Method of Extracting Speech Characteristics of Bugis Regional Language

Sriwijanaka Yudi Hartono[1)] Saidah Suyuti[2)]
[1,2] Electrical Engineering Study Program Fakulty of Engineering Muslim University of Indonesia
JL. Urif Sumoharjo Km 05 Makassar South Sulawesi Indonesia ,90231
Sriwijanaka.hartono@umi.ac.id[1)] saidahsuyuti06@gmail.com[2]

Abstract *--- his research focuses on developing an application to extract speech signal characteristics of Bugis regional speakers using Matlab and Wavesurfer software. The study centers on feature extraction, specifically Short Time Energy (STE), Autocorrelation Cepstral for pitch extraction, and Cepstrum Linear Predictive Coding (LPC) for formant analysis. Results indicate that male voices have greater signal energy than female voices. The pitch values for male speakers range from 52.29 Hz to 135.76 Hz, while female speakers' pitch ranges from 73.39 Hz to 242.42 Hz. The formant analysis showed that male and female speakers have distinct formant frequencies, with male speakers having lower formant frequencies. For instance, the first formant (f1) in male speakers ranges from 433.0 Hz to 590.3 Hz, while in female speakers it ranges from 273.9 Hz to 452.8 Hz. The pitch of male voices is concentrated between 50 Hz and 250 Hz, whereas female voices fall between 120 Hz and 500 Hz. Additionally, differences in formants (f1 to f5) were observed in the word "na'baca," with female speakers generally having higher formant frequencies. This analysis provides valuable insights into the acoustic characteristics of the Bugis regional language. The developed application aims to enhance speech signal processing technologies, supporting applications like speech recognition, linguistic analysis, and voice synthesis, contributing to the preservation of the Bugis language through signal processing techniques.*
Keywords*: characteristic extraction, short time energy, pitch, formant*

I. Introduction

Regional languages hold significant cultural and identity value within communities, but with the rapid advancement of technology and globalization, many regional languages are at risk of extinction. Bugis, a language spoken in South Sulawesi, is one such language with unique linguistic features that are vital to the culture and identity of its speakers. However, due to limited research and technological tools available for the analysis of regional languages, the preservation and understanding of Bugis language through modern technology remain underdeveloped. Therefore, the need to explore and analyze the acoustic characteristics of Bugis speech becomes essential, especially as regional languages face the challenge of disappearing in the digital age.

This research is motivated by the necessity to develop a technological solution that can assist in the preservation and analysis of Bugis language through speech signal processing. By focusing on feature extraction methods such as Short Time Energy (STE), pitch measurement using Autocorrelation Cepstral, and formant analysis with Cepstrum Linear Predictive Coding (LPC), this study aims to offer a deeper understanding of the Bugis language's acoustic properties. These techniques are fundamental in understanding speech signals and can be applied in various fields such as speech recognition, linguistic analysis, and language synthesis.

Several studies have investigated speech feature extraction methods for language processing. For example, Hartono and Sayuti (2024) explored pitch extraction for regional languages, particularly focusing on Bugis language using Autocorrelation algorithms [4]. Meanwhile, other research has focused on the development of speech feature extraction systems for general applications, such as those by Wang et al. (2019) on speech feature extraction for oral English [2] and by Mazumder and Salam (2019) on various feature extraction techniques for speech processing [3]. Additionally, Gill (2016) provides a review of different speech processing techniques, which contribute to understanding the broader context of feature extraction methods in speech analysis [5].

This research is expected to provide a comprehensive overview of the acoustic characteristics of the Bugis language, leveraging advanced speech signal processing techniques to aid in its preservation. The developed application will not only assist in linguistic analysis but also contribute to the enhancement of speech signal processing technologies, supporting the effort to maintain and preserve the Bugis language in a technological context.

II. Research Methodology

its relevance in speech signal processing and its ability to extract meaningful features from speech signals in the context of the Bugis language.

1. ***Data Collection*** *The first step in this research involves collecting speech samples from native Bugis speakers. The samples are selected from different regions in South Sulawesi to ensure a diverse set of voices. Both male and female speakers are included in the dataset to analyze potential gender-related differences in speech characteristics. Each participant is asked to pronounce specific words or sentences that represent typical speech patterns in Bugis, including phrases such as /"na'baca"/ (to read), /'ma'bicara ogi"/ (to speak clearly), and /"taro ada taro gau"/ (put here, put there). These phrases are chosen for their relevance in daily conversations and their representation of common linguistic features in Bugis speech. The recordings are made using high-quality microphones to capture clear and accurate sound data, ensuring*

that the speech signals are suitable for the analysis process.

2. ***Preprocessing*** *The collected audio samples are processed to ensure that the signals are suitable for analysis. Preprocessing includes noise reduction, signal normalization, and segmentation to isolate individual speech units (e.g., words or syllables). This step is crucial for ensuring the accuracy of feature extraction in the subsequent stages.*
3. ***Feature Extraction*** *Feature extraction is conducted in three main stages:*
   a. ***Short Time Energy (STE)****: This technique is applied to analyze the energy of the speech signal over short time intervals. STE is used to determine the strength or amplitude of the speech signal, allowing for the identification of speech segments, such as voiced or unvoiced sounds.*
   b. ***Pitch Extraction using Autocorrelation Cepstral****: The pitch of the speech signal is extracted using the autocorrelation method, which identifies the periodicity of the speech signal. The Cepstral technique is applied to refine the extraction of pitch, which is important for understanding the prosody and intonation of the Bugis language. The pitch values are recorded for both male and female speakers to observe the differences in pitch ranges.*
   c. ***Formant Analysis using Cepstrum Linear Predictive Coding (LPC)****: Formants are key frequencies that shape the vocal tract's resonances. LPC is used to model the speech signal and extract formant frequencies, which help in identifying the vowel characteristics and other phonetic features of the Bugis language. Formant analysis helps compare male and female speech characteristics, focusing on the differences in formant frequencies between genders.*
4. ***Analysis and Interpretation*** *After feature extraction, the results are analyzed to identify patterns and differences between the speech characteristics of male and female speakers. The pitch values and formant frequencies are compared to assess the phonetic differences between the two groups. Additionally, the STE values are analyzed to determine the overall energy distribution in the speech signal. The results of these analyses are used to draw conclusions about the acoustic properties of the Bugis language.*

Figure 1 shows the block diagram of the conceptual framework in this research process.

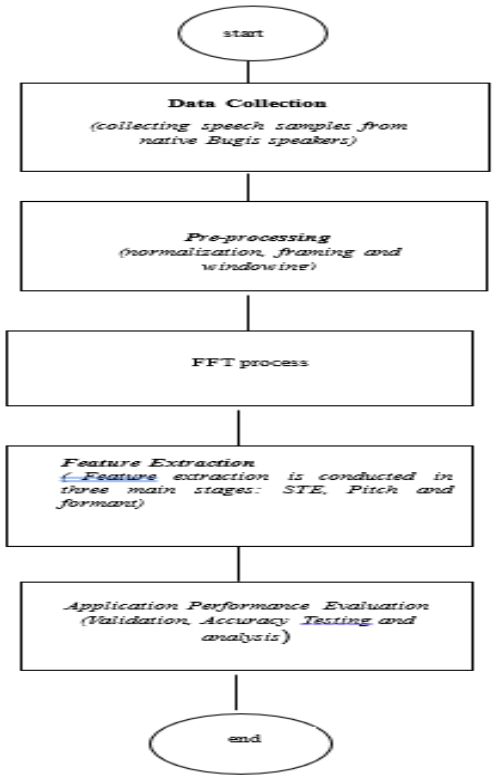

Figure 1 the block diagram of the research process

As for the feature extraction process,it can be seen in figure 2 below :

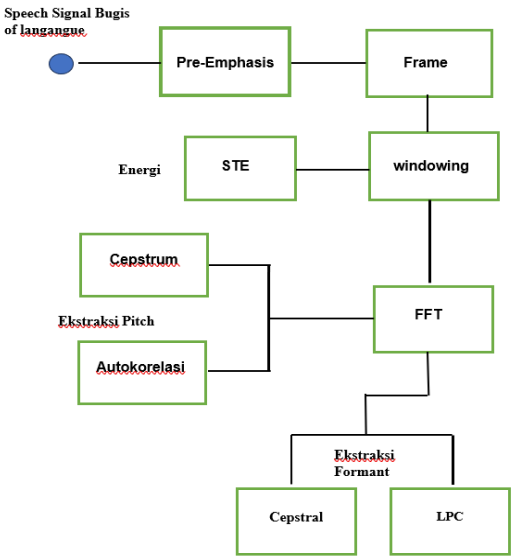

Figure 2. the block diagram future Extraction process speech Bugis regional of language.

III. Results and Discussion

The feature extraction method performed in this research focuses on extracting speech signals based on the

Bugis local language. The initial stage involves **recording**, which generates the original speech signals from utterances (words/sentences) in the Bugis language to serve as the input data for the database. The utterances in the Bugis language used for the recording are: /”nabaca”/, /”ma’bicara ogi”/, and /”taro ada taro gau”/. The recordings were made using MATLAB software and WavSurfer software, with settings configured for a sampling frequency of 12 kHz, mono channel, line 1, and a resolution of 16-bit, which resulted in audio files in the ".wav" format.

The recorded speech signals are then processed in the pre-processing stage, which aims to optimize the quality of the speech signal and eliminate unwanted disturbances such as noise or background sounds. Pre-processing helps minimize interference and highlight the relevant information in the speech signal that will be used for feature extraction.

Next, the framing and windowing processes are applied. The purpose of this step is to divide the speech signal into several frames, making it easier to analyze the speech signal. The framing process is designed in such a way that the speech signal does not experience significant changes over time, which facilitates the observation of temporal characteristics. Meanwhile, the windowing process is used to reduce the effects of discontinuity caused by segmenting the signal into small frames. Windowing also ensures that the spectral analysis conducted on the signal frames is accurate and helps in performing the Fast Fourier Transform (FFT).

A. Extraction The Short Time Energy (STE)

step is the **Short Time Energy (STE)** process, which is a graph that illustrates the variation in energy within the speech signal over time. Each point on the graph represents the energy value calculated for a specific time window. Below is one of the test results for STE on the Bugis language utterance /”na’baca”/ in the table 1. As seen in the results, the **energy values** fluctuate based on the speaker's gender and regional origin. For instance, **Bugis 05** (male from Bulukumba) had the highest energy value of 1.827 watts, while **Bugis 04** (female from Pinrang) had the lowest energy value of 0.369 watts.

The differences in energy values across various speakers may indicate different speaking styles, speech rates, and regional phonetic characteristics. The **higher energy values** from male speakers in regions like **Bulukumba** could suggest more intense speech, possibly due to regional dialect differences or speech emphasis. In contrast, **female speakers** tended to have lower energy values, which could be attributed to speech rate and pitch characteristics.

Table 1. Energy values for Bugis language utterance /“nabaca”/:

No.	Signal Data Name	Speaker’s Gender	Origin of Bugis Speaker’s Speech	Speech Signal Energy (watt)
1	Bugis 01	Laki-laki	Sidrap	0,949
2.	Bugis 02	Perempuan	Bone	0,734
3	Bugis 03	Laki-laki	Sinjai	1,103
4	Bugis 04	Perempuan	Pinrang	0,369
5	Bugis 05	Laki-laki	Bulukumba	1,827
6.	Bugis 06	Perempuan	Sidrap	0,433

B. Pitch Extraction Results Using Cepstrum Autocorrelation Method

The pitch extraction experiment was conducted using the cepstrum autocorrelation method with the aid of MATLAB programming, utilizing a Bugis language speech signal database in .wav format. The speech signals were recorded from native Bugis speakers, with utterances n Bugis language example process such as /“nabaca” (reading). In using the Autocorrelation Cepstrum to display pitch value information, this is done because speech signals are not always in a periodic condition, but instead fluctuate over time, and the signal is mixed with noise. Below is Figure 3., showing the cepstrum autocorrelation process of one speech data containing a word/sentence in the Bugis regional language."

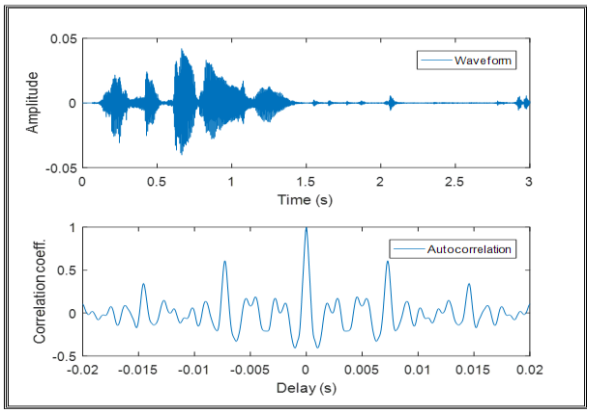

Figure 3.Showing the cepstrum autocorrelation process of speech data containing a word/sentence in /”nabaca”/.

Nest, show in table 2. presents the pitch frequency values (in Hz) extracted from the Bugis language utterance /“nabaca”/ using the cepstrum autocorrelation method.

Table 2. Showing the cepstrum autocorrelation process of one speech data containing a word/sentence /”nabaca”/in the Bugis regional language

No.	Signal Data Name	Speaker's Gender	Origin of Bugis Speaker's Speech	Pitch extraction Frequency (Hz)
1	Bugis 01	Laki-laki	Sidrap	56,93
2.	Bugis 02	Perempuan	Bone	74,42
3	Bugis 03	Laki-laki	Sinjai	52,29
4	Bugis 04	Perempuan	Pinrang	95,80
5	Bugis 05	Laki-laki	Bulukumba	68,89
6.	Bugis 06	Perempuan	Sidrap	73, 39

Pitch extraction using the cepstrum autocorrelation method provided valuable insights into the fundamental frequency of Bugis language speech. Pitch frequencies represent the vocal fold vibration rate and are critical for understanding prosodand voice characteristics.

- Male speakers generally had lower pitch frequencies, such as Bugis 03 (from Sinjai) with 52.29 Hz, while female speakers exhibited higher pitch values, such as Bugis 02 (from Bone) with 74.42 Hz. This finding is consistent with known differences in voice pitch between male and female speakers.
- The pitch values for Bugis 01 (from Sidrap) and Bugis 05 (from Bulukumba) showed some regional variations, with Bugis 01 having a pitch of 56.93 Hz and Bugis 05 having 68.89 Hz. These differences in pitch frequencies may suggest regional variations in pronunciation, which is an important factor in the linguistic analysis of Bugis.

The pitch extraction results emphasize the role of fundamental frequency in characterizing speech rhythm and intonation,

C. Formant Extraction Results Using Cepstral LPC Method

The last feature extraction is the extraction of formants from speech signals in the Bugis regional language using Matlab programming. Formants are peaks in the frequency spectrum produced by the resonance of the vocal tract when a person produces sound. Each human voice has a unique series of formants that give it its distinctive characteristics. These formant frequencies are closely related to the resonance patterns that occur as sound passes through the mouth cavity, throat, and other vocal tract passages.

By using the Cepstral LPC method, a series of formant frequency values can be obtained from the speech signal. Figure 4. below will display one of the wav database files that has undergone formant extraction using the cepstral LPC method, which is one of the wav files from the speech signal in the Bugis regional language."ntence in the Bugis regional language."

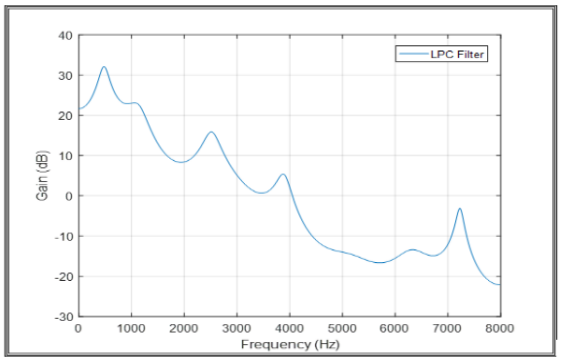

Figure 4. Diisplay formant extraction using the cepstral LPC method in speech signal Bugis regional lamguage

The final step in the feature extraction process is the extraction of formants using the cepstral LPC method, with the help of MATLAB programming. The speech signal used for the formant extraction was from the Bugis language utterance /"nabaca"/ (reading). Below are the results of the formant extraction for the Bugis utterance /"nabaca"/. Table 3. Formant Extraction Results for the Bugis Language Utterance /"nabaca"/ Using Cepstral LPC Method

Tabel 3. Formant Extraction Results for the Bugis Language Utterance **/"nabaca"/** Using Cepstral LPC Method

No.	Signal Data Name speaker's	Speaker's Gender	Origin of Bugis Speaker's Speech	*Formant* (f, Hz)				
				f.1	f.2	f.3	f.4	f.5
1.	Bugis 01	Laki-laki	Sidrap	522,2	639,9	1680,1	2434,6	3723,7
2.	Bugis 02	perempuan	Bone	273,9	583,0	1850,2	2674,8	4049,2
3.	Bugis 03	Laki-laki	Sinjai	433,0	456,9	1329,4	2391,3	3845,8
4.	Bugis 04	Perempuan	Pinrang	452,8	868,4	1688,1	2330,5	3311,5
5.	Bugis 05	Laki-laki	Bulukumba	590,3	693,4	1548,0	2669,3	3090,6
6.	Bugis 06	Perempuan	Sidrap	314,0	868,4	1938,8	2564,1	3409,3

Formants, which are the resonant frequencies of the vocal tract, were extracted using the cepstral LPC method. The formant frequencies provide a deep insight into the articulatory characteristics of speech and are vital for understanding speech perception.

- The formant frequencies for male speakers generally had higher frequencies for the higher formants (f.4 and f.5) compared to female speakers. For example, Bugis 01 (male from Sidrap) showed f.1 = 522.2 Hz, f.2 = 639.9 Hz, while Bugis 02 (female from Bone) exhibited f.1 = 273.9 Hz, f.2 = 583.0 Hz.
- The regional influence on formant frequencies was also evident, with Bugis 05 (from Bulukumba) showing formants such as f.1 = 590.3 Hz and f.2 = 693.4 Hz, while Bugis 06 (from Sidrap) had f.1 = 314.0 Hz and f.2 = 868.4 Hz. This suggests that regional dialectal differences affect both the vocal tract configuration and the articulatory patterns of Bugis speakers.

The formant extraction results underscore the importance of these frequencies in analyzing the vocal tract shape and articulatory properties of the speaker. The differences between male and female speakers, as well as regional variations, highlight the diversity of Bugis language phonetics and provide a foundation for speech synthesis and language preservation technologies.

IV. Conclusion

Based on the results of the research, simulations, and analyses conducted, the following conclusions can be drawn :

1. Differences in Speech Signal Characteristics Between Male and Female Speakers: This study found significant differences in the parameters of signal energy, pitch, and formants between male and female speakers. The energy of the male speech signal was higher compared to female speakers, with the pitch range for male speakers between 52.29 Hz – 135.76 Hz and for female speakers between 73.39 Hz – 242.42 Hz. Additionally, the formant frequencies of male speakers were generally lower than those of female speakers, as observed in the analysis of various words and sentences tested.
2. Contribution to the Development of Speech Signal Processing Technology for the Bugis Regional Language: The application developed in this research is expected to make a significant contribution to speech signal processing for applications such as voice recognition, linguistic analysis, and speech synthesis. Furthermore, this study aims to support the preservation of the Bugis language through a technology-based approach, as well as open opportunities for furthe

Suggestions:

1. Further Development of the Application: It is suggested that the speech signal feature extraction application for the Bugis language be further developed by improving the accuracy of feature extraction and expanding the variety of words or sentences tested. Further development could include testing with different dialects of the Bugis language to increase the application's flexibility.
2. Application in Speech Signal Processing for Other Purposes: It is hoped that the results of this research can be implemented in various speech signal processing technologies, such as automatic speech recognition, speech synthesis systems, or more in-depth linguistic analysis for the Bugis language, in order to support the preservation and development of the language.
3. Collaboration with Other Researchers: It is recommended that this research be continued through collaboration with researchers from various fields, such as linguistics and information technology, to produce a more comprehensive and detailed application for processing speech signals of the Bugis language.

Acknowledgement

The author would like to express sincere gratitude to the following distinguished individuals and institutions:

1. My parents and family, whose prayers and unwavering support were invaluable throughout the research process.
2. The Chairman of the Yayasan Badan Wakaf, University of Muslim Indonesia.
3. The Rector of the University of Muslim Indonesia.
4. The Chairman and Staff of LP2S, University of Muslim Indonesia.
5. The Dean of the Faculty of Engineering, University of Muslim Indonesia.
6. The Head of the Electrical Engineering Program, Faculty of Engineering, all colleagues and staff members Together The research team members, students, and respondents/informants who generously dedicated their time and support to assist in this research.

It is hoped that this research will provide valuable contributions and encourage future researchers in the field.

References

[1]. Hartono , Y.S, Basalamah A. Ekstraksi Pitch Pada Suara Laki-laki Dan Perempuan Dengan Metode Algoritma Autokorelasi Dan AMDF Untuk Penentuan Tipe Jenis Jangkauan Suara Berdasarkan Tipe Suara. Makassar Sulawesi Selatan; 2016.

[2]. Umar R, Riadi I, Hanif A. Analisis Bentuk Pola Suara Menggunakan Ekstraksi Ciri Mel-Frequency Cepstral Coefficients (MFCC) Voice Pattern Form Analysis Using Feature Extraction Mel-Frequency Cepstral Coefficients (MFCC). 2018;4(2):294–304.

[3]. Wang J, Li Y, Xue Z, Li C. Research on Speech Feature Extraction System in Oral English. 2019;347(Icedem):9–11.

[4]. Mazumder MA, Salam RA. Feature extraction techniques for speech processing: A review. Int J Adv Trends Comput Sci Eng. 2019;8(1.3 S1):285–92.

[5]. Hartono, S.Y.,, Suyuti S. Ekstraksi Pitch Sinyal Wicara (Speech) Bahasa Daerah Bugis Dengan Algoritma Autokorelasi. 2024;9(April):125–34.
[6]. Gill AS. A Review on Feature Extraction Techniques for Speech Processing. 2016;5(10):18552–6.
[7]. R A, K L, Shivakumar DM. Feature Extraction of Speech Signal using LPC. Ijarcce. 2016;5(12):189–90.
[8]. Hartayu R, Anam C, Aziz DA, Studi P, Elektro T, Teknik F, et al. Simulasi Ekstraksi Fitur Suara menggunakan Mel-Frequency Cepstrum Coefficient. 2022;8:80–7.
[9]. Athira G, Anjusha VK. Analysis of Feature Extraction Methods for Speech Recognition. 2017;4(4):328–33.
[10]. GIANNAKOPOULOS T. Introduction to AUDIO ANALYSIS: A MATLAB Approac. First Edit. Amsterdam, New York: Elsevier; 2014. 33–38 hal.
[11]. Thakuria L, Pran P, Talukdar H. Analysis and Detection of Emotional states of a person by extracting Pitch and Formants with respect to Assamese Language. 2014;5(2):890–3.
[12]. Pitch Estimation using Autocorrelation Method Pitch Estimation using Autocorrelation Method and. 2016;(March).
[13]. Caesarendra, W., Ariyanto M. Panduan Belajar Mandiri MATLAB. Jakarta: PT.Elex media Kompetindo; 2016. page 37-64.