SOUND QUALITY CLASSIFICATION OF WOOD USED FOR SARAWAK TRADITIONAL MUSICAL INSTRUMENT-SAPE

WONG TEE HAO

FACULTY OF ENGINEERING
UNIVERSITI MALAYA
KUALA LUMPUR

2024

SOUND QUALITY CLASSIFICATION OF WOOD USED FOR SARAWAK TRADITIONAL MUSICAL INSTRUMENT- SAPE

WONG TEE HAO

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF ENGINEERING
UNIVERSITI MALAYA
KUALA LUMPUR

2024

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Wong Tee Hao

Matric No: 17202487/1

Name of Degree: Doctor of Philosophy

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Sound Quality Classification of Wood used for Sarawak Traditional Musical Instrument- Sape

Field of Study: Energy (NEC 521: Mechanics and Metal Work)

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the Universiti Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 5 June 2024

Subscribed and solemnly declared before,

Witness's Signature Date: 5 June 2024

Name:

Designation:

SOUND QUALITY CLASSIFICATION OF WOOD USED FOR SARAWAK TRADITIONAL MUSICAL INSTRUMENT- SAPE

ABSTRACT

Sape, a traditional musical instrument in Malaysia, is meticulously handcrafted through a complex process. Each Sape, crafted by various makers, differs in size, materials, and design, leading to variations in their quality. Despite individual methods employed by Sape makers to assess quality during production, a standardized guideline for quality inspection remains absent. This research aims to delineate the primary factors influencing Sape quality, employing both qualitative and quantitative methodologies. Initial stages involved gathering insights from seasoned Sape makers and players through questionnaires and focus group discussions, revealing material as the foremost quality determinant in the Sape. Subsequently, the focus shifted to investigating common woods used in Sape construction, specifically Adau, Tapang, and Merbau, representing light, medium, and heavy hardwood categories, respectively. Rectangular wood samples simulating Sape soundboards were created, and sound data was recorded through flexural vibration tests. Expert evaluations of the sound quality were conducted via listening tests. Utilizing MATLAB's MIRToolbox, 18 acoustic properties were extracted from the wood samples. Statistical analyses were employed to identify the most reliable quality ratings. To address dataset imbalances, Synthetic Minority Oversampling Technique was used, enhancing dataset quality before training 40 machine learning classification algorithms. Among these, the Gaussian-kernel Support Vector Machine stood out, achieving remarkable performance with 88.18% validation and 93.37% test accuracies. This model was employed to build a MATLAB-based Sape sound quality classifier. Utilizing the Shapley Additive Explanations interpretation method, the analysis emphasized the importance of selected features in predicting wood acoustic quality, highlighting "Spectral Roll-off 85%" as the most crucial predictor of sound quality. Finally, a user-friendly Graphical User Interface was developed to aid Sape makers in assessing wood quality objectively, enhancing the process of selecting high-quality Sape instruments.

Keywords: Sape, traditional musical instrument, soundboard, machine learning, quality of musical instrument.

PENGKLASIFIKASIAN KUALITI BUNYI KAYU YANG DIGUNAKAN UNTUK ALAT MUZIK TRADISIONAL SARAWAK - SAPE

ABSTRAK

Sape, sebuah alat muzik tradisional dari Malaysia, dihasilkan secara teliti melalui proses yang kompleks. Setiap Sape yang dihasilkan oleh pembuat yang berbeza mempunyai perbezaan dalam saiz, bahan, dan reka bentuk, yang menyebabkan variasi dalam kualitinya. Walaupun setiap pembuat Sape menggunakan kaedah individu untuk menilai kualiti semasa pengeluaran, panduan piawai untuk pemeriksaan kualiti masih belum ada. Kajian ini bertujuan untuk mengenal pasti faktor-faktor utama yang mempengaruhi kualiti Sape, dengan menggunakan kaedah kualitatif dan kuantitatif. Fasa awal melibatkan pengumpulan pandangan daripada pembuat dan pemain Sape berpengalaman melalui soal selidik dan perbincangan kumpulan fokus, yang menunjukkan bahan sebagai penentu kualiti utama dalam Sape. Seterusnya, tumpuan bertukar kepada penyelidikan kayu yang biasa digunakan dalam pembinaan Sape, iaitu Adau, Tapang, dan Merbau, mewakili kategori kayu keras ringan, sederhana, dan berat masing-masing. Sampel kayu segi empat yang mensimulasikan papan bunyi Sape dicipta, dan data bunyi direkod melalui ujian getaran lentur. Penilaian pakar terhadap kualiti bunyi dijalankan melalui ujian mendengar. Dengan menggunakan MIRToolbox MATLAB, 18 sifat akustik diekstrak dari sampel kayu tersebut. Analisis statistik digunakan untuk mengenal pasti penilaian kualiti yang paling dipercayai. Bagi menangani ketidakteraturan dataset, Synthetic Minority Oversampling Technique digunakan untuk meningkatkan kualiti dataset sebelum melatih 40 algoritma pengelasan pembelajaran mesin. Antara ini, Gaussian-kernel Support Vector Machine menonjol dengan prestasi yang luar biasa dengan ketepatan pengesahan 88.18% dan ketepatan ujian 93.37%. Model ini digunakan untuk membina pengelas kualiti bunyi Sape berdasarkan MATLAB. Dengan menggunakan kaedah tafsiran Shapley Additive

Explanations, analisis menekankan kepentingan ciri-ciri tertentu dalam meramalkan kualiti akustik kayu, dengan menonjolkan "Spectral Roll-off 85%" sebagai peramal yang paling penting bagi kualiti bunyi. Akhirnya, Graphical User Interface yang mesra pengguna dibangunkan untuk membantu pembuat Sape menilai kualiti kayu secara objektif, meningkatkan proses memilih instrumen Sape berkualiti tinggi.

Kata Kunci: Sape, alat muzik tradisional, papan bunyi, pembelajaran mesin, kualiti alat muzik.

ACKNOWLEDGEMENTS

I extend my deepest gratitude to God for guiding me throughout my PhD pursuit. All praise and glory belong to Him. I would like to express my heartfelt appreciation to my wife, Sheue Li, and my daughter, Jane, for being my constant motivation and unwavering support. Their encouragement has been my driving force, and I am profoundly grateful for their presence in my life. I am also thankful for the support of my family for their consistent and unwavering encouragement and understanding during my academic journey. Their enduring support has been my source of strength and inspiration.

My sincere thanks go out to the University of Technology Sarawak for their generous sponsorship, which made my PhD studies possible. I am sincerely thankful to my supervisors, Associate Professor Ir. Dr. Rahizar Ramli, Associate Professor Dr. Ahmad Saifizul Abdullah, and Ts. Dr. Soong Ming Foong for their invaluable guidance, expert advice, and continuous encouragement. Associate Professor Ir. Dr. Rahizar Ramli's expertise in acoustic research and machine learning significantly shaped the direction and feasibility of my research. Associate Professor Dr. Ahmad Saifizul Abdullah's counsel on surveys, focus group discussions, and statistical analysis was immensely helpful. Ts. Dr. Soong Ming Foong's guidance in machine learning and unwavering support were instrumental in my academic journey.

A special acknowledgment is extended to Mr. Mathew Ngau Jau, a living national heritage and distinguished Sape maestro, whose invaluable contributions significantly enriched this research. As a National Living Heritage, Mr. Mathew's esteemed status underscores his exceptional expertise and commitment to preserving the cultural

heritage of the Sape. His role as a mentor and cultural guardian has added immeasurable value to this study, as reflected in his generous provision of feedback, evaluations, and insights. Mr. Mathew's willingness to share his profound knowledge and dedicate time to this project has not only enhanced its authenticity but has also deepened our understanding of the Sape instrument and its cultural significance. This acknowledgment is a testament to Mr. Mathew's pivotal role in making this research a meaningful exploration of Sape sound quality and cultural preservation.

Finally, I would also like to acknowledge the crucial contributions of Mr. Ng Jia Sheng, Mr. Matthew Chan Shen Zhe, and Mr. James Tok Chun Sean to my research. Their support was invaluable. Grateful acknowledgments are extended to the Sape community for their vital role in this study. Their feedback and insights greatly contributed to the successful completion of my research project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

Symbols

c	:	Speed of sound
E	:	Dynamic elastic modulus
$E(n)$	:	Energy of the nth frame
$\bar{E}$	:	Mean energy across all frames
f_n	:	Natural frequency
f_R	:	Fundamental frequency
F0	:	Fundamental frequency
F1, F2, F3	:	Partials/ Harmonics
$\text{Fluctuation}(f)$	:	Fluctuation strength at frequency f
G	:	Shear modulus
h	:	Thickness of the wood sample
$H(t)$	:	Harmonic spectrum at time t
$H(z)$	:	System
k	:	Number of nearest neighbours in KNN
K	:	Number of frequency bins
$K(x, x')$	:	Radial basis kernel function
L	:	Length of the wood sample
m	:	Mass of the wood sample
n	:	Mode number
N	:	Percentile
$o(n)$	:	Output
p	:	Pressure
Q	:	Quality factor

R	:	Acoustic radiation damping coefficient
T	:	Total number of samples
t_0	:	Onset time
$\tan\delta$	:	Internal friction
$u(n)$	:	Input
v	:	Velocity
V	:	Volume of the wood sample
x, x'	:	Vector in any fixed dimensional space
$x(t)$	:	Amplitude of the signal
$X(k)$	:	Magnitude of the spectrum at bin k
$X[n,k]$	:	Magnitude of the k-th bin of the Fourier transform at frame n
$\lvert X[n,k]\rvert$	:	Normalized magnitude spectrum of the nth frame
$\lvert X[n-1,k]\rvert$	:	Normalized magnitude spectrum of the previous frame
z	:	Acoustic impedance
α	:	Krippendorff's alpha
σ	:	Gaussian kernel's width
ρ	:	Density
β_n	:	Coefficient of vibration
$\delta(\cdot)$	:	Indicator function

Abbreviations

AC	:	Autocorrelation Coefficients
ACE	:	Acoustic Conversion Efficiency
ADSR	:	Attack, Decay, Sustain, and Release
AM	:	Amplitude Modulation
ANN	:	Artificial Neural Networks

ANOVA	:	Analysis of Variance
ASE	:	Audio Spectrum Envelope
ASS	:	Audio Spectrum Spread
AT	:	Attack Time
BTS	:	Bartlett Test of Sphericity
CNC	:	Computer Numerical Control
DCT	:	Discrete Cosine Transform
EFA	:	Exploratory Factor Analysis
FFT	:	Fast Fourier Transform
FGD	:	Focus Group Discussion
FNR	:	False-Negative Rate
GMM	:	Gaussian Mixture Models
GUI	:	Graphical User Interface
HCDF	:	Harmonic Change Detection Function
HMM	:	Hidden Markov Models
HOS	:	Higher-Order Statistics
I	:	Inharmonicity
IAR	:	Intensity of Acoustic Radiation
ISMIR	:	International Society of Music Information Retrieval Conference
KMO	:	Kaiser-Meyer-Olkin
KNN	:	K-Nearest Neighbours
LAT	:	Log Attack Time
LPC	:	Linear Prediction Coefficients
LPCC	:	Linear Prediction Cepstral Coefficients
LSF	:	Line Spectral Frequencies
MANOVA	:	Multivariate Analysis of Variance

MATLAB	:	Matrix Laboratory
MFCCs	:	Mel-Frequency Cepstral Coefficients
MIR	:	Music Information Retrieval
MIREX	:	Music Information Retrieval Evaluation eXchange
ML	:	Machine Learning
MOE	:	Modulus of Elasticity
MPEG-7	:	Multimedia Content Description Interface
MRMR	:	Minimum Redundancy Maximum Relevance
NBC	:	Naive Bayesian Classifiers
PCA	:	Principal Component Analysis
RBF	:	Radial Basis Function
RH	:	Relative Humidity
RMS	:	Root Mean Square
SB	:	Spectral Bandwidth
SC	:	Spectral Centroid
SF	:	Spectral Flux
SHAP	:	Shapley Additive Explanations
SLM	:	Sound Level Meter
SMOTE	:	Synthetic Minority Oversampling Technique
SMS	:	Short Message Service
SPSS	:	Statistical Package for the Social Sciences
STFT	:	Short-Time Fourier Transform
SVM	:	Support Vector Machine
TPR	:	True-Positive Rate
ULH	:	Upper Limit of Harmonicity
USB	:	Universal Serial Bus

WAV	:	Waveform Audio File
ZCR	:	Zero-Crossing Rate

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background

Malaysia is a multiracial and multicultural nation which is richly blessed with nature diversity and unique heritage. Sarawak, which is a largest state in Malaysia, is recognised for having indigenous populations made up of more than 40 sub-ethnic groups and a wide variety of cultures. Traditional cultural heritage is what Sarawak is famous for and traditional musical instruments are one of them. There are more than 40 traditional musical instruments in Malaysia which comprises of percussion, string, and wind instruments. Among many, Sape is undoubtedly one of the most seen traditional musical instruments in Sarawak. The Sape, a traditional lute of the *Orang Ulu* people, sometimes known as the upriver people, of central Borneo. It was formerly only used in healing rites held in *rumah panjang* (longhouses), later evolved into a social instrument used for entertainment.

1.1.1 Sape

Sape or Sambe is one of the most popular traditional plucked lute musical instruments among the locals in Sarawak, Malaysia particularly the *Orang Ulu* (Kayan, Kenyah, Kelabit). The Kenyah are the renowned exponents of this instrument. Throughout the years, the Kayan, Kelabit, Iban, and Penan have also adopted this traditional musical instrument. The word “Sambe” means to ‘brush lightly with the fingers’ is the description of the technique used by the Sape players to Sape characteristic ornamentation (Chong, 2014). However, the term “Sape” is now widely used. Sape is categorised as a chordophone family of the instrument and originated from the Long Nawang, Kabupaten Bulungan which is located at the border between Sarawak, Malaysia, and Kalimantan, Indonesia.

It is believed that the idea of Sape came from a *Kenyah* man, who was looking for medicinal herbs in the forests for his sick wife. While he was resting under a tree, he fell asleep, and in his dream, the ancestors told him that he needed to find transportation to save his wife from the spirit world. The man, therefore, carved a wooden Adau wood into the instrument with the shape of a boat. The instrument is played to pave the way for his wife's return to the human world (Edward, 2018). Initially, the use of Sape was restricted to ritual, healing, or death activities. However, today it is also used in celebration, self-entertainment, and musical performances (Lim et al., 2020).

Sape is a short-necked, plucked string-type instrument and has a shape like a guitar. It is carved from a single bole of wood, usually the Adau, Meranti, and Merdang (Gorlinski, 1989). The elongated body is hollowed out from the back and functions as a resonator. The strings were made from sago tree or rattan originally, but now these have been replaced by nylon or guitar steel strings. Originally, the Sape is a two-stringed instrument as what is described in the past (MacDonald, 1956; Myers, 1914; Roth & Low, 1896; Shelford, 1904). One string is used for the melody and the three movable frets are placed beneath it. The other one, without the frets, plays the rhythmic accents or the drone. The moveable frets will be adjusted according to the repertoire of the songs played. The frets are made of bamboo or rattan and are usually glued to the body of the Sape using beeswax.

Today, traditional Sape has evolved into three, four, or five strings. Due to this, Sape is now capable of playing a wider note range of up more than three octaves. The fret arrangement follows the western major scale. The number of frets used for the melody strings has also increased in number compared to the traditional Sape back in the former times. More frets are added to the traditional Sape so that the pentatonic scale can be played. However, additional frets could be added to produce a diatonic scale. The diatonic scale is often used in contemporary Sape (Lim & Abdul Rahman, 2016).

The strings are attached to the tuning pegs inserted into the head of the Sape and at the other ends, the strings are held in position by the bridge and the pins. At the head, the strings run over the piece of wood or rattan which is called zero frets as shown in Figure 1.1. Pick-up, volume knob, mono jack socket, and earth grounding are also installed in the contemporary Sape. This allows the Sape to amplify the sound played, especially during the performances. The body of the Sape is painted with intricate decorations upon the completion of the Sape. The decorations are typical for the *Orang Ulu* ethnic group and are based on the maker's imagination and are usually painted in black and red as shown in Figure 1.2 (Pilo, 2018). Some Sape makers will also do some decorative carvings on the body including carving the head into a shape like the Borneo hornbill.

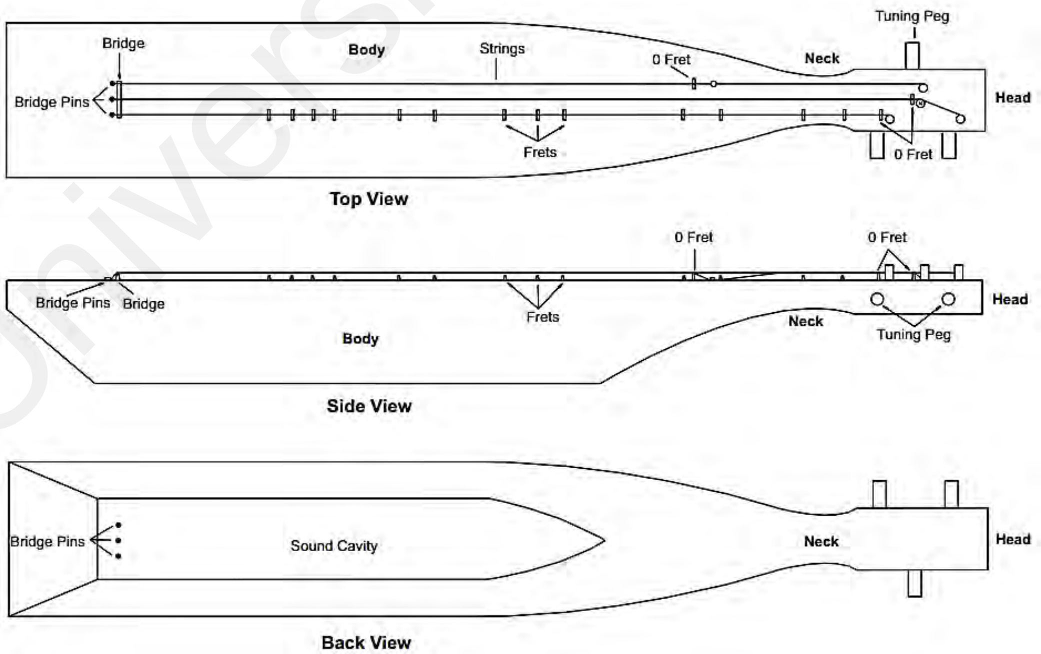


Figure 1.1: Schematic drawing of Sape

Figure 1.2: Traditional musical instrument, Sape

1.1.2 Process of making a Sape

The process of crafting a Sape musical instrument involves several meticulous steps deeply rooted in tradition and craftsmanship. To begin, the selection of wood is a critical aspect, often utilizing special wood types like Adau (*Elmerrillia mollis dandy*), Tapang (*Koompassia Excelsa*), or Merbau (*Intsia palembanica*). Adau stands out as an optimal choice due to its carvability, durability, resilience against cracks, and its capacity to produce a resonant humming sound.

Traditional Sape production necessitates various specialized tools such as *beliung*, *bikong*, *bikong sulok*, and *asai*. The woodcutting phase is pivotal, with no standardized size set, as tuning factors and standards differ. Therefore, consistency is maintained by sourcing wood from the same tree trunk, cutting it to identical measurements. This wood, harvested from forests, undergoes months of air-drying before the crafting commences.

Upon drying, a 5-foot-tall, 2-foot-diameter wood block is halved and marked for the Sape outline. The outer circle guides the carving process, where an axe (*asai*) is used to cut approximately an inch outside the drawn line. Subsequently, the exposed back is perforated with *bikong*, enabling further crafting stages.

Creating uniformity in shape involves using templates to mark the Sape's design on the wood. The body formation spans multiple stages, shaping both sides, hollowing the back for a sound hole, and smoothing the front surface. Sound holes are meticulously pierced using *beliung*, influencing the Sape's volume and softness, while side thickness also contributes.

Refinement and smoothing follow shaping, employing long-handled knives and modern tools like electric sanders to achieve an even, flawless surface. Drilling precise holes for tuning pegs is crucial, dictating string installation and fret placement. Traditional tuning pegs, often handcrafted from hard wood, or guitar tuning pegs are used, requiring precision in fitting.

The artistic touch emerges through drawing or carving motifs, typically using templates for guidance. Once chosen, the motif is traced or carved, followed by painting. String installation precedes pitch fret attachment, with materials evolving from traditional rattan and creeper strips to stainless steel or guitar strings for enhanced sound quality.

Fret placement, made from materials like palm wood or bamboo, adheres to fixed scales, either pentatonic or diatonic. This intricate process, spanning wood selection, carving, drilling, and artistic embellishments, embodies the rich tradition and skilled artistry behind the creation of the captivating Sape musical instrument (Lim et al., 2020).

The time taken to make a Sape can vary a lot. It might take from one week to several weeks, depending on how much experience the maker has and how much time they can spend on it each day. Some makers work on making Sape in their free time after finishing their regular jobs. This flexible schedule means that each maker takes their own time to finish a Sape. It shows their dedication and passion for crafting these instruments. Overall, the process of making a Sape can be quite different for each maker, based on their skills and the time they can devote to it, highlighting the personal touch and commitment involved in creating these culturally important musical instruments.

1.2 Problem Statement

Traditionally, the Sape has no standardized measurement or standard tuning. This musical instrument is fabricated based on the size of the tree log. Due to the increased use and popularity of Sape especially its involvement in the orchestra, band, and others ensemble groups, Sape are now usually tuned based on open strings A=440 which is the tuning standard for the musical note of A above middle C. This allows the Sape made with different materials and sizes to be able to play together with other musical instruments (Lim & Abdul Rahman, 2016).

However, Sape made by different Sape makers which vary in wood, dimension, and design will produce different sound characteristics. For instance, the depth of the sound cavity will affect the sound volume and effects of Sape. The thickness of the sides also contributes to the sound effects of the Sape (Lim et al., 2016). The sound characteristics that differ can be explained by the term “timbre”. According to the definition by the oxford university press dictionary, the timbre of the sound refers to "the quality of sound that is produced by a particular voice or musical instrument" (Colman, 2015). Sound timbre is independent of pitch and loudness, from which its source or manner of production can be inferred. It depends on the relative strengths of the components of different frequencies, which are determined by resonance. Therefore, different Sape with the same tuning will result in different sound timbre being produced. The differences may be subtle for not only the non-musician to notice but even the experienced Sape maker and player.

The diverse crafting approaches for Sape instruments give rise to a significant problem—the absence of standardized guidelines for ensuring consistent sound quality. With divided opinions among Sape makers and players, a generalized conclusion from Sape experts could provide valuable insights. Factors such as Sape specifications and preferred materials among experts might shed light on producing the highest quality Sape instruments.

Additionally, the absence of prior research on the acoustic and timbre features of high-quality Sape instruments poses a considerable challenge. Specifically, inquiries into the sound characteristics valued by Sape makers or players to achieve superior quality remain unexplored. Moreover, to ensure the reliability of identified main factors, it is essential to investigate the correlation between timbre features and perceived sound quality in Sape instruments. The current lack of comprehensive research on this correlation creates a significant gap in our understanding. Therefore, addressing this gap requires an investigation into the intricate relationship between timbre characteristics and the perceived sound quality of Sape instruments.

Moreover, there is a deficiency in the user-friendly sound quality evaluation system for inspecting Sape musical instruments. This gap, affecting both novice and experienced Sape makers and players, underscores the necessity of developing a comprehensive system to address these challenges. The reliance on subjective assessments by individual Sape makers and players further highlights the need for an automated approach. The lack of a standardized and automated system poses challenges, leading to variations in evaluations across different instruments. This subjectivity, coupled with the lack of systematic guidelines during the crafting process, emphasizes the demand for a more objective and automated method of sound quality assessment. As the Sape gains prominence in various musical settings, the growing need for a reliable method beyond individual perspectives becomes evident, requiring a standardized and efficient evaluation process.

The above problems form the focal points of investigation in this thesis and are framed into the research questions in the next subsection.

1.3 Research Questions

This study aims to address the complexities surrounding Sape instrument production and sound quality evaluation through the exploration of the following research questions:

i) Factors Influencing Sound Quality in Sape Instruments

- What are the key influential factors that significantly contribute to achieving optimal sound quality in Sape instruments, as perceived by experts?
- Are there specific materials or design specifications that most Sape experts prefer for producing superior sound quality?

ii) Correlation between Timbre Features and Perceived Sound Quality

- How do variations in dimensions, materials, and designs of Sape instruments affect their sound characteristics and timbre?
- What is the significant relationship between the perceived sound quality of Sape instruments and their timbre features?

iii) Development of a Sound Quality Evaluation System

- What are the essential criteria and parameters that should be included in a standardized evaluation system for assessing the sound quality of Sape instruments?
- How can a user-friendly and standardized sound quality evaluation method be developed to aid Sape makers and players?

These questions delve into crucial aspects of Sape instrument construction and sound quality. The study aims to uncover key factors influencing optimal sound quality, explore the relationship between instrument characteristics and timbre, and contribute to a standardized evaluation system. The overarching goal is to enhance our understanding of Sape instruments and develop improved sound evaluation methods.

1.4 Research Objectives

The primary aim of this research is to devise a machine learning-based quality classification model capable of effectively categorizing the sound quality generated by the Sape musical instrument. To achieve this goal, this work aims to accomplish the following objectives:

i) To identify the key influential factors that affect the sound quality of Sape.
ii) To determine the significant relationship between the perceived sound quality and the timbre features.
iii) To develop an automated sound quality classification system.

These objectives collectively aim to lay the foundation for constructing a sophisticated machine learning-based classification model that accurately evaluates and categorizes Sape sound quality. The comprehensive exploration of influential factors and their relationship to sound perception will contribute significantly to the development of an automated classification system.

1.5 Research Scope

This research aims to determine the fundamental factors contributing to the sound quality of the Sape traditional musical instrument. The primary scope involves developing a sound quality classification model for the Sape through a comprehensive integration of diverse methodologies, including quantitative and qualitative approaches, flexural vibration tests, statistical analyses, and machine learning methods.

The research comprises six stages: focus group discussions, questionnaires, sound quality evaluation, parameter analysis, sound quality classification, and Graphical User Interface (GUI) development. Initially, focus group discussions and questionnaires will gather perceptions from Sape makers and players regarding potential quality factors of the Sape. The most significant factor identified will be further investigated. Subsequently, a flexural vibration test will be conducted to capture sound recordings from various wood-based soundboard samples. Expert Sape makers will then assess the recorded sound quality through listening tests. The sound signals will undergo analysis in both time and frequency domains to extract distinctive features from different soundboards. Statistical analysis will help select the best sound quality features for employment in machine learning classification. Lastly, a GUI will be developed, serving as a platform for future Sape makers to analyze soundboard quality during Sape fabrication.

This research is focused on the sound quality production of the Sape musical instrument. The sound is generated by plucking the string, the vibration is then transferred to the bridge and the body of the Sape. The sound radiation is then amplified by the hollowed body of the Sape (Wong et al., 2022). Due to the complexity of this sound production, this research is focused only on the most significant factor influencing the sound quality of this musical instrument. To determine the significant factors, a quantitative and qualitative study is carried out by collecting the perceptual opinions of the Sape players and Sape makers. Most of the participants selected in this study are from Sarawak, Malaysia, and have experience in playing and making the Sape.

The results in Chapter 3 showed that the wood type is the most significant factor influencing the sound quality of the Sape. Therefore, this research focused only on the woods used in making the Sape. There are many types of wood used to make the Sape, however, this research selected one type of wood from each of the wood categories which are light hardwood, medium hardwood, and heavy hardwood. These selected woods are the common woods found and used by the Sape makers in Sarawak.

For simplicity and convenience of the study, the wood sample used in this research is cut into a rectangular soundboard imitating the actual dimension of the Sape body. Even though the actual Sape body shape is not the exact rectangular shape as shown in Figure 1.3, it is believed that the vibration, acoustic, and physical properties of this Sape soundboard sample could represent the actual characteristics of the Sape body. It is worth noting that the Sape is fully handmade by the Sape maker and therefore, no Sape is 100% similar in terms of size, material, performance, etc. The research is therefore limited and focused only on the soundboard. The soundboard (top plate) is chosen due to its important role in sound generation and radiation, especially in string musical instruments (Yoshikawa & Waltham, 2014). This research excludes the contribution of other parts of the Sape, such as the side body, neck, and strings.

Figure 1.3: The soundboard sample and the actual Sape

As the focus of this research is on sound quality, the sound produced by the soundboard therefore needs to be evaluated by the Sape experts. It is therefore needed to get the experienced Sape makers to listen and evaluate the quality of sound produced by the soundboard. The rating given by the Sape maker is then used in the machine learning classification while the best features of the good sound quality are selected based on the statistical analysis. This research explores the potential of employing machine learning in the classification of the sound quality of Sape. The outcome obtained from this research will shed light on the improvement of the Sape by collecting precious opinions from the experienced Sape community, understanding the important quality factors, and incorporating the machine learning method in determining and evaluating the sound quality production of the Sape.

The research process depicted in Figure 1.4 provides a visual representation of the study's progression. The initial phases encompass comprehensive background study and literature review, followed by an exploration of quality factors employing both quantitative and qualitative methodologies. Subsequently, the process involves the preparation of soundboard samples, data collection, and rating. The utilization of machine learning involves feature extraction and selection to train the chosen classification model. Post-training, the development of a GUI is undertaken. This visual illustrates the overall workflow encapsulating the sequential steps within this research.

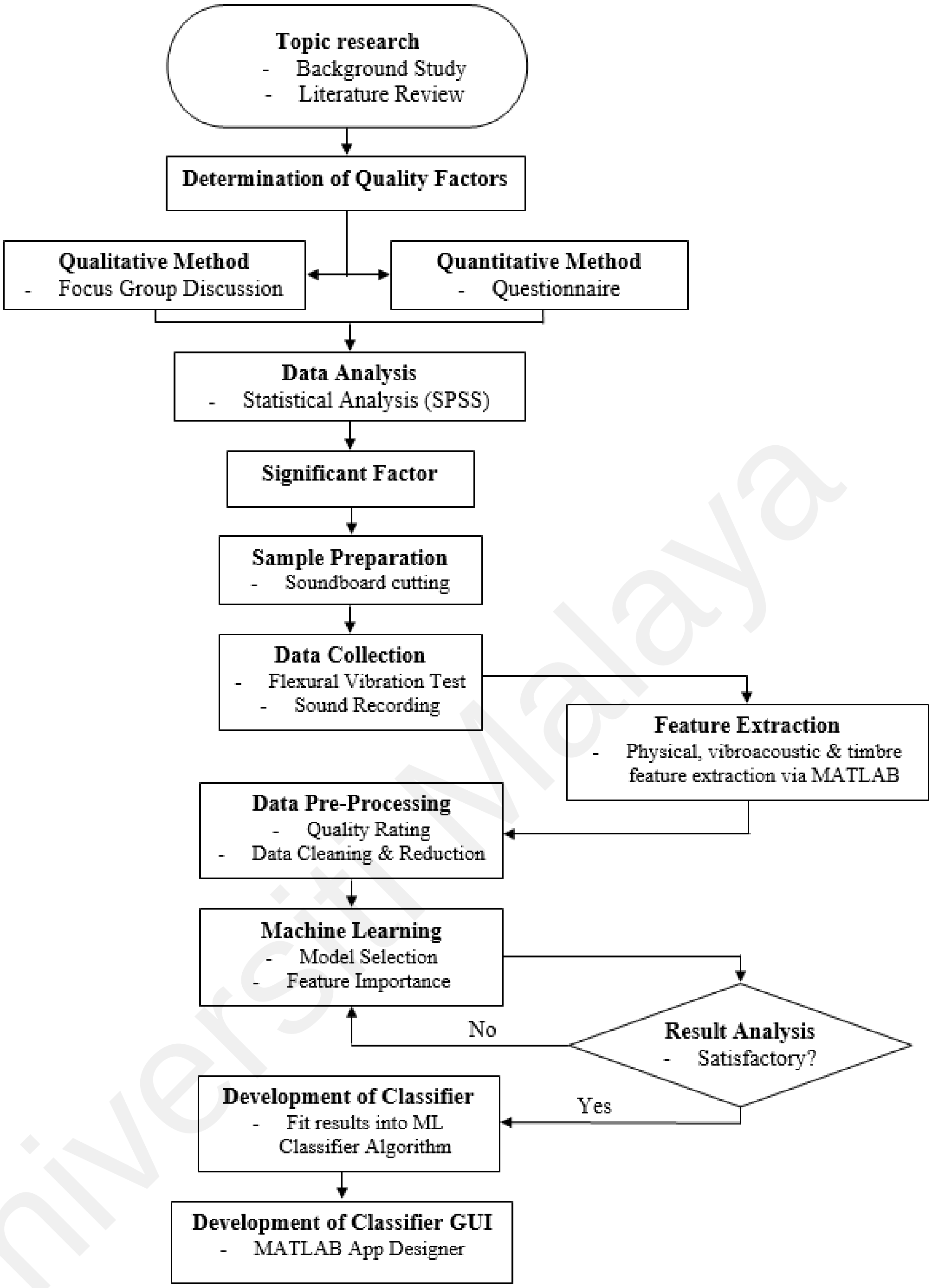


Figure 1.4: Research flowchart

CHAPTER 2: LITERATURE REVIEW

2.1 Previous Studies on Sape

The study on Sape musical instruments and Sape music is not new. Chieng (2011) studied Sape music and found an unexpected rhythmic feature in the music recording played by Tusau Padan. The results of this research on the inequality in time structuring in Sape music have significance for playing styles, and the musical traditions of Sape musicians. In an effort to preserve the music of Sape, Musib (2015) researched the contextual sound preservation of Sape music with the aim of high-quality audio archiving.

A study was also conducted on the performance, practice, and repertoire of Sape music to understand the differences in the melody, harmony and rhythm (Sudom, 2016). The innovation and the development of Sape from the traditional longhouses to the current contemporary style are also explained in the studies by Hashim (2017) and Lim and Abdul Rahman (2016). Other than music, the process of making the Sape is documented by Karlina et al. (2018). With the documentation, the entire process of Sape making and the tools used is recorded.

Wong et al. (2017) and Wong and Dayou (2019) started the work on Sape from the sound and vibration perspective. The study focused on the effects of string tension and plucking force on the fundamental frequency of sound and vibration of the Sape. Results showed no significant effect of plucking force on the fundamental frequency produced. On the other hand, the string tension did show effects on the fundamental frequency. The authors continued their work on the effect of the fretting arrangement and derived a general correlation equation of the fretting distance to the fundamental frequency production (Wong et al., 2022).

Although several studies have been conducted on Sape, it is still underexplored. To the author's knowledge, there is no study done on the sound quality of the Sape. As this study intended to measure the sound quality production of Sape, the quality of string musical instruments is studied for the understanding of the quality measurement principle.

2.2 Sound Quality of Musical Instruments

According to Campbell (2013), the studies in musical acoustics consist of three general categories. It consists of the physics of musical instruments and other sources, the transmission of the sound source to the listener and the psychoacoustics of musical sound perception. The work in evaluating the quality of musical instruments has been done for many years. For example, pianos (Fletcher et al., 1962), violins (Fritz, Blackwell, et al., 2012), and brass instruments (Kemp et al., 2010). Early investigations on musical instrument quality mostly focused on the frequency spectra of the steady-state components of the radiated sound. The focus is now shifting towards the interaction between the players and the instrument which is termed "playability". Regardless of the study's primary objectives, one common goal is to identify a certain instrument's objectively quantifiable characteristics that closely match performers' perceptions of timbre or its perceived quality. The term "timbre" is defined as "everything that is not loudness, pitch, or spatial perception" by the American National Standard Institute (ANSI, 1973). Campbell (2014) described "timbre" as the quality of sound that distinguished a note played on a clarinet from a note of the same pitch and loudness played on a trumpet.

The challenge in studying the quality of the musical instrument is linking players' evaluations to objective scientific measurements. Chaigne and Kergomard (2016) and Masullo et al. (2021) explain how a musical instrument is perceived, there are two key factors to consider. First, to comprehend the factors, or so-called playability of the instrument, that affect how the player perceives the instrument. The second is to determine the factors that affect how listeners perceive sound quality. Since Sape is still underexplored, this study intended to determine the factors influencing the sound quality of Sape from the perspective of the players and makers.

2.3 Music Information Retrieval (MIR)

Two of the various approaches often used in studies on musical instruments are the acoustical characterization and sound recognition system. Scientists many years ago started to discover the acoustical characteristics of different types of musical instruments by using various techniques. The initial techniques used are modal analysis and acoustic radiation. Over the years, there are many other new parameters developed and introduced from these fundamental techniques. The common acoustical characterization parameters are mechanical admittance and impedance, sound radiation coefficient, the intensity of the acoustic radiation, and anti-vibrational, and transmission parameter to name a few.

Mechanical admittance is defined as the ratio of the velocity, v to the force, F. This characteristic is useful in understanding the body vibration of the musical instrument. The study which used admittance on musical instrument vibration measurements can refer (Daltrop et al., 2010; Pölkki et al., 2003). Reciprocal to the admittance, driving point mechanical impedance on the other hand is defined as the ratio of the applied force, F to the velocity, v produced by the instrument body. Measurement is done by applying the force to the instrument body and the resulting velocity is measured with the accelerometer (Meyer, 2009).

Sound radiation characteristics of different musical instruments have been extensively studied as well. Sound radiation coefficient which is defined as the ratio of the material's speed of sound, c to its density, ρ describes how much the sound radiation of the musical instrument body is damped. It can be measured by the vibrational response of the instrument soundboard for a given force (Meyer, 2009). The intensity of the Acoustic Radiation (IAR) parameter is introduced by Tronchin in 2005 on the kettledrums. It is defined as the product of the space-averaged amplitude of the cross-spectrum sound pressure, p, and the velocity, v generated from the surface vibration. As the name suggests, it is a parameter related to acoustic intensity and acoustic radiation (Tronchin, 2005, 2020).

Studies were also carried out in determining the sound characteristics of the woods used in musical instruments. Various kinds of wood are tested and analysed based on the anti-vibration and transmission parameters. The anti-vibration parameter is the reciprocal of the sound radiation ratio produced by the woods. It is the ratio of the longitudinal wave speed, c to the density, ρ of the wood. On the other hand, the transmission parameter is the product of the longitudinal wave speed, c, and the quality factor, Q. The results are then used in the acoustical classification and comparison of the woods used in different categories of musical instruments (Yoshikawa, 2007).

On the other hand, sound recognition systems started to get more attention due to the growth of digital music. Music information retrieval (MIR), which is the subset of the broader field of sound recognition, is known to be the field that contributes to the solutions of the musically related task. Sound recognition is a multi-disciplinary field that includes speech recognition, information retrieval, music information retrieval, environmental sound retrieval, etc. Figure 2.1 illustrates the general taxonomy of the sound classification scheme introduced by (Gerhard, 2003). Under the field of MIR, there are various tasks. For instance, music genre recognition, song identification, mood classification, music annotation, tempo, fingerprinting, etc. One of the tasks is musical instrument classification.

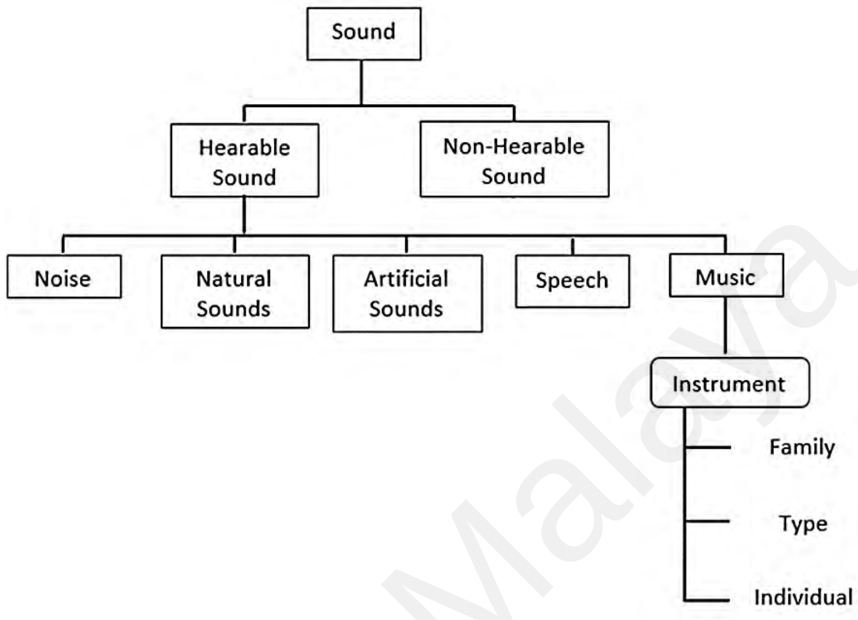


Figure 2.1: Taxonomy of sound (Gerhard, 2003)

The application of MIR can help in the identification of the individual musical instrument, its type, and its family. For instance, the Sape, falls under the family of string instruments, specifically categorized as plucked strings, and individually identified as a lute. MIR is gaining popularity among researchers, musicians, and acousticians in the efforts of getting a better understanding of the sound produced by musical instruments. As we are currently living in the digital world, where vast amounts of musical databases are made available online. The demands are there for the development of computational tools for the analysis, summarization, classification, and indexing of those musical data (Bhalke et al., 2016). These demands have inspired a growing research attempt in the automatic classification of the sound produced by the different types of musical instruments.

2.4 Sound Recognition System

As mentioned earlier, MIR is a subtask of the audio recognition system. The task is dealing with the automatic audio recognition of music signals which at the end will extract the information or characteristics of the music content. Musical instrument class is one of the characteristics that could be obtained by the analysis. The application of an audio recognition system in musical instrument classification is not a new thing as there have been numerous attempts by researchers on it in recent years. Most of the research done in musical instrument classification has adopted the technique used in speech recognition and speaker identification system. This is because a few features from the speech recognition system can be directly applied to solve the musical instrument classification problem (Kaminsky & Materka, 1995).

Generally, the musical instrument classification system consists of three steps, preprocessing, feature extraction, and classification as shown in Figure 2.2. Most of the research on musical instrument classification emphasized feature extraction which is vital in getting the correct characteristics of the sound processed.

Figure 2.2: Process of audio recognition system in musical instrument classification

In the first step, the audio input that is captured by the microphone will go through a windowing process by segmenting the audio into shorter signal chunks. A musical audio signal is usually long and may contain a large number of samples given that the sampling rate is higher than 10 kHz. The audio sample, therefore, couldn't be analysed directly and needed to go through the pre-processing step. This is because the audio signal is constantly changing. To simplify it, the audio signal is split into a continuous sequence of finite frames of samples. The frames with short scales are then assumed to not change much. This process converts the non-stationary audio signal into a stationary signal over a short period (Fu et al., 2010). Typically, the segmented frame length is in between 10 to 50 milliseconds and will be overlapping with the adjacent frames for about 25 to 50% (Bhalke et al., 2016; Deng et al., 2008). This is to ensure that there are no missing signals during the segmentation process. The frame size, however, is related to the length of the processed sound signal (Alías et al., 2016).

In the pre-processing step, some research will remove the noise or silence part of the audio input before proceeding to the next step (Marques & Moreno, 1999). It can help in reducing the computational complexity of the recognition system. For instance, the zero-crossing rate (ZCR) or the energy threshold value is used in the research done to eliminate the unwanted silence part of the audio signal. Other than that, they also applied the pre-emphasis which serves the purpose of compensating for the suppressed high-frequency formants during sound production by the musical instruments (Bhalke et al., 2016).

The next step of sound recognition is the feature extraction of the audio signal. To classify the audio input into any musical instrument class, it is very crucial to identify the characteristics of the sound produced by each musical instrument. This process is also called parameterization which eventually will build the feature vectors that best represent the musical sound. The built feature vectors contain the most significant characteristics or parameters of the musical sound. This will then be very useful in the classification process. There are various methods to extract the characteristics or features from the audio inputs, which will be discussed later.

The significant parameters of the musical instruments' sound constructed in the feature extraction will be used as a descriptor to represent a similar type of musical instrument or to distinguish between different types of musical instruments. This could be done through the classification process based on various techniques or machine learning algorithms called the classifier. There are many classifiers available currently but the choice of the suitable one depends on the goal of the classification system, the accuracy of the classifier, and avoiding overfitting. In general, the classification algorithm consists of two phases: the training phase and the testing phase. In the training phase, the machine learning algorithm under supervised conditions will build representative acoustic models that best represent the sound class that the system wants to recognize. This is done by taking multiple sound samples of the same musical instrument if the musical instrument type is the goal of the machine learning system. After the algorithm is trained, it will then be tested in the testing phase. The unknown sound samples will be imported into the system for classification. The algorithm will classify the incoming sound signal into different classes based on the information acquired in the previous phase (Alías et al., 2016).

The effectiveness of the sound classification system is the main concern of the researcher. It is measured by comparing the accuracy of different features or classifiers used in the sound classification system. Today, researchers are still trying to get the best feature set or classifier that could be used in musical instrument classification. Since 2014, there has been an annual competition organized by the MIR community called Music Information Retrieval Evaluation eXchange (MIREX). This event lets the participant test their music classification system in a few categories such as genre, musical instrument, music, mood, and artist classification (Fu et al., 2010). Other than that, the MIR community has been organizing meetings through the International Society of Music Information Retrieval Conference (ISMIR) every year since 2014.

2.5 Feature Extraction

Feature extraction and classifier are important components of the classification system. Feature extraction determines the features to be used for the machine learning system. The problems of classifying the sound samples into different classes based on feature vectors will be addressed. The feature vectors represent the similarities between the sound samples. The features extracted may be redundant and irrelevant. This will cause a burden on the computation time. Therefore, some of the features will be discarded and only a subset of the features will be used at the end. This process is called feature selection. Both feature extraction and feature selection are very crucial in machine learning. It can ease the computation time by selecting only the useful and relevant features particularly when the dataset is too large (Murty & Devi, 2015).

There are several approaches to categorise the feature extraction of the audio signal in the machine learning system. Due to the manifold nature of audio features, there is no general taxonomy that could be applied to all fields of research. Hence, it is usually designed according to the research field and purpose of the study. Fu et al. (2010) unified the taxonomies of audio features by Scaringella et al. (2006) into a single hierarchical taxonomy. The taxonomy consists of low-level features and mid-level features with the top-level providing information on the human's perception towards music through the semantic labels. The low-level features in this taxonomy are divided into timbre and temporal features. As for the mid-level features, it contains information on rhythm, pitch, and harmony. The taxonomy is grouped into short-term and long-term features.

Alías et al. (2016) extended the taxonomy introduced by Mitrović et al. (2010) in their review of feature extraction techniques on speech, music, and environmentally sound. Taxonomy is classified into physically based and perceptually based approaches.

These two approaches are then further divided into different parameterization domains such as time, frequency, wavelet, image, cepstral, etc. This is different from the taxonomy by Mitrović et al. (2010) which listed the parameterization domain on the first level of taxonomy and the physically-based and perceptually-based features are put under the frequency domain.

In this section, the taxonomy in Figure 2.3 will be adopted and the features extraction techniques in the literature for the classification of musical instruments will be reviewed. It is noted that some of the domains may not be relevant in the review of the musical instrument classification therefore they will not be covered in this section. Only the relevant domains such as time, frequency, cepstral, and wavelet domain are covered.

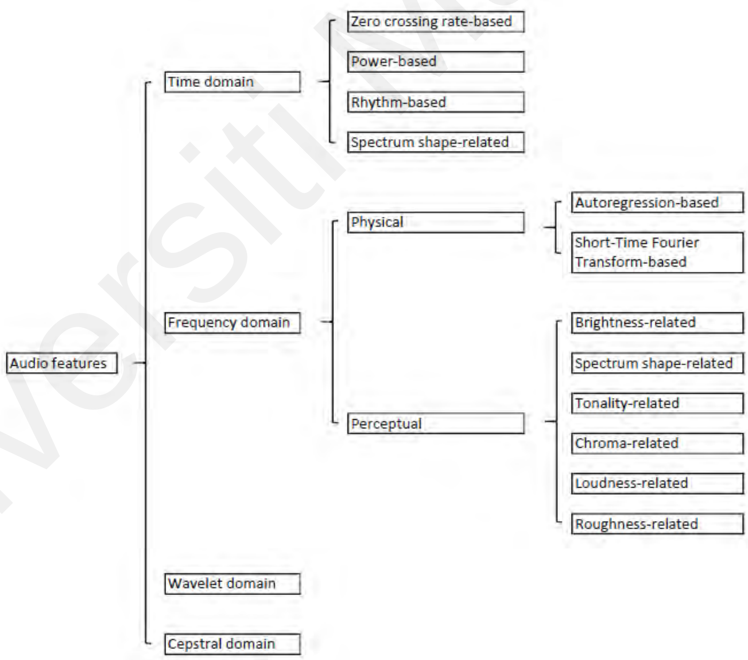


Figure 2.3: Taxonomy of audio features in musical instrument classification

2.5.1 Time Domain

Also called a temporal domain, the time domain is perhaps the most basic domain for audio signals. It is not complex and easy to extract audio features from. It can be displayed directly from the raw audio signal without further transformation. There are four classes of physical time-domain audio features: zero crossing-based, amplitude-based, power-based, and rhythm-based features.

2.5.1.1 Zero-Crossing Rate-Based Features

The technique used here is based on the analysis of the rate of change of the sound signal. It is a simple but effective method commonly used in MIR.

(a) Zero-Crossing Rate (ZCR)

Known to be one of the easiest features to get from the audio signal. The zero-crossing rate is defined as the number of times the audio signal waveform passed the zero-amplitude level within one second. This feature is widely used in audio classification and machine learning systems. It is measured based on the rate of change of the audio signal and is probably the simplest way for feature extraction. Kedem (1986) and El-Maleh et al. (2000) mentioned in their papers that the ZCR can provide a rough estimation of the dominant frequency and the spectral centroid in the signal. ZCR is quite popular in the musical instrument classification field.

2.5.1.2 Power-Based Features

Power-based features are extracted based on the audio signal power. A few relevant features are described below.

(a) Energy

Using the frame-based procedure, the energy feature summarizes the energy distribution of each frame over time. Mitrovic et al. used the term short-time energy to represent this feature (Mitrović et al., 2010). The researchers used this feature for finding the energy distribution in each frame and tried to find the differences between the instruments. Bhalke et al. (2016) used time-domain energy as a feature in their musical instrument recognition paper.

(b) ***Temporal Centroid***

Temporal centroid gives the time average over the signal envelope in seconds. It represents the instant moment in time that contains the largest average energy of the signal. The temporal centroid has been used as a time-domain audio feature. It may also be classified as a MPEG-7 feature in the musical instrument classification field (Deng et al., 2008).

(c) ***Log Attack Time (LAT)***

The log attack time characterizes the attack of the sound signal. Musical instruments can produce either instant or smooth transitions of musical sounds. It is computed as the logarithm of the time taken from the start to the first significant local peak (Deng et al., 2008).

(d) ***Root Mean Square (RMS)***

Also named as the volume is the review by Mitrović et al. (2010), RMS is computed by finding the root mean square of the waveform magnitude within the frame (Liu et al., 1997). The RMS feature is utilized to quantify the energy content of a signal. It is calculated by taking the square root of the mean of the squares of the signal values. This feature is crucial for capturing the dynamic energy variations within the acoustic signals, providing valuable insights into the overall energy distribution and intensity of the soundboard vibrations, which are essential indicators of soundboard quality.

2.5.1.3 Rhythm-Based Features

Rhythm is a relevant characteristic of musical sound that characterizes the sonic events' structural organization (Alías et al., 2016). Feature derived under this taxonomy is discussed here.

(a) Periodicity

Periodicity or tempo is the measure of the rhythmic strength or repetitive structures of audio signals (Lu et al., 2001). Periodicity is obtained by applying the autocorrelation function to acquire the mean value of the maximum peaks through all the signal frames.

2.5.1.4 Spectrum Shape-Related Features

The spectrum shape of the audio signal is another relevant feature that could be employed in the task of musical instrument classification. Spectrum shape-related features are described in the following paragraphs.

(a) Attack, Decay, Sustain, and Release (ADSR) Envelope

The temporal envelope of musical instrument sounds is characterized by attack time, decay time, sustain time, and release time as shown in Figure 2.4. Attack time is the time taken for the sound signal to rise from zero to the peak. The decay time is the subsequent time to run down the signal level from the peak to the sustained level. Sustain time is the main sequence where the signal level remained the same and lastly, the release time represents the time taken for the signal to decay back to zero levels. ADSR combined to form a signal envelope that could be extracted as a feature in vector form in the musical instrument classification task (Bhalke et al., 2016).

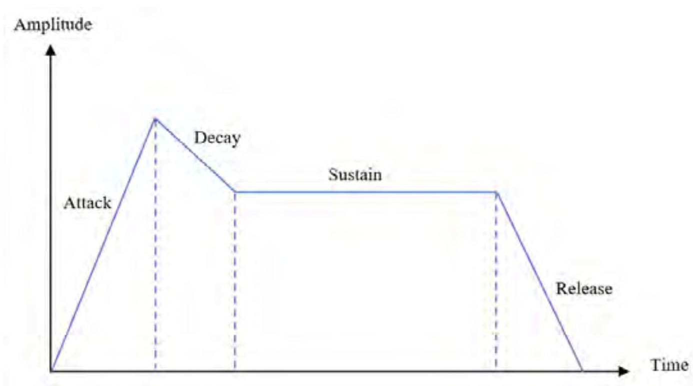


Figure 2.4: The ADSR envelope

(b) Amplitude Modulation (AM)

Amplitude modulation (AM) features are extracted from the audio signal for the peaks which correspond to the frequency of amplitude modulation. AM has measured over two spectral ranges 4 to 10 Hz and 10 to 40 Hz (Eronen, 2001).

(c) Autocorrelation Coefficients (AC)

Autocorrelation coefficients (AC) represent the overall shift of the spectrum (Peeters, 2004). Brown reported that AC is useful in musical instrument identifications (Brown et al., 2001).

(d) Temporal Kurtosis

Temporal kurtosis shows the spikiness of the audio envelope. It is used in measuring the variation of the transients of the audio signal over successive frames (Klapuri & Davy, 2007).

2.5.2 Physical Frequency Domain

The frequency domain is also named the spectral domain. According to Mitrović et al. (2010), audio features in the spectral domain form the largest set of audio features. They are acquired from autoregression analysis or Short-Time Fourier Transform (STFT). This paper employed the approach by Mitrović et al. (2010) in further dividing the frequency domain into two subsets: physical features and perceptual features. In this section, features extracted in the physical frequency domain for the musical instrument classification task will be discussed first.

2.5.2.1 Autoregression-Based

Autoregression-based features use linear prediction analysis on signal processing. The linear predictor captures the spectral predominance of audio signals (Alías et al., 2016). Commonly used autoregression-based features are discussed below.

(a) ***Linear Prediction Coefficients (LPC)***

Linear prediction coefficients capture the spectral envelope of the audio signal, such as formant frequencies that could be found in the vocal tract. It has been used extensively in speech recognition applications. The application of LPC in musical instrument classification could be found in the works by Marques and Moreno (1999). The prediction model used is shown in Figure 2.5. It consists of the input *u(n)* which is the periodical sound produced by the musical instrument, *H(z)* which represents the musical instrument system, and the output *o(n)* represents the musical instrument class.

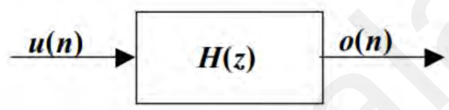


Figure 2.5: A linear prediction model for musical instrument sound production (Marques & Moreno, 1999)

(b) ***Line Spectral Frequencies (LSF)***

Line Spectral Frequencies are also called Line Spectral Pairs (LSP). It is obtained by finding the root phases of the two polynomials that are decomposed from the LPC (Campbell, 1997). LSF is proved to be more robust when compared to LPC as they provide statistical properties.

2.5.2.2 Short-Time Fourier Transform-Based Frequency Features

Short-Time Fourier Transform or STFT-based audio features are obtained from the signal spectrogram that is employed by STFT computation. According to Mitrović et al. (2010), there are two ways to yield the STFT features, either from the spectrogram envelope or from the STFT phase. The application of STFT-based features in musical instrument classification is found to be mostly, if not all, from the spectrogram envelope. These features are widely employed by researchers and are discussed below.

*(a) **Spectral Flux***

Spectral flux (SF) is defined as the 2-norm of the frame-to-frame spectral amplitude difference vector by Scheirer and Slaney (1997). SF measures the changes in the spectrum shape over time. Signals without much variation like noise will show low SF, while the high SF indicates sudden changes that are useful in detecting certain information like the onset of sound.

*(b) **Spectral Peaks***

As defined by Wang (2003), spectral peaks are the constellation maps that display the most significant local peaks in the time-frequency signal distortions. The advantage of this feature is that it is highly robust to noise since the significant peak frequencies are usually free from noise disturbance. This feature is used by Wang (2003) in the Shazam search engine.

*(c) **Audio Spectrum Envelope***

Audio spectrum envelope (ASE) is defined as the log-spectrum frequency power spectrum that produced a reduced spectrogram of the original audio signal. ASE consists of coefficients that describe the power spectrum density within a series of frequency bands. Categorized as a MPEG-7-based low-level descriptor, it is suitable for automatic musical sound recognition (Kim et al., 2004).

2.5.3 Perceptual Frequency Domain

Another division of frequency-based features is the perceptual domain. Perceptual features have a semantic meaning as the human auditory perception. In this section, several perceptual features will be included and discussed.

2.5.3.1 Brightness-Related Perceptual Frequency Features

The brightness of an audio signal characterizes the frequency spectrum distribution, indicating how much high-frequency content is present. An audio signal is considered bright when it is dominated by high frequencies, typically in the range of 2 kHz to 20 kHz. Brightness is also defined as the balancing point of the signal energy, reflecting the concentration of spectral energy at higher frequencies (Scheirer & Slaney, 1997).

(a) ***Spectral Centroid***

Spectral centroid (SC) is one of the commonly used features. It describes the centre of gravity (centroid) of spectral energy. It can also be defined as the first moment which is the frequency position of the mean value of the spectrum (Tzanetakis & Cook, 2002). Deng et al. (2008) in their work on musical instrument classification defined that the SC measures the average frequency weighted by the sum of spectrum amplitude within each frame.

(b) ***Sharpness***

Even though it is often treated to be similar to the spectral centroid, sharpness is computed based on the specific loudness instead of the spectrum magnitude. The sharpness of a sound increases as the strength of the high frequencies of the spectrum increases (Zwicker & Fastl, 2013).

2.5.3.2 Spectrum Shape-Related Perceptual Frequency Features

Spectrum shape is considered one of the popular and widely used approaches in MIR. The relevant set of spectrum-shape-related features is listed below.

(a) ***Bandwidth***

Bandwidth is also called a centroid width. It shows the weighted average of the deviations between the spectral components with the spectral centroid (Wold et al., 1996). It is the second-order statistic of the spectrum which could distinguish tonal sounds and noise-like sounds. Bandwidth can be defined from the logarithmic approach or the power spectra (Liu et al., 1997). Alternatively, it could also be computed from the entire spectrum or within the spectrum subbands (Ramalingam & Krishnan, 2006). According to the MPEG-7 standard, bandwidth is defined as the audio spectrum spread (ASS) which is obtained by computing the standard deviation of the signal spectrum.

(b) ***Spectral roll-off point***

Spectral roll-off point is defined as the *N*% percentile of the power spectral distribution. *N* is set at the 95th percentile by (Scheirer & Slaney, 1997). It's a measurement of the skewness of the spectral shape.

(c) ***Spectral flatness***

Spectral flatness measures the flatness of the frequency distribution of the power spectrum. It is calculated by taking the ratio between the geometric and the arithmetic mean of a subband in the power spectrum (Ramalingam & Krishnan, 2006). Spectral flatness can differentiate between noise-like sounds and tonal sounds. Noise-like sounds and tonal sounds are high and low in ratio, respectively. This is beneficial in the musical instrument classification task.

(d) ***Spectral crest factor***

This feature is the contrast of spectral flatness. The spectral crest factor measures the spikiness of the power spectrum. It can be obtained by finding the ratio of the maximum power spectrum and the mean power spectrum of a subband. Opposite to the spectral flatness, noise-like sounds will show a low spectral crest factor while tonal sounds give a higher spectral crest factor. Eronen and Klapuri applied crest factors in their research on musical instrument classification (Eronen & Klapuri, 2000).

(e) ***Entropy***

Another measurement of spectral flatness is entropy. It is used in measuring the noisiness of the audio signal. Shannon entropy is usually computed in different subbands (Ramalingam & Krishnan, 2006).

(f) ***Spectral slope***

Spectral slope is a measurement of the inclination of the spectrum shape by applying the linear regression method (Morchen et al., 2005).

(g) ***Spectral skewness and kurtosis***

Spectral skewness is defined as the asymmetry of the spectral distribution around the spectral centroid. Spectral kurtosis, on the other hand, tells the spikiness of the frequency spectrum. The value of spectral kurtosis is high if the spectrum is spikier and low if it is flatter (Klapuri & Davy, 2007).

2.5.3.3 Tonality-Related Perceptual Frequency Features

The review by Alías et al. (2016) put the features under the tonality category differently from the review by Mitrović et al. (2010). According to Alías et al. (2016), tonality features are related to the fundamental frequency which is defined as the lowest frequency of the stationary harmonic sound signal. Tonality describes the structure of the sounds that constitute the fundamental frequency and its partials. Tonality-related features that are widely used in musical instrument classification will be listed and discussed below.

(a) ***Fundamental Frequency***

Denoted as "F nought" or F0, the estimation of fundamental frequency could be done with several approaches, such as spectral methods, autocorrelation methods, or cepstral methods. In the review by Mitrović et al. (2010) and some other literature, the fundamental frequency is denoted as the pitch of the audio signal. Work by Eronen (2001) extracted fundamental frequency as a feature in instrument recognition.

(b) ***Harmonicity***

Also called partials, harmonics are the integer multiples frequencies of the fundamental frequency. They are often denoted as F1, F2, F3, etc. Harmonicity features can distinguish between periodic and non-periodic sound signals and are commonly employed in recognizing musical instruments. There are two measurements of harmonicity according to the MPEG-7 standard. The first one is the Harmonic ratio which measures the proportion of harmonic components in the power spectrum. The other one measures the upper limit of harmonicity (ULH) which estimates the frequency beyond the spectrum that no longer contains harmonic structure (Zhang & Zbigniew, 2007).

(c) ***Inharmonicity***

Fundamental and its subsequent harmonics may not always show perfect harmonicity (integer multiples of F0) in real situations. The actual location of the harmonics may deviate away from its ideal location. This is called inharmonicity and is one of the features extracted in musical instrument timbre classification (Agostini et al., 2003).

(d) ***MPEG-7 Spectral Timbral Descriptors***

Several features are closely related to the harmonic structure of the sound according to the MPEG-7 standard. They are found to be suitable for the discrimination of musical instrument sounds. The features are harmonic centroid, harmonic deviation, harmonic spread, and harmonic variation. The harmonic centroid is the amplitude-weighted average of the harmonic frequencies which is related to the sharpness and brightness. Harmonic deviation measures the deviation of the harmonic peaks from their neighbouring harmonic peaks. The harmonic spread is the power-weighted root-mean-square deviation of the harmonic peaks obtained from the harmonic centroid. It is related to the bandwidth of the harmonic frequencies. Lastly, harmonic variation describes the correlation between the two adjacent harmonic peak amplitudes. It represents the harmonic variability of the harmonic structure over time. The application of these features could be found in the work by (Deng et al., 2008).

(e) ***Jitter***

Jitter determines the deviations of the cycle-to-cycle fundamental frequency. Barbedo and Tzanetakis (2010) in their work on the classification of musical instruments describe jitter as the measurement of the stability of the partials over time.

2.5.3.4 Chroma-Related Perceptual Frequency Features

The chroma-related feature is considered as the perceptual feature by Mitrović et al. (2010) and is mainly used in musical information retrieval as it could describe the octave invariance of the sound signal. Chroma is normally ranged into 12 pitch classes, with each class one note of the twelve-tone equal temperament (Shepard, 1964). Two notes with a separation of one or more octaves are said to have the same chroma. The same chroma means that the notes will produce the same effect on human auditory perception.

(a) ***Chromagram***

Chromagram is computed from a logarithmic Short-Time Fourier Transform to the spectrogram that represents the energy of the 12 pitch classes. It maps all spectral audio information into one octave which results in spectral compression. This could be used in describing the harmonic musical sound signals.

2.5.3.5 Loudness-Related Perceptual Frequency Features

Loudness is one of the perceptual features that the human auditory system can sense in listening to the sound signal. Loudness-related perceptual features aim to simulate human hearing ability in the audio retrieval system. Peeters et al. defined loudness as the subjective impression of the sound intensity (Peeters, 2004).

(a) ***Loudness***

Loudness is computed from the normalized power spectrum of the input frame which subtracts an approximation of the absolute threshold of hearing. It is then filtered by gammatone filter banks and the frequencies across are summed to obtain the power of each auditory filter. These powers which represent the internal excitations will be compressed, scaled, and summed across the filters to extract the loudness estimation (McKinney & Breebaart, 2003).

(b) ***Specific Loudness Sensation***

Specific loudness sensation is a measurement of loudness in a sone unit. Sone units are defined as a perceptual scale for loudness measurement according to Peeters (2004). Pampalk et al. (2002) computed this feature by merging the spectral masking effect and the Bark-scale frequency analysis.

2.5.3.6 Roughness-Related Perceptual Frequency Features

Roughness is a fundamental hearing sensation that measures the sensory dissonance of sound signals. According to Daniel and Weber (1997), the amplitude variations which change rapidly will cause unpleasantness and reduce the noise quality, hence deducing that the sound is rough. The computation of roughness can refer to the work by McKinney and Breebaart (2003) and Zwicker and Fastl (2013). The application of roughness as a feature in musical instrument classification can see (Barbedo & Tzanetakis, 2010).

2.5.4 Wavelet Features

The application of wavelet is based on the division of the continuous-time signal or given function into different scale components (Alías et al., 2016). Wavelet transform can extract the desired time-frequency components of the musical sound signal. The wavelet is decomposed into sub-bands which will be further analysed. The characteristics information of the particular musical sound signal can then be obtained. According to Mallat (1989), compared to the Fourier transform, the wavelet transform has advantages in showing functions consisting of discontinuities and sharp peaks. It is also good for constructing and deconstructing finite non-stationary signals.

2.5.4.1 Daubechies Wavelet coefficient histogram features

Proposed by Li et al. (2003) in their study on music genre classification, Daubechies wavelet coefficient histogram is applied by decomposing the audio signal by Daubechies wavelet. Histograms are built from the wavelet coefficients obtained for each subband. The histograms estimate the waveform variation of each subband. Wavelet features are obtained by computing the first three statistical moments and the energy of the coefficients subband.

2.5.5 Cepstral Features

Introduced by Bogert (1963) with the concept of "cepstrum", cepstral features represent the smoothed frequency based on the logarithmic magnitude. It was first employed in speech analysis by Davis and Mermelstein (1980) and is now widely used in various fields of audio information retrieval.

2.5.5.1 Perceptual Filter Bank-Based Features

Perceptual filter banks-based features are computed based on the cepstral domain. The sound signal is first Fourier transformed; the magnitude is then converted into the logarithmic scale. Discrete Cosine Transform will be performed on the previous result to decorrelate the output data.

(a) Mel-frequency cepstral coefficients (MFCCs)

Also called MFCC, this feature is very well known in automatic speech recognition and audio content classification. The mel scale is a perceptual scale of pitches judged by listeners to be equal in distance from one another. To extract MFCC features, the audio signal is first framed into short frames, and the periodogram estimate for each frame is computed. The power spectra are then mapped onto the mel scale, which approximates the human ear's response more closely than the linearly spaced frequency bands used in the normal Fourier transform. The mel frequencies emphasize frequencies that are more relevant to human perception, particularly in the range where human hearing is most sensitive. After mapping to the mel scale, the energy in each filter is summed. The filterbank energies are then logarithmized and decorrelated using the Discrete Cosine Transform (DCT). Only 8-13 DCT coefficients will be used to represent the spectral shape of the audio signal. The first DCT coefficients represent the spectrum’s mean power. The second coefficient represents the spectral centroid. Higher-order coefficients are related to spectral details like pitch (Mitrović et al., 2010). Figure 2.6 shows the MFCCs obtained from the flute musical instrument.

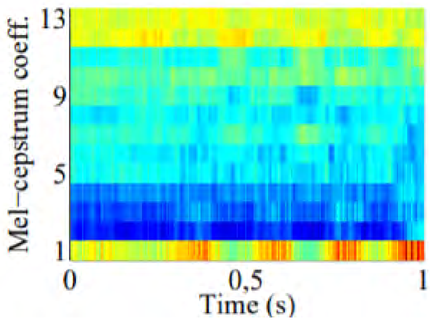


Figure 2.6: MFCCs for the flute musical instrument (Mitrović et al., 2010)

2.5.5.2 Autoregression-Based Features

Autoregression analysis is often used in signal processing. This technique uses linear prediction analysis that can predict the value of every signal sample by the linear combination of previous values (Tremain, 1982).

(a) Complex Cepstrum

According to Oppenheim et al. (1968), complex cepstrum is the inverse Fourier transform of the logarithm of the signal's Fourier transform. Application of the complex cepstrum on musical instrument recognition can be read in the work by Brown (1999).

(b) Linear Prediction Cepstral Coefficients (LPCC)

Linear prediction cepstral coefficients are the alternative for linear prediction coefficients (LPC) discussed earlier above. They are obtained by the inverse Fourier transform of the log magnitude frequency response of the linear prediction spectral envelope (Wu et al., 1997). In comparison to LPC, LPCC is more robust in representing the spectral envelope.

2.6 Classification

After feature extraction and selection, classifiers are used in the machine learning system to classify the isolated musical sounds into the instrument and its family. In this section, several techniques commonly used in automatic musical instrument classification will be discussed. It is worth noting that the accuracy or effectiveness of the classifier is affected by many factors (number of samples, combination with different features, number of samples used in the testing phase and training phase, etc.). Therefore, the classifiers in the following paragraphs will not include the accuracy obtained by each piece of literature reviewed in this paper.

2.6.1 K-Nearest Neighbours

Also denoted as KNN, this classifier is one of the popular machine learning algorithms. In the training phase, it will store the feature vectors from all the training samples and then use them in classifying the new test samples. By referring to the set of k nearest training samples in the feature set, the new sample will be assigned to the class with the most examples in the set. The system uses the Euclidean distance measurement method. Details of how the classification process goes can be referred to Figure 2.7 below.

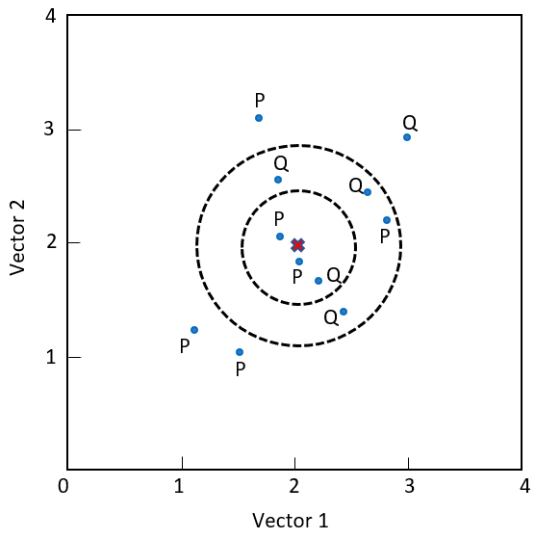


Figure 2.7: Design of the KNN technique

From Figure 2.7, the cross is the target of the classification. If $k = 3$ is selected (the inner circle), then the cross would be categorized as class P as the three nearest neighbours to that cross are mostly from class P. However, if $k = 7$ is selected (outer circle), the cross would now be categorized as class Q with Q being the majority neighbour.

k-Nearest Neighbour is a simple algorithm that is widely used in the automatic machine learning system, but some downsides are to be considered when implementing this technique. According to Mitchell (1997), this algorithm is lazy and requires storing all the training samples in the memory to generate a decision for the new sample. It is also highly sensitive to the irrelevant features which could dominate the distance metrics. Heavy computational load is another drawback of this algorithm.

2.6.2 Support Vector Machine

Another popular classifier used is the support vector machine (SVM). It is based on the statistical learning theory developed by Vapnik (1998). The working principle of SVM is looking for the optimal linear hyperplane which gives the lowest generalization errors when classifying the unknown test sample. The linear hyperplane is mapped so that the margin between the different categories is separated as wide as possible. It serves as the borderline between the categories. The new test samples could be categorized based on which side they fall on when they are mapped into space. The hyperplane is a linear line where the features can be separated into 2 dimensions. It will become a 2D plane when it is displayed in three-dimensional space. This approach can be used when a linear hyperplane can't separate the data in 2-dimensional space and requires higher dimensional space to do so. This is achieved by applying the so-called "kernel trick" as illustrated in Figure 2.8. The kernel trick transforms the low dimensional input space to a higher dimensional space so that the segregation (hyperplane) could take place.

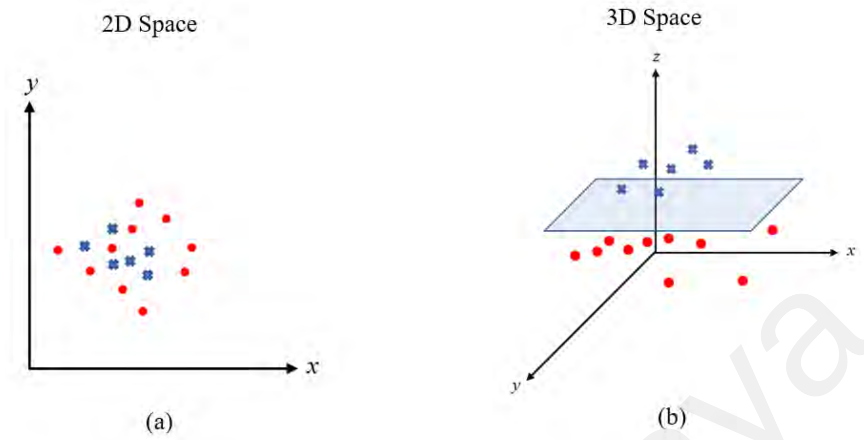


Figure 2.8: (a) Inseparable data in 2D space (b) Hyperplane separating the data in 3D space with the kernel trick.

Even though SVM is a popular algorithm used by many in their research, SVM still gives some drawbacks. In the multiclassification task, SVM needs to perform a series of interconnections between the classes. Computation-wise, it is an intensive process to work on. Also, there is a risk of selecting the less optimal kernel function during the process.

2.6.3 Decision Trees

Decision trees have been pervasively implemented in classification tasks and machine learning systems. This technique attempts to focus on the relevant features and abandons irrelevant ones in the construction of the tree. A decision tree is built top-down that begins with the most informative root node as shown in Figure 2.9. Usually, two branches will split from the root which represents different descriptor values or attributes. Each node in the tree represents the test of the samples' attributes, and the descendant node represents the result of the test. The complete tree is built by repeating the training process recursively with the training samples. After that, pruning work will be carried out to avoid overfitting. The decision tree is commonly used in supervised learning methods which produce high accuracy, and stability, and are easy to interpret.

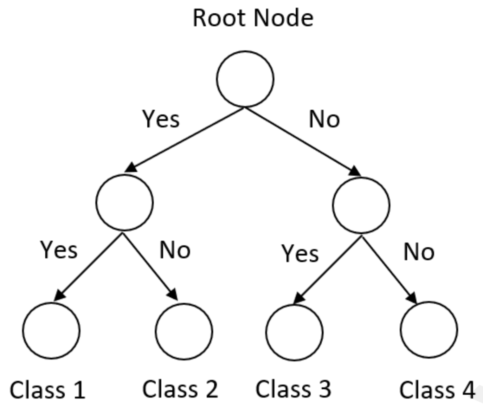


Figure 2.9: Decision tree in classifying 4 classes of musical instruments.

In the musical instrument classification task, the decision tree can also help in identifying the best feature in discriminating instruments. From the literature, the common decision tree algorithm used is J48 or also known as C4.5. This algorithm is also called a statistical classifier which is developed by Quinlan (2014).

2.6.4 Naive Bayesian Classifiers

Naïve Bayesian Classifier (NBC) is the classification technique based on the Bayes' Theorem. This technique uses a conditional probability model in the prediction of classes. Naïve Bayes classifier assumes that the classification features are independent, hence it is called "naïve". Like the other classifiers, NBC will be trained by collecting enough training samples. The probabilities of different classes and features will be obtained by counting the frequencies of their occurrence in the training phase. A new sample can then be classified based on conditional probabilities. NBC is one of the easy and fast algorithms one can use in the classification task. It requires less training data and is not computationally intensive. However, this algorithm is known as a bad estimator. It is also too "naïve" by assuming that the features are completely independent in real life. Deng et al., in their works on the feature analysis for musical instrument classification, used the NBC technique as one of the classifiers (Deng et al., 2008).

2.6.5 Artificial Neural Networks

Artificial neural networks (ANN) are inspired by biological neural networks. It is constructed based on a large collection of interconnected artificial neurons. These neurons are arranged into layers in which the transmission of signal happens from the input layer to the output layer by the connection called edges. These edges have a weight that tells the strength between the connecting layers. The weight may change during the learning process. With sufficient training samples, the network becomes capable of predicting the outcome from the input. This learning process can be done either supervised or unsupervised. The prediction accuracy of ANN gets better when more examples are processed. It keeps on learning and refining the weight of every sample processed. Implementation of ANN in musical instrument classification can be found in (Kaminsky & Materka, 1995).

2.6.6 Hidden Markov Models

Abbreviated as HMM, Hidden Markov Model is a statistical Markov model that contains two components. The first is a set of hidden variables that is unobservable directly from the data while the second is another set of variables that are conditional on the first set of hidden variables (Herrera-Boyer et al., 2003). HMM is used in predicting a sequence of hidden variables from a set of observed variables. This allows the model to generate a random measurement in each state from a variety of distributions.

2.6.7 Gaussian Mixture Models

The Gaussian mixture model is a probabilistic model representing the subpopulation that is normally distributed within the overall population. Without needing to know which data point belongs to, this allows GMM to automatically learn the subpopulation.

This model can do the clustering of groups of data mixed. This is done by the computation of the three parameters which are the mean, covariance, and the mixing probability of the Gaussian mixture. Due to this, GMM is unsupervised learning. This classifier has been used in speech recognition, image pattern recognition, and musical instrument classification. GMM is one of the popular classifiers used intensively in instrument classification. For instance, refer to (Marques & Moreno, 1999).

2.6.8 Discriminant Analysis

Discriminant analysis is a technique used in machine learning to find the linear, quadratic, or logistic functions of the features that characterize or separate samples into two or more predefined classes. Discriminant analysis is related to the multivariate analysis of variance (MANOVA) and regression analysis. This technique could determine the most discriminative features of each class and the most similar or dissimilar classes. Martin and Kim (1998) used linear discriminant analysis in their research on musical instrument identification.

2.6.9 Higher-Order Statistics

Higher-order statistics (HOS) is the technique that uses the sample function with cubic power or higher. Conventional techniques (lower-order statistics) are functions with constant, linear, or quadratic functions. Mean and variance are examples of lower-order statistics. HOS in the analysis of musical signals used skewness and kurtosis as the estimation of the shape parameters.

2.7 Chapter Summary

This chapter gives an overview of the literature and efforts made by past researchers on Sape musical instruments. Most of the works were conducted on exploring and understanding the music, repertoire, and culture. In more recent works, the studies focused on the sound and vibrations of the Sape. This study aims to measure the sound quality or timbre of the Sape musical instrument, hence the quality of the musical instrument and the factors influencing the musical instrument is included in this chapter. The works done by previous researchers on the quality of musical instruments often involved the perceptual opinions of the instrument's experts. The studies desire to be able to provide proper guidance in practical music making which will benefit the instrument players, makers, teachers, as well as engineers and the designer (Campbell, 2013). The factors considered by the instrument experts are also very important. The instrument users are the ones familiar with the instrument and can provide useful input for the researchers. Past works of literature showed that there are a few factors to be considered which include shape, size, materials, environment, etc.

The chapter continued with a reviewing of the literature on MIR and sound recognition in which the general process of sound recognition is explained. The two important steps in the process, which are feature extraction and classification, are reviewed. The application of the MIR in classifying musical instruments into different families or individual instruments is gaining wide interest from researchers and musicians. Different approaches have been used and they harvested different results. The effort is to obtain the best feature set which contains either individual or a combination of temporal, spectral, cepstral, and other properties of sound in the classification task. Choosing a good classifier is also important in that it can better identify the subtle characteristics of different instruments or families. Significantly, this body of literature lays the foundation for the realization that machine learning can play a pivotal role in the classification of sound quality, providing a promising avenue for future exploration.

From the literature review, it can be observed that various approaches in the features and classifiers used in classifying monophonic instrumental sounds were discussed. While the review is not exhaustive, it is apparent that there is no specific feature or classifier which can be considered the best in the musical instrument classification task. Most of the works in the literature review are on the comparison of the accuracy obtained by the different combinations of features and classifiers. Different combinations display varying accuracies, showcasing both advantages and disadvantages. The selection of the appropriate features or classifier is dependent on the specific task of classification. For instance, the complexity of the learning phase, database size, real-time limitations, etc. However, it can be concluded that fewer features used in the sound recognition system will usually achieve better accuracy and reduces the computational burden (Yang et al., 2022).

It can be noticed that certain literature worked on traditional musical instruments. These efforts sparked the interest of this study in working towards the classification of the sound of the local traditional musical instrument, the Sape. However, in this thesis, the research interest is not in the classification of the individual instrument or families.

To the best of the author's knowledge, the ability of the sound classification system in identifying the sound quality produced by the traditional musical instrument is yet to be explored. It is agreeable that "no instruments are 100% alike" and this is true in the case of Sape. Hence, the quality of each instrument might differ from one another, and this is something worth studying. The subtle differences in sound quality might create another tough challenge in the MIR field, but it is worth exploring and this study attempts to fill in the gap.

CHAPTER 3: DETERMINATION OF THE QUALITY FACTORS OF SAPE THROUGH QUALITATIVE AND QUANTITATIVE APPROACH

3.1 Introduction

Assessing the sound quality with the expertise of Sape musicians holds paramount importance in the crafting process. Nonetheless, a standardized evaluation framework tailored for skilled Sape players remains conspicuously absent. This observation finds its origins in insightful dialogues conducted with experienced Sape experts. These discussions underscore a discernible void within the assessment procedure. Importantly, this issue extends beyond the realm of Sape, encompassing musical instruments more broadly. The challenge lies in the inherently subjective nature of musical instrument evaluation, which can diverge among experts. For instance, in the realm of Chinese traditional music, assessing musical instruments involves multiple expert musicians quantifying attributes like brightness, smoothness, and harmony, based on national benchmarks. However, this approach has become less tenable due to the waning number of Chinese music experts (Li et al., 2019).

The existing body of literature on musical instrument sound quality reveals the concerted efforts of researchers, acousticians, and musicians to establish connections between the subjective and objective properties of all kinds of musical instruments. Such analyses have been applied to various musical instruments, such as piano, trombone, French horn, trumpet, flute, cello, oboe, electric guitar, guitar, and violin (Fritz & Dubois, 2015). Objective properties are usually related to an instrument's physical, mechanical, or acoustic properties (Spycher et al., 2008). On the other hand, subjective properties are related to perceptual or psychological input by instrument makers and/or players (Fritz & Dubois, 2015; Schmid, 2015). The review by Fritz and Dubois (2015) highlighted a state-of-the-art method for correlating instrument sound quality for evaluation by experts. Such studies require a perceptual opinion of reputable makers and/or players. Thus, perceptual opinion plays an important part in an evaluation as it reflects the subjective properties needed in assessing musical instrument quality.

Considering the comprehensive literature review presented in Section 3.2, which extensively explores the Sape musical instrument and includes key studies by Wong et al. (2017), Wong and Dayou (2019), and Wong et al. (2022) focusing on Sape instrument quality, there is currently no existing research, to the best of the authors' knowledge, that effectively incorporates the insights of Sape experts. While prior investigations have predominantly emphasized the objective aspects of Sape sound quality, the subjective dimensions have remained underexplored. To address this gap, our study's primary objective is to identify the core factors significantly influencing the sound quality of the Sape. Diverging from previous approaches, our research actively engages Sape experts, employing both qualitative and quantitative methodologies. By gathering insights from instrument players and makers through methods such as questionnaires and focus group discussions, we aim to provide a comprehensive perspective on the Sape's sound quality. Ultimately, our research seeks to capture diverse subjective viewpoints, with the overarching goal of advancing the instrument's quality and refining its manufacturing processes.

The remainder of this chapter is organised as follows: Section 3.2 presents a background study concerning the Sape, traditional musical instruments and machine learning. Section 3.3 explains the methodological approach used in this study to analyse and determine the Sape’s significant sound quality factors. Section 3.4 presents and discusses the results of the analysis. Finally, Section 3.5 offers the study’s concluding remarks.

3.2 Background Study

3.2.1 Previous Studies on Sape, Traditional and String Musical Instruments

The study of the traditional musical instruments in the Southeast Asia region is not new. For instance, Slamet and Kusumaningtyas (2020) and Slamet et al. (2021) studied the Indonesian traditional musical instrument, Gamelan. The studies compared the effects of forging and post-cast heat treatment, and tin composition on the microstructure and mechanical, and acoustic properties of gamelan. Aditanoyo (2018) studied another Indonesian traditional musical instrument, Angklung. His research focused on the correlation between the acoustic characteristics and mechanical vibrations of Angklung, a popular traditional Indonesian bamboo musical instrument, using measurements of force, acceleration, and sound pressure to calculate mechanical admittance and acoustic sound pressure level, and finding that the tubes have inharmonic frequency overtones. Apart from that, a study on the assessment of the Kledi musical instrument is done by Ghozali (2018) using qualitative methods and concludes that it is made up of natural materials and has a homophone playing technique with a distinctive tone and that its enculturation process occurred through informal means.

In Malaysia, several studies were conducted on traditional musical instruments. Batahong et al. (2018) examined the anatomy, physical, and acoustical properties of Sundatang, a traditional musical instrument of Rungus and Kadazandusun Ethnic groups in Sabah, Malaysia, by measuring its dimensions, sound frequencies produced from tuned strings, and the effect of fretting on the sound, with the findings providing important information for the instrument's advancement study. Siswanto and Syiddiq (2018) and Siswanto et al. (2018) studied the Malay traditional musical instrument, Kompang. Their research focused on the membrane vibration of Kompang in which they investigated the effect of different levels of humidity on the tension and vibration frequency of the goat skin and x-ray film membrane of the traditional musical instrument.

Among the existing studies concerning musical instruments, string instruments, such as guitar and violin, have been among the most researched. String musical instruments are one of the families under the Western classification system of instruments. In this classification system, instruments are divided into the woodwind, string, brass, and percussion families. String musical instruments can be excited by either bowing, plucking, or hammering (Bucur, 2016). Usually, a string instrument is made with a hollow body. The excited string vibration will be transmitted to the instrument body and then to the enclosed air volume, radiating the sound more audibly to the audience and the player.

Sape is a traditional string musical instrument from Sarawak. Its shape is similar to a guitar. It is carved from a single bole of wood, and the body is hollowed out and functions as a resonator. The hollowed sound cavity is at the back of the body as opposed to the guitar. Originally the Sape's strings were made from the sago tree, but now these have been replaced by steel strings. The traditional Sape has three or four strings as shown in Figure 1.2, while the contemporary Sape has five or six strings. The additional strings on the contemporary Sape allow for a wider range of notes and chords to be played, making it a more versatile instrument than its traditional counterpart. The traditional Sape has an important place in the cultural heritage of Sarawak, the contemporary Sape's additional strings and design elements make it a more versatile instrument that can be used in a wider variety of musical genres.

Over the past few years, several studies have been conducted on the Sape, covering a diverse range of topics, including an introduction to the instrument (Hartanto et al., 2021; Lim et al., 2020), an investigation into its evolution and transformation over time

(Lim & Abdul Rahman, 2016), and an exploration of melody, harmony, and rhythm in Sape repertoires (Lin & Lim, 2021; Sudom & Naili, 2022).

Additionally, investigations are focusing on the effects of fret setting, string tension, and plucking force on the fundamental frequency of the Sape (Wong et al., 2022; Wong et al., 2017; Wong & Dayou, 2019). These studies have shed light on the influence of fret configuration and string tension on the generated fundamental frequency while indicating that plucking force has no impact on this aspect.

Despite the wealth of research on various aspects of the Sape, the authors have not come across studies specifically examining the factors influencing its quality. Addressing this knowledge gap, the present research employs qualitative and quantitative approaches to identify the significant factors that contribute to the quality of the Sape. The study aims to contribute to the broader understanding and preservation of the invaluable local cultural heritage of Sarawak, Malaysia.

3.2.2 Timbre Quality of Musical Instruments

The timbre quality of musical instruments is a crucial aspect of auditory perception, often referred to as "tone colour." It distinguishes instruments playing the same note at different volumes, going beyond simple pitch and loudness (ANSI, 1973). The harmonics, overtones, and resonances that define the acoustic identity originate from a complex combination of intrinsic and exogenous elements.

The merger of auditory events produces timbre, a complex auditory property. Its multidimensional perceptual representation has dimensions related to the spectral, temporal, and spectrotemporal characteristics of the audio input. Timbral dimensions are also influenced by an understanding of the sound source's mechanical processes, including shape and material composition. Another viewpoint is provided by spectromorphology, which includes time-varying frequency, amplitude behaviours, and spectrum modulations. The richness of timbre is further enhanced by the relationships it has with fundamental frequency, playing effort, pitch, and dynamic level (McAdams, 2019).

The exploration of timbre quality carries significance across the realms of music, engineering, and craftsmanship. Delving into the interplay between a musical instrument's physical attributes and its perceived timbre offers a valuable avenue for enhancing design and production. For instance, the evaluation of violin sound quality entails comprehensive perceptual assessments by musicians Saitis et al. (2017) and Fritz et al. (2012b), alongside the investigation of correlations between timbral attributes, acoustic properties, and construction parameters (Fritz et al., 2012a).

Beyond the violin, the quest to identify quality parameters for various musical instruments has been pursued. Instruments like the saxophone and clarinet (Gazengel et al., 2010; Gazengel & Petiot, 2013; Pinard et al., 2003), as well as the didgeridoo (Smith et al., 2007), have also undergone comprehensive investigations. Moreover, innovative approaches have surfaced, such as the introduction of alternative methods in wind musical instrument mouthpiece production, accompanied by expert evaluations to gauge quality (Bacciaglia et al., 2020). These collective efforts contribute synergistically to a deeper understanding of the intricate role timbre plays in shaping the quality of musical instruments.

As inferred from Fritz and Dubois (2015) review of the perceptual evaluation of musical instruments, the understanding of how musicians assess instrument quality emerges as a pivotal stride towards identifying physical quality parameters. This comprehension holds the potential to drive enhancements and novel avenues for instrument design and manufacturing.

3.2.3 Factors Influencing Sound Quality in String Musical Instruments

This study's literature review was focused on string musical instruments, which relate to the targeted instrument in this study. A string instrument's sound produced by the strings alone is not loud enough to be audible. Hence, the strings must be coupled with the instrument's body, which acts as a resonator to produce an audible sound. A connecting bridge will transmit the vibration of the plucked or bowed strings to the soundboard. The enclosed air cavity will then amplify the sound produced.

According to Wegst (2006), an instrument's body shape and material greatly influence the sound quality of a string instrument. For example, the design of the sound holes and the shape of a guitar or violin, which is round, make it aesthetically pleasing and prevent stress concentrations created by sharp corners. The soundboard, back, and side plate thickness are important parameters in string instrument design. The thickness of the plate can affect the input impedance and resonance frequency produced, further influencing the sound quality of the instrument (Yoshikawa & Waltham, 2014). The design and coupling between the string and the bridge component also strongly influence the vibration produced, as Jansson (1990) indicated on the violin.

Material-wise, the woods used for soundboards (spruce and maple) usually have a low characteristic impedance. Low impedance is helpful, especially for the sound radiation from an instrument to the surrounding air. According to Gore (2011), the material properties of wood for guitars, soundboards, braces, backplates, necks, and bridges played an important role. The material properties include density, Young's modulus, stability with humidity variation, hardness, and heat bendability. The wood selection for different instrument parts is due to different vibration and radiation characteristics requirements. Softwoods are usually selected as the material for soundboards. In contrast, hardwoods are used for frame boards or backboards (Yoshikawa, 2007; Yoshikawa & Waltham, 2014).

Other than that, humidity also plays a part in the sound quality of an instrument. Researchers have experimentally hypothesised and proved that playing regularly in intermediate or high humidity increases stiffness and decreases the loss coefficient (Hunt & Balsan, 1996). The result was believed to answer common beliefs that regular playing and the wood's age would improve the instrument's quality. Some other examples of the factors studied are varnish (Schelleng, 1968), relative densities of early and late growth layers in wood (Stoel & Borman, 2008), chemical treatments of the wood (Barlow et al., 1988), the role of haptic cue (Saitis et al., 2018), and plate tuning methods (Hutchins et al., 1960).

From the literature above, the objective properties of musical instruments have been studied in depth to understand their effect on the instrument's sound quality. Studies were also conducted to discover the correlation between subjective properties (i.e., psychological evaluations) and objective properties (including but not limited to acoustic properties, mechanical measurements, or physical modelling). Taking acoustic properties as an example, timbre, attack behaviour, loudness, and the degree of possible timbre variation are noted as the four primary features determining the sound quality of an instrument (Bader & Hansen, 2008). Dünnwald (1991) categorised 700 violins into "good" and "bad" classes by measuring the level of the first signature mode from the sound output. Bissinger (2008) performed vibration and radiation measurements on 17 violins and ranked them according to the quality ratings that a professional player and himself provided. The A0 vibration mode and strong radiation were significantly correlated to an instrument's good sound quality. In the studies by Yang et al. (2017a); Yang et al. (2017b) the authors tried to find the correlation between the vibroacoustic properties of the wood used in lute and Yueqin musical instruments with experts' subjective evaluations.

As mentioned by Fritz and Dubois (2015), it is clear that the perceptual evaluation by the instrument makers or players is important in the study of a musical instrument's quality. Understanding how they evaluate the quality of an instrument is one necessary step in the efforts to understand and improve the design and manufacturing of the instrument. From the literature above, it can be seen that the sound quality of an instrument in the perceptual opinions of the experts is not affected by only a single factor. However, little of the literature mentioned above has been examined, which is the most significant factor among all the factors studied. From the expert's point of view, which factor will most significantly affect the quality of the instrument? This study presents data that fills this gap.

3.3 Methodology

The study's design incorporated both qualitative and quantitative research methodologies, specifically integrating focus group discussions (FGDs) and a structured questionnaire. These research tools were thoughtfully selected based on their suitability for delving into the multifaceted dimensions of the subject under investigation (Kline, 2013; Krueger & Casey, 2000).

To elaborate further, the utilization of FGDs deserves an in-depth explanation. FGDs constitute a qualitative research approach wherein small groups of individuals sharing common characteristics convene to engage in detailed discussions about a specific topic. This methodological choice was made due to its inherent capacity for exploring the intricate facets of participants' viewpoints, experiences, and insights related to the focal area of Sape quality. These discussions provided a platform for participants, who were seasoned Sape makers in this case, to articulate their tacit knowledge and contextual factors that influence Sape quality.

Complementing this qualitative component, a structured questionnaire was meticulously designed to procure quantitative data. Questionnaires are recognized for their ability to systematically gather standardized responses from a diverse and broader sample of participants. The questionnaire was designed with precision, encompassing a series of well-crafted questions. This approach enabled the collection of specific, quantifiable information of various aspects of Sape quality. By employing this quantitative approach, the aim was to quantify participants' opinions, preferences, and perceptions, offering a more comprehensive and statistically supported perspective on the subject matter.

It is imperative to underscore that the fusion of both qualitative and quantitative methodologies in this research design was a deliberate choice. The acknowledgement is that Sape quality is a multifaceted phenomenon shaped by various factors, some deeply ingrained in the expertise and experiences of Sape makers. Qualitative methods, such as FGDs, were indispensable for exploring these intricate, context-specific elements. Concurrently, the quantitative data garnered through questionnaires allowed for the generalization of findings to a wider population of Sape enthusiasts, enhancing the comprehensiveness of the study.

In essence, the research design was thoughtfully structured to harness the strengths of both qualitative and quantitative approaches. This comprehensive methodology was strategically selected to ensure a holistic grasp of the factors impacting Sape quality, aligning with the nuanced nature of the research objective.

3.3.1 Focus Group Discussion

According to Thomas et al. (1995), a focus group is a technique involving in-depth group interviews in which participants are selected because they are purposive, although not necessarily representative, sampling of a specific population. Therefore, focus group discussion participants are selected because they know a study area (Burrows & Kendall, 1997). In this study, the focus group discussion was conducted to explore, comprehend, and discuss the issue thoroughly and comprehensively. This method allowed Sape makers to express their thoughts and opinions as it involved an open discussion session.

3.3.1.1 Sample

According to Krueger and Casey (2000), smaller focus groups hold greater potential for gathering insightful data. Therefore, a focus group comprising six to eight participants is recommended. For this study, participants who have at least five years of experience in Sape-making and are currently active in the industry were targeted. They also needed to be well-known within the Sape community. To recruit the participants, a total of 15 invitations were distributed via email, text messages, and social media platforms. The limited size of the Sape maker community was cited as the reason for the low response rate, as just five people responded and agreed to take part. One of the reasons given was that some Sape makers lived in rural places, making transportation difficult. Other logistical problems were also mentioned. There are only about 15 known Sape makers in Sarawak, according to the Sape maker.

Five male participants, ranging in age from 28 to over 70, participated in the focus group, which was held in Kuching, Sarawak, and Kuala Lumpur. All participants identified as expert Sape makers, with two having over ten years of experience and the remaining three having over five years of experience. All participants work full-time as

Sape makers and are highly regarded within the Sape-making community. One of the participants in the focus group has been recognized as a national living heritage by the Malaysian government, indicating his outstanding knowledge and skills in the traditional arts and crafts of the country. In addition, this participant is actively involved in teaching Sape classes, which demonstrates his commitment to preserving and promoting the cultural heritage of the Sape. As a mentor to some Sape makers, this participant has played an important role in passing on the knowledge and techniques of Sape making to the next generation. Overall, the expertise and experience of this participant have enriched the discussions in the focus group and provided valuable insights into the cultural significance and technical aspects of the Sape. The focus group aimed to elicit relevant perceptual and physical attributes concerning the sound quality of the Sape, which would inform the study's questionnaire. To ensure clear communication, the discussions were conducted in English and the national language.

3.3.1.2 Questions

The focus group questions were created based on potential factors influencing the sound quality of the Sape. Adapting several variables from Barbosa et al. (2015), it included: subjective, materials, design, playability, typology of the maker, etc. Several questions under each variable were asked. The questions are presented in Table 3.1.

Table 3.1: Focus group questions

Variables	Questions
Subjective	What factors may affect the quality of the Sape? What is the most significant factor that affects the quality of the Sape? What other things are there that may not have been discussed as influencing factors of the Sape quality?
Materials	Do you think wood type influences Sape quality? What is the best wood for making a Sape? Please rate from 1(best) to 4(worst), the best wood for Sape making from the category of wood below. Softwood Light Hardwood Medium Hardwood Heavy Hardwood Do you think that the age of wood influences Sape quality? The air-dried wood is better compared to the freshly cut wood in making a Sape. What do you think about it?
String or frets	The Sape commonly used steel strings. Therefore, steel-string is better compared to nylon string in the quality of sound produced. What do you think? Do you think that the material of the frets could affect the sound quality produced?
Design	Do you think the design of the Sape influences Sape quality? The dimension or size of the Sape could affect the Sape quality.
Maker	Do you think the quality of the Sape depends on the Sape maker? Experience in Sape making is important in determining the quality of the Sape produced. What do you think about it?
Appearance	Do you think that a Sape with a pleasing aesthetical appearance looks better in quality?
Environment	The playing environment is important as it can make the Sape player plays better. What do you think about it? Do you think that the temperature and humidity of the environment can affect the instrument quality?
Control	The good Sape allows me to control it when I am playing. What do you think about it?
Past experiences	Can you recall when is the last time you make Sape? Have you been making Sape in the past 12 months?
Intention & Recommendation	In the future, do you think that you will keep on making Sape? In the future, do you think that you will change the way/materials in making Sape? Let's say, in the future if there's a system that could automatically determine the sound quality of Sape, will you use it? What do you think the government can do to improve the Sape production quality? What can you do to improve the current Sape production?
Maker typology	How would you classify yourself as a Sape maker? Expert Proficient Competent Advanced Beginner Novice

3.3.1.3 Data Analysis

The analysis of data from FGDs entails a multi-step process that commences with the transcription of recordings and data organization. Subsequently, the transcribed data undergoes cleaning to eliminate extraneous material. Next, the data is coded by assigning labels or codes to data segments, which enables the identification of themes and patterns in the data. After coding, meaningful categorization of the codes into themes follows, facilitating structured data analysis. Ultimately, the themes and patterns in the data are interpreted to derive findings (Krueger & Casey, 2000).

In this research, a comprehensive approach was employed to analyze the information gleaned from the FGDs, which included anecdotes shared by participants. Four analytical tools—word frequency analysis, word cloud analysis, mind map analysis, and anecdotal analysis—were utilized to extract insights from the rich dataset.

The most frequently used words and captivating anecdotes were visualized in a word cloud, major themes and sub-themes were identified in a mind map, and patterns and trends were discovered through word frequency analysis. These three tools, in conjunction with the inclusion of anecdotes, facilitated the identification of prevalent themes and topics in the discussion, the understanding of relationships between various topics, and the detection of patterns and trends in the data. This comprehensive approach enhanced the ability to fully comprehend the outcomes of the discussion, enriched by the unique perspectives and experiences shared by the participants.

3.3.2 Questionnaire

3.3.2.1 Sample

In this study, Sape players and makers actively playing, teaching, and making musical instruments were invited to participate in the survey. The potential respondents were contacted through the Sape community WhatsApp group, Facebook messenger, and a mass mailing list. Of 250 invitations, 48 responded by filling out the questionnaire online, and one person responded manually due to their limited reading ability. The questionnaire response rate was 19.2%.

3.3.2.2 Instrumentation

Sape players and makers were asked to complete the Sape sound quality survey questionnaire online. The questionnaire was set via Google Forms with the topic "Physical and Perceptual Characteristics Affecting the Sound Quality of the Sape". The questionnaire consisted of four sections as shown in Table 3.2. Section A described the characteristics of the primary Sape instrument respondents owned and played the most. Section B collected Sape-playing experiences from the respondents. Section C asked the respondents their opinions concerning the sound quality of the Sape. In this section, 24 questions were asked, and participants were asked to indicate their level of agreement or disagreement. A 4-point Likert scale was employed to elicit stronger viewpoints from respondents and minimize ambiguity in the results, while also reducing indecisiveness among participants. Finally, section D collected the demographic backgrounds of the respondents. The questionnaire was prepared in English and *Bahasa Melayu*.

Table 3.2: Sound quality variables

C##	Sound quality variables	Opinion			
		Strongly disagree	Disagree	Agree	Strongly agree
C1	The wood type used can affect the sound quality of Sape	1	2	3	4
C2	The nice aesthetic appearance of Sape tends to make it looks more quality	1	2	3	4
C3	The string material used can greatly affect the sound quality production of Sape	1	2	3	4
C4	The number of strings used can affects the sound quality of Sape	1	2	3	4
C5	The weight of the Sape can affect the sound quality produced	1	2	3	4
C6	The dimension/size of the Sape can affect the sound quality produced	1	2	3	4
C7	The quality of the Sape is very much dependent on the expertise of the Sape maker	1	2	3	4
C8	Sape with a higher selling price is better in quality	1	2	3	4
C9	The sound quality produced by Sape is very much dependent on the skills of the player	1	2	3	4
C10	The temperature and humidity of the playing environment will affect the playing quality of the Sape	1	2	3	4
C11	The Sape sounds better when I am in a good mood	1	2	3	4

Table 3.2: Sound quality variables (continued)

C##	Sound quality variables	Opinion			
		Strongly disagree	Disagree	Agree	Strongly agree
C12	The playing environment will affect the sound quality of Sape	1	2	3	4
C13	The bad quality of the Sape will limit my playing ability	1	2	3	4
C14	Good Sape can produce good sound quality even with limited playing skills	1	2	3	4
C15	Hardwood is better than softwood in making the Sape as it will produce a better sound quality	1	2	3	4
C16	Steel strings are better than nylon strings in playing a good Sape sound	1	2	3	4
C17	The number of frets will affect the melody and sound quality produced by Sape	1	2	3	4
C18	Sape produced by different makers will give different quality	1	2	3	4
C19	Contemporary Sape sounds better than the traditional Sape	1	2	3	4
C20	The installation of pickup, volume knob, mono jack socket, and earth grounding on the Sape makes it sounds better	1	2	3	4
C21	Sape with diatonic scaling is better than pentatonic scaling as it can play with more pitches	1	2	3	4
C22	The thickness of the Sape body can affect the sound quality produced	1	2	3	4
C23	The materials used for the frets can affects the sound quality of Sape	1	2	3	4
C24	Large Sape sounds better compared to the small Sape	1	2	3	4

3.3.2.3 Data Analysis

Data collected from the online questionnaire were imported into an SPSS (Statistical Package for the Social Sciences) database. In this study, Exploratory Factor Analysis (EFA) (Comrey & Lee, 2013a) and Principal Component Analysis (PCA) were conducted to scrutinise the manifest variables and find groups of latent variables that were highly intercorrelated. Each group represented common underlying factors (Everitt & Skrondal, 2010). The EFA resulted in fewer significant factors than variables. Thus, the EFA reduced the number of variables to reveal the underlying pattern of the variation of variables in the sample concerning the Sape's sound quality.

The extraction of the number of factors followed the Kaiser criterion (Kaiser, 1960). Factors were extracted if the variable had a quality score or Eigenvalue greater than 1. An orthogonal rotation (varimax) was used when the factors were not correlated. However, an oblique rotation (promax) determined the final factor solution if the factors were nonorthogonal and correlated. Oblique rotation was more precise than orthogonal rotation when performing the EFA, as it showed better flexibility in searching for patterns regardless of their correlation (Rummel, 1988).

In this paper, the factor loading cut-off was set at 0.40, following the suggestion by Swisher et al. (2004) that cut-offs between 0.30 and 0.60 should be typically considered. The total factors extracted were about one-fifth as many as the variables. They accounted for at least 60% of the total variance. This outcome explained the reliability of the number of factors extracted in this study. On the other hand, if a variable had more than one substantial factor loading, the variable was retained on the factor with the highest loading.

A reliability test was conducted on the factor sets of the variables to check analytical rigour. Ideally, surveys with Cronbach's Alpha values approaching the value of 1 are reproducible and consistent (Peter, 1979). The Cronbach's Alpha value obtained should show a minimum value of 0.70 for reliability and use in further analysis (Nunnally, 1978). The Bartlett Test of Sphericity (BTS) and the Kaiser-Meyer-Olkin (KMO) Index of Sampling Adequacy were then performed to examine the correlation between the variables. A minimum KMO value of 0.5 is required, and it is even better if it is closer to 1.0, indicating a strong correlation. The BTS was used to test the null hypothesis that the correlation matrix was an identity matrix. A Chi-square signification value less than 0.05 would reject the null hypothesis. Rejection of the null hypothesis means that the variables were related and fit for factor analysis.

3.4 Results and Discussion

3.4.1 Focus Group Discussion

Due to the small community of Sape makers, the focus group was only conducted twice, as explained in Section 3. Five participants were recruited, and they described themselves as experts. The participant demographics are shown in Table 3.3 below.

Table 3.3: Participants demographics

Participant	Age	Gender	Race	Marital Status	Employment	Experience
1	71	Male	Kenyah	Married	Retiree	> 20 years
2	71	Male	Iban	Married	Retiree	> 10 years
3	53	Male	Iban	Married	Retiree	> 5 years
4	31	Male	Kenyah	Married	Self-employed	> 5 years
5	28	Male	Iban	Single	Self-employed	> 5 years

The focus group discussion was recorded and manually transcribed due to the mix of English and *Bahasa Melayu* spoken. These transcripts underwent thorough manual analysis to identify recurring themes and patterns that emerged. Anecdotes, providing a personalized perspective from the participants, were thoughtfully incorporated. Through these compelling anecdotes, a vivid narrative emerged. The seasoned Sape makers collectively echoed a resounding sentiment—wood, the foundational material, reigns supreme in shaping Sape quality.

Across the FGD sessions, participants uniformly emphasized the pivotal role of wood in Sape quality. "Adau" was a name that resonated among the participants as the wood of choice for crafting exceptional Sape instruments. The collective preference for "Adau" was rooted in its characteristics—it strikes a balance between hardness and softness, allowing for easy crafting and ensuring that the Sape sustains its distinct timbre over time. This consensus on wood selection underlines the reverence that Sape makers hold for this specific material.

Beyond wood type, participants acknowledged that the age of wood significantly influences Sape quality. Older wood, they observed, possesses qualities that enhance the Sape's tonal richness and longevity. Furthermore, the participants unanimously agreed that dried wood, as opposed to freshly cut wood, is superior in the Sape-making process. The anecdotal evidence points to the Sape makers' profound understanding of how wood's age and drying process contribute to the overall quality of the instrument.

When it came to wood hardness, participants provided intriguing insights. Their rankings, ranging from light hardwood to heavy hardwood, unveiled their nuanced perception of how wood density affects Sape quality. This ranking reflects their experiential wisdom—a testament to how Sape makers consider even the subtlest variations in wood properties. The narrative woven by the participants in the FGDs underscores the Sape makers' role as artisans with unique preferences and craftsmanship. They emphasized that each Sape maker brings their distinct touch to the instrument, reminding us that Sape quality is shaped not only by the material but also by the maker's skill and individuality.

While wood took centre stage, participants acknowledged the importance of strings, dimensions, and frets. Their varied opinions on these aspects indicate that there might be room for exploration and variation within the Sape-making tradition. The differences in perspectives could stem from individual preferences and regional practices, offering insights into the diversity within Sape craftsmanship. The acknowledgement of environmental factors, such as temperature, humidity, and moisture, impacting Sape quality underscores the holistic approach that Sape makers take. It's a reminder that crafting a Sape isn't limited to the workshop; its surroundings and care post-creation matter. This aligns with the broader understanding of instrument maintenance and preservation.

The anecdotes also unveiled both shared perspectives and divergent opinions among the participants regarding the factors influencing Sape quality. There was a clear consensus among them regarding the significance of wood type, quality, age, and the drying process in shaping the Sape's quality. Similarly, the dimensions, size, and fret configurations of the Sape were collectively acknowledged as pivotal determinants of its quality. However, differences in viewpoints surfaced when it came to aspects like the Sape's design, painting, and the materials used for strings and frets. While some participants did not consider these elements as decisive factors in assessing Sape quality, others argued that they indeed played a role.

In essence, the anecdotes and insights shared by these Sape experts paint a multifaceted picture of Sape quality. They underscore the essential role of wood, aligning with established principles of material selection in instrument craftsmanship. Simultaneously, they highlight the nuanced considerations, creativity, and individuality of Sape makers. The differences in opinions regarding certain factors reveal the richness and diversity within Sape craftsmanship. Overall, this discussion deepens our understanding of Sape quality and its multifactorial nature.

To further analyze the data collected from the focus group discussion, a word cloud was generated as shown in Figure 3.1. The prominent words that emerged from the cloud include "wood," "string," "shape," and "fret." These words provide a summary of the key quality factors that were frequently mentioned by the participants and offer an understanding of the factors that influence Sape quality. The term "wood" is of particular interest, as it was the most prominent word in the cloud. This suggests that the participants placed a high value on Sape quality and that it was a central topic of discussion during the focus group. The other prominent words, such as "string," "shape," and "fret," are related to the construction and design of the Sape and highlight the importance of these factors in determining the overall quality of the instrument.

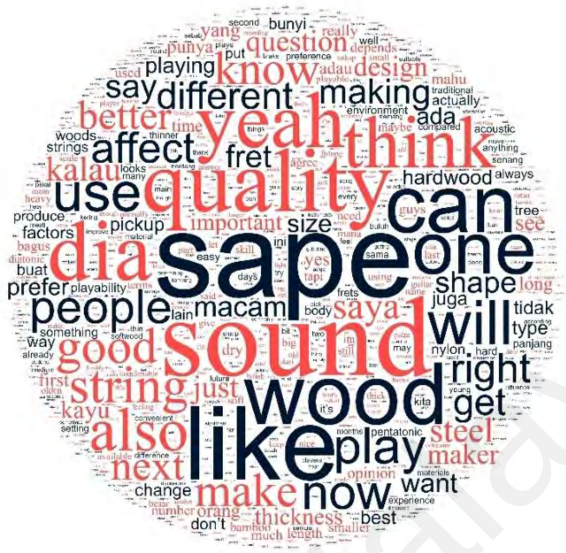


Figure 3.1: Word cloud generated from the FGD transcribe.

Mind mapping was performed to summarise the factors influencing the sound quality of the Sape using the word cloud output. The mind map is shown in Figure 3.2. Several variables emerged: material, design, maker, control, and environment. It was clear from the mind mapping that the construction material was the most significant factor influencing the sound quality of the Sape. According to the participants' excerpts, choosing the correct wood was the most important aspect of making a good quality Sape. To receive this information, the moderator asked, "What is the most significant factor affecting the quality of the Sape?". The responses received were consistent among all the participants who responded that wood was the most significant factor. The second most significant factor was found to be the design of the Sape. For instance, it is discussed that the Sape’s body was made with a thickness of between 1cm to 2cm to get the optimum sound production.

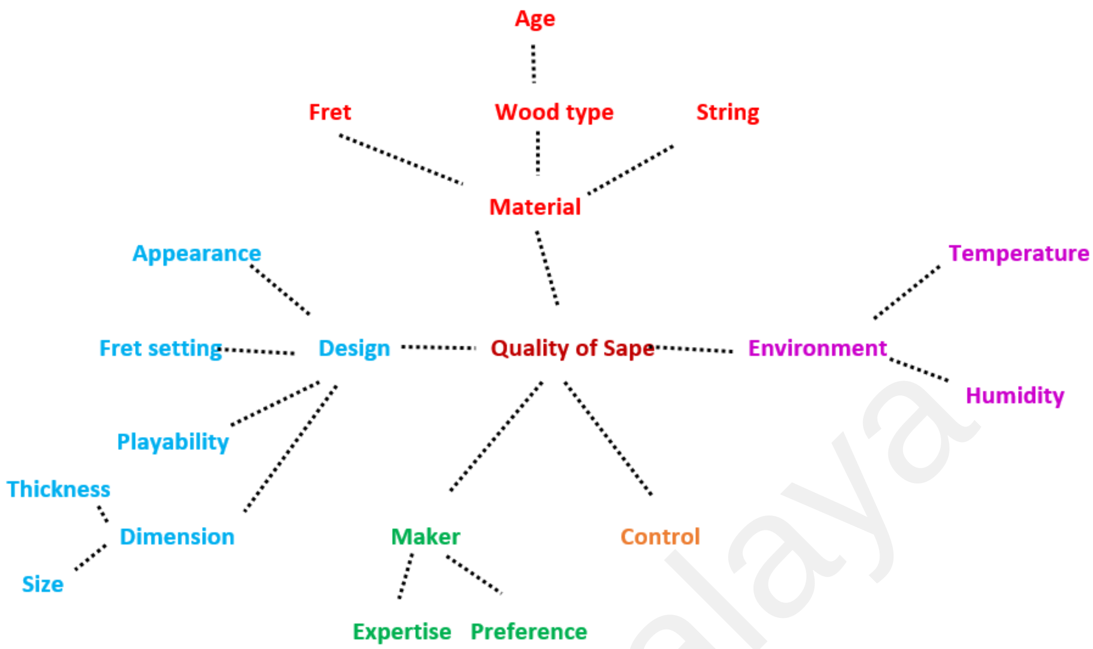


Figure 3.2: Mind mapping from participants' excerpt

3.4.2 Questionnaire

This survey had several known limitations regarding the instrumentation, sampling, and response rate, affecting the study's internal validity. The sampling strategy used in the primary data collection was at the convenience of the authors. The survey was conducted fully online, so the targeted respondents were contacted using online platforms (Facebook Messenger) and text messages (Whatsapp, SMS). The response rate limited the sample size to 48, around 19% of the active Whatsapp Sape community group. A larger sample would have been better regarding the accuracy of population estimates. It is generally recommended that a bare minimum of 10 observations per variable are necessary to avoid computational difficulties (Comrey & Lee, 2013b). However, due to the small Sape community which is spread across the regions in Sarawak without much convenient internet access, the response rate is limited. The demographics of the sample population are included in Appendix A. The population consisted of more males than females. Most of the population were 31 years old and below; they originated from Kuching and Miri (two major cities in Sarawak) and were from the Iban, *Orang Ulu*, or Bidayuh races (major races in Sarawak).

The factor analysis with an orthogonal rotation (varimax) with five factors is shown in Table 3.4. The first, second, third, fourth, and fifth factors accounted for 19.3%, 12.4%, 11.8%, 9.6%, and 8.3% of the variance, respectively. These five factors added up to a cumulative 62% of the variance. The factor analysis was repeated with an oblique rotation (promax), assuming the factors were correlated. However, the factor correlation was found to be non-significant. The highest correlation coefficients occurred between Factors 2 and 4, with a value of 0.141. The result from the orthogonal rotation was therefore retained.

Table 3.4: Total variance explained for orthogonal rotation statistics.

Factor	Eigenvalue	% of variance	Cumulative %
1	3.864	19.321	19.321
2	2.481	12.407	31.728
3	2.357	11.785	43.513
4	1.913	9.567	53.079
5	1.667	8.335	61.415

Table 3.5 presents each variable's factor loading, attribute, and questionnaire item. The factors were interpreted and labelled as Material, Environment, Player/Maker, Design, and Size/Weight. The study characterized Factor 1 as having high loadings on questions related to the material and design of the Sape instrument. The attributes that had significant loadings on this factor included material (C3: 0.787 and C23: 0.541), design (C22: 0.746 and C6: 0.713), and the type of strings used (C16: 0.427), which suggest that the quality of the materials used, the thickness of the Sape body, and the dimension/size of the Sape are critical factors that influence the sound quality produced.

Factor 2 was primarily defined by the attributes related to the environment and the player/maker factors that influence the quality of the Sape instrument. The essential attributes that loaded heavily on this factor were environment (C10: 0.822 and C12: 0.622) and player/maker (C7: 0.779), indicating that the quality of the playing environment, including temperature and humidity, and the expertise of the Sape maker are critical factors that influence the sound quality produced. Factor 3 was characterized as having high loadings on questions related to the player/maker factor and mood. The most important attributes that loaded heavily on this factor were player/maker (C14: 0.764 and C11: 0.763) and the Sape produced by different makers (C18: 0.730), indicating that the skill of the player and the mood of the player are crucial factors that influence the sound quality produced, as well as the quality of the Sape produced by different makers.

Table 3.5: Factor loadings for oblique rotation by attribute and questionnaire item

Factor	Factor Score	Attribute	Questionnaire Item
1	0.787	Material	C3: The string material used can greatly affect the sound quality production of Sape
	0.746	Size/Weight	C22: The thickness of the Sape body can affect the sound quality produced
	0.713	Size/Weight	C6: The dimension/size of the Sape can affect the sound quality produced
	0.575	Player/Maker	C13: Bad quality of Sape will limit my playing ability
	0.562	Player/Maker	C9: Sound quality produced by Sape is very much dependent on the skills of the player
	0.541	Material	C23: The materials used for the frets can affect the sound quality of Sape
	0.427	Material	C16: Steel strings are better than nylon strings in playing a good Sape sound
2	0.822	Environment	C10: The temperature and humidity of the playing environment will affect the playing quality of the Sape
	0.779	Player/Maker	C7: The quality of the Sape is very much dependent on the expertise of the Sape maker
	0.622	Environment	C12: The playing environment will affect the sound quality of Sape

Table 3.5: Factor loadings for oblique rotation by attribute and questionnaire item (continued)

Factor	Factor Score	Attribute	Questionnaire Item
3	0.764	Player/Maker	C14: Good Sape can produce good sound quality even with limited playing skills
	0.763	Player/Maker	C11: The Sape sounds better when I am in a good mood
	0.730	Player/Maker	C18: Sape produced by different makers will give different quality
4	0.708	Design	C17: The number of frets will affect the melody and sound quality produced by Sape
	0.677	Design	C4: The number of strings used can affect the sound quality of Sape
	0.575	Design	C19: Contemporary Sape sounds better than the traditional Sape
	0.449	Design	C2: The nice aesthetic appearance of the Sape tends to make it looks more quality
5	0.659	Size/Weight	C5: The weight of the Sape can affect the sound quality produced
	0.622	Size/Weight	C24: Large Sape sounds better compared to the small Sape
	0.491	Material	C1: The wood type used can affect the sound quality of Sape

Factor 4 was identified as having high loadings on questions related to the design of the Sape instrument. The most important attributes that loaded heavily on this factor were the number of frets (C17: 0.708), strings used (C4: 0.677), and the contemporary design of the Sape (C19: 0.575). These attributes suggest that the number of frets and strings used are critical factors that influence the melody and sound quality produced, as well as the preference for contemporary design over traditional design. Factor 5 was identified as having high loadings on questions related to the weight and size of the Sape instrument. The most important attributes that loaded heavily on this factor were the weight of the Sape (C5: 0.659) and the size of the Sape (C24: 0.622). These attributes suggest that the weight and size of the Sape are critical factors that influence the quality of the sound produced.

The Kaiser-Meyer-Olkin (KMO) measure was found to be 0.51, indicating an adequate sample size for factor analysis. Additionally, Bartlett's Test of Sphericity (BTS) was significant with a p-value less than 0.05, suggesting that the correlations between items were sufficiently large for performing factor analysis. For internal consistency reliability, Cronbach's alpha was calculated for each factor. Alpha values of 0.7 and above are required. However, according to Pallant (2013), for a factor with items lower than 10, an alpha value higher than 0.5 is acceptable. As shown in Table 3.6, factors 1, 2, and 3 exhibited good to acceptable internal consistency reliability, indicating that the items comprising these factors are closely associated and measure the same underlying construct. In contrast, factor 4, which pertains to Design, displayed a slightly lower Cronbach's alpha value of 0.521, yet it remained within the acceptable range.

The Size/Weight factor (factor 5) had the lowest Cronbach's alpha value of 0.386, which falls below the acceptable threshold of 0.5. This indicates that the items comprising this factor are less consistent and may not be measuring the same underlying construct as effectively as the other factors. It is conceivable that some of the items within this factor may not be as relevant or may not fit as cohesively with the other items, which could contribute to the diminished overall reliability of the factor. Another possibility is that the items in factor 5 may not be worded well or may not be relevant to the construct being measured. For example, the item "The weight of the Sape can affect the sound quality produced" (C5) may be seen as somewhat redundant with the concept of size, as heavier instruments tend to be larger. Additionally, the item "Large Sape sounds better compared to the small Sape" (C24) may not be relevant to all Sape players or enthusiasts and may be influenced by personal preferences or biases.

Table 3.6: Cronbach's alpha coefficient for the resulting factors

Factor	#Items in Scale	Factor Name	Alpha	Standardized Item Alpha
1	7	Material	0.743	0.756
2	3	Environment	0.697	0.716
3	3	Player/Maker	0.674	0.704
4	4	Design	0.514	0.521
5	3	Size/Weight	0.355	0.386

This finding suggested that the material of the Sape was the main factor in determining the sound quality of the Sape. Sape are made from many different types of wood; the most popular is Adau wood. It is believed that the very first Sape was made from this wood. Due to the wide variety of wood available in Sarawak, other types of wood have also been widely used, for example, Merbau, Tapang, Meranti (*Shorea spp*.), Jati (*Tectona grandis*), etc. It is known that different types of wood can produce different acoustic and vibrational properties. Yoshikawa (2007) researched the traditional woods best suited for string instruments based on the wood's physical properties. The woods most commonly used in Western instruments fall within the same regression line of vibrational and anti-vibrational parameters. However, the wood used in Western string instruments may not be the same as in the East. In his study, Mulberry wood was widely used to make Japanese traditional musical instruments. The Japanese Biwa did not fall in the same regression line as Western instruments. It was explained that the unique properties of Asian traditional instruments have unique physical properties requirements compared to Western instruments. It is noted that the findings from both the focus group and questionnaire showed that the material or wood used in making the Sape is the most significant factor in determining the quality. This finding is worth further research work on the Sape woods' physical or vibroacoustic properties.

The second most important factor appeared to be environmental effects on the sound quality of the Sape. Two environmental items explained that the Sape's sound quality could be influenced by the surrounding conditions, such as temperature and humidity. This outcome demonstrated that Sape players perceived that they would play the Sape better if the environment were perfect for them. The Sape is usually played as a hobby, during traditional events, performances, and free time. Therefore, such an event's surroundings seemed to play an important role for players. The humidity of the surroundings would also affect the playing quality of the instrument. The feedback from one of the respondents of the focus group mentioned that the humidity would change the moisture content of the Sape, and it would affect the sound quality of the Sape. An environment that is too dry would also dry up the Sape body and it will crack. Therefore, the solution of the Sape maker nowadays is to apply a layer of varnish to the Sape body as a layer of protection against the change in humidity.

Player/maker emerged as the third most important factor in the factor analysis. The sound quality of the Sape in the eyes of the experts depended on the Sape player's or maker's expertise. An experienced Sape maker may produce better instrument sound quality due to their vast experience. A good Sape player may produce good playing quality from the instrument. This outcome suggests that the effects of the expert's skills could be significant in determining product quality. The fourth factor in the factor analysis appeared to be the design effects. The items in this attribute refer to the musical instrument's design, appearance, etc. It is known that different makers will produce different designs of Sape. From the traditional design to contemporary design, three strings to six strings, pentatonic to diatonic frets setting, it appeared that the differences in terms of the design could produce different quality. The fifth factor was identified and named the size/weight factor, as it pertains to the size and weight of the Sape instrument. The findings suggest that the weight and size of the Sape may also have an impact on its overall quality.

The FGD participants and questionnaire results both agreed on the importance of the material used to make the Sape, indicating that the type and quality of wood are crucial factors in determining its quality. Additionally, both sources agreed that the dimensions and size of the Sape have an impact on sound quality and playability and that the setting of the frets is also significant in determining the quality of the Sape. However, the FGD participants had differing opinions on other factors, such as the influence of design and painting on Sape quality, whereas the questionnaire results suggest that design is a factor that influences the quality of Sape. Another area of difference between the FGD participants and questionnaire results is the use of strings and materials for the frets. Some FGD participants suggested using nickel-coated steel strings, while others recommended fishing steel strings made in Japan. Additionally, the materials for the frets suggested by the FGD participants included bamboo, palm tree, and rotan, which was not covered in the questionnaire.

Overall, the findings of both the FGD and questionnaire suggested that the quality of the Sape was influenced by a variety of factors, including the type and quality of wood, the dimensions and size of the Sape, the setting of the frets, and potentially the design and environment in which the Sape was played. These findings suggested that the Sape musical instrument-making industry could have benefited from considering a range of factors beyond just the type and quality of wood used in Sape production. Manufacturers should have also considered the dimensions and size of the instrument, the setting of the frets, and potentially the design and environment in which the Sape was played. Further research could have explored the impact of different materials for strings and frets on Sape quality and examined how regional and cultural differences influenced Sape production. Ultimately, these findings could have contributed to improving the quality and playability of the Sape musical instrument, as well as enhancing the overall musical experience for performers and audiences alike.

3.5 Chapter Summary

This study was conducted to find the most significant factor affecting the sound quality of the Sape musical instrument. An FGD and survey were conducted to achieve the objective of this study. The results will be very useful in understanding the factors that could affect the sound quality produced by the instrument. Different Sape experts may provide different opinions on the research topic. However, it was necessary to determine which factor was the most significant from a majority point of view.

The results from the FGD indicated that the material used in making the Sape played the most important role in determining the quality of sound produced by a Sape. The analysis of the FGD showed that the wood type used in making the Sape was a highly significant factor. The survey extracted four factors from the factor analysis with acceptable internal consistency reliability. The material appeared to be the first factor with a significant influence on the sound quality of the Sape. The focus group and survey showed the same results in which material seemed to be the most important factor in the opinions of the Sape players and makers.

The findings from this study shed light on the quality determination of the Sape musical instrument which has not been fully understood. Due to this study's valid and reliable perceived attributes, future research should also be conducted to examine these factors' usefulness and determine their ability to predict the sound quality of the Sape. As a starting point, it is worth further investigating the materials used in making the Sape as the woods commonly used by the Sape makers consist of softwood, medium hardwood, and heavy hardwood. To understand better, the physical and vibroacoustic properties of the Sape woods will be studied in future studies to find the correlation to sound quality production. The findings may also contribute to the effort in looking for substitute wood in making Sape as certain woods are facing extinction nowadays.

However, it is essential to acknowledge the study's limitations. One notable limitation is the geographical scope, which focused solely on Sape experts from Sarawak, Malaysia, excluding Sape makers and players from Kalimantan, Indonesia. This limitation restricts the broader regional context of Sape craftsmanship and playing in Borneo.

In summary, while this study has provided a crucial initial step in unravelling the factors that shape Sape sound quality, it represents just the beginning of a more comprehensive exploration of this rich musical heritage. The findings presented here offer a foundation upon which future research can build to further illuminate the nuanced intricacies of Sape craftsmanship and its cultural significance.

CHAPTER 4: ASSESSMENT OF WOOD QUALITY FOR SAPE MAKING: VIBROACOUSTIC ANALYSIS AND MACHINE LEARNING CLASSIFICATION

4.1 Overview

Chapter 3 of this study extensively investigated the multifaceted factors influencing the sound quality of the Sape musical instrument. This comprehensive exploration pinpointed wood as the most substantial factor shaping the instrument's sound quality. This chapter extends this exploration by employing a detailed scientific approach to delve deeper into the influence of wood types on Sape sound production.

Building upon the findings from Chapter 3, this chapter focuses on an empirical evaluation of the quality of three prevalent wood types frequently used in crafting the Sape instrument. The primary aim is to conduct a meticulous assessment of these woods, utilizing physical, vibroacoustic, and timbre-based analyses of rectangular wood samples.

The methodology involves conducting flexural vibration tests to extract essential data on the physical and acoustic properties of the wood samples. Additionally, objective sound quality parameters will be derived from the collected sound data to determine the most influential features that define wood quality. These features will serve as critical inputs in a machine learning framework designed to develop a robust method for classifying wood quality.

This chapter represents a deliberate and structured extension of the preceding chapter's findings, aiming to harness scientific methodologies to further understand the pivotal role of wood in determining Sape sound quality. The outcomes anticipated from this chapter will significantly contribute to the preservation and advancement of Sape craftsmanship, offering valuable insights into maintaining consistent sound quality in this revered musical tradition.

4.2 Introduction

The Sape, a traditional musical instrument of Sarawakian culture, is renowned for its unique and individual character due to its hand-made fabrication process. However, ensuring consistent and high-quality sound across different Sape instruments has become a significant challenge due to the complexity of this process. This study seeks to evaluate the sound quality of Sape instruments to better understand the factors that contribute to their perceived quality. Specifically, the research questions guiding this study are: What are the physical, vibrational, acoustic, and timbre properties of Sape wood that contribute to its sound quality, and can these properties be quantified and used to develop a reliable and accurate method for evaluating Sape sound quality? To answer these questions, a comprehensive evaluation of different types of wood commonly used in Sape instrument construction was conducted. The study measured the physical properties, as well as the vibrational, acoustic, and timbre characteristics of the woods, and analyzed the relationships between these variables and the perceived sound quality of the instruments. The ultimate aim of this research is to provide insights and recommendations that can inform the design, manufacture, and preservation of high-quality Sape instruments and contribute to the preservation of this important cultural tradition.

The quality evaluation of wood used in string musical instruments is a critical aspect that directly impacts the instrument's acoustic properties and overall performance. The anatomical grading of wood used in the construction of musical instruments is a well-established practice, with the final price of the instrument often reflecting the quality of the wood (Dinulică et al., 2021). Research has extensively focused on the acoustical properties of wood, particularly in the context of soundboard wood used in string instruments, such as European spruce, due to its significance in Western classical music (Brémaud, 2012). Additionally, the selection of wood species for making string instruments is a global consideration, with hundreds of wood species available for this purpose (Wegst, 2006). The assessment of resonance wood quality involves the evaluation of physical and histological properties, including density, modulus of elasticity, sound velocity, radiation ratio, emission ratio, and loudness index (Spycher et al., 2008a).

Furthermore, studies have explored the acoustic properties of specific wood species, such as neem wood, which has shown potential for use in the backs and ribs of stringed musical instruments based on its density as a predictor for acoustic properties (Hassan & Tippner, 2019). The thermal modification of resonant wood for string instruments has been investigated, with specific density, damping decrement, and acoustic constant identified as crucial factors for quality evaluation (Danihelová et al., 2022). Moreover, the classification of woods for string instruments distinguishes between soundboard woods and frame-board woods, emphasizing the importance of wood selection for different parts of the instrument (Yoshikawa, 2007). There is also ongoing research into the physical-acoustic properties of various wood species for manufacturing musical instruments, including string instruments like violins and classical guitars (Fedyukov et al., 2019).

Additionally, the acoustic quality of wood has been linked to its vibrational performance, with studies developing classification schemes to discriminate between soundboard wood and frame-board wood traditionally used in string instruments (Yang, Liu, & Liu, 2017). The properties of tropical hardwoods commonly used for fretboards of string instruments have been investigated, highlighting characteristics such as high density, strength, hardness, wear resistance, and dimensional stability (Liu et al., 2020). Furthermore, the influence of wood aging on sound quality in the production of musical instruments has been emphasized, indicating the significance of wood properties in achieving desired sound outcomes (Zoric & Kaljun, 2018). The vibrational and viscoelastic properties of wood have also been identified as essential for obtaining high-quality soundboards in string instruments (Golpayegani et al., 2012).

As far as the authors are aware, no prior research has been conducted on the quality of Sape musical instruments. Consequently, the objective of this study is to address this knowledge gap by evaluating the quality of Sape musical instruments from the materials perspective. Specifically, the study focuses on the evaluation of three different types of wood used in the construction of Sape soundboards. To achieve this objective, free-free flexural vibration is applied to rectangular Sape soundboards, and the resulting data is analyzed to determine various acoustic, vibration, and timbre features. These features serve as objective attributes for the classification of wood types and the grading of sound quality. Furthermore, statistical analysis is performed to identify the most significant characteristics for the automatic classification of soundboard quality. The study also compares the accuracy of different classification models in predicting the sound quality of Sape soundboards. Ultimately, it is expected that the findings of this study will provide valuable insights into the production of high-quality Sape musical instruments, contributing to the preservation and advancement of this traditional art form.

4.3 Materials and Methods

4.3.1 Wood Sample Preparation

In this study, three types of wood commonly used to make Sape, Adau, Tapang, and Merbau were selected to represent light hardwood, medium hardwood, and heavy hardwood, respectively. The botanical names, physical, and mechanical properties such as density and modulus of elasticity (MOE) of the selected woods are listed in Table 4.1.

Table 4.1: Botanical information, physical, and mechanical properties of Adau, Tapang, and Merbau wood (Lembaga Perindustrian Kayu, 2010)

Vernacular Name	Category	Family	Botanical Name	Density $(\mathrm{kg\,mm^3})$	MOE $(\mathrm{N/mm^2})$
Adau	Light Hardwood	Magnoliaceae	*Elmerrillia mollis dandy*	300-705	-
Tapang	Medium Hardwood	Leguminosae	*Koompassia Excelsa*	800-865	17800
Merbau	Heavy Hardwood	Leguminosae	*Intsia palembanica*	515-1040	15400

Adau (*Elmerrillia mollis*), is the vernacular name given by Sarawak natives to the Chempaka wood. Chempaka wood is the Standard Malaysian Name for the timber of the family of Magnoliaceae. This timber is moderately soft and light to moderately heavy with a density of $300 - 705\ \mathrm{kg\,m^{-3}}$ air dry. Adau is classified as light hardwood due to its lightweight and low density. Tapang (*Koompassia excelsa*) which is from the family of Leguminosae is a medium hardwood type. It has the air-dry density ranging from $800 - 865\ \mathrm{kg\,m^{-3}}$. The Standard Malaysian Name for this wood is called Tualang. Its heartwood is red-brown to deep brick-red-brown when fresh and darkens with age to a chocolate brown. On the other hand, Merbau (*Intsia palembanica*) is the Standard Malaysian Name for timber under the same family as Tapang (Leguminosae).

This wood is a heavy hardwood with a density of $515 - 1040\ \mathrm{kg\ m^{-3}}$ air dry. It looks orange-brown when fresh, darkening to brown or red-brown on exposure (Wong, 2002).

It is mentioned in the earlier chapter that no Sape is 100% identical as it is fully handmade by the Sape makers. The complication in the making process makes it impossible for machine production. Therefore, the research on the wood in this study started with the soundboard or the top body of the Sape. For simplicity, the soundboard wood samples are prepared in a rectangular shape. The length, thickness and width of the soundboard followed the common dimensions of Sape that can be seen in the market (refer Figure 1.3). The wood samples were planned by Computer Numerical Control (CNC) machine to the final dimensions of $16 \times 165 \times 700\ \mathrm{mm^3}$ $(\mathrm{radial} \times \mathrm{tangential} \times \mathrm{longitudinal})$. Three samples were prepared for every type of wood giving a total of 9 samples. The woods are kept in the laboratory with $60 \pm 2\%$ relative humidity (RH) and at the 30 ± 1°C temperature. The wood samples are then labelled "A" for Adau wood, "T" for Tapang wood, and "M" for Merbau wood as shown in Figure 4.1.

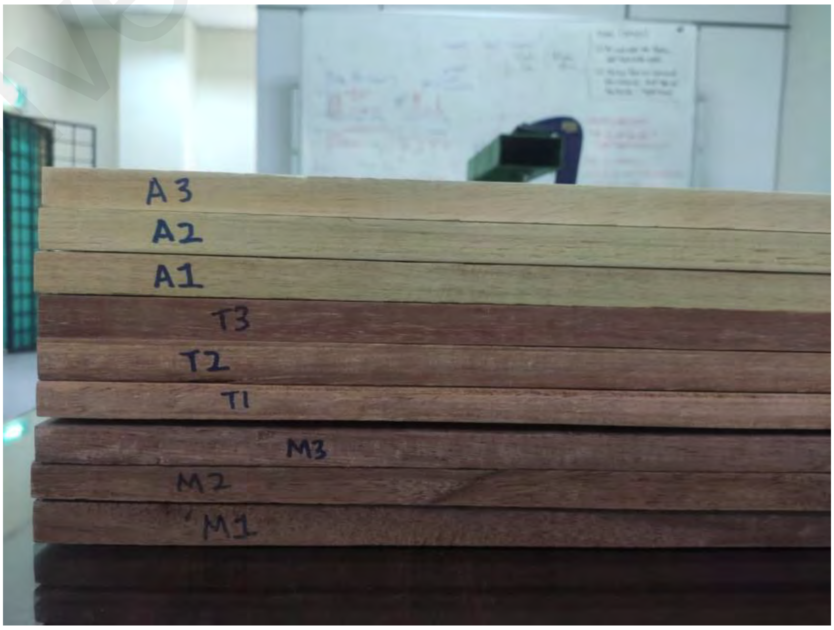


Figure 4.1: Labelled wood samples

4.3.2 Methods

4.3.2.1 Flexural Vibration Test Setup

The quality of a musical instrument is often linked to the acoustic vibration properties of the wood used. In particular, the vibration characteristics are crucial for soundboards of string instruments, as they transfer and radiate the vibration of the strings and create a unique timbre. The important properties of wood vibration characteristics include natural frequencies (f_n), density (ρ), sound propagation speed (c), internal friction ($\tan\delta$), dynamic elastic modulus and shear modulus ratio (E/G), acoustic impedance (z), acoustic radiation damping coefficient (R), and specific dynamic elastic modulus (E/ρ) (Brémaud, 2012; Fletcher & Rossing, 2012).

To assess the vibration properties of the soundboard, a flexural vibration test is performed to determine the soundboard quality. A frame made of an aluminium profile, measuring 85 cm in length, 35 cm in width, and 25 cm in height, with two elastic threads attached, was used to hold the soundboard sample. The sample was placed on the two elastic thread supports to allow for free vibration. Table 4.2 presents the nodal points pertinent to the fundamental node in the flexural vibration test. Positioned at 0.224 times the length from both ends of the sample, these locations align with the nodal points for the fundamental frequency in the flexural vibration test (Roohnia, 2019). The schematic diagram of the experimental set-up is shown in Figure 4.2 and Figure 4.3.

Table 4.2: Location of the nodal position in the flexural vibration test (Roohnia, 2019)

Mode Number	1	2	3	4
	$0.244L$	$0.132L$	$0.073L$	$0.277L$
	$0.776L$	$0.868L$	$0.927L$	$0.723L$

L = length of the specimen

To induce vibration, a stainless-steel ball measuring 15mm in diameter and weighing 13.9g was utilized in this experiment. To ensure repeatability of force excitation, the ball is released after being lifted to a 50° angle from its resting position and hits the wood sample after being deflected by a fixed horizontal rod, which maintains a consistent impact force. The use of a fixed horizontal rod ensures that the ball hits the sample in a consistent manner, reducing variability in the results due to differences in the force of impact. The sound produced by the impact is recorded by a sound level meter positioned 5 cm above the sample. The experimental setup provides a reliable method for inducing vibration and measuring the resulting sound, but it may not capture all aspects of real-world scenarios. The repeatability of force excitation is established through the use of the fixed horizontal rod, which ensures that the ball consistently hits the sample during each impact. The experiment is carried out in the lecture hall at a mean temperature of 20°C with a relative humidity of 54%.

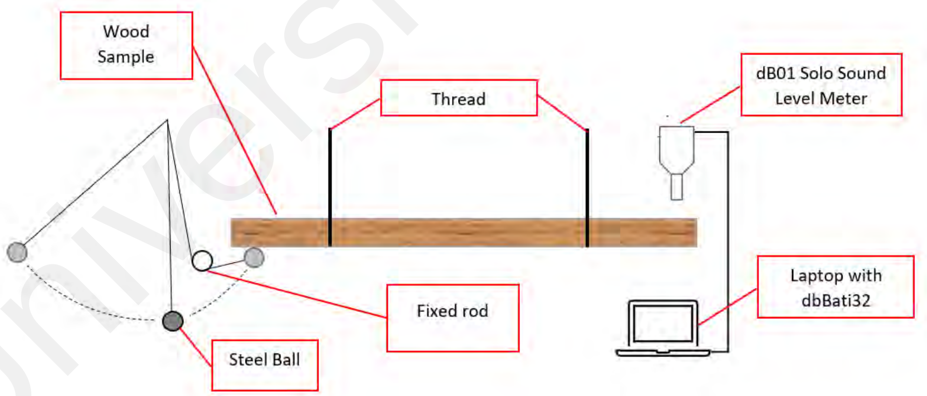


Figure 4.2: Schematic diagram of the experimental set-up

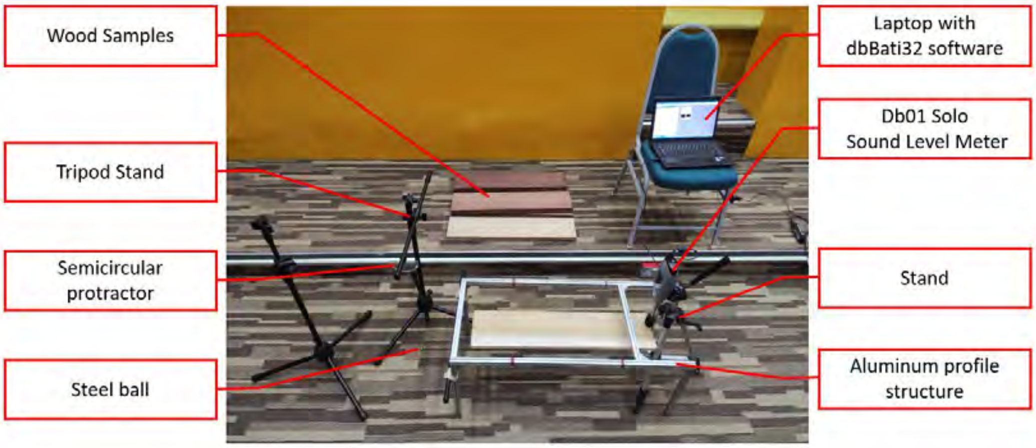


Figure 4.3: Experiment setup for flexural vibration test

The experiment employed an 01dB Metravib Solo data logging integrating sound level meter equipped with an MCE215 microphone transducer featuring a sensitivity of 50 mV/Pa. Before commencing the experiment, calibration of the sound level meter was carried out using a Cirrus CR:515 calibrator. For data capture, the 01dB SLM was linked to a computer via a USB cable, and the dBBATI32 software facilitated the process. The SLM, functioning as a signal transducer, transferred real-time sound signals to the computer upon selecting the record feature in the acquisition tab. Sound data was sampled at 51.2 kHz with 16,384 spectral lines and recorded over a 5-second duration.

Each wood sample underwent 40 repetitions of data collection, leading to a total of 360 sound recordings. Analysis of the recorded data was performed using MATLAB software, employing the fast Fourier transform (FFT) to ascertain resonant frequencies. The experiment took place in a lecture hall with an average temperature of 20℃ and relative humidity of 54%. The sound data collected during the flexural vibration test was stored in WAV format and subsequently analyzed on a computer using MATLAB. Notably, each wood sample generated 40 sound data recordings, with natural frequencies determined through FFT analysis using MATLAB.

Figure 4.4 illustrates the 01dB Metravib Solo data logging integrating sound level meter employed in the experiment, utilizing an MCE215 microphone transducer with a sensitivity of 50 mV/Pa. The SLM was calibrated using the Cirrus CR:515 calibrator as depicted in Figure 4.5, setting the reference value to 94 dB for a frequency of 1000 Hz. The 01dB SLM, connected via USB to a computer, operated in real-time as a signal transducer when the record feature was activated in the dBBATI32 software. The acquisition parameters are detailed in Table 4.3.

Figure 4.4: 01dB solo octave sound level meter

Figure 4.5: Cirrus CR:515 calibrator

Table 4.3: Flexural vibration acquisition parameters

Sample	Sampling frequency (kHz)	Frequency Resolution (Hz)	Acquisition duration (s)
All 9 wood samples	51.2	0.2	5

4.3.2.2 Data Processing

(a) Signal Processing in MATLAB

The sound data collected from the flexural vibration test using the sound level meter is saved in the wav format. The sound data is then transferred to the computer equipped with MATLAB software. From here, the sound file is renamed to the format 'A1 (1)' in which 'A' stands for Adau wood, 'M' stands for Merbau wood and 'T' stands for Tapang wood. The number in the bracket represents the number of tests carried out. There is a total of 40 sound data collected from each wood sample. After performing FFT using MATLAB software, the natural frequency is obtained. The acoustic vibration properties or timbre features extracted in this study followed the literature and are listed in Table 4.4.

Table 4.4: Vibrational, acoustical, and timbre features

No	Features	References
1	Fundamental frequency, f	(Brémaud, 2012)
2	Density, ρ	(Brémaud, 2012)
3	Dynamic Elastic Modulus, E	(Brémaud, 2012)
4	Acoustic Radiation Damping Coefficient, R	(Brémaud, 2012; Yang et al., 2017a)
5	Acoustic Impedance, z	(Wegst, 2006; Yang et al., 2017a)
6	Internal Friction, $\tan\delta$	(Brémaud, 2012)
7	Acoustic Conversion Efficiency, ACE	(Brémaud, 2012)
8	Speed of sound, c	(Wegst, 2006)
9	Spectral Centroid, SC	(Aramaki et al., 2007)
10	Spectral Bandwidth, SB	(Aramaki et al., 2007)
11	Spectral Flux, SF	(Aramaki et al., 2007)
12	Attack time, AT	(Aramaki et al., 2007)
13	Inharmonicity, I	(Aramaki et al., 2007)

i Density, ρ

One of the most crucial factors in classifying wood is the density of the wood samples. According to (Lembaga Perindustrian Kayu, 2010), the density of the wood at a 15% moisture level is a major factor in determining how it is classified in Malaysia. The density range for heavy hardwood typically ranges from 800 to $1120\ \mathrm{kg\ m^{-3}}$, medium hardwood from 720 to $880\ \mathrm{kg\ m^{-3}}$ and light hardwood from 400 to 720

$kg\,m^{-3}$. According to the interview conducted by Hashim (2017), one of the characteristics that Sape makers use to categorize or select the Sape soundboard is based on its density. Sape makers will choose a denser wood over a less dense wood because, in their opinion, denser wood generally produces better sound quality. Therefore, one of the characteristics to identify the wood to build a Sape soundboard is density.

The density, ρ of the wood sample is calculated by measuring the mass, m and volume, V of each sample as shown in the Equation 1:

$$\boldsymbol{\rho = \frac{m}{V}} \tag{4.1}$$

ii Dynamic Modulus of Elasticity, E

One of the most crucial measurements to watch the behaviour of the wood when applying a force is the dynamic modulus of elasticity, E, also known as Young's modulus. By employing formula and calculation, the dynamic modulus of elasticity may demonstrate the mechanical and acoustical characteristics of the wood. The mechanical characteristics that demonstrate a material's tensile and compressive stiffness are known as its modulus of elasticity. Normally, the bending test is used to determine the modulus of elasticity; however, Liu et al. (2006) claimed that this method is time-consuming and unworkable, and instead suggested using a vibration method based on a Fast Fourier Transform analysis of hammered sound to determine the dynamic modulus of elasticity of solid wood.

Calculating the wood's fundamental frequency is a crucial step in determining the modulus of elasticity. Higher rigidity correlates with a higher modulus of elasticity, which is further correlated with more force being required to generate a given deformation (Rosato, 2003). A higher wood elasticity modulus is required when creating a soundboard for a musical instrument (Brémaud, 2012). Because less energy is lost when moving through the wood samples' grain, which has a higher modulus of elasticity, the material is a good resonator (Buksnowitz et al., 2012). The dynamic modulus of elasticity, E can be computed as shown in Equation 4.2:

$$\boldsymbol{E} = \frac{48\pi^2 L^4 \rho f_n^2}{\beta_n^4 h^2} \tag{4.2}$$

where L denotes the length of the wood sample (m), ρ denotes the density of the wood (kg/m^3), f_n denotes the natural frequency (Hz), β_n denotes the coefficient of vibration, h denotes the thickness of the wood sample (m), and n denotes the mode number. Since only fundamental frequency will be used, according to Yoshikawa (2007), for the fundamental mode in which $n = 1$, the coefficient $\beta_1 = 4.73$.

iii *Acoustic Radiation Damping Coefficient, R*

The acoustic radiation damping coefficient is one of the factors to determine the acoustic quality of the wood to create the musical instrument's soundboard (Brémaud, 2012; Wegst, 2006). The amount of body vibration that is dampened by sound radiation is indicated by the acoustic radiation damping coefficient. The ratio of the speed of sound, c, to the material's density, ρ, can also be used to define the acoustic radiation damping coefficient mathematically. The acoustic radiation damping coefficient, R, can also be determined by obtaining the young modulus, E, and density, ρ as shown in Equation 4.3:

$$\boldsymbol{R} = \frac{c}{\rho} = \sqrt{\frac{E}{\rho^3}} \tag{4.3}$$

The value of R will simply be the average amplitude or loudness of the wood sample when the acoustic radiation damping coefficient is calculated (Brémaud, 2012). A soundboard with a greater R-value will therefore often produce a louder sound.

iv Acoustic Impedance, z

According to studies by Wegst (2006), Yang et al. (2017a), and Brémaud (2012), one of the acoustic factors used to categorise the type of wood used to make soundboards is acoustic impedance, z. The capacity of a substance to transmit vibrations is known as its acoustic impedance. Mathematically speaking, the acoustic impedance is also equal to the product of the material's density and sound speed. The density, ρ, and the young modulus, E, can also be used to compute the acoustic impedance using the Equation 4.4:

$$\boldsymbol{z = c\rho = \sqrt{E\rho}} \tag{4.4}$$

The higher the acoustic impedance, the higher the resistance to the propagation of the sound waves towards the wood tissue. Each tissue of a distinctive wood type has its unique acoustic impedance (Suzuki et al., 2019). Since the speed of sound is constant, deciding the acoustic impedance of different wood types depends on thickness. Thus, lower-density wood will regularly have a lower acoustic impedance which shows lower resistance to the transmission of sound waves towards wood tissue in this way creating louder sound.

v Internal Friction, $tan\ \delta$

The degree of resistance obstructing the flow of vibration is known as internal friction, also known as the damping coefficient or loss factor, $\tan\delta$. Internal friction is one of the most crucial factors to produce a decent soundboard, claim Ahmed and Adamopoulos (2018). The Equation 4.5 below, which is based on Brémaud (2012), can be used to calculate internal friction. Internal friction can also be expressed as Q^{-1}, in which the Q denotes the quality factor (see Figure 4.6).

$$\tan \delta = \frac{\Delta f}{f_R} \approx Q^{-1} \tag{4.5}$$

where $\Delta f = f_2 - f_1$ = the bandwidth of the vibration at half-power (or at -3dB)

f_R = Fundamental frequency

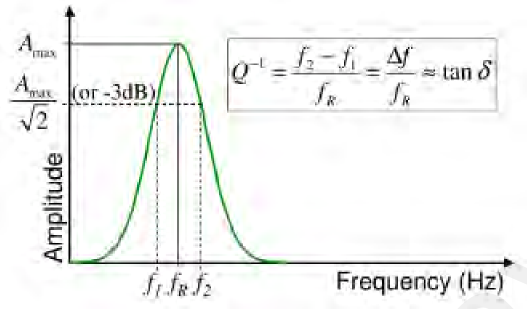


Figure 4.6: Bandwidth method to obtain bandwidth (Brémaud, 2012)

In general, lower internal friction is preferable when creating soundboards for musical instruments. The purpose of the soundboard is to transmit vibration from one end to the other. Less energy will be lost due to resistance due to the low internal friction, which allows for more vibration transmission.

vi Acoustic Conversion Efficiency, ACE

The ratio of the musical instrument's radiated acoustic energy to the energy provided by the string can be used to determine the efficiency of an acoustic system (Sedik et al., 2010). The ratio of sound radiation coefficient, R to the internal friction, $\tan \delta$ is a mathematical definition of the acoustic conversion efficiency. The Equation 4.6 is displayed below.

$$\mathrm{ACE} = \frac{R}{\tan \delta} = \frac{\sqrt{\frac{E}{\rho^3}}}{\tan \delta} \tag{4.6}$$

ACE measures how effectively vibrational energy is converted into sound energy. Consequently, better wood has a higher ACE because more vibrational energy is transformed into acoustic energy (Ahmed & Adamopoulos, 2018).

vii Speed of Sound, c

The speed of sound within the material is one of the most important acoustical properties for material selection for musical instruments. The speed of sound travels through a material can be computed as the root of the material's Young modulus, E divided by the material density, ρ as shown in the Equation 4.7 (Wegst, 2006).

$$\boldsymbol{c} = \sqrt{\frac{\boldsymbol{E}}{\boldsymbol{\rho}}} \tag{4.7}$$

According to Wegst (2006), the speed of sound is independent of wood species. However, it varies with the grain direction. The speed of sound also depends on the temperature or moisture content in the wood. If the temperature or moisture content increases, the speed of sound will decreases (Green, 1999). For string musical instrument soundboards, a high speed of sound is preferred as the speed of sound facilitates the transmission of vibrational energy (Ahmed & Adamopoulos, 2018).

viii Attack Time, AT

The vibration's attack time can provide certain timbral features of the wood. The time it took for the signal from the woods to reach its peak is indicated by its temporal duration. The time required for the vibration of the wood to reach its peak is calculated using the formula mirattacktime as shown in Figure 4.7 (Lartillot, 2021). The outcomes can alternatively be characterised as an increase in energy at the sound's start.

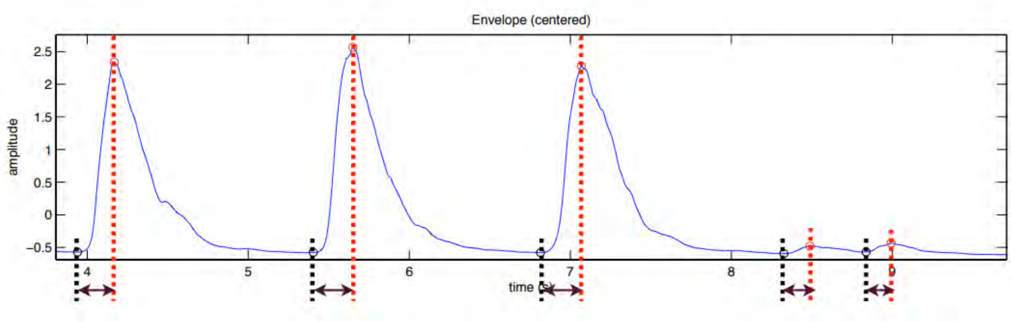


Figure 4.7: Graph of amplitude over time to calculate the attack time (Lartillot, 2021)

ix Spectral Bandwidth, SB

Spectral bandwidth is used to measure the spread of the spectral components around the spectral centroid (Figure 4.8). Using MIRToolbox, mirspread is used to calculate the spectral bandwidth. Spectral bandwidth represents the standard deviation of the data.

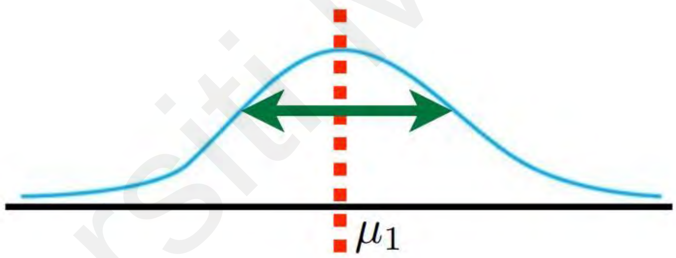


Figure 4.8: Spectral bandwidth (Lartillot, 2021)

x Spectral Centroid, SC

The time average over the signal envelope, measured in seconds, is the temporal centroid, also known as the spectral centroid. The outcomes show the period during which the produced sound signal has the highest average energy (Mazarakis et al., 2006). Knowing the precise instant when the sound energy reaches its peak is a crucial element for understanding the form of distribution. The centroid is represented by the dashed line in Figure 4.8.

xi *Spectral Flux, SF*

A spectro-temporal descriptor called spectral flux figures out how the spectrum changes over time. The distance between the spectrums of each subsequent frame can likewise be used to characterise spectral flux as shown in Figure 4.9. The local spectrum representation of the signal's modulus is used to determine the mean Pearson correlation value, which provides the spectral flux (Aramaki et al., 2007).

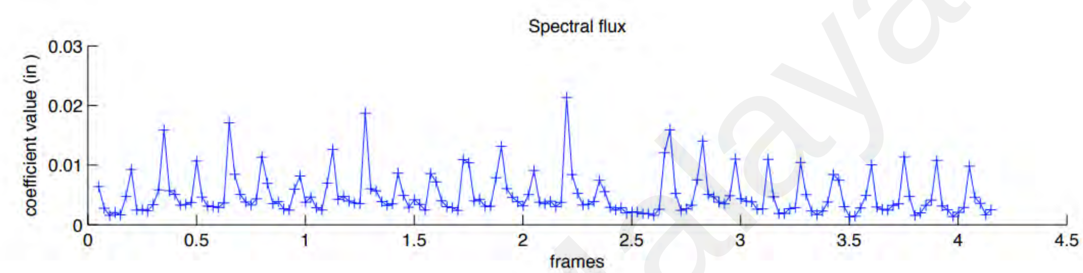


Figure 4.9: Spectral flux (Lartillot, 2021)

xii *Inharmonicity, I*

Inharmonicity is a property that is used to assess the number of partials that are not multiples of the fundamental frequency, f_0 or the natural frequency as shown in Figure 4.10. The quantity of energy that is outside the optimum harmonic series range will be considered when determining the inharmonicity (Lartillot, 2021). As a result, a low inharmonicity rating will likely be preferred when evaluating the wood's quality because it will result in a soundboard that produces a more harmonic sound.

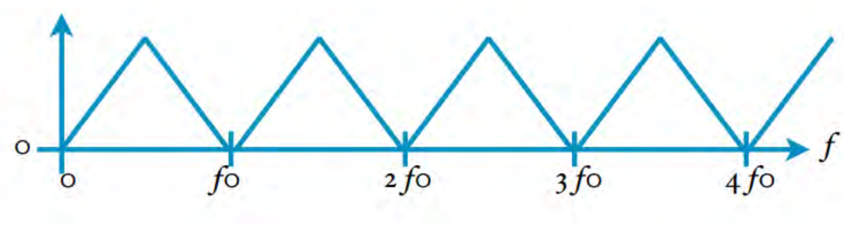


Figure 4.10: Fundamental frequency and its partials (Lartillot, 2021)

(b) Quality Grading

To rate the quality of the wood samples in this study, the study conducted by Yang et al. (2017a) is referred to. The soundboard is graded objectively by the experts based on the acoustic quality which includes the sound loudness, dynamic range, sound length and tone. According to Wegst (2006), soundboard woods are the best round radiators of all. This is beneficial for the musical instrument to produce sufficient loudness by the transmission of sound from the soundboard into the air.

The methods used by Sape makers to choose high-quality wood can serve as additional evidence for the qualities stated. From the focus group discussion conducted, the Sape makers' rating of wood is based on hearing, touching, and seeing. Based on the volume and length of the wood's vibration, they will examine it. Therefore, loudness and period characteristics are chosen as the criteria for grading the wood samples.

i Loudness

The subjective concept of loudness is sound pressure, which can be expressed by the signal's amplitude. Sape makers frequently select loudness as the characteristic to determine the quality of a wood sample (Hashim, 2017). The reason for this is that the soundboard serves as a platform for the sound to resonate and magnify to produce a loud sound. According to Yang et al. (2017a), one of the crucial characteristics to characterise soundboard wood for making Yueqin is loudness. Hence, a Sape soundboard with higher loudness typically has higher resonance. A steady force is applied to the wood samples to determine their loudness, and the amplitude of the signal reflects their loudness. The loudness can be determined by measuring the amplitude of the signal. The amplitude is a measure of the strength of the sound signal and is commonly expressed in decibels (dB). We used MATLAB to process the sound samples and extract the amplitude values, which we then converted to dB values. These dB values represent the loudness of the sound samples, with higher dB values indicating louder sounds.

ii Temporal Duration

The MIRtoolbox in MATLAB, notably the mirduration function, was used to extract the amount of time that passed between the beginning and end of an audio event. According to earlier studies (Peeters, 2004), this function determines the attack and decay phases of each event and calculates the section of the curve between the onset and offset times that exceeds 40% of the maximum amplitude between the attack and decay times.

It should be noted that the duration of an audio event is significantly influenced by its damping coefficient, which affects how well different wood samples produce sound. A sound that is sustained and more pleasant over a longer duration of time is often a hallmark of high-quality wood. Therefore, in previous studies, the use of duration as an evaluation metric has been widely accepted because it provides a precise and quantitative assessment of sound quality. In conclusion, the MIRtoolbox offers a reliable and effective approach for extracting the duration parameter, which provides a measure of the calibre of sound produced by wood samples, while also taking into account the influence of the damping coefficient.

(c) ***Machine Learning***

Machine learning algorithms are a powerful and popular tool in the sound recognition system. The application of machine learning in the musical field is not a new thing. Aiming to produce new genres of music or new musical interactions, music researchers and musicians continually explore the possibilities of new algorithms to carry out the learning. By utilising improvements in processing data resources, machine learning enables us to address increasingly complicated musical contexts. (Fiebrink & Caramiaux, 2016).

There are two main types of machine learning which are supervised learning and unsupervised learning. In supervised learning, the algorithm builds a model from the labelled data while unsupervised learning is dealing with unlabeled data. Other than these two, there are also other types of learning available such as semi-supervised learning and reinforcement learning. In semi-supervised learning, the training data includes labelled and unlabeled data which has the advantage of reducing the time consumption in data labelling. On the other hand, reinforcement learning is a complete multi-step algorithm with clearly defined rules that allows machine learning to decide which steps to take.

In this study, supervised learning is used as the sound data collected is labelled. The features selected to be used are the input while the output is the label of the data. This algorithm will learn and diagnose the meaningful relationship between the input data and output data and build a model of that relationship. The data will be divided into the training set and the testing set. The training set consists of all labelled data used in building the model while the testing set is used to test the accuracy of the model as shown in Figure 4.11.

MATLAB software (version R2021a) is used to process and extract the feature from the sound data collected from the free-free flexural vibration test. With the aid of the add-on available in MATLAB, it allows us to process and analyze sound data easily. The signal processing toolbox, classification learner app, MIRToolbox, and FilterDesigner are useful add-ons from MATLAB that enable users to carry out analyzing data, feature extraction, designing filters, and machine learning. With the use of MATLAB and its add-on, the following tasks can be carried out smoothly and easily.

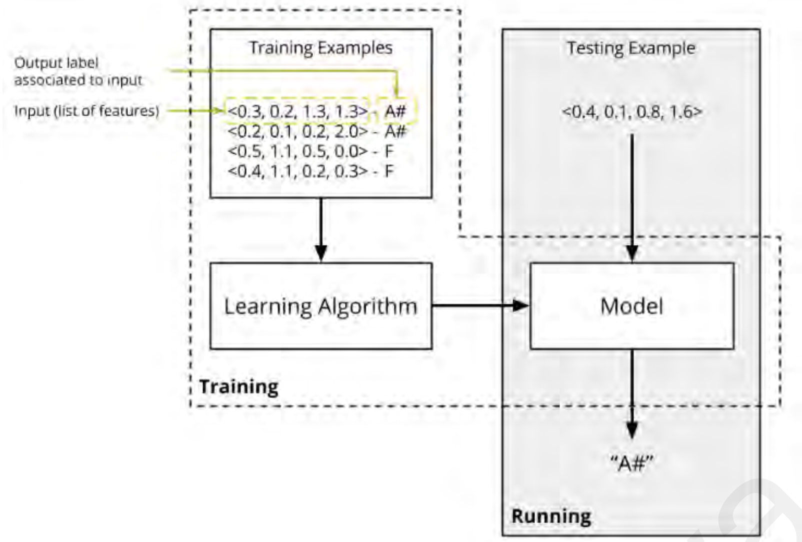


Figure 4.11: Supervised learning training and testing process (Fiebrink & Caramiaux, 2016)

Several tasks, including binary classification, regression modelling, ensembling, and multi-class classification, can be accomplished using the supervised machine learning method used in this study. Data must be divided into two categories using binary classification, between more than two types of classes using multi-class classification, between continuous values using regression modelling, and between discrete values using ensembling to provide correct predictions. In the MATLAB classification learner app, there are eight different classifier types, including decision trees, discriminant analysis, logistic regression analysis, naive Bayes classifier, support vector machine, closest neighbours classifier, ensemble classifier, and neural network classifier.

Feature extraction and selection processes were executed utilizing MIRToolbox and formula calculations. Subsequently, the classification task was conducted via the classification learner app available in MATLAB, providing various classifier options such as decision trees, discrimination analysis, logistic regression analysis, naïve Bayes, support vector machine, and others.

All retrieved features were reviewed and split into two distinct sets, comprising 70% as training data and 30% as test data, once the MATLAB code and system were successfully developed. Of the total of 360 recorded datasets, 252 sets (70%) were employed for model training, while the remaining 108 sets (30%) were reserved for testing purposes. Regarding the training data, a portion was utilized in cross-validation, employing a five-fold approach.

The analysis involved processing the collected sound data in WAV format. A system was developed in MATLAB to extract features from the sound data, and the corresponding programming code was written. Upon initial testing, if the system failed to extract characteristics accurately and successfully, the MATLAB code underwent scrutiny for refinement.

4.4 Results and Discussion

Using fundamental frequency and the density obtained, the dynamic modulus of elasticity (E), acoustic radiation damping coefficient (R), acoustic impedance (z), internal friction ($\tan\delta$), acoustic conversion efficiency (ACE), and speed of sound (c) can be computed. It is categorised under physical and vibroacoustic features. Other than that, there were 5 timbre features which include attack time (AT), spectral bandwidth (SB), spectral centroid, (SC), spectral flux (SF), and inharmonicity, (I), density and fundamental frequency were used as the inputs while the three types and grades of the wood sample were the outputs. All the data collected is shown in Table 4.5 and Table 4.6.

Table 4.5: Physical and vibroacoustic features

Sound File	f_r	ρ	E	z	R	$\tan\delta$	ACE	c
A1	177.80 ± 0.00	468.07 ± 0.00	41322061.54 ± 0.00	139203.24 ± 0.00	0.63 ± 0.00	0.0040 ± 0.0001	160.03 ± 4.19	296.85 ± 0.00
A2	180.67 ± 0.09	481.49 ± 0.00	42509367.48 ± 0.00	143202.97 ± 0.00	0.62 ± 0.00	0.0038 ± 0.0001	160.86 ± 3.80	296.85 ± 0.00
A3	178.00 ± 0.00	474.24 ± 0.00	43307862.61 ± 0.00	143434.93 ± 0.00	0.64 ± 0.00	0.0041 ± 0.0001	156.53 ± 2.73	301.93 ± 0.00
T1	144.80 ± 0.00	1276.68 ± 0.00	1760170998.43 ± 0.00	1378713.65 ± 0.00	1.18 ± 0.00	0.0021 ± 0.0015	813.12 ± 529.79	1276.68 ± 0.00
T2	149.61 ± 0.04	1279.82 ± 0.00	1672866178.48 ± 0.00	1307108.36 ± 0.00	1.25 ± 0.00	0.0019 ± 0.0015	925.92 ± 439.94	1279.82 ± 0.00
T3	123.39 ± 0.05	1120.12 ± 0.00	1231178695.79 ± 0.00	1099148.49 ± 0.00	1.14 ± 0.00	0.0018 ± 0.0007	767.68 ± 394.35	1120.12 ± 0.00
M1	180.81 ± 0.04	926.30 ± 0.00	584381710.54 ± 0.00	736533.81 ± 0.00	0.85 ± 0.00	0.0014 ± 0.0006	771.28 ± 395.97	793.42 ± 0.00
M2	177.61 ± 0.03	918.40 ± 0.00	559834748.48 ± 0.00	717951.19 ± 0.00	0.85 ± 0.00	0.0013 ± 0.0007	850.10 ± 523.30	779.77 ± 0.00
M3	173.83 ± 0.07	914.99 ± 0.00	550445740.49 ± 0.00	710607.12 ± 0.00	0.84 ± 0.00	0.0014 ± 0.0007	776.26 ± 404.30	774.61 ± 0.00

Table 4.6: Timbre features

Sound File	AT	SB	SC	SF	I
A1	0.02 ± 0.0003	2523.78 ± 122.00	1823.73 ± 106.92	0.98 ± 0.09	0.43 ± 0.010
A2	0.02 ± 0.0002	2613.04 ± 123.66	1707.90 ± 68.54	0.97 ± 0.07	0.42 ± 0.006
A3	0.02 ± 0.0000	2490.36 ± 90.73	1782.65 ± 51.47	0.98 ± 0.05	0.43 ± 0.005
T1	0.02 ± 0.0003	2356.69 ± 76.11	2492.47 ± 116.35	0.88 ± 0.05	0.46 ± 0.006
T2	0.02 ± 0.0003	2382.71 ± 47.59	2461.33 ± 108.25	0.90 ± 0.06	0.45 ± 0.003
T3	0.02 ± 0.0002	2492.12 ± 91.65	2365.12 ± 234.00	0.81 ± 0.07	0.46 ± 0.011
M1	0.02 ± 0.0003	2552.10 ± 97.31	2444.84 ± 129.37	0.80 ± 0.07	0.43 ± 0.030
M2	0.02 ± 0.0003	2364.88 ± 60.84	2282.58 ± 90.06	0.89 ± 0.05	0.45 ± 0.004
M3	0.02 ± 0.0003	2401.37 ± 70.28	2336.47 ± 129.68	0.88 ± 0.06	0.45 ± 0.014

In this study, the woods were objectively ranked based on their performance measured by the loudness and duration of the sound signals they produced. The loudness and duration were averaged across all nine wood samples and presented in Table 4.7, while the distribution of the wood samples was illustrated in Figure 4.12. The combined score in the table is calculated based on the normalized average duration and loudness values, each given a weight of 0.5. This provides a balanced measure of the wood samples' sound quality, considering both parameters. Adau wood had the highest average combined score of 41.96, indicating the best overall sound quality among the tested woods. Merbau wood had a lower average combined score of 36.88, earning the second rank. Tapang wood had an average of 36.69, slightly lower than Merbau wood, indicating poorer sound quality, and was ranked third. These rankings were determined solely based on the objective measures of duration and loudness, which are commonly used to assess sound quality in musical instruments.

Table 4.7: Sound quality evaluation results for different wood samples

Wood	Loudness (dB)	Duration (s)	Combined Score	Rank	Cluster
A1	42.25	0.0271	41.47	3	Good
A2	42.13	0.0282	42.23	1	Good
A3	42.34	0.0280	42.19	2	Good
T1	37.82	0.0237	36.70	7	Poor
T2	40.00	0.0221	36.59	8	Medium
T3	39.63	0.0226	36.78	6	Medium
M1	39.72	0.0228	36.98	5	Medium
M2	40.35	0.0226	37.14	4	Medium
M3	38.21	0.0232	36.52	9	Poor

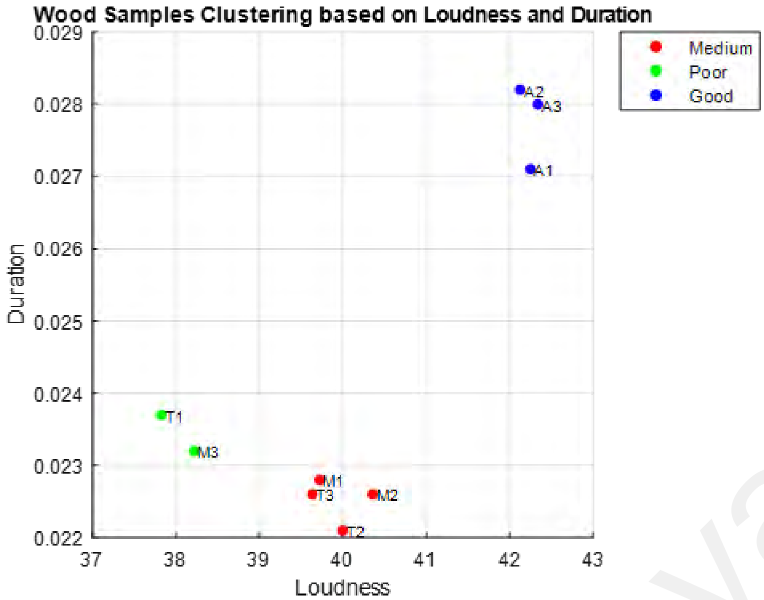


Figure 4.12: Distribution of wood samples based on loudness and duration (A=Adau, T=Tapang, M=Merbau)

The k-means clustering analysis was performed based on the combined score of loudness and duration, to identify groups of wood samples that exhibited similar acoustic characteristics. Based on the clustering analysis, the 9 wood samples were divided into three clusters as shown in Figure 4.12. The first cluster included Adau woods A1, A2, and A3, which can be considered as the "good" cluster as they exhibited the highest combined scores among all the samples. The second cluster included Tapang woods T2 and T3, and Merbau woods M1 and M2, and can be categorized as the "medium" cluster as their combined scores fell in between the good and poor clusters. The third cluster included Tapang wood T1 and Merbau wood M3 and can be categorized as the "poor" cluster with the lowest combined scores among all the samples.

It is noteworthy that while most of the samples within the same wood type were clustered together, some samples exhibited different performances compared to the majority of their respective wood types. For instance, Tapang wood and Merbau wood were clustered in the medium and poor clusters, respectively, indicating that the acoustic characteristics of a wood sample can vary due to various factors such as differences in wood density, moisture content, or even variations in the manufacturing process. Overall, these findings provide insights into the acoustic properties of different types of wood and suggest that the combined score of loudness and duration can be an effective metric for evaluating the quality of sound produced by musical instruments. However, further research is needed to validate these findings and explore the relationship between wood type and sound quality.

The results of this study demonstrate that machine learning can be used to cluster the sound quality of wood samples based on their acoustic properties. Additionally, the clustering of wood types can also potentially be used to classify sound quality. The results from k-means clustering and ranking based on the combined score showed that all Adau woods are clustered as good, while Tapang and Merbau woods are clustered as both medium and poor. This suggests that the use of machine learning in classifying wood types can also be used as a classification of sound quality. This study highlights the potential of using machine learning to improve the classification of sound quality in wood samples.

To further investigate this, the 13 features selected in this study will be used as the input data of the different learning algorithms available in the MATLAB classification learner app. Thirty per cent of the data is set aside as a testing data set, while the remaining 70% is used as training data. For validation, the training data is divided into 5-folds cross-validation. No specific algorithm or classifier model is pre-selected in this study, but it will all be tested and the accuracy of each will be compared.

To reduce the chance of overfitting, the Minimum Redundancy Maximum Relevance (MRMR) algorithm is used to identify the key features for optimal classification (Peng et al., 2005). Each feature can be ranked according to how important it is to the target variable, and the ranking procedure can also take their overlap into account. An "excellent" feature achieves the optimal balance between minimal internal redundancy and high relevance to the target variable (Cai et al., 2012). The 13 features will be selected based on their importance scores. From the MRMR algorithm results shown in Table 4.8, the 4 highest features are selected for the classification which consists of acoustic radiation damping coefficient, spectral flux, spectral centroid, and inharmonicity. The features are applied to all classifiers and the accuracy results is displayed in Table 4.9. The highest accuracy model is achieved by fine tree, medium tree and bagged tree classifier at 98.1%. It is followed by coarse tree, quadratic SVM, medium Gaussian SVM, and wide neural network at 86.1%.

Table 4.8: Feature importance scores using MRMR algorithm.

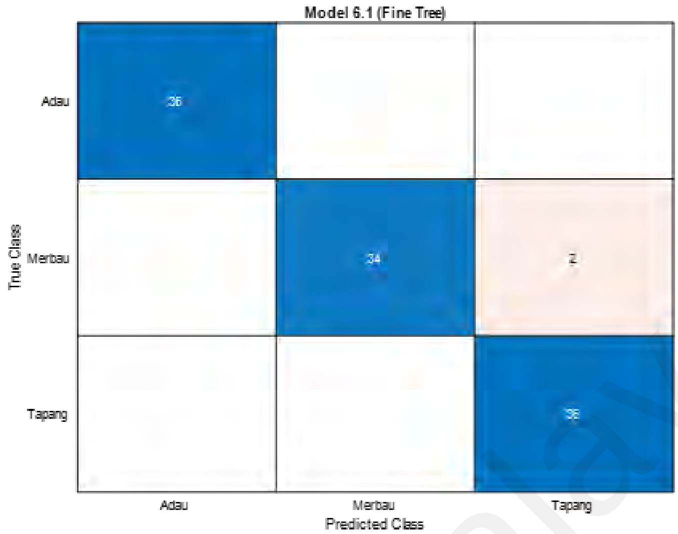
Features	Feature Importance Score
z	0.9939
I	0.8653
$\tan\delta$	0.8536
SC	0.7621
SB	0.6734
R	0.5637
SF	0.5481
AT	0.4927
ACE	0.4313
f_r	0.3909
c	0.3190
ρ	0.3075
E	0.2911

Table 4.9: Training and testing accuracy comparison between all features and MRMR top 4 features.

Classifier		MRMR Top 4 Features	
		Training Accuracy	Testing Accuracy
Decision Tree	Fine Tree	90.5	98.1
	Medium Tree	90.5	98.1
	Coarse Tree	84.5	86.1
Discriminant Analysis	Linear Discriminant	79.4	79.6
	Quadratic Discriminant	80.2	79.6
Naïve Bayes	Gaussian Naïve Bayes	13.8	76.9
	Kernel Naïve Bayes	77.4	77.8
Support Vector Machine	Linear SVM	79.0	79.6
	Quadratic SVM	82.5	86.1
	Cubic SVM	84.5	80.6
	Fine Gaussian SVM	84.1	81.5
	Medium Gaussian SVM	86.5	86.1
	Coarse Gaussian SVM	78.2	79.6
Nearest Neighbor	Fine KNN	84.9	82.4
	Medium KNN	82.9	84.3
	Coarse KNN	69.8	71.3
	Cosine KNN	78.2	78.7
	Cubic KNN	82.1	84.3
	Weighted KNN	85.3	84.3
Ensemble	Boosted Trees	42.5	33.3
	Bagged Trees	92.1	98.1
	Subspace Discriminant	78.6	79.6
	Subspace KNN	69.8	71.3
	RUSBoosted Trees	42.5	33.3
Neural Network	Narrow Neural Network	80.6	85.2
	Medium Neural Network	82.1	85.2
	Wide Neural Network	84.5	86.1
	Bilayered Neural Network	82.5	84.3
	Trilayered Neural Network	79.8	83.3
Kernel Approximation	SVM Kernel	71.8	71.3
	Logistic regression Kernel	71.8	63.9

The decision tree is found to be good in classifying the wood types in this study with a high accuracy of 98.1%. Decision tree classifiers are renowned for providing a more comprehensive view of performance results. As decision tree classifiers use enhanced tree pruning algorithms and optimised splitting parameters, it is commonly used by many data classifiers. In the comparison study done by Charbuty and Abdulazeez (2021), the decision tree is applied in many areas such as medical disease analysis, text classification, and image classification and showed the best accuracy compared to other classifiers. Decision trees underlying efficient collection rule is simple to understand which made it preferable. The bagging tree is under ensemble methods which combine decision trees to produce a better classification model rather than just one decision tree. It combined several weak learners to form a strong learner. Bagging trees is one of the techniques of an ensemble in which the goal is to reduce the variance of a decision tree. The benefit of a bagging tree is that it can handle data with high dimensionality with missing values while maintaining its accuracy of it.

The results of the classification from the classification learner app from MATLAB showed that the decision trees classifiers are best in predicting the wood type or wood grade in this study. From the test confusion matrix shown in Figures 4.13(a) and 4.13(b), the error of classification occurred between Merbau and Tapang wood. For fine tree and medium tree, two Merbau wood were misclassified as Tapang wood while two Tapang wood were misclassified as Merbau wood in the bagged tree classifier. The classification of Adau wood showed no error with all the data being correctly predicted. The classification result proves the strong ability of machine learning in predicting the wood types or the grade of the Sape soundboard. The results show that the training result has poorer accuracy than the testing accuracy. This counterintuitive outcome can be attributed to the potential presence of noise or overfitting in the training dataset, where the model learns the noise and nuances specific to the training data, resulting in lower accuracy. Additionally, the testing data might better represent the true distribution of data, highlighting the model's generalization capability more accurately.



(a)

(b)

Figure 4.13: (a) Confusion matrix of fine tree and medium tree (b) Confusion matrix of bagged tree

4.5 Chapter Summary

Summarizing Chapter 4, this study aimed to assess the quality of commonly used woods in Sape instrument production. Rectangular wood samples resembling Sape soundboards underwent flexural vibration tests, generating essential sound data. Analyses of physical, vibroacoustic, and timbre features, particularly focusing on loudness and duration performance, revealed that Adau wood exhibited superior quality, followed by Merbau and Tapang wood. The MRMR algorithm identified four key features—acoustic radiation damping coefficient, spectral flux, spectral centroid, and inharmonicity—used to train diverse classifiers in MATLAB. Impressively, the decision tree classifier achieved 98.1% accuracy. The findings highlighted the potential of machine learning in objectively classifying Sape wood quality, offering a practical model for predicting wood quality in Sape instrument crafting.

Moving forward, future research endeavors should encompass subjective evaluations from Sape makers in the evaluation process. By incorporating the subjective assessments of Sape experts, the next chapter aims to delve deeper into the qualitative evaluation of Sape instrument quality. This inclusion will provide a more comprehensive understanding, considering both objective and subjective perspectives, thereby enriching the overall assessment of Sape instrument craftsmanship.

CHAPTER 5: ENHANCING SAPE INSTRUMENT QUALITY ASSESSMENT: INTEGRATING EXPERT EVALUATIONS

5.1 Overview

Chapter 5 serves as a direct extension of the preceding research conducted in Chapter 4, which meticulously evaluated the quality of woods commonly employed in the construction of Sape musical instruments. The analysis conducted in the previous chapter provided valuable insights into the objective assessment of wood quality, employing machine learning techniques and crucial acoustic features. The findings illuminated the potential of machine learning models in precisely categorizing wood quality for Sape instrument crafting.

The previous chapter primarily focused on an objective evaluation of wood quality, leveraging machine learning techniques and crucial acoustic features. However, this evaluation lacked the subjective ratings of sound quality by experienced Sape makers, signifying a significant research gap in the assessment process. Consequently, this chapter aims to address this void by integrating subjective evaluations from experts in the field, further enriching the Sape instrument quality assessment. The objectives in this section encompass a detailed analysis of both the collected data and the machine learning methodology utilized. Furthermore, the chapter endeavors to delve deeper into the qualitative assessment of Sape instrument quality by incorporating the invaluable perspectives and evaluations of experts in the field.

In pursuit of a holistic evaluation of Sape instrument quality, this chapter serves a dual purpose: to rigorously analyze acquired data using statistical and machine learning methodologies, and to integrate subjective evaluations from seasoned experts in the field. By amalgamating objective and subjective assessments, this chapter aims to provide a comprehensive understanding of Sape instrument quality. Subsequent sections will delineate the process of data analysis, outline the machine learning techniques employed, and showcase insights drawn from subjective evaluations offered by experienced Sape experts.

The overarching goal of this chapter is to establish guidelines for identifying high-quality soundboard wood in Sape instrument construction, leveraging machine learning algorithms and acoustic feature analysis. This involves scrutinizing the relationship between acoustic attributes and quality ratings provided by proficient Sape makers, thereby establishing stringent criteria for superior soundboard wood in Sape crafting. The research is structured into two core segments: the design and training of a classifier to predict soundboard wood quality based on sound samples (Part 1), and the exploration of the significance of acoustic features in determining Sape soundboard wood quality, utilizing insights derived from the trained classifier (Part 2). Through bridging this research gap and offering a structured framework for identifying top-tier soundboard wood, this study promises significant advancements in refining the material selection process within Sape instrument making.

5.2 Methodology

5.2.1 Quality Rating

For consistent and unbiased quality rating evaluations, audio samples collected in the previous experiment went through two post-processing steps using Audacity software. The first step normalized the volume to standardize loudness, eliminating recording environment-related inconsistencies. The second step focused on noise reduction to enhance audio clarity by removing any white noise present in the recordings. These techniques were used to ensure audio consistency and reduce emotional bias during quality rating evaluations. The 360 audio samples were subjected to quality rating evaluations through listening tests conducted by experts in Sape instrument, utilizing a 5-point Likert scale. The scale ranged from "Very Poor" to "Excellent" to assess the perceived quality of the audio samples.

The candidate pool for the evaluation consisted of five experienced Sape makers from Sarawak, each with a minimum of five years of expertise in their craft as shown in Table 5.1. Demographic analysis revealed that although 60% of the sample population had not received formal music training, it is important to note that Sape, as a Borneo traditional instrument, lacks formal musical education and is primarily passed down through generations within the tribe. Therefore, the absence of formal training does not hinder their professionalism in Sape crafting. The expertise of the candidates for the listening test was primarily determined based on their years of experience in crafting Sape.

Table 5.1: Demographic information of candidates

Candidate	A	B	C	D	E
Age	32	28	42	29	71
Gender	M	M	M	M	M
Ethnicity	Kenyah	Iban	French	Iban	Kenyah
Formal musical training	1 year	None	2 years	None	None
Years of experience	9	8	7	7	30+

To ensure consistency in the perceived audio quality, several control procedures were implemented. The initial hardware settings of the listening test were maintained consistently throughout the rating process. Flexibility in the time constraint of the listening test was provided to minimize the potential impact of psychological stress. Additionally, the candidates were not provided with prior information regarding the corresponding wood type of each audio sample.

One aspect that needs consideration is the subjective nature of the quality rating and the approach used to analyze the ratings. The rating measure employed in the evaluation is based on a Likert scale, which has an ordinal nature and does not inherently provide information about the "distance" between successive categories of the quality attribute. Although the scale values are numerical, their interpretation lacks a clear objective basis beyond their established order. To address this, qualitative descriptions were provided for each rating category, helping to standardize and normalize the data obtained from each candidate.

In terms of the reliability of the rating scores, it is important to consider the niche nature of the targeted Sape community, which may introduce potential biases in the sampling distribution. However, based on the earlier demographic analysis, all candidates are considered experts in the Sape instrument field, capable of representing the overall community's perception of Sape sound quality. Due to the substantial workload involved in rating 360 sound samples, each candidate conducted the listening test only once, which limits the availability of information to validate the reliability of an individual's rating score based on classical test theory. Nonetheless, inter-rater reliability can be analyzed using non-parametric statistical tests.

5.2.2 Acoustic Feature Extraction

In this study, relevant functions from the MATLAB MIRToolbox were selected for feature extraction. In the initial phase of feature reduction, the mirfeatures() function extracted numerous complete features, resulting in a high number of data points. To address this, frame decomposition was used to analyze temporal signals within short-term windows, reducing the feature count to 148. In the second phase, the dataset was further reduced by selecting the statistical descriptor with the most variability for each feature. The final dataset comprised 27 acoustic features, and the mean values were chosen as representatives. This reduced the dimensionality to a more manageable level for analysis.

The selected functions aligned with the research objectives and are listed in Table 5.2. These functions covered different aspects of the Sape sound quality, including dynamics, rhythm, timbre, and pitch/tonality. The dynamics dimension focused on the root-mean-square (RMS) as a measure of the global energy of the signal. For the rhythm category, fluctuation was estimated using spectrogram computation and Fast-Fourier Transform (FFT). Parameters such as peak and centroid values characterized the behavior of fluctuations. Attack and decay parameters, which examined the time elapsed and the slope of the attack or decay, were also considered.

Table 5.2: Shortlisted functions for feature extraction.

Dynamics	• Dynamic RMS
Rhythm	• Fluctuation peak and centroid • Attack time and slope • Decay time and slope
Timbre	• Spectral centroid, brightness, spread, skewness, kurtosis, roll-off, entropy, flatness, roughness, and irregularity • Timbral zero-cross, low energy and spectral flux
Pitch & Tonality	• Tonal chromagram peak and centroid • Tonal key strength, mode and HCDF

Timbre analysis concentrated on the statistical description of spectral distribution, encompassing parameters such as flux, centroid, spread, skewness, kurtosis, roll-off, entropy, flatness, roughness, and irregularity. Timbral zero-cross and low energy were indicators of noisiness and focused on evaluating signal characteristics related to amplitude changes and energy levels. Pitch and tonality analysis involved the tonal chromagram, which described the energy distribution along pitches or pitch classes.

Parameters such as peak and centroid were used to characterize the chromagram. Tonal key strength and mode provided measures of tonality, and the Harmonic Change Detection Function (HCDF) detected fluctuations in tonal centroid measure. These selected functions allowed for a comprehensive analysis of the Sape sound quality, covering aspects such as energy dynamics, temporal characteristics, spectral properties, and tonal characteristics.

5.2.3 Data Preprocessing

To ensure the suitability of the dataset for a specific algorithm, pre-processing steps were conducted on the quality rating and acoustic feature data frames. These steps aimed to address assumptions and constraints associated with linear models, including linearity, normality, equal variance, and independence (Kumari & Yadav, 2018).

Linearity, which refers to the linear relationship between predictor variables and the target variables, was assessed using scatter plots. It was considered that the quality rating scores are ordinal and may not exhibit perfect linearity. To examine normality assumptions, histograms were plotted for each predictor variable, considering the expectation that the residuals follow a normal distribution with a mean of zero. Although the dataset size was limited to 360 data points, the overall shape of the distributions was considered, even if they might not be completely normal.

Homoscedasticity, which ensures equal variance within the error of the data, was not evaluated as the dataset had unity weight across all features. Independence, indicating the absence of multicollinearity among predictors, was assessed through a correlation analysis among the 27 acoustic features. To address varying magnitudes across different features, standardization and normalization techniques were applied. Standardization scaled the values to have a mean of zero and a standard deviation of one, while normalization scaled the values between zero and one (Bhandari, 2020).

A train-test split was then performed with a ratio of 70% for training data and 30% for testing data to evaluate the performance of the machine learning algorithm (Vabalas et al., 2019). To evaluate the predictive accuracy of the trained ML model on limited data samples, a cross-validation scheme called k-Fold Cross Validation with k = 5 was employed. This approach minimizes bias and provides a more realistic estimate of model performance (Brownlee, 2018). The dataset was shuffled and divided into 5 folds. For each fold, the current fold served as the validation set while the remaining 4 folds served as the training set. The model was fitted on the training set and evaluated using the validation set. The process was repeated for each unique fold, and the best performing model was selected based on the evaluation scores.

5.2.4 Machine Learning & Model Interpretability via SHAP

MATLAB Classification Learner application plays a huge part in this stage as it provides a complete array of classifiers to be trained and compared in parallel. A total of 40 classification models are selected to be trained on the quality rating classification problem, of which 31 models are simple algorithms with basic hyperparameter initialization fixed and the remaining 9 models are optimizable algorithms with varying hyperparameters.

The Bayesian optimization algorithm was chosen to perform hyperparameter optimization in the 9 models mentioned above. The core idea is to achieve global minima with the acquisition function of expected improvements per second plus within 100 iterations, without limit on the model training time. Number of iterations of 30, 1000 and 10,000 are also tested briefly and the model accuracies does not improve significantly (1~2%), hence 100 iterations are justified to be sufficient while maintaining time efficiency in terms of identifying the best hyperparameter combination.

Model interpretability refers to the extent to which we understand the inner workings of a machine learning (ML) algorithm and its decision-making process (Gilpin et al., 2018). While most ML models are considered "black box models" due to their counterintuitive representations, their interpretability decreases as complexity increases, despite their high predictive power (Molnar, 2020). Interpretability is necessary when problem formalization is incomplete, as predictions alone may not be sufficient. Understanding the rationale behind the model's decisions helps address the problem comprehensively and detect biases, leading to greater robustness. In this research, the aim is to estimate and understand raters' behaviors in ranking soundboard quality for cultural preservation purposes.

This study employed the SHapley Additive exPlanations (SHAP) approach for model interpretability. SHAP utilizes Shapley values, derived from cooperative game theory to attribute credit optimally and assess the relative contribution of each predictor in making predictions. Lundberg and Lee (2017) introduced additive feature attribution methods, including SHAP, which offer accurate estimations of feature importance within machine learning models.

However, the limitations of MATLAB prevent the use of SHAP for global interpretability in this study, restricting it to local interpretability. As a result, the computation of SHAP values is performed in a Python environment, leveraging the best classifier's hyperparameter configuration determined through MATLAB. By incorporating SHAP, this research aims to enhance the interpretability of the model, providing insights into the significance and impact of individual predictors on the predictions made.

5.3 Results and Discussion

5.3.1 Quality Grading

Five sets of quality rating scores were collected, revealing a conservative tendency in assigning the lowest score and suggesting no wood samples were rated as "very poor" quality. Median rating scores showed a parabolic trend for hardwoods labeled 2 and 3 across candidates, while hardwoods labeled 1 exhibited a consistent linear trend with one outlier. Candidate 3 consistently provided slightly lower ratings on average, indicating varying interpretations of sound quality. Assessing ranked mean scores with statistical tests is necessary for meaningful conclusions.

Since the quality rating scores are ordinal data, as established in the previous section, non-parametric statistical tests are necessary to analyze categorical (ordinal) data. The conducted tests include the Kruskal-Wallis H Test, Krippendorff's alpha, and Kendall's tau-b to compare the relationship between the rating scores. The candidates are denoted as A to E, corresponding to numbers 1 to 5. The detailed results of these statistical tests can be found in Appendix B.

The Kruskal-Wallis H test was used to compare rating scores, assuming independence and without assuming normality. Color-coded p-values (red for < 0.05, green for ≥ 0.05) were used. No consistent significant similarity was found between rating score combinations of any two candidates in terms of ranked means. At least 5 out of 9 wood samples showed significant differences in rating scores among the candidates. This suggests that personal preferences contribute to statistically significant variations in Sape wood quality ratings.

Krippendorff’s α is based on the observed disagreement corrected for disagreement expected by chance, with 1.0 indicates perfect agreement, 0 indicates no agreement beyond chances and negative value indicates inverse agreement. Colors indicate agreement (green), random guessing (yellow), and disagreement (red). Consistency requires green color in 5 out of 9 wood samples per candidate combination. In pairwise comparisons, candidate A's ratings disagree with others, while candidate B's ratings disagree with candidate E (except for Adau 1 scores). No strong correlations are found in other combinations, suggesting random guessing. Comparing three or more candidates shows insignificance due to score variances. However, pairwise comparisons consistently exhibit mutual disagreement across all wood samples, consistent with the Kruskal-Wallis H test.

Kendall's tau-b is a non-parametric correlation technique. Red color indicates p-values ≥ 0.05, while green color indicates p-values < 0.05. No candidate pairs show significant correlation across all wood samples. Most pairs have correlated ratings for Tapang 3, indicating low variance. "Nan" appears due to constant rating scores, preventing correlation calculation. This suggests Mr. Mathew Ngau's consistency in rating similar wood samples, showcasing his expertise in identifying soundboard quality.

In summary, the results of all three non-parametric statistical tests consistently revealed statistically significant differences among the sets of rating scores, underscoring a notable lack of inter-rater reliability. Given this, the critical task of selecting the most reliable set of rating scores comes to the forefront. The ratings given by each candidate is summarized and shown in Appendix C. From the summary table, Candidate E, Mr. Mathew Ngau, emerges as the preferred source for rating scores, substantiated by compelling reasons. Notably, Mr. Mathew Ngau boasts an unparalleled wealth of experience, surpassing 30 years, a considerable margin compared to the other four candidates (see Table 5.1 for demographic information). His extensive expertise and authority in the realm of Sape playing, coupled with his prestigious position as a National Living Heritage, underscore the robustness of his consistent ratings. Furthermore, Mr. Mathew Ngau's role as a mentor adds an additional layer of credibility to his evaluations. In contrast to the varying degrees of experience exhibited by other candidates, his seasoned proficiency positions him as an invaluable source for soundboard wood quality classification.

5.3.2 Acoustic Feature Analysis

To assess the eligibility of the acoustic feature dataset for linear model analysis, several assumptions were tested, including the examination of data distribution through histograms and scatter plots, as well as the identification of multicollinearity using correlation mapping. The histogram and scatter plot analyses revealed that the distribution of acoustic feature data points generally followed a Gaussian distribution, although some skewness was observed, meeting the normality assumption (refer Appendix D and Appendix E).

During the correlation mapping analysis, it was observed that certain pairs of acoustic features displayed strong correlations, surpassing a threshold of ±0.8. To mitigate the issue of multicollinearity, a pruning process was implemented to remove highly correlated features. As a result, the number of remaining features was reduced to 18, as indicated in Table 5.3. The details of the 18 acoustic features can be found in Appendix F.

Table 5.3: Selected 18 acoustic features.

Acoustic Feature / Descriptor	
Attack Time	Spectral Irregularity
Fluctuation Peak Position	Timbre Zero-cross
Fluctuation Peak Magnitude	Timbre Low Energy
Fluctuation Centroid	Timbre Spectral Flux
Dynamic RMS	Tonal Chromagram Peak Position
Decay Time	Tonal Chromagram Centroid
Spectral Roll-off 85%	Tonal Key Clarity
Spectral Flatness	Tonal Mode
Spectral Roughness	Tonal HCDF

5.3.3 Machine Learning Model Training

After performing normalization, the dataset was then combined with the rating scores provided by Mr. Mathew. All the data consisted of numeric variables, except for the rating scores, which were treated as categorical variables. The datasets were used to train selected models with the following setup: the train-test portion was 252/107 observations, Bayesian optimization with 100 iterations was used as the optimizer, and 5-fold cross-validation was employed for validation.

Table 5.4 presents the ranking of trained models based on their validation accuracies, with a minimum threshold of 70%. KNN, Ensemble, and SVM continue to be among the top-performing models, demonstrating that the acoustic feature dataset is well-suited for non-linear models. However, Neural Network has been outperformed by other models. This suggests that the selected acoustic features exhibit non-linear relationships with the rating scores, which adds complexity to estimating the scores. The discrepancy between validation and test accuracies varies across models, with some experiencing a significant gap of approximately 5%. This suggests potential overfitting, where the models lack flexibility to accurately predict new observations. However, the remaining models still exhibit satisfactory accuracies ranging from 70% to 76%.

Table 5.4: Accuracies of trained models

Model Type	Category	Validation		Test	
		Accuracy %	Total Cost	Accuracy %	Total Cost
Gaussian Kernel	SVM	76.19	60	71.96	30
Subspace KNN	Ensemble	76.19	60	71.03	31
Optimizable ver.	KNN	75.79	61	72.90	29
Quadratic SVM	SVM	75.79	61	72.90	29
Optimizable ver.	Ensemble	73.81	66	76.64	25
Optimizable ver.	Kernel	73.81	66	70.09	32
SVM Kernel	Kernel	73.41	67	69.16	33
Boosted Trees	Ensemble	72.62	69	70.09	32
Wide Neural Network	NN	72.22	70	67.29	35
Cubic SVM	SVM	71.43	72	71.03	31
Bagged Trees	Ensemble	71.03	73	71.03	31
Optimizable ver.	Tree	70.63	74	65.42	37
Optimizable ver.	NN	70.63	74	72.90	29
…	…	…	…	…	…

5.3.4 SMOTE Technique to Tackle Imbalanced Classification

While accuracy is one of the key performance indicators of a trained classification algorithm, it is also important to check the confusion matrix and evaluate whether the model's decision is biased or not. The frequency plot of rating scores by Mr. Mathew in Figure 5.1 reveals that scores of 3 and 4 are the most common, while scores of 2 and 5 are less frequent, indicating a significant gap between the majority and minority categories. This imbalanced dataset is common in the field of social sciences, which the data generated is highly dependent on inherent biasness and can exhibit skewed distribution.

The distribution raised a concern that the model can possibly manipulate the performance accuracies by simply guessing randomly among majority category, achieving high accuracy without truly learning the underlying patterns and making meaningful predictions. Hence, normalized confusion matrices of the best models are analysed. Figure 5.2 shows that the True-Positive Rate (TPR) and False-Negative Rate (FNR) for each class sum up to 100%. The models performed well in predicting class 3, 4, and 5, but struggled with class 2. Within class 2, the TPR and FNR were relatively evenly distributed, indicating that the models often misclassified observations that should belong to class 2 as class 3 instead. This suggests a potential bias in classifying class 2 observations.

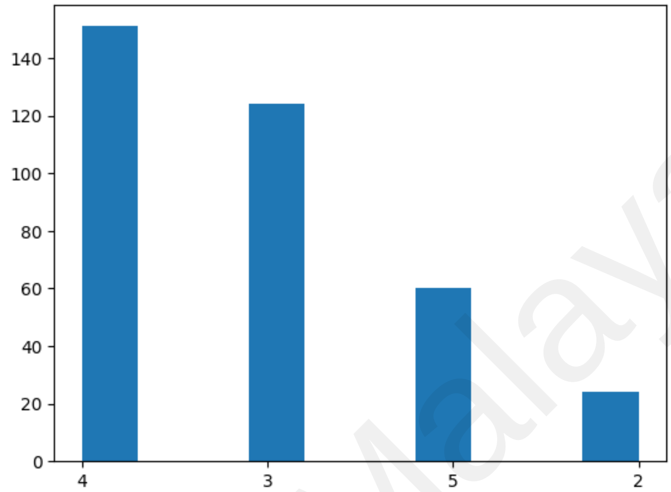


Figure 5.1: Frequency plot of rating scores dataset (360 observations)

To address the issue of class imbalance, resampling techniques were employed. This approach helps mitigate bias that may arise from the imbalanced representation of classes in the dataset, improving the performance and accuracy of the classification models. In this study, the Synthetic Minority Oversampling Technique (SMOTE) was used as the resampling technique. SMOTE generated synthetic samples for class 2, 3, and 5 to match the observation count of class 4. By identifying positive instances within the minority class and creating new observations using an interpolation technique, SMOTE increased the dataset to 604 observations. This data augmentation approach expanded the dataset without introducing exact duplicates, mitigating the risk of overfitting associated with random oversampling techniques.

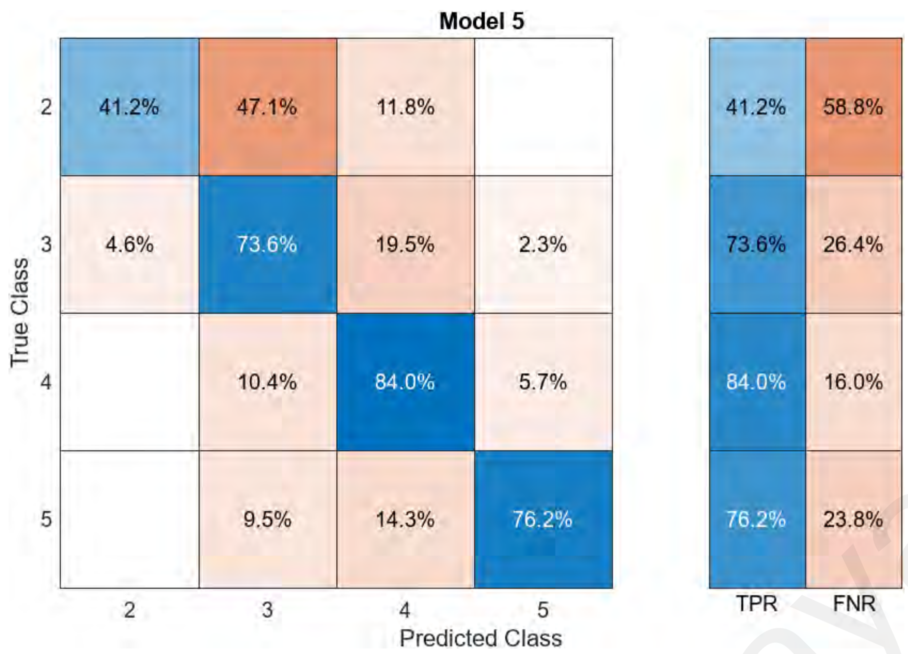


Figure 5.2: Normalized confusion matrix of optimized SVM model against dataset.

Table 5.5 ranks the trained models based on their validation accuracies, with a threshold of 80%. SVM and Ensemble models outperformed other models, indicating that the acoustic feature dataset is well-suited for non-linear models. The drawback of information loss was compensated for by the positive impact of data augmentation using the SMOTE technique.

The confusion matrix of the best model is analyzed for potential bias in formulating decision process. It is observed that SVM is well balanced with TPR overwhelmed FNR as shown in Figure 5.3. The highest FNR across all classes is around 24% on class 4 misclassification and the major contributor is from misclassification of class 3. It is also observed that class 3 has similar FNR against class 4 as well, suggesting that the models are slightly weaker in distinguishing between class 3 and class 4, though the overall performance is still commendable. Note that while the data pool increases in volume, class 3 and 4 practically received less non-synthetic observations post-SMOTE, yet the misclassification FNR dropped as compared to that of pre-SMOTE data pool. This suggests that the models' predictive power against majority class improved regardless, and this improvement is favorable.

Table 5.5: Accuracies of trained models with SMOTE-modified dataset

Model Type	Category	Validation		Test	
		Accuracy %	Total Cost	Accuracy %	Total Cost
Gaussian Kernel	SVM	88.18	50	93.37	12
Optimizable ver.	Ensemble	86.99	55	86.74	24
Quadratic SVM	SVM	86.29	58	87.85	22
Cubic SVM	SVM	86.05	59	92.27	14
Subspace KNN	Ensemble	85.34	62	90.06	18
SVM Kernel	Kernel	85.34	62	88.95	20
Med Gaussian SVM	SVM	84.63	65	88.95	20
Optimizable ver.	KNN	84.40	66	89.50	19
Wide NN	NN	84.16	67	90.61	17
Optimizable ver.	Kernel	84.16	67	87.85	22
Optimizable ver.	NN	83.92	68	87.85	22
Fine KNN	KNN	83.69	69	87.85	22
Boosted Trees	Ensemble	83.69	69	80.11	36
Medium NN	NN	83.22	71	86.19	25
…	…	…	…	…	…

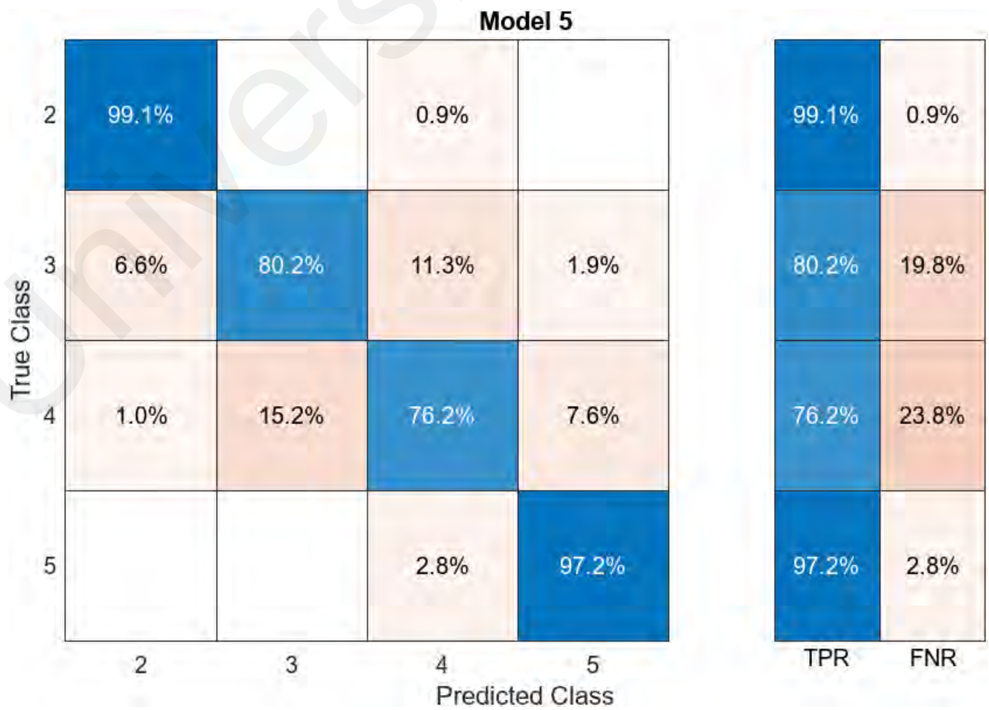


Figure 5.3: Normalized confusion matrix of Gaussian Kernel SVM model against SMOTE-modified dataset.

It is observed that both best performing models are of Gaussian-kernel based SVM which justify the effectiveness of implementing Gaussian-kernel SVM to tackle the classification problem of Sape soundboard wood acoustic quality. The mechanism of both SVMs is based on radial basis kernel which transforms input of n-dimensional to a higher m-dimensional plane such that the dot product of vectors can be computed efficiently. Equation 5.1 describes radial basis function kernel where x and x' are vectors in any fixed dimensional space:

$$\boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{exp}\left(-\frac{\left\|x-x'\right\|^2}{2\sigma^2}\right) \tag{5.1}$$

where σ is the kernel width and $||x - x'||$ is the Euclidean distance between two points x and x'.

If the exponential is expanded, both x and x' will be raised to infinite power, as e^x is an infinite series and the polynomial terms within keep expanding to calculate an exact solution. Hence, radial basis function (RBF) splits the datapoints to hyperplane of infinite dimensions, which provides a strong radial-fitted curve to better distinguish between classes.

Note that for any two points x and x', $||x - x'||^2$ is the Euclidean distance between both points. The distance metric describes dissimilarity between 2 data points such that the further apart, the more dissimilarity between them. The RBF kernel reaches its maximum value of 1 when the Euclidean distance between points is 0, indicating identical points. In this case, there is no distance, signifying a high degree of similarity. Conversely, when points are distant, the kernel value diminishes toward 0, suggesting dissimilarity as the points become more separated. Determining the appropriate kernel width, σ, is crucial in defining the threshold for considering points as similar, and its selection depends on the specific characteristics of each dataset. Finding the optimal σ is vital and can be achieved through hyperparameter tuning methods such as Grid Search Cross Validation and Random Search Cross Validation. This mechanism suggests that RBF kernel is similar to KNN algorithm in terms of identifying boundaries between classes, but RBF overcomes the complexity problem.

5.3.5 Feature Importance by SHAP Interpretation

As indicated previously, MATLAB's built-in functionalities primarily support local interpretability, where SHAP (Shapley Additive Explanations) values are computed for individual observations. Figure 5.4 illustrates the local interpretation of two distinct query points. It's notable that each observation results in a unique hierarchy of feature significance for predicting the quality score. Every observation possesses its own combination of SHAP values contributing to the final prediction. To gain insights into the model's underlying behavior, aggregating SHAP value computations across the entire dataset becomes necessary. This collective computation offers an overarching view of the weight distribution across each feature, elucidating the Gaussian-kernel SVM's estimated feature importance in determining quality score predictions.

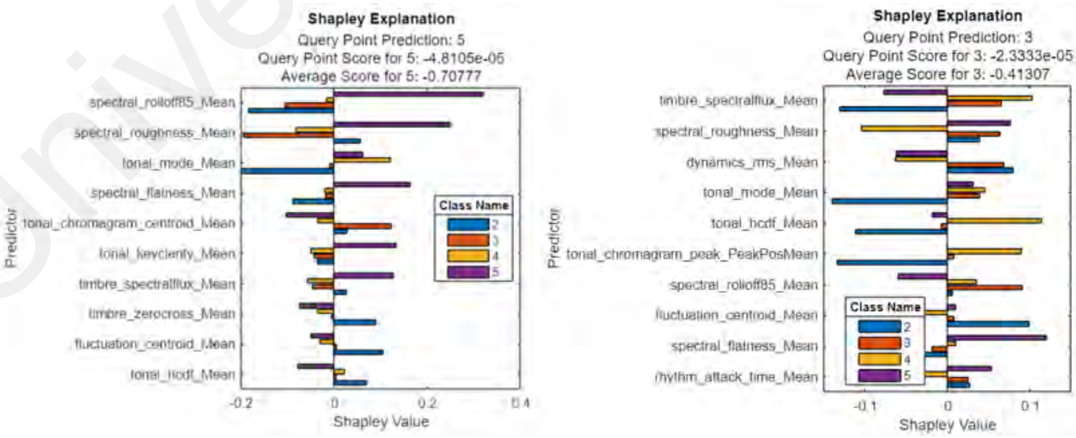


Figure 5.4: Local interpretation by MATLAB built-in SHAP method

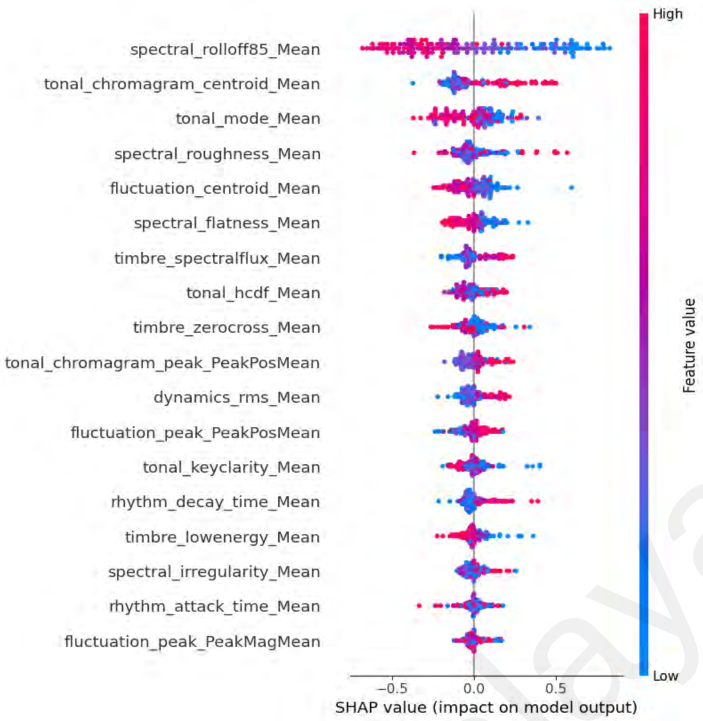


Figure 5.5: Summary plot of collective SHAP computation by Python scikit-learn.

Figure 5.5 presents an ordered representation of feature importance, showcasing a color-coded visual guide where red represents higher variable values and blue indicates lower values for specific observations. This color scheme helps in interpreting the correlation between feature values and the associated quality rating. For instance, the feature "roll-off 85%" exhibits a prominent negative correlation with the quality rating, where higher values (depicted in red) are clustered on the left side of the plot, indicating a negative impact. Conversely, lower values (depicted in blue) are clustered on the right side, suggesting a positive impact. This visual representation offers a clear understanding of how specific features contribute to the overall quality assessment.

Within this ordered arrangement, "roll-off 85%" emerges as the most critical feature for quality prediction, followed by "chromagram centroid," "tonal mode," "spectral roughness," and others, signifying their hierarchical importance. Furthermore, the SHAP result plot highlights that features related to timbre, pitch, and tonality hold more weight in predicting sound quality compared to features associated with dynamics and rhythm.

The prominence of timbre, pitch, and tonality in Asian traditional musical instruments is deeply rooted in the emphasis on these elements, a notion consistent with the findings in this study on Sape instruments. Traditional Chinese music, for instance, is characterized by a harmonious blend of music elements and voice timbres, emphasizing the preservation of high-pitched voices and traditional Chinese music elements (Hu, 2022). This resonance aligns with the Sape's specific attention to timbre, pitch, and tonality, as indicated in the feature importance list. The similarity between the feature importance list of Sape and the emphasis on timbre, pitch, and tonality in other Asian traditional instruments underscores the shared significance of these elements.

Additionally, the importance placed on tonality and pitch in Asian traditional musical instruments is evident in the preservation of traditional music elements and the focused consideration of sound properties like pitch and timbre in musical listening experiences (Clarke, 2001). This mirrors Sape's emphasis on these aspects, as found in this study.

To supplement this analysis, further statistical examinations were performed utilizing the built-in tool in MATLAB, accessible through the Classification Learner app. These analyses aid in discerning the importance of features in predicting the classification output. The outcomes are detailed in Table 5.6, offering a comprehensive summary of feature importance rankings derived from various statistical methods.

Table 5.6: Top 4 features importance across different statistical analyses

SHAP	MRMR	Chi2	ReliefF	ANOVA	Kruskal Wallis
Spectral Roll-off 85	Spectral Roll-off 85	Spectral Roll-off 85	Tonal Mode	Spectral Roll-off 85	Spectral Roll-off 85
Tonal Chromagram Centroid	Tonal Mode	Tonal HCDF	Spectral Roll-off 85	Dynamic RMS	Spectral Flatness
Tonal Mode	Tonal HCDF	Spectral Irregularity	Spectral Roughness	Timbre Spectral Flux	Timbre Spectral Flux
Spectral Roughness	Fluctuation Centroid	Dynamic RMS	Tonal Chromagram Peak	Spectral Flatness	Dynamic RMS

Based on the outcomes, "roll-off 85%" stands out as the most crucial feature for predicting quality, a consistent observation across various statistical analyses, including SHAP, MRMR, Chi2, ANOVA, and Kruskal Wallis. The recurrent identification of "roll-off 85%" as a pivotal factor in the feature importance rankings from diverse statistical methods and SHAP analysis highlights its pivotal role in determining the sound quality classification of Sape soundboards. The rapid decline of high-frequency spectral content, specifically in "roll-off 85%," seems to be a critical factor in discerning different levels of sound quality in Sape instruments. This discovery holds implications for future research, instrument crafting, and quality assessment methodologies concerning Sape soundboards.

5.4 Chapter Summary

In conclusion, this research effectively achieved its objectives by meticulously examining the acoustic features of Sape soundboard wood samples, gathering quality rating scores via listening tests, and establishing a significant relationship between these scores and individual perceptions. Leveraging Mr. Mathew's ratings for predictions, the research diligently trained and optimized machine learning models, addressing dataset imbalances through SMOTE techniques. The study unveiled a notable non-linear correlation between acoustic features and quality ratings, highlighting the Gaussian-kernel SVM as the most effective model, boasting a validation accuracy of 88.18% and a testing accuracy of 93.37%.

In the assessment of feature importance, the computation of SHAP values reveals that "roll-off 85%" holds the utmost significance in predicting quality, with "chromagram" and "tonal mode" following closely. These results resonate with the overarching trends observed in Asian traditional musical instruments, underscoring the fundamental importance of tonality, timbre, and pitch in the evaluation of sound quality. These insights offer valuable guidance for refining quality assessment methodologies in the Sape manufacturing industry, potentially influencing critical decision-making processes.

Looking ahead to Chapter 6, the research will extend its scope by utilizing the well-suited trained model within MATLAB to develop a GUI. This GUI will serve as a user-friendly tool for future Sape makers, seamlessly incorporating the trained machine learning model. The upcoming work in Chapter 6 aims to provide an accessible and efficient platform that integrates advanced technology with traditional craftsmanship, furthering the evolution and enhancement of Sape instrument production.

CHAPTER 6: ADVANCEMENT IN GUI FOR AUTOMATED QUALITY CLASSIFICATION OF SAPE SOUNDBOARD

6.1 Overview

Chapter 6 serves as an extension of the comprehensive research undertaken in earlier chapters, aimed at bridging the existing gap within Sape instrument craftsmanship. The previous chapters have successfully dissected acoustic features, collected quality ratings through listening tests, and integrated machine learning models to predict Sape soundboard wood quality. However, a notable void persists within the domain of user-friendly tools for Sape makers to evaluate soundboard quality during the fabrication process.

In response to this gap, Chapter 6 embarks on the development of a GUI tailored explicitly for Sape instrument makers. The primary objective of this chapter is to provide an accessible and intuitive platform that amalgamates the evaluative insights of both expert Sape makers, and the machine learning model trained in prior research. The GUI seeks to streamline the evaluation process by seamlessly incorporating the expertise of Sape instrument makers with the predictive capabilities of the machine learning model.

This chapter's pivotal goal is to equip Sape instrument craftsmen with a tool that amalgamates traditional expertise with contemporary technological advancements. By amalgamating the qualitative evaluation of soundboard wood by experts with the data-driven predictions of the trained machine learning model, the GUI aspires to offer a robust and user-friendly solution. The envisioned GUI promises to revolutionize the Sape manufacturing landscape by providing an efficient and effective means of evaluating soundboard wood quality during the intricate fabrication process.

6.2 Methodology

The dataset comprising 360 sound samples underwent initial preprocessing steps, including volume normalization and noise reduction. Experienced Sape makers assessed these samples using a 5-point Likert scale. Subsequently, seven crucial acoustic features were extracted using MIRToolbox. A correlation mapping analysis was conducted to detect potential multicollinearity among these features, confirming their minimal correlation and ensuring independence for subsequent analyses.

Upon combining the dataset with Mr. Mathew's rating scores, classification using SVM commenced. It's important to note that while accuracy is significant, examining the confusion matrix is equally vital to identify biases in the model's decisions. Notably, Mr. Mathew's ratings exhibited a prevalence of scores 3 and 4, raising concerns about potential imbalances in class representation.

To mitigate class imbalance issues, the Synthetic Minority Oversampling Technique (SMOTE) was employed. SMOTE helped address unequal class representation by generating synthetic samples for less represented classes (2, 3, and 5), aligning their observations with the more common class 4. This technique expanded the dataset to 604 observations, effectively enhancing its diversity and reducing the risk of model bias associated with imbalanced datasets.

6.2.1 Support Vector Machine in Machine Learning

SVM stands as a widely recognized classifier rooted in the statistical learning theory pioneered by Vapnik (1998). Its fundamental principle revolves around identifying an optimal linear hyperplane that minimizes generalization errors when classifying unknown test samples. This hyperplane, serving as the boundary between distinct categories, is strategically placed to maximize the margin, ensuring a clear separation between various classes of data. However, in instances where a linear hyperplane struggles to effectively segregate data in two-dimensional space, SVM leverages a higher-dimensional space through a technique known as the 'kernel trick.'

In this chapter, the utilization of SVM, particularly the Gaussian kernel variant, emerges as a key strategy for classifying the quality of Sape audio samples based on evaluations by seasoned Sape makers. The Gaussian kernel SVM model specifically finds relevance due to the promising results unveiled in Chapter 5, where it exhibited the highest accuracy in classifying the quality of produced sound.

Given SVM's reliance on distance metrics for classification, ensuring uniformity in the data's scale becomes crucial. In this study, the input data's 18 features will be normalized between the values of 0 to 1 before the classification process. Data normalization involves scaling all features to a uniform range, thereby preventing individual features from disproportionately influencing distance calculations. This standardization is particularly essential as SVM considers the Euclidean distance between data points for classification (Koo et al., 2021). By scaling features to a consistent range, normalization mitigates the risk of bias that might arise due to varying feature scales and ensures fair treatment of each feature's contribution during the classification process.

The Gaussian kernel, as applied in this context, assesses the similarity between data points without explicitly mapping them into higher dimensions. It operates by evaluating similarity based on the Euclidean distance between points, with a parameter, σ, representing the kernel's width. Moreover, to mitigate overfitting, a box constraint value of one was carefully chosen within the SVM model. This parameter plays a crucial role in optimizing the classifier's performance by balancing the goal of maximizing the margin while minimizing classification errors.

The chapter's SVM-based quality classification process was executed in MATLAB 2022b. Following methodology by Vabalas et al. (2019), a standard 70/30 train-test split was applied to the dataset to assess the model's predictive capabilities. Additionally, to ensure robustness and evaluate the model's generalizability with limited data, a k-Fold Cross Validation approach was employed with $k = 5$, aligning with the recommendations outlined by Brownlee (2018).

6.2.2 Graphical User Interface (GUI)

The research methodology adopted in this study combines various computational techniques and machine learning methodologies to evaluate and classify soundboard quality in Sape instruments. Central to this methodology was the development and implementation of a GUI using MATLAB app designer. The GUI, carefully crafted within the MATLAB environment, served as the primary tool for feature extraction and classification of soundboard quality. Leveraging MATLAB's extensive signal processing capabilities, the GUI facilitated the extraction of seven fundamental acoustic features crucial for soundboard quality assessment.

Furthermore, the GUI seamlessly integrated a previously trained SVM model. This SVM model, trained on a diverse and comprehensive dataset encompassing various soundboard qualities, utilized the extracted acoustic features as inputs for classification. The model underwent meticulous optimization to discern and categorize soundboard qualities based on the uploaded sound files. Upon uploading sound files representative of Sape soundboard samples, the GUI executed a series of algorithms designed to systematically extract the essential acoustic features. These algorithms were tailored to capture intricate nuances and characteristic elements relevant to Sape soundboard quality assessment.

The user-friendly interface of the GUI ensures a simple and intuitive process for users to upload sound files. To maintain consistency and accuracy in soundboard quality assessment, it is advised that users adhere to the recording guidelines depicted in Figure 4.2 and Figure 4.3. These guidelines are essential to ensure that the recorded sound aligns with the predetermined settings, promoting a standardized and reliable input for the classification process. Once the sound file is uploaded and processed, the GUI provides immediate classification results, indicating the predicted quality of the Sape soundboard. This seamless interaction streamlines the assessment process, allowing for efficient and precise evaluation of soundboard quality based on the extracted acoustic features.

6.3 Results and Discussion

6.3.1 Support Vector Machine in Machine Learning

The classification results from the Gaussian Kernel SVM model using the normalized dataset are presented in Table 6.1. The model demonstrated notable performance in evaluating the quality of Sape audio samples. During validation, the model displayed an accuracy of 90.3%, indicating its ability to effectively categorize samples based on evaluations provided by experienced Sape makers. When tested on unseen data, the model sustained a strong accuracy of 87.8%, highlighting its capacity to generalize well to new samples. These results signify the model's proficiency in accurately assessing the quality of Sape audio, suggesting its potential for practical application.

Table 6.1: Accuracies of Gaussian Kernel SVM model.

Model Type	Category	Validation		Test	
		Accuracy %	Total Cost	Accuracy %	Total Cost
Gaussian Kernel	SVM	90.3	41	87.8	22

Assessing misclassifications or errors made by the model, referred to as the 'Total Cost', revealed promising outcomes. The reduction in total cost from the validation phase (41) to the testing phase (22) following the optimization of the Gaussian kernel SVM is indicative of an improved model performance on unseen data. Through careful fine-tuning of key parameters, the model demonstrated enhanced generalization, effectively capturing underlying patterns in Sape audio samples. This optimization process, which focused on the box constraint level and kernel scale, contributed to a more balanced trade-off between maximizing the margin and minimizing misclassifications.

The adjusted box constraint level (445.7698) played a crucial role in achieving this balance, influencing the model's ability to accurately classify Sape audio samples. By finding an optimal point, the model reduced both false positives and false negatives, thereby lowering the overall total cost. Simultaneously, the optimization of the kernel scale parameter (0.6360) contributed to the creation of an optimized decision boundary. This boundary, influenced by the width of the Gaussian kernel, became more effective in separating different classes within the Sape audio samples.

Furthermore, the optimization process addressed concerns related to overfitting, ensuring that the model's complexity was well-suited to the characteristics of the Sape audio data. This preventive measure helped the model avoid memorizing the training data while enabling it to learn intricate patterns that generalize effectively to new, unseen data. The specific parameter values chosen through optimization reflected a configuration tailored to the nuances of the Sape audio samples, enhancing the model's accuracy and reducing misclassifications during the testing phase. In summary, the reduction in total cost underscores the effectiveness of the optimized Gaussian kernel

SVM in accurately classifying Sape audio sample quality, demonstrating improved generalization and robustness on previously unseen data.

Figure 6.1 displays the confusion matrix of the SVM model, revealing effective predictions for classes 2 and 5 but encountering challenges with classes 3 and 4. The satisfactory accuracy demonstrated by the Gaussian Kernel SVM model after training reflects its capability to predict the soundboard quality of the Sape instrument. With its validated performance, this model holds promise as a reliable tool for evaluating and predicting soundboard quality within the realm of Sape instrument crafting. The model's proficiency in distinguishing and classifying various soundboard qualities reinforces its potential utility in this specialized domain. This suggests that the model's learned patterns and decision-making capabilities make it a suitable candidate for assessing and predicting the Sape's soundboard quality, thereby contributing to the enhancement of traditional instrument craftsmanship.

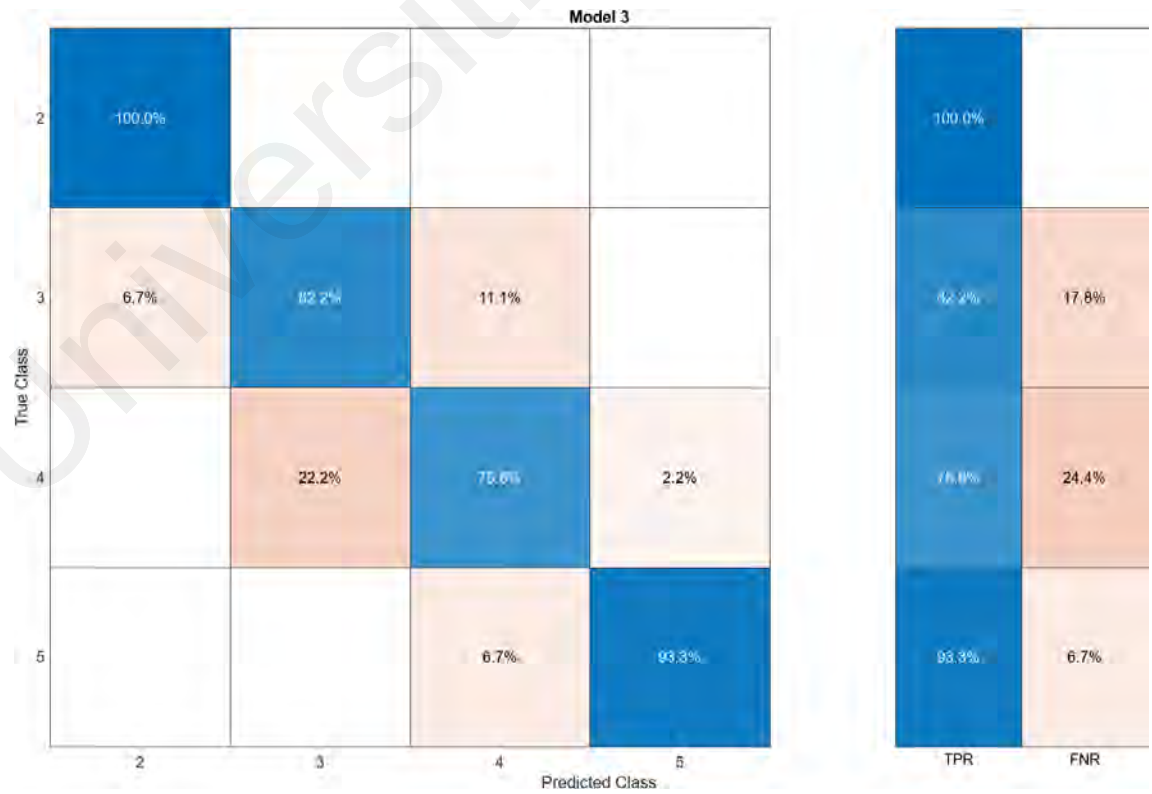


Figure 6.1: Confusion matrix of Gaussian Kernel SVM model.

6.3.2 Limitation in Feature Extraction for GUI Development

During the implementation of the SVM model in MATLAB for integration into the GUI, a notable constraint arose in the direct extraction of features required for sound prediction. It was observed that, out of the 18 features initially utilized, only a subset of features could be directly extracted. The features accessible for direct extraction were dynamic rms, dynamic attack time, decay time, spectral flatness, spectral roughness, timbre low energy, and timbre spectral flux. This limitation stemmed from the inherent constraints within MATLAB's feature extraction capabilities for certain feature types.

The decision to incorporate 7 out of the 18 features into the GUI development was driven by this limitation. Attempting to extract the complete set of features directly in MATLAB proved to be impractical due to the associated technical challenges. The alternative approach involving MIR feature extraction and manual data transfer was considered cumbersome and time-consuming.

While this limitation led to a reduced feature set for direct extraction, it is essential to acknowledge that the selected features maintain significance and relevance to the overall predictive accuracy of the SVM model. This limitation, therefore, underscores the need for further advancements in feature extraction methodologies within MATLAB or alternative tools for seamless integration of comprehensive feature sets in future research endeavors.

6.3.3 Graphical User Interface

The GUI, developed using MATLAB's App Designer, prominently features the utilization of the trained Gaussian Kernel SVM model as the core classification tool. This model serves as the underlying engine for soundboard quality prediction within the Sape instrument manufacturing process. The GUI is thoughtfully designed to offer a seamless experience to potential users, providing an intuitive platform to upload their unique sound files as shown in Figure 6.2. Empowered by an 'Upload Sound File' button, users can effortlessly upload their sound files directly onto the interface. Upon selection, the GUI promptly displays essential details such as the file location and name, ensuring transparency and easy reference throughout the process.

Integral to the interface is a 'Prediction' button that, when activated, triggers the Gaussian Kernel SVM model to perform real-time quality classification of the uploaded sound file. Subsequently, the predicted result is swiftly exhibited in a dedicated box within the interface, providing immediate feedback on the soundboard quality. This functionality empowers users involved in Sape instrument manufacturing by enabling swift and informed decision-making based on the model's predictions.

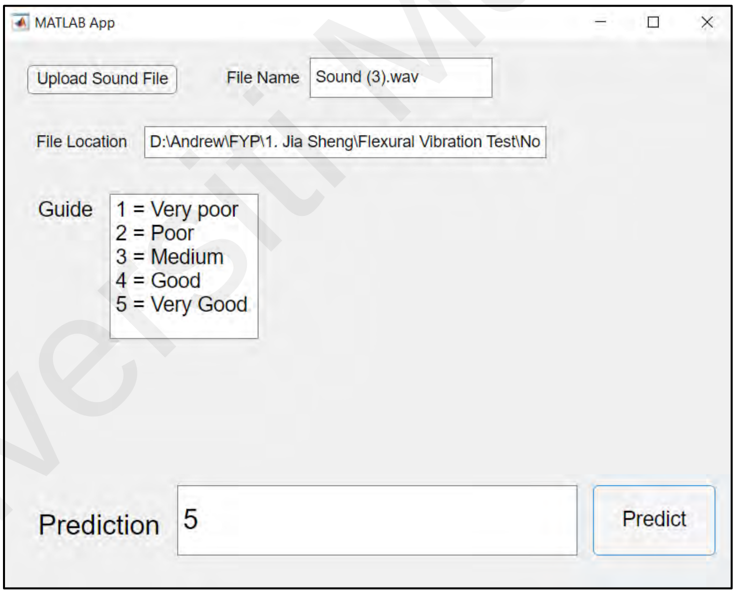


Figure 6.2: GUI for soundboard quality classification by MATLAB App Designer.

Furthermore, GUI is enriched with comprehensive quality guidelines, serving as a reference point for users. This guideline encompasses crucial criteria and information utilized by the Gaussian Kernel SVM model during the classification process. By integrating the trained model and user-friendly functionalities, this GUI streamlines the soundboard quality assessment during the Sape manufacturing process, offering an accessible and efficient tool for manufacturers and craftsmen.

6.4 Chapter Summary

In summary, this study combined acoustic analysis and machine learning techniques to devise a robust tool for assessing soundboard quality in Sape instrument crafting. Beginning with the extraction and analysis of essential acoustic features, bolstered by Mr. Mathew's expert ratings, the research addressed issues of bias and class imbalance through adept resampling techniques like SMOTE. Leveraging a developed Gaussian Kernel SVM model, a user-friendly GUI was meticulously crafted, enabling users to swiftly predict soundboard quality by uploading their own sound files. This integrated approach, bridging traditional craftsmanship with advanced technology, presents a pioneering solution poised to revolutionize Sape instrument manufacturing, offering craftsmen an efficient means to evaluate and predict soundboard quality, thus preserving and advancing the heritage of the Sape instrument.

CHAPTER 7: CONCLUSION AND RECOMMENDATIONS

7.1 Conclusion

This research presents a platform for Sape makers in evaluating the sound quality of soundboard by developing a GUI incorporating machine learning trained with the acoustic features of the sound produced by the soundboard of Sape. This research was driven by the motivation to comprehensively understand and evaluate the factors influencing the sound quality in Sape musical instruments. Numerous factors, including wood type, dimensions, and string material, can influence the quality of musical instruments. Despite their potential impact, there has been a lack of studies comprehensively examining these factors. With this goal in mind, this research first conducted a quantitative and qualitative study to determine the most prominent factors influencing the quality of Sape soundboard. From the survey, data were collected from the online questionnaire where the respondents are the Sape players and makers that are still actively playing, teaching, and making the Sape. The response rate was close to 20%. The data collected were then analyzed using Exploratory Factor Analysis (EFA) and Principal Component Analysis (PCA). The results of EFA and PCA extracted 5 factors with the material being the most prominent factor, followed by the environment, player/maker, design, and size/weight.

On the other hand, the exploration of sound quality factors in Sape instruments involves qualitative investigations. The study into Sape instrument quality, specifically through focus group discussions with experienced Sape makers, showcased a unanimous focus on wood as the core factor shaping Sape quality. The participants highly regarded "Adau" wood due to its balanced properties, emphasizing its significant role in crafting exceptional Sape instruments. Age and the drying process of wood emerged as crucial factors for enhancing the sound quality and longevity of Sape instruments. Moreover, the makers' detailed classification of wood hardness revealed their deep understanding of subtle differences in wood properties that impact Sape quality. Beyond wood, discussions covered various elements like dimensions, strings, frets, and environmental factors. Differing opinions on these aspects hinted at possible room for exploration within Sape-making practices. While certain elements received unanimous acknowledgment for their importance, such as wood and specific structural features, varying viewpoints emerged regarding design elements, painting, and materials used for strings and frets. The insights shared by Sape experts highlighted the complexity of determining Sape quality. While there was a consensus on the influence of wood type, quality, age, and drying processes, differences in perspectives showcased the diversity in Sape craftsmanship. These discussions provided a broader understanding of the diverse considerations, creativity, and regional variations inherent in Sape instrument making. Visual representations like word clouds and mind mapping offered a clear summary of the discussions, emphasizing the significance of wood selection and design while acknowledging the multiple factors affecting Sape quality. With both the qualitative and quantitative approaches showing that the material used in Sape making is the most significant factor, the research proceeds with the study on the material used in Sape-wood.

Chapter 4 involved the creation of nine rectangular samples resembling Sape soundboards, crafted from three distinct wood types, utilizing a CNC machine. Rectangular wood samples were fashioned and tested for physical, vibroacoustic, and timbre characteristics using a flexural vibration test. The objective was to identify the most prominent features influencing the sound quality and develop a reliable method for classifying wood quality. Results showed that Adau wood had the highest quality, followed by Merbau and Tapang wood. Four key features, including acoustic radiation damping coefficient, spectral flux, spectral centroid, and inharmonicity, were selected by the Minimum Redundancy Maximum Relevance (MRMR) algorithm and used to train and test various classifiers in MATLAB. The decision tree classifier achieved 98.1% accuracy in predicting wood quality. This study demonstrates the potential of using machine learning to classify Sape wood quality and provides a useful guide for Sape production. However, this chapter exclusively focused on objective measures. To further enhance and expand the findings, Chapter 5 addresses the research gap by integrating the subjective evaluation of sound quality, assessed by expert Sape makers. This inclusion of subjective evaluations aims to provide a more comprehensive understanding and classification of Sape wood quality beyond solely objective measurements.

In chapter 5, the research expanded the evaluation of the sound data to the subjective measurements. It is done by the evaluation of the experienced Sape makers. Several non-parametric statistical tests were conducted to analyze the quality rating including Kruskal-Wallis H test, Krippendorff's alpha, and Kendall's tau-b to compare the relationship between the rating scores. Collectively, these three non-parametric statistical tests consistently pointed to the sets of rating scores being statistically different from each other, underscoring the lack of inter-rater reliability in the evaluations. Consequently, candidate E emerged as the reliable source for sound quality classification. In this chapter, 18 acoustic features were extracted from the sound data after eliminating the features that are highly correlated. To enhance the accuracy and robustness of the machine learning training, the SMOTE resampling technique is implemented to address class imbalances and mitigate associated biases. The ML results revealed that the Gaussian-kernel SVM emerged as the top-performing model with a validation accuracy of 88.18% and a testing accuracy of 93.37%. These findings offer valuable insights for quality assessment in Sape manufacturing, potentially informing decision-making processes within the industry.

In Chapter 6, the research continues with the development of GUI with the help of the MATLAB app designer. Utilizing a sophisticated Gaussian Kernel SVM model, a user-friendly GUI was intricately designed, facilitating users to promptly assess soundboard quality by uploading their sound files. This innovative integration, merging conventional craftsmanship with cutting-edge technology, introduces a pioneering solution set to transform Sape instrument production. It provides makers with an effective tool to appraise and anticipate soundboard quality, thereby conserving and elevating the Sape instrument's heritage.

This study successfully achieved three main goals. Firstly, it found and studied the key factors that affect Sape quality. Secondly, it used machine learning to evaluate wood quality objectively and subjectively. Lastly, a user-friendly GUI has been developed to provide an automated sound quality classification system. By accomplishing these aims, the research combined Sape makers' knowledge with machine learning, creating a thorough way to judge and predict Sape soundboard quality. This approach helps maintain set traditions and sets a strong base for future studies in this field.

Lastly, it is important to acknowledge that the tools developed in this study, such as the classification model and GUI, aim to assist Sape makers rather than impose strict guidelines. These tools are intended to support the expertise of Sape artisans, enhancing their abilities rather than prescribing standardized methods. Each Sape maker has their unique style and approach, and these technological aids should complement and support their individual expertise. The goal of this study is to provide opportunities for improvement while respecting and preserving the artisanal skill and creativity of Sape makers, thereby maintaining the traditional craft while embracing technological advancements.

7.2 Contributions of Current Research

Some of the contributions made by this research are outlined below:

- This research provides an overview of studies focusing on the classification of musical instruments using machine learning systems. It delves into the two primary stages involved in the automatic classification process: feature extraction and categorization. The classification of musical instruments aligns with the Hornbostel-Sachs system. Within feature extraction, it enumerates pertinent features frequently employed in research and arranges them into a taxonomy based on computational domains. Additionally, it discusses and evaluates various classification methodologies commonly utilized in these studies. This review is presented in Chapter 2.

- Additionally, this research identified the significant factors influencing the sound quality of Sape based on the qualitative and quantitative approaches. The results from both the focus group discussion and questionnaires indicated that the quality of the Sape instrument was affected by several elements. These comprised the type and quality of wood, the dimensions and size of the Sape, the arrangement of frets, and possibly, the instrument's design and the environment in which it was played. These findings imply that the Sape instrument manufacturing sector could enhance its practices by considering a wider array of factors beyond solely the type and quality of wood employed in Sape production. This work is presented in Chapter 3.

- Chapter 4 lays out the initial steps taken in utilizing machine learning for sound quality classification, specifically concerning the Sape instrument. This segment introduces the application of machine learning techniques in categorizing Sape sound quality by integrating physical, vibroacoustic, and timbre features.

- Finally, this research outlines a methodology for integrating sound acoustic features into machine learning, as discussed in Chapter 5. Additionally, Chapter 6 focuses on creating a GUI that integrates the trained Gaussian Kernel SVM model. This advancement has the potential to aid Sape makers in their instrument-making process, potentially reducing both times spent, and errors encountered during production.

7.3 Recommendations for Future Works

This study laid the groundwork for determining the quality of the Sape musical instrument. Yet, it encountered some limitations that need addressing in future research. In upcoming research, it would be beneficial to expand the range of wood samples beyond the three types examined in this study. By incorporating a wider variety of wood commonly used in Sape construction, a more comprehensive understanding of how different woods influence sound quality could be obtained. Moreover, extending the assessment beyond the soundboard to include other parts of the Sape instrument, such as the body or neck, could reveal additional factors affecting overall quality. This broader evaluation could offer insights into aspects beyond sound production, providing a more comprehensive view of Sape construction.

Furthermore, as the current GUI involves pre-recorded sound with a similar experimental setup, a potential future direction could focus on enhancing the usability of the app to enable on-the-spot recording. This improvement aims to facilitate real-time assessment, potentially capturing nuances that may be lost in pre-recorded files. The advantage of such a real-time assessment feature is the ability to evaluate Sape sound quality without the need for an extensive and time-consuming experimental setup. This enhancement could streamline the assessment process, making it more accessible and efficient for both Sape makers and players, ultimately contributing to the advancement of the field.

Another area for improvement lies in the accuracy of sound quality classification, particularly in addressing the low accuracy observed in certain categories (below 80% training or testing accuracy). Specifically, the classifier's inability to identify Class 1 (very poor sound samples) due to a lack of training data needs to be addressed. Future work should consider including the ratings of Class 1 quality by experts to bolster the training data for this category. By doing so, the classifier could achieve a more accurate and comprehensive identification across all quality classes. This enhancement would significantly improve the reliability and utility of the ML model in assessing the sound quality of the Sape instrument.

REFERENCES

Aditanoyo, T. (2018). Vibro-acoustic characterization of Angklung bamboo tube. *Journal of Physics: Conference Series, 1075*, 012024. https://doi.org/10.1088/1742-6596/1075/1/012024

Agostini, G., Longari, M., & Pollastri, E. (2003). Musical Instrument Timbres Classification with Spectral Features. *EURASIP Journal on Advances in Signal Processing, 2003*(1). https://doi.org/10.1155/s1110865703210118

Ahmed, S. A., & Adamopoulos, S. (2018). Acoustic properties of modified wood under different humid conditions and their relevance for musical instruments. *Applied Acoustics, 140*, 92–99. https://doi.org/10.1016/j.apacoust.2018.05.017

Álías, F., Socoró, J. C., & Sevillano, X. (2016). A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds. *Applied Sciences, 6*(5), 143. https://doi.org/10.3390/app6050143

American National Standards Institute. (1973). *American National Standard Psychoacoustical Terminology (Vol. 20)*. American National Standards Institute.

Aramaki, M., Baillères, H., Brancheriau, L., Kronland‐Martinet, R., & Ystad, S. (2007). Sound quality assessment of wood for xylophone bars. *Journal of the Acoustical Society of America, 121*(4), 2407 – 2420. https://doi.org/10.1121/1.2697154

Bacciaglia, A., Ceruti, A., & Liverani, A. (2019). Evaluation of 3D printed mouthpieces for musical instruments. *Rapid Prototyping Journal, 26*(3), 577–584. https://doi.org/10.1108/rpj-07-2019-0187

Bader, R., & Hansen, U. (2008). Modeling of musical instruments. In *Handbook of signal processing in acoustics* (pp. 419-446). Springer.

Barbedo, J. G. A., & Tzanetakis, G. (2010). Musical instrument classification using individual partials. *IEEE Transactions on Audio, Speech, and Language Processing, 19*(1), 111–122. https://doi.org/10.1109/tasl.2010.2045186

Barbosa, J., Malloch, J., Wanderley, M. M., & Huot, S. (2015). What does " Evaluation " mean for the NIME community? *HAL (Le Centre Pour La Communication Scientifique Directe)*. https://hal.inria.fr/hal-01158080

Barlow, C. Y., Edwards, P. P., Millward, G. R., Raphael, R. A., & Rubio, D. J. (1988). Wood treatment used in Cremonese instruments. *Nature, 332*(6162), 313. https://doi.org/10.1038/332313a0

Batahong, R. Y., Wong, T. H., & Dayou, J. (2018, 27-28 Mac 2018). *Anatomy, Physical and Acoustical Properties of Sundatang* [Paper presentation]. Seminar Kebangsaan Bahasa, Budaya dan Sastera Masyarakat Kadazan Dusun, IPG Kampus Kent.

Bhalke, D. G., Rao, C. B. R., & Bormane, D. S. (2016). Automatic musical instrument classification using fractional fourier transform based- MFCC features and counter propagation neural network. *Journal of Intelligent Information Systems, 46*(3), 425–446. https://doi.org/10.1007/s10844-015-0360-9

Bhandari, A. (2020). Feature scaling standardization vs normalization. *Analytics Vidhya*.

Bissinger, G. (2008). Structural acoustics of good and bad violins. *Journal of the Acoustical Society of America, 124*(3), 1764–1773. https://doi.org/10.1121/1.2956478

Bogert, B. P. (1963). The quefrency alanysis of time series for echoes : Cepstrum pseudo-autocovariance, cross-cepstrum, and saphe cracking. *Time Series Analysis*, 209–243. https://ci.nii.ac.jp/naid/10027136573

Brémaud, I. (2012). Acoustical properties of wood in string instruments soundboards and tuned idiophones: Biological and cultural diversity. *Journal of the Acoustical Society of America, 131*(1), 807–818. https://doi.org/10.1121/1.3651233

Brown, J. C. (1999). Computer identification of musical instruments using pattern recognition with cepstral coefficients as features. *Journal of the Acoustical Society of America, 105*(3), 1933–1941. https://doi.org/10.1121/1.426728

Brown, J. C., Houix, O., & McAdams, S. (2001). Feature dependence in the automatic identification of musical woodwind instruments. *Journal of the Acoustical Society of America, 109*(3), 1064–1072. https://doi.org/10.1121/1.1342075

Brownlee, J. (2018). *A gentle introduction to k-fold cross-validation*. Machine learning mastery, 2019.

Bucur, V. (2016). *Handbook of materials for string Musical Instruments*. In Springer eBooks. https://doi.org/10.1007/978-3-319-32080-9

Buksnowitz, C., Evans, R., Müller, U., & Teischinger, A. (2012). Indented rings (hazel growth) of Norway spruce reduce anisotropy of mechanical properties. *Wood Science and Technology, 46*(6), 1239–1246. https://doi.org/10.1007/s00226-012-0480-0

Burrows, D., & Kendall, S. (1997). Focus groups: what are they and how can they be used in nursing and health care research? *Social Sciences in Health, 3*, 244-253.

Cai, Y., Huang, T., Hu, L., Shi, X., Xie, L., & Li, Y. (2012). Prediction of lysine ubiquitination with mRMR feature selection and analysis. *Amino Acids, 42*(4), 1387–1395. https://doi.org/10.1007/s00726-011-0835-0

Campbell, D. M. (2014). Evaluating musical instruments. *Physics Today, 67*(4), 35–40. https://doi.org/10.1063/pt.3.2347

Campbell, J. J. (1997). Speaker recognition: a tutorial. *Proceedings of the IEEE, 85*(9), 1437–1462. https://doi.org/10.1109/5.628714

Campbell, M. (2013). Objective evaluation of musical instrument quality: A grand challenge in musical acoustics. *Proceedings of Meetings on Acoustics*. https://doi.org/10.1121/1.4801037

Chaigne, A., & Kergomard, J. (2016). Acoustics of musical instruments. In *HAL (Le Centre pour la Communication Scientifique Directe)*. https://hal.science/hal-01338980

Charbuty, B., & Abdulazeez, A. M. (2021). Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends, 2*(01), 20–28. https://doi.org/10.38094/jastt20165

Chieng, J. (2011). Unheard Sound: Inequality in time structuring in sampé music. *關渡音樂學刊, 14,* 93 – 112. https://doi.org/10.29873/gmj.201106.0005

Chong, P. L. (2014). *Kenyah recreational songs and ther significance to music education*. https://repository.up.ac.za/handle/2263/40240

Clarke, E. (2001). Meaning and the specification of motion in music. *Musicae Scientiae, 5*(2), 213-234. https://doi.org/10.1177/102986490100500205

Colman, A. M. (2015). *A dictionary of psychology*: Oxford University Press.

Comrey, A. L., & Lee, H. B. (2013a). *A first course in factor analysis*. In Psychology Press eBooks. https://doi.org/10.4324/9781315827506

Comrey, A. L., & Lee, H. B. (2013b). *Interpretation and application of factor analytic results.* In Psychology Press eBooks (pp. 252–274). https://doi.org/10.4324/9781315827506-16

Daltrop, S., Kotlicki, A., & Waltham, C. (2010). Vibro-acoustic characteristics of an Aoyama Amphion concert harp. *Journal of the Acoustical Society of America, 128*(1), 466–473. https://doi.org/10.1121/1.3442576

Daniel, P., & Weber, R. (1997). Psychoacoustical roughness: Implementation of an optimized model. *Acta Acustica united with Acustica, 83*(1), 113-123.

Danihelová, A., Vidholdová, Z., Gergeľ, T., Kružlicová, L. S., & Pástor, M. (2022). Thermal modification of spruce and maple wood for special wood products. *Polymers, 14*(14), 2813. https://doi.org/10.3390/polym14142813

Davis, S. B., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 28*(4), 357–366. https://doi.org/10.1109/tassp.1980.1163420

Deng, J. D., Simmermacher, C., & Cranefield, S. (2008). A study on feature analysis for musical instrument classification. *IEEE Transactions on Systems, Man, and Cybernetics, 38*(2), 429–438. https://doi.org/10.1109/tsmcb.2007.913394

Dinulică, F., Stanciu, M. D., & Savin, A. (2021). Correlation between Anatomical Grading and Acoustic–Elastic Properties of Resonant Spruce Wood Used for Musical Instruments. *Forests, 12*(8), 1122. https://doi.org/10.3390/f12081122

Dünnwald, H. (1991). Deduction of objective quality parameters on old and new violins. *Catgut Acoust. Soc. J, 1*(7), 1-5.

Edward, C. (2018, June 23). A passion for making Sape. *Borneo Post Online*. https://www.theborneopost.com/2018/06/24/a-passion-for-making-Sape/

El-Maleh, K., Klein, M., Petrucci, G., & Kabal, P. (2000, June). Speech/music discrimination for multimedia applications. In *2000 IEEE international* *conference on acoustics, speech, and signal processing. Proceedings (Cat. No. 00CH37100)* (Vol. 4, pp. 2445-2448). IEEE.

Eronen, A. (2001). Comparison of features for musical instrument recognition. *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575).* https://doi.org/10.1109/aspaa.2001.969532

Eronen, A., & Klapuri, A. (2000). Musical instrument recognition using cepstral coefficients and temporal features. In *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 00CH37100).* https://doi.org/10.1109/icassp.2000.859069

Everitt, B. S., & Skrondal, A. (2010). The Cambridge Dictionary of Statistics. https://doi.org/10.1017/cbo9780511779633

Fedyukov, V. I., Saldaeva, E. Y., Chernova, M. S., & Chernov, V. Y. (2019). Research into Dendro-Acoustic Properties of Intro-duced Clones' Wood as Material for Manufacturing Musical Instruments. *South-east European Forestry, 10*(2), 173–179. https://doi.org/10.15177/seefor.19-18

Fiebrink, R., & Caramiaux, B. (2016). *The Machine learning algorithm as creative musical tool*. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1611.00379

Fletcher, H., Blackham, E. D., & Stratton, R. (1962). Quality of piano tones. *Journal of the Acoustical Society of America, 34*(6), 749–761. https://doi.org/10.1121/1.1918192

Fletcher, N. H., & Rossing, T. D. (2012). *The physics of musical instruments*. Springer Science & Business Media.

Fritz, C., & Dubois, D. (2015). Perceptual Evaluation of Musical Instruments: State of the art and Methodology. *Acta Acustica United With Acustica, 101*(2), 369–381. https://doi.org/10.3813/aaa.918833

Fritz, C., Blackwell, A. F., Cross, I., Woodhouse, J., & Moore, B. C. J. (2012a). Exploring violin sound quality: Investigating English timbre descriptors and correlating resynthesized acoustical modifications with perceptual properties. *Journal of the Acoustical Society of America, 131*(1), 783–794. https://doi.org/10.1121/1.3651790

Fritz, C., Curtin, J., Poitevineau, J., Morrel-Samuels, P., & Tao, F. (2012b). Player preferences among new and old violins. *Proceedings of the National Academy of Sciences of the United States of America, 109*(3), 760–763. https://doi.org/10.1073/pnas.1114999109

Fu, Z., Lu, G., Ting, K. M., & Zhang, D. (2010). A Survey of Audio-Based Music Classification and Annotation. *IEEE Transactions on Multimedia, 13*(2), 303–319. https://doi.org/10.1109/tmm.2010.2098858

Gazengel, B., & Petiot, J. (2013). Objective and subjective characterization of saxophone reeds. *HAL (Le Centre Pour La Communication Scientifique Directe)*. https://hal.science/hal-00782949

Gazengel, B., Petiot, J., & Brasseur, E. (2010). Vers la définition d'indicateurs de qualité d'anches de saxophones. *HAL (Le Centre Pour La Communication Scientifique Directe)*. https://hal.science/hal-00534625

Gerhard, D. (2003). *Audio signal classification: History and current techniques*. Regina, SK, Canada: Department of Computer Science, University of Regina.

Ghozali, I. (2018). Enculturation of Kledi Dayak kebahan Penyelopat (Inheritance, Organology Study, and playing techniques of Traditional Musical Instrumentin remote area of West Kalimantan). *TEROB, 10*(1), 1-27.

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M. A., & Kagal, L. (2018). Explaining Explanations: An Overview of Interpretability of Machine Learning. *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*. https://doi.org/10.1109/dsaa.2018.00018

Golpayegani, A. S., Brémaud, I., Gril, J., Thévenon, M., Arnould, O., & Pourtahmasi, K. (2011). Effect of extractions on dynamic mechanical properties of white mulberry (Morus alba). *Journal of Wood Science, 58*(2), 153–162. https://doi.org/10.1007/s10086-011-1225-7

Gore, T. (2011). Wood for guitars. *Proceedings of Meetings on Acoustics.* https://doi.org/10.1121/1.3610500

Gorlinski, V. K. (1989). *Sampeq of the Kenyah of East Kalimantan, Indonesia: A Case Study of the Recreational Music Tradition*. University of Hawai'i at Manoa.

Green, D. W. (1999). *Wood Handbook Wood as an engineering material*. http://www. fpl. fs. fed. us/documnts/fplgtr/fplgtr113/ch04. pdf, 4.

Hartanto, C. K., Darmawan, D. R., Manalu, C. R., & Lenny, A. (2021). Alat Musik Tradisional Di Masa Modern (Sape'Dayak Kayaan Dalam Kajian Nilai Budaya). *Gondang: Jurnal Seni Dan Budaya, 5*(2), 182-192.

Hashim, H. (2017). *Fusion Sape' with Build in Stand, Speaker and Battery* [Thesis]. Universiti Malaysia Sarawak.

Hassan, K., & Tippner, J. (2019). Acoustic properties assessment of neem (Azadirachta indica A. Juss.) wood from trees irrigated with secondarily treated wastewater. *Bioresources, 14*(2), 2919–2930. https://doi.org/10.15376/biores.14.2.2919-2930

Herrera, P., Peeters, G., & Dubnov, S. (2003). Automatic classification of musical instrument sounds. *Journal of New Music Research, 32*(1), 3–21. https://doi.org/10.1076/jnmr.32.1.3.16798

Hu, M. (2022). Features of singing in chinese pop and traditional music: the influence of the music genre on vocal music. *Revista Música Hodie*, 22. https://doi.org/10.5216/mh.v22.73137

Hunt, D. G., & Balsan, E. (1996). Why old fiddles sound sweeter. *Nature, 379*(6567), 681. https://doi.org/10.1038/379681a0

Hutchins, C. M., Hopping, A. S., & Saunders, F. A. (1960). Subharmonics and plate tap tones in violin acoustics. *Journal of the Acoustical Society of America, 32*(11), 1443–1449. https://doi.org/10.1121/1.1907934

Jansson, E. V. (1990). Experiments with the violin string and bridge. *Applied Acoustics, 30*(2–3), 133–146. https://doi.org/10.1016/0003-682x(90)90042-s

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kaminsky, I., & Materka, A. (1995). Automatic source identification of monophonic musical instrument sounds. In *Proceedings of ICNN'95-International Conference on Neural Networks*. https://doi.org/10.1109/icnn.1995.488091

Karlina, A., Muniir, I., & Muniir, A. (2018). Proses Pembuatan Alat Musik Sape Di Desa Capkala Kabupaten Bengkayang. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa (JPPK), 7*(8).

Kedem, B. (1986). Spectral analysis and discrimination by zero-crossings. *Proceedings of the IEEE, 74*(11), 1477–1493. https://doi.org/10.1109/proc.1986.13663

Kemp, J., Logie, S., Chick, J., Smith, R. A., & Campbell, M. (2010). Analysis of transients for brass instruments under playing conditions using multiple microphones. *HAL (Le Centre Pour La Communication Scientifique Directe).* https://hal.archives-ouvertes.fr/hal-00539676

Kim, H., Moreau, N., & Sikora, T. (2004). Audio classification based on MPEG-7 spectral basis representations. *IEEE Transactions on Circuits and Systems for Video Technology, 14*(5), 716–725. https://doi.org/10.1109/tcsvt.2004.826766

Klapuri, A., & Davy, M. (2007). *Signal processing methods for music transcription*. In Springer eBooks. https://doi.org/10.1007/0-387-32845-9

Kline, P. (2013). *Handbook of Psychological Testing*. In Routledge eBooks. https://doi.org/10.4324/9781315812274

Koo, B., Kim, J. M., Nam, Y., & Kim, Y. (2021). The performance of Post-Fall detection using the Cross-Dataset: feature vectors, classifiers and processing conditions. *Sensors, 21*(14), 4638. https://doi.org/10.3390/s21144638

Krueger, R. A., & Casey, M. A. (2000). *A practical guide for applied research*. A practical guide for applied research.

Kumari, K., & Yadav, S. (2018). Linear regression analysis study. *Journal of the Practice of Cardiovascular Sciences, 4*(1), 33. https://doi.org/10.4103/jpcs.jpcs_8_18

Lartillot, O. (2021). *MIRtoolbox 1.8.1 user's manual*. Oslo: University of Oslo

Lembaga Perindustrian Kayu, M. (2010). *100 Malaysian timbers : 2010 edition*. Malaysian Timber Industry Board.

Li, K., Chen, Y., Jiang, H., & Han, Y. (2019). Chinese Traditional Musical Instrument Evaluation Based on a Smart Microphone Array Sensor. *Multidisciplinary Digital Publishing Institute Proceedings, 15*(1), 40. https://doi.org/10.3390/proceedings2019015040

Li, T., Ogihara, M., & Li, Q. (2003). A comparative study on content-based music genre classification. In *Proceedings of the 26th Annual International ACM SIGIR*

Conference on Research and Development in Informaion Retrieval. https://doi.org/10.1145/860435.860487

Lim, C. K. N., & Abdul Rahman, M. F. (2016). Evolution of Sape: From longhouse to the international stage. *Journal of Borneo Kalimantan, 2*(1), 42-52. https://ir.unimas.my/id/eprint/13741

Lim, C. K. N., Rashidi, N. H., & Yahya, S. A. (2016). *Sape: alat muzik tradisional Sarawak. Dewan Bahasa dan Pustaka*. https://books.google.com.my/books?id=M4bsnQAACAAJ

Lim, C. K. N., Rashidi, N. H., & Yahya, S. A. (2020). *Sarawak Traditional Musical Instrument Sape (First Edition ed.)*. Dewan Bahasa dan Pustaka.

Lin, C. P., & Nie, C. L. K. (2021). The Indigenous Music of Sarawak and Its Transmission Over the Last 60 Years with a Special Focus on the Music of the Kenyah and the Lun Bawang. *Sinophone and Taiwan Studies*, 165–198. https://doi.org/10.1007/978-981-16-4473-3_7

Liu, M., Peng, L., Lyu, S., & Lyu, J. (2020). Properties of common tropical hardwoods for fretboard of string instruments. *Journal of Wood Science, 66*(1). https://doi.org/10.1186/s10086-020-01862-7

Liu, Z., Huang, J., Wang, Y., & Chen, T. (1997, June). Audio feature extraction and analysis for scene classification. In *Proceedings of First Signal Processing Society Workshop on Multimedia Signal Processing* (pp. 343-348). IEEE. https://doi.org/10.1109/mmsp.1997.602659

Liu, Z., Liu, Y., & Shen, J. (2006). Advances in study and research on acoustic property of wood for soundboard of musical instrument. *Journal of Northwest Forestry University, 21*(3), 124.

Lu, L., Jiang, H., & Zhang, H. (2001). A robust audio classification and segmentation method. In *Proceedings of the Ninth ACM International Conference on Multimedia*. https://doi.org/10.1145/500141.500173

Lundberg, S., & Lee, S. (2017). A unified approach to interpreting model predictions. *Neural Information Processing Systems, 30*, 4768–4777. https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf

MacDonald, M. (1956). *Borneo people*.

Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*(7), 674–693. https://doi.org/10.1109/34.192463

Marques, J., & Moreno, P. J. (1999). A study of musical instrument classification using gaussian mixture models and support vector machines. *Cambridge Research Laboratory Technical Report Series CRL, 4*, 143.

Martin, K. D., & Kim, Y. E. (1998). Musical instrument identification: A pattern-recognition approach. *Journal of the Acoustical Society of America, 104(3_Supplement)*, 1768. https://doi.org/10.1121/1.424083

Masullo, M., Yamauchi, K., Dan, M., Cioffi, F., & Maffei, L. (2021). Intercultural differences in the perception of hvac sound quality in car cabins: from conventional to electric vehicles. *Applied Sciences, 11*(23), 11431. https://doi.org/10.3390/app112311431

Mazarakis, G., Tzevelekos, P., & Κουρουπέτρογλου, Γ. (2006). *Musical instrument recognition and classification using time encoded signal processing and fast artificial neural networks*. In Springer eBooks (pp. 246–255). https://doi.org/10.1007/11752912_26

McAdams, S. (2019). *The perceptual representation of timbre*. In Springer handbook of auditory research (pp. 23–57). https://doi.org/10.1007/978-3-030-14832-4_2

McKinney, M., & Breebaart, J. (2003). Features for audio and music classification.

Meyer, J. (2009). *Acoustics and the performance of music: Manual for acousticians, audio engineers, musicians, architects and musical instrument makers*. Springer Science & Business Media.

Mitchell, T. M. (1997). *Machine learning. 1997*. Burr Ridge, IL: McGraw Hill, 45(37), 870-877.

Mitrović, D., Zeppelzauer, M., & Breiteneder, C. (2010). Features for content-based audio retrieval. In *Advances in computers* (Vol. 78, pp. 71-150). Elsevier. https://doi.org/10.1016/s0065-2458(10)78003-7

Molnar, C. (2020). *Interpretable Machine learning*. https://openlibrary.org/books/OL28325212M/Interpretable_Machine_Learning

Mörchen, F., Ultsch, A., Thies, M., & Lohken, I. (2005). Modeling timbre distance with temporal statistics from polyphonic music. *IEEE Transactions on Audio, Speech, and Language Processing, 14*(1), 81–90. https://doi.org/10.1109/tsa.2005.860352

Murty, M. N., & Devi, V. S. (2015). *Introduction to pattern recognition and machine learning (Vol. 5)*. World Scientific.

Musib, A. F. (2015). *Contextual sound preservation of selected local string instruments* [Doctoral dissertation, Universiti Putra Malaysia, Malaysia].

Myers, C. S. (1914). A Study of Sarawak Music. Sammelbände der Internationalen Musikgesellschaft, 15(H. 2), 296-308.

Nunnally, J. (1978). *Psychometric Theory 2nd edition* (New York: McGraw).

Oppenheim, A. V., Schafer, R. W., & Stockham, T. G. (1968). Nonlinear filtering of multiplied and convolved signals. *Proceedings of the IEEE, 56*(8), 1264–1291. https://doi.org/10.1109/proc.1968.6570

Pallant, J. (2013). SPSS survival manual: a step by step guide to data analysis using IBM SPSS. *Australian and New Zealand Journal of Public Health, 37*(6), 597–598. https://doi.org/10.1111/1753-6405.12166

Pampalk, E., Rauber, A., & Merkl, D. (2002). Content-based organization and visualization of music archives. *Proceedings of the Tenth ACM International Conference on Multimedia*. https://doi.org/10.1145/641007.641121

Peeters, G. (2004). A large set of audio features for sound description (similarity and classification) in the CUIDADO project. *CUIDADO IST Project Report, 54*(0), 1-25.

Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27*(8), 1226–1238. https://doi.org/10.1109/tpami.2005.159

Peter, J. P. (1979). Reliability: A review of psychometric basics and recent marketing practices. *Journal of Marketing Research, 16*(1), 6-17. https://doi.org/10.2307/3150868

Pilo, W. (2018, December 10). Sape master still going strong. *DayakDaily*. https://dayakdaily.com/Sape-master-still-going-strong/

Pinard, F., Laine, B., & Vach, H. (2003). Musical quality assessment of clarinet reeds using optical holography. *Journal of the Acoustical Society of America, 113*(3), 1736–1742. https://doi.org/10.1121/1.1543586

Pölkki, J., Erkut, C., Penttinen, H., Karjalainen, M., & Välimäki, V. (2003). New Designs for the Kantele with Improved Sound Radiation. In *Proceedings of the Stockholm Music Acoustics Conference.* https://research.aalto.fi/en/publications/new-designs-for-the-kantele-with-improved-sound-radiation(8cf912d4-d16c-48fa-ab5d-249e24409063).html

Quinlan, J. (2014). *C4. 5: programs for machine learning*. Elsevier.

Ramalingam, A., & Krishnan, S. (2006). Gaussian mixture modeling of Short-Time Fourier Transform features for audio fingerprinting. *IEEE Transactions on Information Forensics and Security, 1*(4), 457–463. https://doi.org/10.1109/tifs.2006.885036

Roohnia, M. (2019). *Wood: vibration and acoustic properties*. In Elsevier eBooks. https://doi.org/10.1016/b978-0-12-803581-8.01996-2

Rosato, D. V. (2003). *Plastics engineered product design*. Elsevier.

Roth, H. L., & Low, H. B. (1896). *The natives of Sarawak and British North Borneo: Based chiefly on the mss. of the late HB Low, Sarawak government service (Vol. 1)*. Truslove & Hanson.

Rummel, R. J. (1988). *Applied factor analysis*. Northwestern University Press.

Saitis, C., Fritz, C., Scavone, G. P., Guastavino, C., & Dubois, D. (2017). Perceptual evaluation of violins: A psycholinguistic analysis of preference verbal descriptions by experienced musicians. *Journal of the Acoustical Society of America, 141*(4), 2746–2757. https://doi.org/10.1121/1.4980143

Saitis, C., Järveläinen, H., & Fritz, C. (2018). *The role of haptic cues in musical instrument quality perception.* In Springer series on touch and haptic systems (pp. 73–93). https://doi.org/10.1007/978-3-319-58316-7_5

Scaringella, N., Zoia, G., & Mlynek, D. (2006). Automatic genre classification of music content: a survey. *IEEE Signal Processing Magazine, 23*(2), 133–141. https://doi.org/10.1109/msp.2006.1598089

Scheirer, E., & Slaney, M. (1997, April). Construction and evaluation of a robust multifeature speech/music discriminator. In *1997 IEEE international conference on acoustics, speech, and signal processing* (Vol. 2, pp. 1331-1334). IEEE.

Schelleng, J. C. (1968). Acoustical effects of violin varnish. *Journal of the Acoustical Society of America, 44*(5), 1175–1183. https://doi.org/10.1121/1.1911243

Schmid, G.-M. (2015). *Evaluation of musical instruments from the musician's perspective* [Master Thesis, University of Basel, Switzerland].

Sedik, Y., Hamdan, S., Jusoh, I., & Hasan, M. (2010). Acoustic properties of selected tropical wood species. *Journal of Nondestructive Evaluation, 29*(1), 38–42. https://doi.org/10.1007/s10921-010-0063-7

Shelford, R. (1904). An illustrated catalogue of the ethnographical collection of the Sarawak Museum. *Journal of the Straits Branch of the Royal Asiatic Society, 40*, 1–59. https://www.biodiversitylibrary.org/part/184776

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America, 36*(12), 2346–2353. https://doi.org/10.1121/1.1919362

Siswanto, W. A., & Syiddiq, M. (2018a). Demonstration of Comparison between Goat Skin and X-Ray Film Membranes on Traditional Musical Instrument Kompang. *Pertanika Journal of Science and Technology, 26*(2).

Siswanto, W. A., Abdullah, M. S., & Darmawan, A. S. (2018b). Effect of humidity on the membrane vibration of musical instrument kompang. *Int J Mech Eng Tech, 9*(6), 1233-1240.

Slamet, S., Kusumaningtyas, I., & Suyitno, S. (2020). Comparative study of bonang gamelan musical instrument between hot forging and Post Cast Heat Treatment/PCHT on microstructure and mechanical properties. *Journal of Physics: Conference Series, 1430*(1), 012054. https://doi.org/10.1088/1742-6596/1430/1/012054

Slamet, S., Suyitno, S., Kusumaningtyas, I., & Miasa, I. (2021). Effect of high-tin bronze composition on physical, mechanical, and acoustic properties of Gamelan materials. *Archives of Foundry Engineering, 21*.

Smith, J. R. L., Rey, G., Dickens, P., Fletcher, N. H., Hollenberg, L. C. L., & Wolfe, J. (2007). Vocal tract resonances and the sound of the Australian didjeridu (yidaki). III. Determinants of playing quality. *Journal of the Acoustical Society of America, 121*(1), 547–558. https://doi.org/10.1121/1.2384849

Spycher, M., Schwarze, F. W. M. R., & Steiger, R. (2008). Assessment of resonance wood quality by comparing its physical and histological properties. *Wood Science and Technology, 42*(4), 325–342. https://doi.org/10.1007/s00226-007-0170-5

Stoel, B. C., & Borman, T. M. (2008). A Comparison of Wood Density between Classical Cremonese and Modern Violins. *PLos One, 3*(7), e2554. https://doi.org/10.1371/journal.pone.0002554

Sudom, G. R. (2016). *An Exploration of Melody, Harmony and Rhythm in Sape repertoires in Kenyah Ethnic in Sarawak* [Thesis]. Universiti Teknologi MARA.

Sudom, G. R., & Naili, R. (2022). An Exploration of Melody, Harmony and Rhythm in Selected Sape Repertoires in Kenyah Ethnic in Sarawak. *Environment-Behaviour Proceedings Journal, 7*(SI9), 413-419. https://doi.org/10.21834/ebpj.v7isi9.4290

Suzuki, S., Gerner, P., & Lirk, P. (2019). *Local anesthetics*. In Pharmacology and Physiology for Anesthesia (pp. 390-411). Elsevier.

Swisher, L. L., Beckstead, J. W., & Bebeau, M. J. (2004). Factor analysis as a tool for survey analysis using a professional role orientation inventory as an example. *Physical Therapy, 84*(9), 784–799. https://doi.org/10.1093/ptj/84.9.784

Thomas, L., MacMillan, J., McColl, E., Hale, C., & Bond, S. (1995). Comparison of focus group and individual interview methodology in examining patient satisfaction with nursing care. *Social Sciences in Health, 1*(4), 206-220.

Tremain, T. E. (1982). The government standard linear predictive coding algorithm: LPC-10. *Speech Technology*, 40-49.

Tronchin, L. (2005). Modal analysis and intensity of acoustic radiation of the kettledrum. *Journal of the Acoustical Society of America, 117*(2), 926–933. https://doi.org/10.1121/1.1828552

Tronchin, L. (2020). Special issue on musical instruments: acoustics and vibration. *Applied Sciences, 10*(9), 3294. https://doi.org/10.3390/app10093294

Tzanetakis, G., & Cook, P. R. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing, 10*(5), 293–302. https://doi.org/10.1109/tsa.2002.800560

Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PLos One, 14*(11), e0224365. https://doi.org/10.1371/journal.pone.0224365

Vapnik, V. (1998). Statistical learning theory Wiley. *New York*, 1, 624.

Wang, A. (2003). *An Industrial Strength Audio Search Algorithm*. ISMIR. http://dblp.uni-trier.de/db/conf/ismir/ismir2003.html#Wang03

Wegst, U. G. K. (2006). Wood for sound. *American Journal of Botany, 93*(10), 1439–1448. https://doi.org/10.3732/ajb.93.10.1439

Wold, E., Blum, T., Keislar, D., & Wheaten, J. (1996). Content-based classification, search, and retrieval of audio. *IEEE MultiMedia, 3*(3), 27–36. https://doi.org/10.1109/93.556537

Wong, T. H., Batahong, R. Y., & Dayou, J. (2022). Fretting Effect on the Fundamental Sound Frequency of Sape. *Borneo Journal of Sciences & Technology, 4*(1), 18-25. https://doi.org/http://doi.org/10.3570/bjost.2022.4.1-04

Wong, T. H., Chang, J. H. W., Che, F. P., & Dayou, J. (2017). Effects of String Tension to Fundamental Frequency of Sound and Body Vibration of Sape. *Transactions on Science and Technology, 4*(4), 437-441.

Wong, T. H., & Dayou, J. (2019). Effects of Plucking Force to Fundamental Frequency of Sound and Body Vibration of Sape. *e-Bangi, 16*(3).

Wong, T. M. (2002). *A dictionary of Malaysian timbers*. Forest Research Institute Malaysia (FRIM).

Wu, Q., Jou, I., & Lee, S. (1997). On-line signature verification using LPC cepstrum and neural networks. *IEEE Transactions on Systems, Man, and Cybernetics, 27*(1), 148–153. https://doi.org/10.1109/3477.552197

Yang, F., Jiang, Y., & Xu, Y. (2022). Design of bird sound recognition model based on lightweight. *IEEE Access, 10*, 85189-85198.

Yang, Y., Liu, Y., Liu, Z., & Shi, S. Q. (2017a). Prediction of Yueqin acoustic quality based on soundboard vibration performance using support vector machine. *Journal of Wood Science, 63*(1), 37–44. https://doi.org/10.1007/s10086-016-1598-8

Yang, Y., Liu, Z., & Liu, Y. (2017b). Prediction of lute acoustic quality based on soundboard vibration performance using multiple choice model. *Journal of Forestry Research, 28*(4), 855–861. https://doi.org/10.1007/s11676-016-0365-4

Yoshikawa, S. (2007). Acoustical classification of woods for string instruments. *Journal of the Acoustical Society of America, 122*(1), 568–573. https://doi.org/10.1121/1.2743162

Yoshikawa, S., & Waltham, C. (2014, July). Woods for wooden musical instruments. In *Proceeding of ISMA* (pp. 281-286).

Zhang, X., & Raś, Z. W. (2007). Analysis of Sound Features for Music Timbre Recognition. *2007 International Conference on Multimedia and Ubiquitous Engineering (MUE'07)*. https://doi.org/10.1109/mue.2007.85

Zorič, A., & Kaljun, J. (2018). The influence of the acoustic properties of wood for the production of electric solid body guitars. In *DAAAM international scientific book* .. (pp. 195–210). https://doi.org/10.2507/daaam.scibook.2018.18

Zwicker, E., & Fastl, H. (2013). *Psychoacoustics: Facts and models (Vol. 22)*. Springer Science & Business Media.