a spectrum analyser, voltage-based triggers, and the ability to save/load waveforms to a disk.

The Peking was placed to where the sound could be captured with minimum interference. In our work the recording was done in the University Malaysia Sarawak (UNIMAS) Faculty of Applied and Creative Art (FACA) Music department studio with full acoustic room facility. The signal produced from PicoScope displayed sharp and distinct fundamental and overtone frequencies peak compared to the low signal level from the background noise. The amplifier (Behringer Powerplay Pro XL, Behringer, China) ensured the sound capture was loud enough to be detected by the signal converter.

In conducting this study, the audio signal derived from the striking of the Peking played by an expert Peking player was recorded. The audio signal was recorded in mono, at 24-bit resolution, 48 kHz sampling rate. The audio signal was recorded with the aid of a digital audio interface in a .wav format. To ensure the recorded audio signal of the Peking was at the optimum level, audio signal calibration of the recording system was carried out.

A test tone of 1 kHz sine wave was used in calibrating the recording system. Here the 'unity' calibration level was at +4dBu or -10dBV and was read by the recording device at '0 VU'. In this regard the EBU recommended the digital equivalent of 0VU is that the test tone generated to the recording device of the experimentation is recorded at -18 dBFS (Digital) or +4dBu (Analog) which is equivalent to 0VU. In this thorough procedure of calibration, no devices are unknowingly boosting or attenuating its amplitude in the signal chain at the time of the recording is carried out. The recording apparatus was the Steinberg UR22 mkII audio interface, Audio-Technica AT4050 microphone, XLR cable (balance), with microphone position on axis (< 20 cm), microphone setting with low cut (flat) 0dB.

Fig. 1. Peking gamelan 1, 2, 3, 5, 6, and 1' (from left to right).

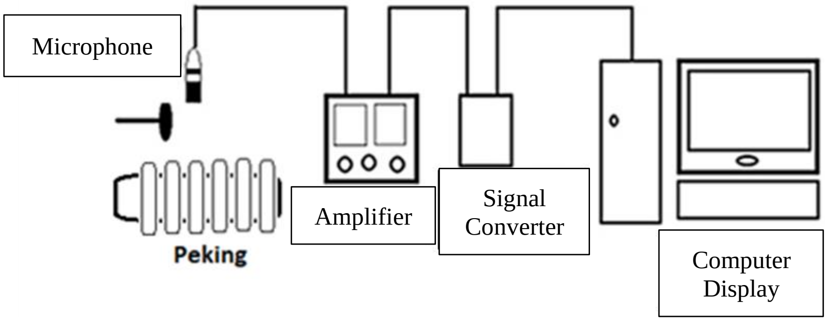


Fig. 2. Diagram scheme for audio recording of Peking Gamelan.

Picoscope software display the real signal which is voltage versus time and dBu versus frequency as shown in Fig. 3. Frequency spectrum characteristics are obtained for each sound from the Peking Gamelan under study. With this FFT spectrum, we can determine the amplitude a_i for each f_i. These descriptors are calculated from the Fast Fourier Transform (FFT) spectra using the Python Programming Language. Through calculation, a frequency table (in kHz) along with normalized amplitudes (relative to the maximum amplitude) can be formed. The FFT spectrum is essentially a discrete collection of N frequencies (f_i) and N amplitudes (a_i).

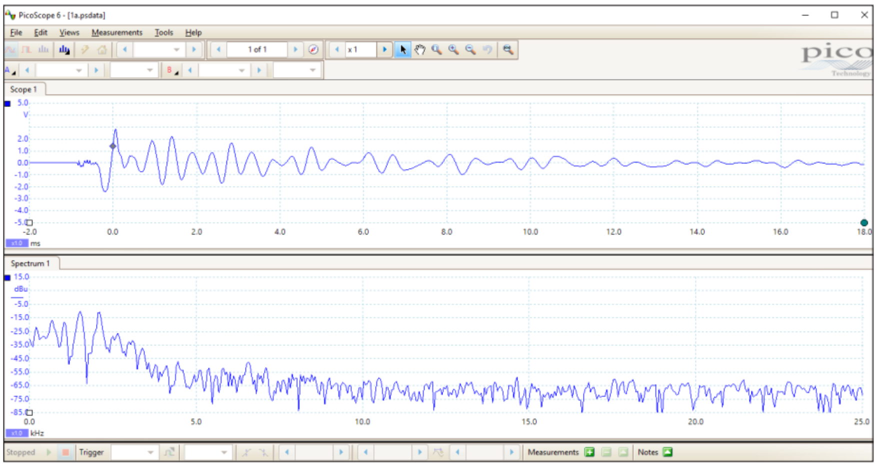
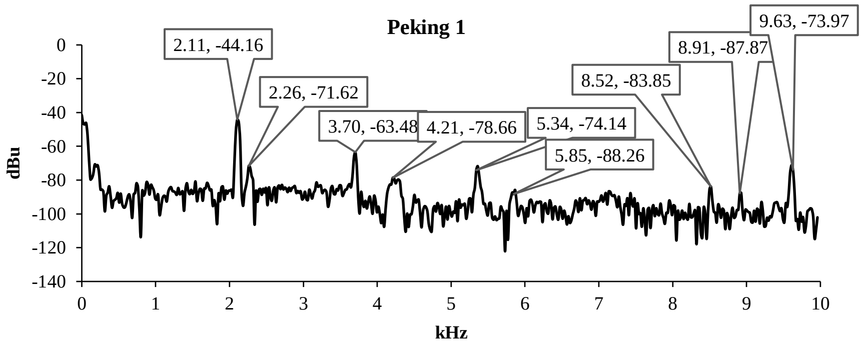
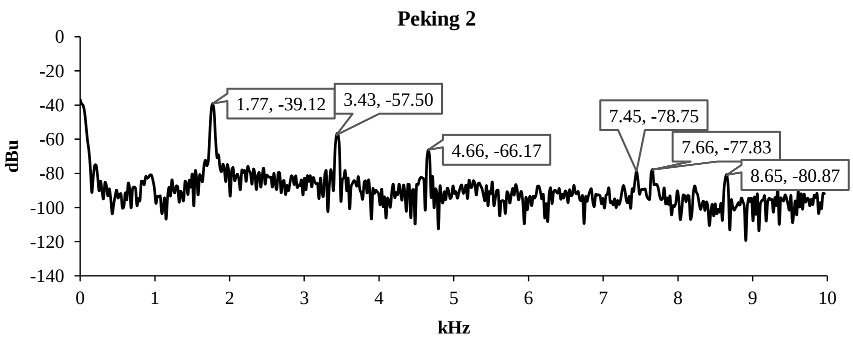


Fig. 3. A typical signal which is voltage versus time and dBu vs. frequency.

3. Results and Discussion

Figure 4 displays the FFT with frequency f_i (x-axis in kHz) against amplitude (y-axis in dBu) recorded using Picoscope. In Fig. 4, the frequency and intensity of the fundamental and overtone peaks from the Peking signal is clearly shown by the distinct peak. The fundamental frequencies of Peking 1, 2, 3, 5, 6 were 1066Hz (C6), 1178Hz (D6), 1342Hz (E6), 1599Hz (G6) and 1793Hz (A6) respectively while Peking 1'was 2123Hz (C7), i.e., one octave higher than Peking 1 [15].





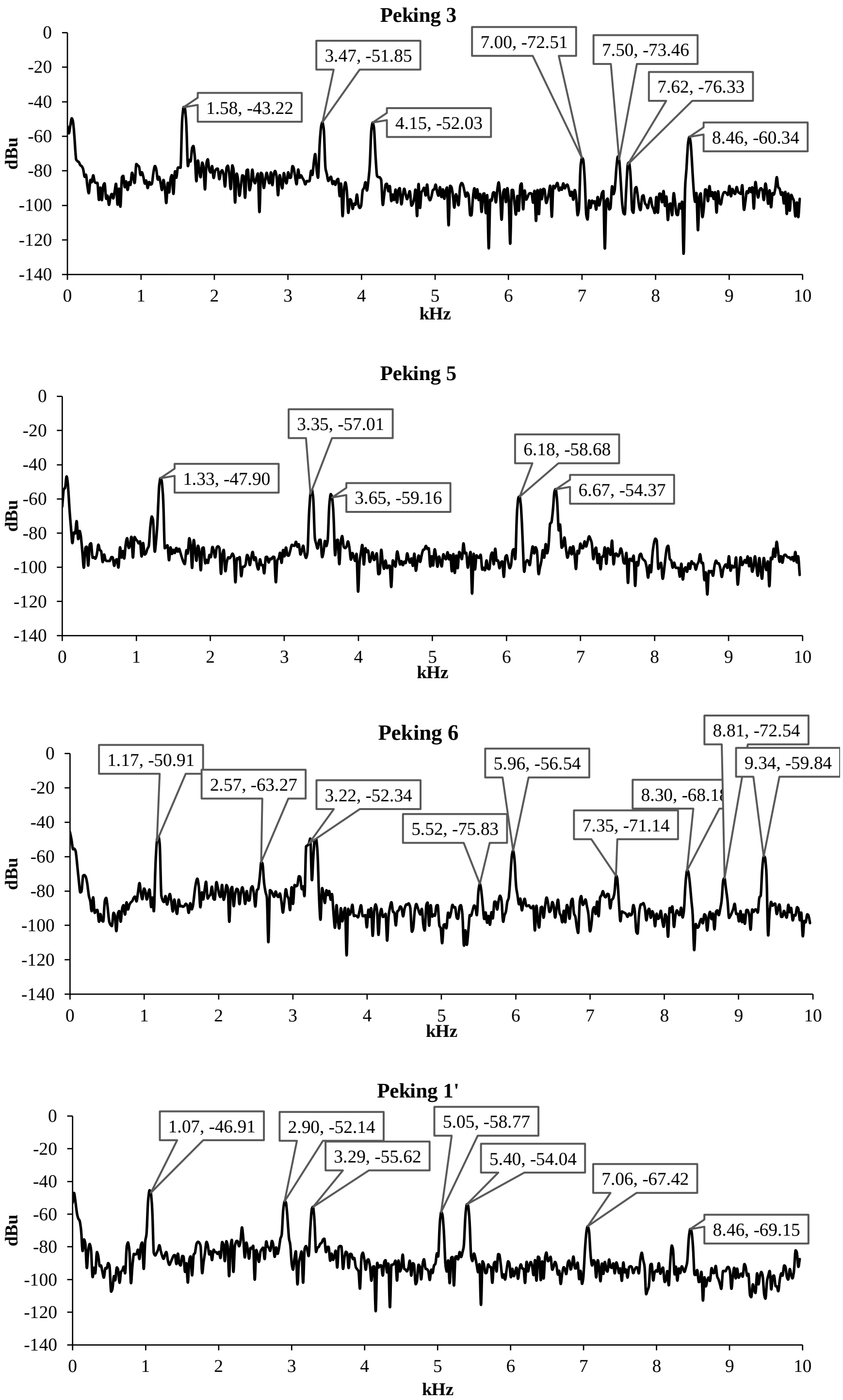


Fig. 4. The spectra of Peking 1, 2, 3, 5, 6, and 1'.

Table 2 displays the frequency f_i (in kHz) along with amplitude ai (dBu) recorded from Fig. 5. In Table 2 the variation of the amplitude a_i is shown together with the corresponding f_i (kHz)**.** Table 3 presents the frequency f_i (kHz) along with normalized amplitudes a_i using data from Table 2. In Fig. 5 the variation of the normalised amplitude is plotted against the frequency obtained from Table 3. In Table 3 the amplitude a_i from Table 2 is divided by the amplitude of the fundamental peak to produced normalized amplitudes a_i. Table 4 presents 6 descriptors calculated using data from Table 3. In Table 4 the 6 descriptors for Peking gamelan was calculated using the equations for sound characteristics resulting from 3 aspects of 2 variables in FFT, namely amplitude and frequency provides from Table 1 [5].

Calculation results are included in the appendix. The data obtained using Python Programming Language in the appendix was double check manually using excel spread sheet. The fundamental frequencies f_o for Peking 1, 2, 3, 5, 6, and 1' are 2.11, 1.77, 1.58, 1.33, 1.17, and 1.07 kHz, respectively, while the centroids for Peking 1, 2, 3, 5, 6, and 1' are 4.74, 4.06, 4.82, 3.27, 4.88, and 3.39 kHz, respectively.

Table 2. The frequency f_i (kHz) along with amplitude a_i (dBu) from Fig. 1.

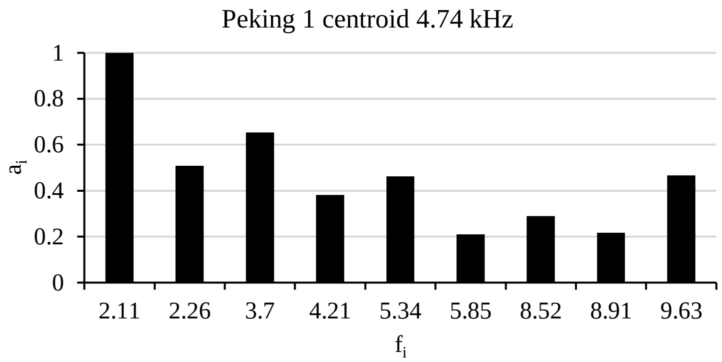
Peking 1		Peking2		Peking 3		Peking 5		Peking 6		Peking 1'	
f_i	a_i	f_i	a_i	f_i	a_i	f_i	a_i	f_i	a_i	f_i	a_i
2.11	-44.16	1.77	-39.12	1.58	-43.22	1.33	-47.9	1.17	-50.91	1.07	-46.91
2.26	-71.62	3.43	-57.5	3.47	-51.85	3.35	-57.01	2.57	-63.27	2.9	-52.14
3.7	-63.48	4.66	-66.17	4.15	-52.03	3.65	-59.16	3.22	-52.34	3.29	-55.62
4.21	-78.66	7.45	-78.75	7	-72.51	6.18	-58.68	5.52	-75.83	5.05	-58.77
5.34	-74.14	7.66	-77.83	7.5	-73.46	6.67	-54.37	5.96	-56.54	5.4	-54.04
5.85	-88.26	8.65	-80.87	7.62	-76.33			7.35	-71.14	7.06	-67.42
8.52	-83.85			8.46	-60.34			8.3	-68.18	8.46	-69.15
8.91	-87.87							8.81	-72.54		
9.63	-73.97							9.34	-59.84		

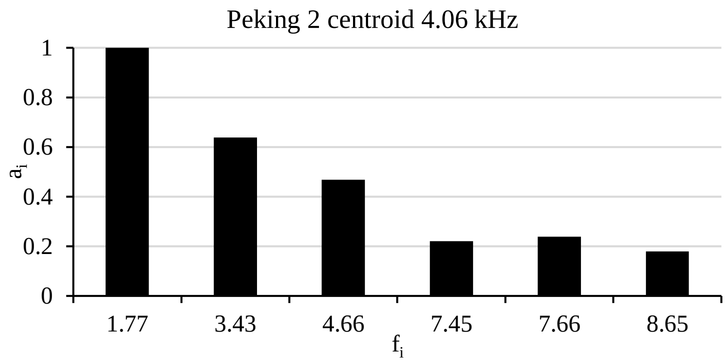
Table 3. The frequency f_i (kHz) along with normalized amplitudes a_i from Table 2.

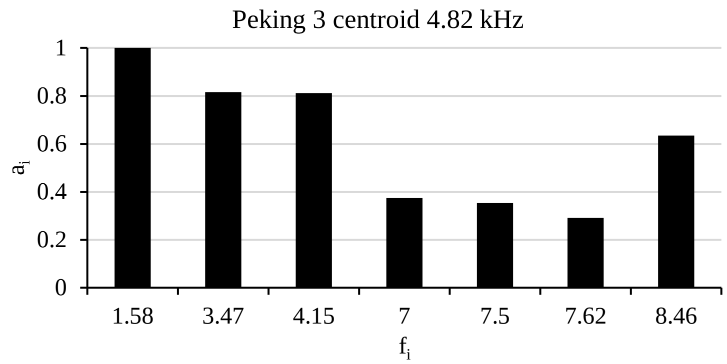
Peking 1		Peking2		Peking 3		Peking 5		Peking 6		Peking 1’	
f_i	a_i	f_i	a_i	f_i	a_i	f_i	a_i	f_i	a_i	f_i	a_i
2.11	1	1.77	1	1.58	1	1.33	1	1.17	1	1.07	1
2.26	0.508	3.43	0.63	3.47	0.81	3.35	0.247	2.57	0.575	2.9	0.773
3.7	0.654	4.66	0.468	4.15	0.811	3.65	0.069	3.22	0.970	3.29	0.622
4.21	0.382	7.45	0.221	7	0.373	6.18	0.109	5.52	0.143	5.05	0.486
5.34	0.463	7.66	0.239	7.5	0.353	6.67	0.465	5.96	0.806	5.4	0.691
5.85	0.210	8.65	0.179	7.62	0.292			7.35	0.304	7.06	0.111
8.52	0.289			8.46	0.634			8.3	0.406	8.46	0.036
8.91	0.217							8.81	0.256		
9.63	0.46[illegible]							9.34	0.692		

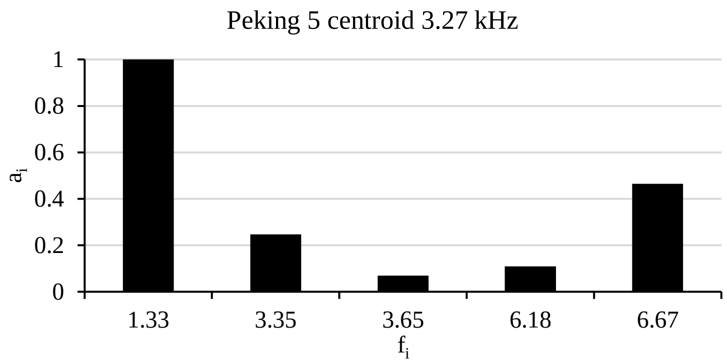
Table 4. The 6 descriptors for Peking Gamelan.

Description	Peking 1	Peking 2	Peking 3	Peking 5	Peking 6	Peking 1’
Centroid	4.7439	4.0683	4.8209	3.2729	4.8828	3.3986
Affinity	2.2483	2.2984	3.0512	2.4608	4.1734	4.1734
Mean contrast	0.6012	0.6506	0.4531	0.7772	0.3967	0.5462
Mean affinity	2.3584	2.5138	2.3655	1.7402	2.6301	2.1873
Monotony	-0.2005	-0.0891	-0.0961	0.0666	0.3084	-0.0601
Sharpness	0.23864	0.3640	0.2335	0.5288	0.1824	0.2686
Harmonicity	3.9478	2.9943	3.1772	1.9248	3.6495	3.0560


Peking 1 centroid 4.74 kHz
a_i
1
0.8
0.6
0.4
0.2
0
2.11 2.26 3.7 4.21 5.34 5.85 8.52 8.91 9.63
f_i


Peking 2 centroid 4.06 kHz
a_i
1
0.8
0.6
0.4
0.2
0
1.77 3.43 4.66 7.45 7.66 8.65
f_i


Peking 3 centroid 4.82 kHz
a_i
1
0.8
0.6
0.4
0.2
0
1.58 3.47 4.15 7 7.5 7.62 8.46
f_i


Peking 5 centroid 3.27 kHz
a_i
1
0.8
0.6
0.4
0.2
0
1.33 3.35 3.65 6.18 6.67
f_i

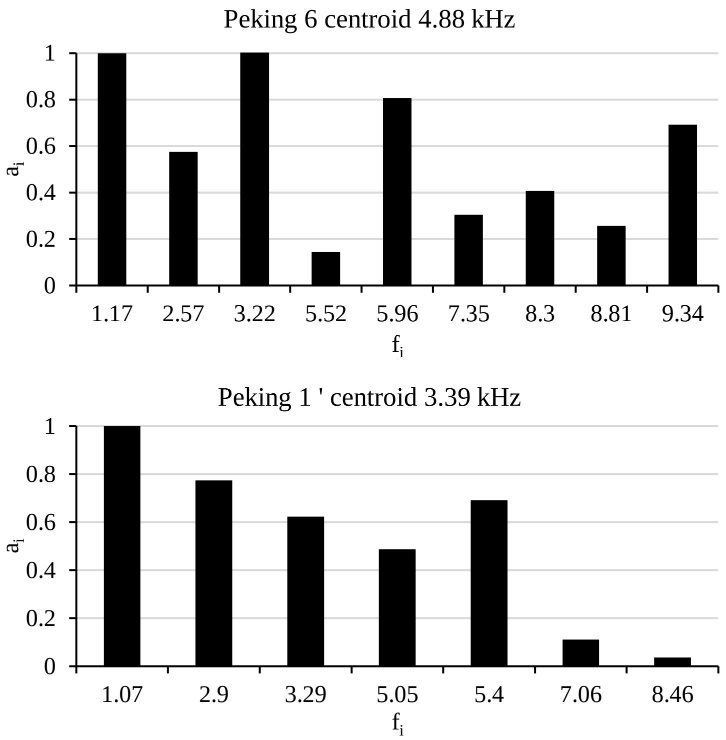


Fig. 5. Frequency distribution (f_i) against normalized amplitude (a_i) with the centroids.

Figure 6 represents the affinity values for Peking 1, 2, 3, 5, 6, and 1'. Affinity A explains the spectrum distance from the ideal case (i.e., how far f_o is from $\bar{f}$). If f_o and $\bar{f}$ are close, then the sound is considered more affinity with $\bar{f}$ where A has a value approaching 1. Peking 1 is more affinity with the centroid $\bar{f}$ (minimum A value). From Fig. 4 (The spectra of Peking 1, 2, 3, 5, 6, and 1') and Figure 5 (Frequency distribution (f_i) against normalized amplitude (a_i)) one can hardly distinguished which Peking is more affinity with the centroid $\bar{f}$. While the spectra of Peking 1, 2, 3, 5, 6, and 1' only displayed the distribution of their intensity with their respective frequency, Fig. 6 highlight that Peking 6 and Peking 1' have less affinity with $\bar{f}$ where A is much higher than 1.

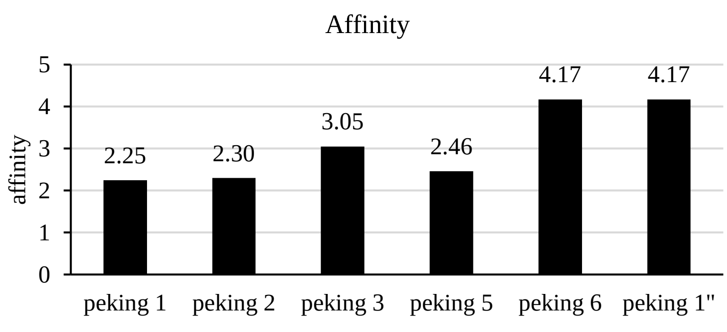


Fig. 6. The distribution of affinity values for Peking 1, 2, 3, 5, 6, and 1'.

Figure 7 represents the brightness values for Peking 1, 2, 3, 5, 6, and 1'. The variation in amplitude a_o relative to the amplitude a_i in the spectrum is evaluated using Sharpness (Brightness) S. S is a measure of relative amplitude perceived from a_o (a_o in the distribution of a_i). S=1 represents pure sound with a dominant maximum and no f_j. In actual conditions, S is always less than 1. Peking 5 has the brightest sound (maximum S value). Mother nature cannot determine the brightness quantitatively. Using the S value, it is proven that Peking 5 have the brightest sound.

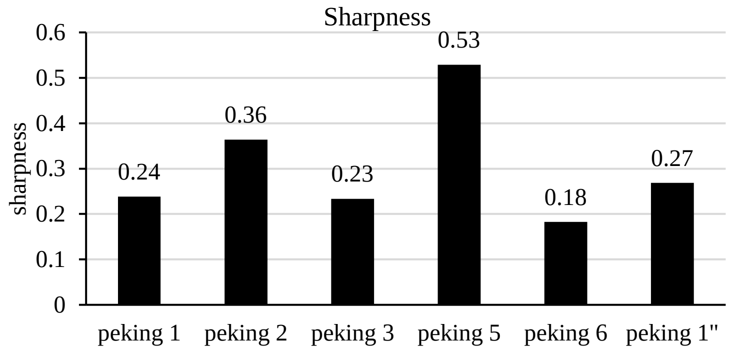


Fig. 7. The distribution of sharpness values (brightness) S for Peking 1, 2, 3, 5, 6, and 1'.

Figure 8 represents the distribution of harmonicity values for Peking 1, 2, 3, 5, 6, and 1'. Sound is considered more harmonic if f_j approaches integer multiples of f_o, and less harmonic if f_j is significantly different from f_o. If all f_j are harmonics of f_o, then H=0. Each time there is one or more f_j that are not harmonics of f_o, H always increases. Peking 5 is the most harmonic (minimum H value). For a harmonic sound, all the f_j is a multiple integer of f_o, which eventually yield H=0. Although Peking 5 is not totally harmonic (where H=1.92), but the minimum values among other Peking proved that Peking 5 is the most harmonic.

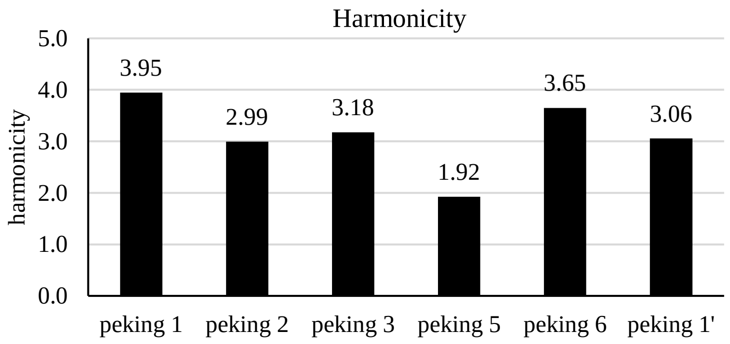


Fig. 8. The distribution of harmonicity values H for Peking 1, 2, 3, 5, 6, and 1'.

Figure 9 represents the distribution of monotony values for Peking 1, 2, 3, 5, 6, and 1'. Monotony M determines the variability of amplitude with frequency, assessing the uniformity of the distribution of f_j. After f_o, the next maxima, and subsequent ones, can either increase or decrease in amplitude, i.e., their amplitude can increase (increasing monotony) or decrease (decreasing monotony). Peking 6 exhibits the greatest increase in amplitude with the largest increase in M value. Peking 1 exhibits the greatest decrease in amplitude with the largest decrease in M value. M indicates whether harmonics are present in most increases (positive M) or most decreases (negative M) consecutively after f_o. The listener's experience a dying sound in Peking 1 compared to Peking 6 since the largest decrease in amplitude occur in Peking 1 as shown in Fig. 5 (Frequency distribution (f_i) against normalized amplitude (a_i) with the centroids). This listener' experience correlate with the monotony M values in Fig. 9.

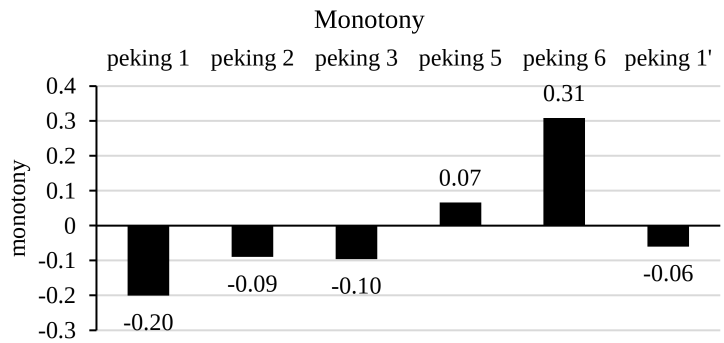


Fig. 9. The distribution of monotony values M for Peking 1, 2, 3, 5, 6, and 1'.

Figure 10 represents the distribution of mean affinity values for Peking 1, 2, 3, 5, 6, and 1'. Clarity or transparency is expressed by Mean Affinity MA. MA assesses the density of frequency distribution relative to the minimum value (including f_o). A low MA value indicates a dense distribution of secondary sound close to $\bar{f}$. The MA value of Peking 5 indicates a dense distribution of secondary sound close to $\bar{f}$. From Fig. 4 and Table 2, Peking 5 had only 3 partials. So, the $\sum_{i=1}^{N}|f_i - \bar{f}|$ is minimum and less overtone sound is heard.

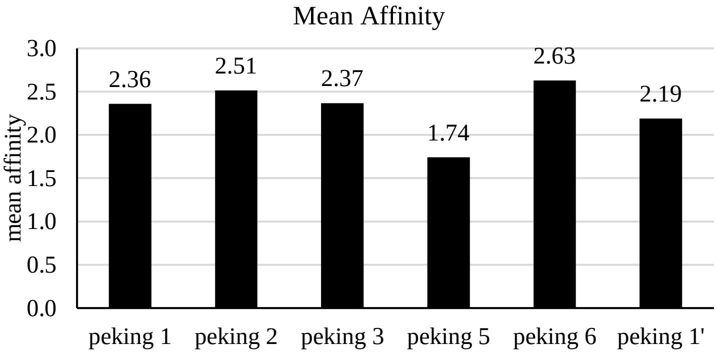


Fig. 10. The distribution of mean affinity values MA for Peking 1, 2, 3, 5, 6, and 1'.

Figure 11 shows the distribution of mean contrast values for Peking 1, 2, 3, 5, 6, and 1'. Mean Contrast (MC) is a coefficient that measures the amplitude of f_j, i.e., a_j relative to amplitude a_o. The MC value for Peking 5 is maximum because its secondary frequencies amplitudes are very small, as shown in Fig. 4 (The spectra of Peking 1, 2, 3, 5, 6, and 1'). A small secondary frequency amplitude yields a big MC. In this study Peking 5 produce overtone that are less heard by the listeners which are quantitatively shown by the MC values.

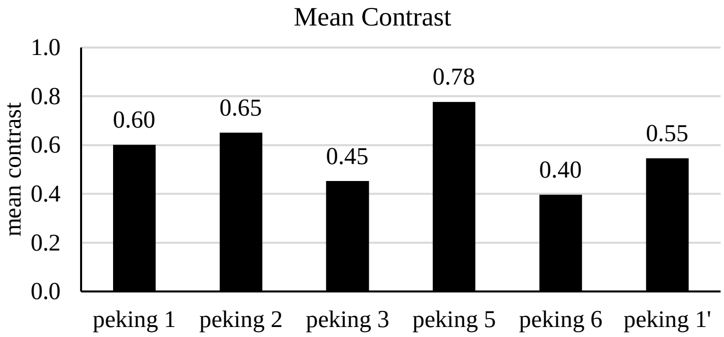


Fig. 11. The distribution of mean contrast values for Peking 1, 2, 3, 5, 6, and 1'.

4. Conclusion

Peking 1 has the highest affinity (minimum A value). Peking 5 has the brightest sound (maximum S value). Peking 5 is the most harmonic (minimum H value). Peking 6 exhibits the greatest increase in amplitude with the highest increase in M value. Peking 1 shows the greatest decrease in amplitude with the highest decrease in M value.

The MA value for Peking 5 indicates a dense secondary sound close to $\bar{f}$. The MC value for Peking 5 is maximum because its secondary frequencies are very low, as shown in Fig. 3. Mean Affinity MA or Mean Contrast MC allows musical instruments to be uniquely identified, as demonstrated by Peking 5. The number of existing harmonic frequencies in a musical instrument does not necessarily mean it has a high harmonicity value.

Peking 5 is the most harmonic (minimum H value) even though its secondary frequencies are not harmonics of the fundamental frequency. Monotony is clearly displayed by Peking 6 (maximum M) and Peking 1 (minimum M). The coefficients MA, MC, H, and M are sufficient to describe the distribution of harmonics in the FFT. The sound descriptors A and S allow us to differentiate between Peking 1 and Peking 5.

Peking 1 is more affinity with the centroid $\bar{f}$ (minimum A value). While the spectra only displayed the distribution of their intensity with their respective frequency, Peking 6 and Peking 1' have less affinity with $\bar{f}$ where A is much higher than 1. Mother nature cannot determine the brightness quantitatively. Using the S value, it is proven that Peking 5 have the brightest sound. For a harmonic sound, all the f_j is a multiple integer of f_o, which eventually yield H=0. Although Peking 5 is not totally harmonic (where H=1.92), but the minimum values among other Peking proved that Peking 5 is the most harmonic. The listener's experience a dying sound in Peking 1 compared to Peking 6 since the largest decrease in amplitude occur in Peking 1. This listener' experience correlate with the monotony M values. The MA value of Peking 5 indicates a dense distribution of secondary sound close to $\bar{f}$. Peking 5 had only 3 partials, so the $\sum_{i=1}^{N}\left|f_i - \bar{f}\right|$ is minimum and less overtone sound is heard. A small secondary frequency amplitude yields a big MC. In this study Peking 5 produce overtone that are less heard by the listeners which are quantitatively shown by the MC values.

In summary, the study's use of FFT and six Time-Frequency Descriptors (TFD) to describe the sound of monophonic music in audio recordings has been successful. Harmonicity H, Monotony M, Mean Affinity MA, and Mean Contrast MC provide essential details about sound properties, efficiently displaying the distribution of harmonics and high pitches. Based on the study's findings, there are several directions in which future research and practical applications could go, especially in the areas of thorough data collection and sound preservation. Building a solid database for 'deep learning' archives is one viable avenue. A thorough database can be established by utilizing the recognized sound descriptors, which will enable the creation of sophisticated algorithms to evaluate and identify different sound features. Additionally, the research recommends possible uses of museum technology, particularly concerning Peking Gamelan. Based on the discovered descriptions, immersive tools for experiencing Peking Gamelan playing can be created. Through these initiatives, the experience of visiting a museum is elevated, and cultural heritage is better appreciated and preserved.

Furthermore, investigating the integration of cutting-edge technology like virtual reality or augmented reality could result in a more engaging and dynamic experience. In summary, the research clarifies the properties of monophonic music sounds. It creates an array of opportunities for incorporating these discoveries into real-world uses, from cutting-edge tools for experiencing Peking Gamelan in a museum setting to deep learning repositories. These directions for further investigation and valuable suggestions can progress museum technology and sound preservation.

Acknowledgements

The authors are grateful to Universiti Malaysia Sarawak for providing the financial and technical assistant.

References

1. American National Standards Institute. (1973). Committee on bioacoustics, S3, Acoustical Society of America. American National Standard Psychoacoustical Terminology.
2. Antoine, A. (2018). *An investigation into the use of artificial intelligence techniques for the analysis and control of instrumental timbre and timbral combinations*. PhD Thesis, University of Plymouth.
3. Caetano, M.; Saitis, C.; and Siedenburg, K. (2019). *Audio content descriptors of timbre*. Springer.
4. Pastor-Garcia, V. (2005). *Estudio y análisis sobre la acústica y organología del Clarinete y su optimización*. PhD Thesis, Universidad Politécnica de Valencia.
5. Gonzalez, Y.; and Prati, R.C. (2022). Acoustic descriptors for characterization of musical timbre using the fast Fourier transform. *Electronics*, 11(9), 1405.

6. Gonzalez, Y.; and Prati, R.C. (2023). Comparative study of musical timbral variations: Crescendo and Vibrato using FFT-acoustic descriptor. *Eng-Advances in Engineering*, 4(3), 2468-2482.
7. Rossing, T.D. (1990). *The Science of sound*. Addison-Wesley.
8. Sethares, W.A. (2005). *Tuning, timbre, spectrum, scale*. Springer Science Business Media.
9. Lartillot, O.; Toiviainen, P.; and Eerola, T. (2008). *A MATLAB toolbox for music information retrieval. Data analysis, machine learning and applications*. Springer.
10. Li, H.; You, H.; Fei, X.; Yang, M.; Chao, K.M.; and He, C. (2018). Automatic note recognition and generation of MDL and MML using FFT. *Proceedings of the 2018 IEEE 15th International Conference on e-Business Engineering* (*ICEBE*), Kuala Lumpur, Malaysia, 195-200.
11. Nagawade, M.S.; and Ratnaparkhe, V.R. (2017). Musical instrument identification using MFCC. *Proceedings of the 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology* (*RTEICT*), Bangalore, India, 2198-2202.
12. Chakraborty, S.S.; and Parekh, R. (2018). *Improved musical instrument classification using cepstral coefficients and neural networks. Methodologies and Application Issues of Contemporary Computing Framework*. Springer.
13. Łetowski. T. (2014). Timbre, tone color, and sound quality: Concepts and definitions. *Archives of Acoustics*, 17(1), 17-30.
14. Adeli, M.; Rouat, J.; Wood, S.; Molotchnikoff, S.; and Plourde, E. (2016). A flexible bio-inspired hierarchical model for analyzing musical timbre. *IEEE/ACM Transaction on Audio, Speech, and Language Processing*, 24(5), 875-889.
15. Hamdan, S.; Musoddiq, I.A.; Musib, A.F.; and Sawawi, M. (2020). Time frequency analysis of Peking gamelan. *Pertanika Journal of Science and Technology*, 28(2), 441-457.

Appendix A

File Name: peking1.csv
Normalized Intensities:
1.0000
0.5082
0.6540
0.3822
0.4631
0.2102
0.2892
0.2172
0.4662
Centroid: 4.743977093777645
Fundamental Frequency: 2.11
Affinity: 2.2483303761979365
Mean Contrast (MC): 0.6012043337499999
Mean Affinity (MA): 2.3584469895802616
Monotony (M): -0.20053126252716172

Harmonicity (H): 3.9478672985782004
Sharpness (S): 0.2386426770097394

File Name: peking2.csv
Normalized Intensities:
1.0000
0.6388
0.4684
0.2211
0.2392
0.1794
Centroid: 4.068328564376346
Fundamental Frequency: 1.77
Affinity: 2.2984907143369186
Mean Contrast (MC): 0.6506289306
Mean Affinity (MA): 2.513890478541218
Monotony (M): -0.08912179409525507
Harmonicity (H): 2.9943502824858754
Sharpness (S): 0.36405266156157734

File Name: peking3.csv
Normalized Intensities:
1.0000
0.8155
0.8117
0.3739
0.3536
0.2922
0.6340
Centroid: 4.820919804197599
Fundamental Frequency: 1.58
Affinity: 3.0512150659478476
Mean Contrast (MC): 0.4531851216666667
Mean Affinity (MA): 2.365582885114629
Monotony (M): -0.09613495015641516
Harmonicity (H): 3.1772151898734173
Sharpness (S): 0.23359632471876576

File Name: peking5.csv
Normalized Intensities:
1.0000
0.2471
0.0694
0.1091
0.4653
Centroid: 3.2729545450368773
Fundamental Frequency: 1.33
Affinity: 2.4608680789750954
Mean Contrast (MC): 0.77727272725
Mean Affinity (MA): 1.7402272729778736
Monotony (M): 0.06664564168737026
Harmonicity (H): 1.9248120300751874

Sharpness (S): 0.5288461538207285

File Name: peking6.csv
Normalized Intensities:
1
0.575
0.970
0.143
0.806
0.304
0.406
0.256
0.693
Centroid: 4.882884386812264
Fundamental Frequency: 1.17
Affinity: 4.173405458813901
Mean Contrast (MC): 0.3967737652647535
Mean Affinity (MA): 2.630149648840357
Monotony (M): 0.30847965786007475
Harmonicity (H): 3.649572649572653
Sharpness (S): 0.18248541495312065

File Name: peking1'.csv
Normalized Intensities:
1.0000
0.7735
0.6228
0.4864
0.6912
0.1117
0.0368
Centroid: 3.3986247824803457
Fundamental Frequency: 1.07
Affinity: 3.1762848434395754
Mean Contrast (MC): 0.5462682258333335
Mean Affinity (MA): 2.1873393167885222
Monotony (M): -0.06019173283926259
Harmonicity (H): 3.056074766355139
Sharpness (S): 0.2686445608128805