Data in Brief 47 (2023) 108936

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Acoustic dataset of coconut (Cocos nucifera) based on tapping system

June Anne Caladcad[a], Eduardo Jr Piedad[b,*]

[a] Department of Industrial Engineering, University of San Carlos, Cebu City 6000, Philippines
[b] Computer Software Division, Advanced Science and Technology, Quezon City 1101, Philippines

ARTICLE INFO

Article history:
Received 29 November 2022
Revised 12 January 2023
Accepted 20 January 2023
Available online 26 January 2023

Dataset link: Acoustic Signal Dataset: Tall Coconut Fruit Species (Original data)

Keyword:
Coconut
Sound
Tapping
Cocos nucifera
Acoustic spectrum
Machine learning

ABSTRACT

During the fruit sample preparation process, coconut fruits classified under the tall coconut variety in their post-harvest period are considered the subject of this article. All samples are pre-classified by local farmers and experts into three maturity levels; premature, mature, and overmature. Each coconut underwent the synchronized tapping and recording process using developed hardware and software. The analog recordings are then converted into digital signals. Sampled frequency and amplitude in discrete-time signals of each sample went through a quantization process. The data presented in this article provides the general differentiation of the coconuts according to their maturity levels through their acoustic properties. This dataset can also be useful in creating an advanced and intelligent classification system of fruits through machine learning and deep learning techniques.

DOI of original article: 10.1016/j.compag.2020.105327

* Corresponding author.
E-mail address: eduardojr.piedad@asti.dost.gov.ph (E.J. Piedad).

https://doi.org/10.1016/j.dib.2023.108936

Specifications Table

Subject	Agriculture Engineering, Computer Science, Hardware and Architecture
Specific subject area	Maturity Levels of Philippine Coconut Fruits, Acoustic Signals, Acoustic Features, Coconut Maturity Level Classification
Type of data	Table Figure
How the data were acquired	Each sample was acquired using developed hardware and software. The hardware is composed of a mechanical knocking arm that was controlled by a 360° servo motor. An ATR2500 USB condenser microphone was used to capture the tapping sounds of the coconut. Both the knocking arm and microphone are connected to the microcontroller (MCU) to control the timing and the duration of the capturing of sound. Data acquisition and the control of the tapping process were facilitated using the software application. The hardware was also built with acoustic foams to conceal, if not, minimize undesirable interference from the surrounding environment.
Data format	Raw Analyzed
Description of data collection	Three acoustic signals were gathered from tapping the three central side ridges of each of the 129 coconuts samples. A total of 387 signals were acquired using the developed hardware and software. The initial recordings gathered were converted through the analog-to-digital converter module of the MCU, which has a 16-bit information length and a sampling frequency of 44.1 kHz following the Nyquist sampling theorem. Each acoustic signal has 132,300 time-series data points.
Data source location	All coconut samples were acquired from: · City/Town/Region: Argao · Country: Philippines The recording process was conducted in: · Institution: University of San Jose - Recoletos · City/Town/Region: Cebu · Country: Philippines
Data accessibility	Repository name: Mendeley Data Data identification number: 10.17632/hxh8kd3snj.1 Direct URL to data: https://data.mendeley.com/datasets/hxh8kd3snj
Related research article	J.A. Caladcad, S. Cabahug, M.R. Catamco, P.E. Villaceran, L. Cosgafa, K.N. Cabizares, M. Hermosilla, E.J. Piedad, Determining Philippine coconut maturity level using machine learning algorithms based on acoustic signal, Computers and Electronics in Agriculture. 172 (2020) 105327. https://doi.org/10.1016/j.compag.2020.105327

Value of the Data

- The data presents and provides insights into the differences in the acoustic properties of coconuts based on their maturity levels; premature, mature, and overmature.
- The data would be useful to agricultural sectors and large coconut exporting countries and companies to address inefficiencies and reduce food loss and waste brought by the use of traditional methods and improve the management aspect of classifying fruits.
- When utilized in machine learning and deep learning, this data will help supplement existing data compilations to establish a better classification system for large commercial use.

1. Objective

The Philippines is one of the largest coconut-exporting countries all over the world [1]. Coconut is also the largest employer of agricultural land and labor in the country [2]. In fact, 69 provinces and about 3.6 million hectares of land are dedicated to planting and growing coconut trees [3]. About 25% to 33% of the country's population relies on the coconut industry as their livelihood [4]. Despite this status, the country still heavily relies on the traditional way of sorting coconuts for commercial trade and export.

Traditionally, coconuts are classified into their maturity levels manually. Traders often use their fingernails, knuckles, or the blunt end of the knife to tap the coconuts before assessing the sounds produced [5]. Such a process could incur drawbacks from human-related constraints like inconsistency, variability, and subjectivity, among others [6–8].

A mechanized system was developed following the traditional methods of classifying coconuts to their maturity levels. Acoustic signals gathered through a synchronized system are the primary dataset presented in this article that was gathered from the 129 coconut samples. These acoustic signals contain the features of coconuts to be used in assessing them for sorting. These datasets were also used as a pre-requisite requirement in developing an intelligent classification system for sorting coconuts using machine learning algorithms based on the acoustic signal in the study of Caladcad et al. [9].

2. Data Description

The raw dataset has a total of 387 samples, which was the result of knocking 129 coconut samples on its three ridges. The acoustic signals acquired from the knocking process are separated per ridge on the associated data of this work. They are labeled as "Ridge A", "Ridge B", and "Ridge C". Each sample is then named with the coconut sample number and their maturity level code. Shown in Fig. 1 is how the data samples are named and in Table 1 are the maturity level codes for each maturity level.

To summarize the dataset presented, the peak amplitudes for each sample are used for brevity. Each sample's peak amplitudes were gathered per maturity level. Using these values, the maximum, minimum and average peak amplitudes were attained to represent the acoustic behavior of the maturity levels. Coconuts under premature classes have the lowest peak amplitudes. Peak amplitudes increase with maturity, with overmature coconuts having the highest amplitudes. This can be clearly seen in the illustration shown in Fig. 2. Table 2 summarizes the peak amplitudes of all samples per maturity level.

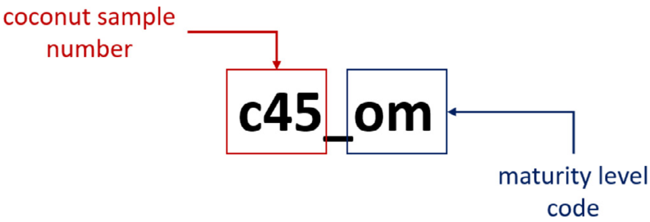

Fig. 1. Data sample naming.

Table 1
Maturity levels and their corresponding code.

Maturity level	Maturity level code
Premature	im
Mature	m
Overmature	om

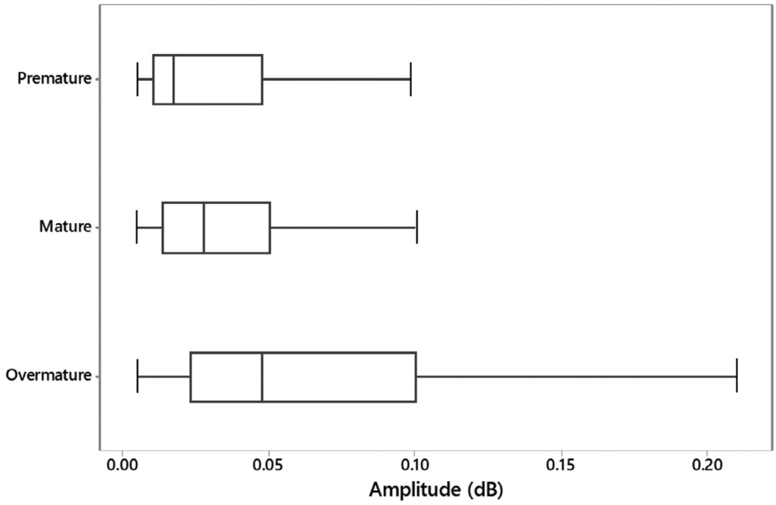

Fig. 2. Peak amplitude range per maturity class.

Table 2
Summary of peak amplitudes of coconut samples.

Maturity Level	Peak amplitude (dB)			
	Minimum	Maximum	Mean*	Standard Deviation
Premature	0.00562	0.09854	0.035841	0.030625
Mature	0.00479	0.24026	0.037322	0.034906
Overmature	0.00534	0.40558	0.075085	0.077525

* Note: The mean peak amplitude represents the average peak amplitude of all the samples per maturity level.

3. Experimental Design, Materials and Methods

3.1. Coconut sample preparation

All coconut samples were acquired from a local farm in Cebu, Philippines. The considered subjects are coconuts classified under the tall coconut variety during their post-harvest period. For practical applications, samples are individually classified into three maturity levels. There are different levels of coconuts identified but most commonly, they are classified into premature, mature, and overmature maturity levels [1,10]. Prior to data gathering, local experts and farmers pre-classified the harvested coconuts according to their maturity level.

A total of 129 coconut samples were gathered randomly, which there are 8, 35, and 85 coconuts for premature, mature, and overmature levels, respectively. These samples were gathered randomly during the harvesting period. This resulted in the imbalance of coconut samples across maturity levels due to the dominance of overmature coconuts. It was also costly to add more samples, especially for premature samples, at another time. Furthermore, identifying maturity levels by vision only is indistinguishable, making it difficult to determine the best time to gather data and add more premature level coconut samples. This imbalance could potentially affect the signal distribution and processes conducted post-sample collection from the lack and/or limited representation of the maturity level with a lesser number of samples. For instance, in classifying tasks in the case of machine learning processes, the learning model might not perform accurately, especially on maturity levels that are underrepresented. To alleviate this issue further, each sample was tapped uniformly at all of their three ridges, multiplying each class three times [5]. There is a total of 387 signals were acquired, wherein each signal has 132,300 time-series data points.

3.2. Hardware specifications

The prototype design of the coconut tapping system was developed primarily for a reliable data acquisition system. An illustration of the hardware is presented in the study of Caladcad et al. [9] while the block diagram of the developed hardware is shown in Fig. 3. The prototype is composed of the following specifications:

- a mechanical knocking arm with a hard rubber end for tapping the coconuts;
- pivot swing foam;
- acoustic foam;
- 360° servo motor;
- ATR25000 USB condenser omnidirectional microphone; and
- a microcontroller.

The analog-to-digital converter module of the MCU has a 16-bit information length with a sampling frequency of 44.1 kHz following the Nyquist sampling theorem [11]. The sampled frequency and amplitude in discrete-time signals underwent the quantization process, and the timing of the tapping process and the computer through the developed software application synchronize recording.

3.3. Software development

The software application will facilitate data acquisition, as well as control the tapping process. It also enables the user to collect information and visualize the acoustic signal recording after the tapping process is done. The designed hardware is synchronized with the software application through the MCU. During the knocking process, a command will be sent to the servo motor to automatically start the tapping process in a consistent motion and force. The software application is also presented in the study of Caladcad et al. [9].

The software has the following features.

- Check the microphone status
- Check automatic recording status (i.e., saving to a local folder, ongoing knocking, generating waveform, converting to .csv format, playing record, and saving to the database)
- Provide options for ridge label (i.e. ‘Ridge A’, ‘Ridge B’, and ‘Ridge C’)
- Initialize the synchronized tapping process with the hardware
- View waveform option
- Convert .wav to .csv option
- Listen to the recording option
- View ridges comparison option
- Starting a new recording

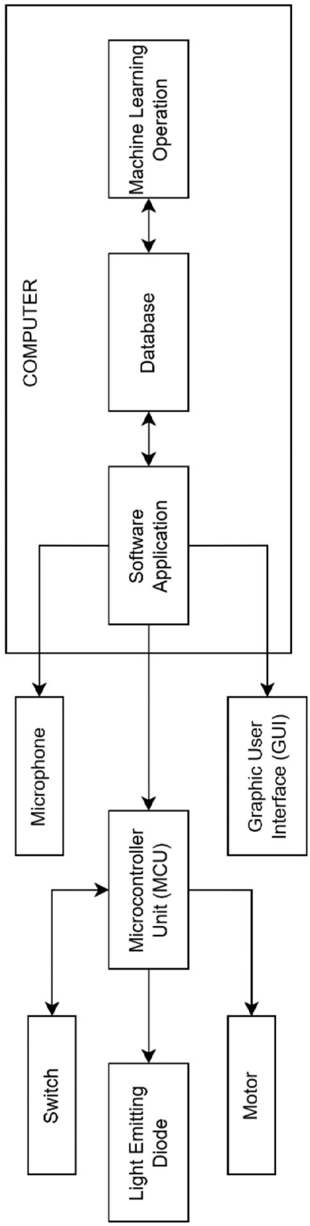

Fig. 3. Block diagram of the developed hardware.

Ethics Statements

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Acoustic Signal Dataset: Tall Coconut Fruit Species (Original data) (Mendeley Data).

CRediT Author Statement

June Anne Caladcad: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing; **Eduardo Jr Piedad:** Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] D.T. Burns, E.-L. Johnston, M.J. Walker, Authenticity and the potability of coconut water - a critical review, J. AOAC Int. 103 (2020) 800–806, doi:10.1093/jaocint/qsz008.

[2] M.L. Moreno, J.K.M. Kuwornu, S. Szabo, Overview and constraints of the coconut supply chain in the Philippines, Int. J. Fruit Sci. 20 (2020) S524–S541, doi:10.1080/15538362.2020.1746727.

[3] Philippine Coconut Authority, Coconut Statistics, https://pca.gov.ph/index.php/resources/coconut-statistics. 2018 (accessed May 2022).

[4] A.M. Pascua, Impact damage threshold of young coconut (Cocos nucifera L.), Int. J. Adv. Agric. Sci. Technol. 4 (2017) 1–9.

[5] M.M. Gatchalian, S.Y. De Leon, Measurement of young coconut (cocos nucifera, l.) maturity by sound waves, J. Food Eng. 23 (1994) 253–276, doi:10.1016/0260-8774(94)90053-1.

[6] B. Zhang, W. Huang, J. Li, C. Zhao, S. Fan, J. Wu, C. Liu, Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: a review, Food Res. Int. 62 (2014) 326–343, doi:10.1016/j.foodres.2014.03.012.

[7] K. Hameed, D. Chai, A. Rassau, A comprehensive review of fruit and vegetable classification techniques, Image Vis. Comput. 80 (2018) 24–44, doi:10.1016/j.imavis.2018.09.016.

[8] G. ElMasry, S. Cubero, E. Molto, J. Blasco, In-line sorting of irregular potatoes by using automated computer-based machine vision system, J. Food Eng. 112 (2012) 30–38, doi:10.1016/j.jfoodeng.2012.03.027.

[9] J.A. Caladcad, S. Cabahug, M.R. Catamco, P.E. Villaceran, L. Cosgafa, K.N. Cabizares, M. Hermosilla, E.J. Piedad, Determining Philippine coconut maturity level using machine learning algorithms based on acoustic signal, Comput. Electron. Agric. 172 (2020) 105327, doi:10.1016/j.compag.2020.105327.

[10] A. Terdwongworakul, S. Chaiyapong, B. Jarimopas, W. Meeklangsaen, Physical properties of fresh young Thai coconut for maturity sorting, Biosyst. Eng. 103 (2009) 208–216, doi:10.1016/j.biosystemseng.2009.03.006.

[11] M. Baker, Demystifying Mixed Signal Test Methods, 2003 Burlington, Massachusetts.