Signal acquisition system based on wireless transmission for environmental sound monitoring system

V N Sulistyawan, S E Widhira, A Fatin and N A Salim

Electrical Engineering Department, Faculty of Engineering, Universitas Negeri Semarang, Gunungpati, Semarang, 50229, Indonesia

veranovianas@mail.unnes.ac.id

Abstract. In today’s technological era, the ability to access information through digital signals, especially voice, requires sophisticated and comprehensive applications that can convert physical signals into electrical signals. Its purpose is to assist humans in displaying and analysing structured and automated data obtained from tools with their unique set of features. In this study, we used signal data for different sounds, such as rock songs, birdsong, acoustic sounds, and conversational sounds. The data is checked using software and undergoes a data acquisition process. The results of the study are expected to reinforce that environmental sound processing is predicted to assist the development of more sophisticated automated monitoring systems capable of combining voice and visual data in a complementary way.

1. Introduction

Sound is a physical phenomenon generated by the vibration of an object or the vibration of an object in the form of an analogue signal with a changing amplitude over time [1]. The experience of "hearing" is strongly linked to sound. Currently, technological advancements are progressing at a rapid pace, and humans not only interact with one another, but also interact with technologies such as computers. The vibration of an object in the form of a signal equivalent to a continually changing amplitude against a unit of time called frequency generates audio (sound) [2]. Speech recognition is a method of converting human speech into machine language. The stimulation of speech signals is classified into two types: Voiceover and Unvoiced, and Mixed, plosive, whisper, and quiet [3]. Dictation and human-computer dialogue systems are the two main applications of speech recognition technology. A variety of experimental systems for data acquisition through spoken dialogue are being probed in the domain of human-computer interaction. Automatic broadcast news transcription is currently being researched in the dictation field. [4]. Although, technically, Speech Recognition is developing a system that maps signals into word sequences. The entire system is designed to extract the spoken word from the input signal automatically. There are two types of audios or sound, analogue and digital, in general. Variations of the original sound pressure are used to create analogues. Variations in air pressure are first converted (by a transducer such as a microphone) into an electrical signal. Whereas digital output converts the physical properties of the original sound into a sequence of numbers, which can then be stored and read back for reproduction. Automatic Speech Recognition performance for people with normal speech function has reached near-perfect accuracy, allowing enabling commercial uses and integration into portable speech-based human-machine interfaces [5].
Humans who are physically challenged, such as the blind and deaf, may now readily converse with machines thanks to continued progress and advancement in the field of speech recognition. [6]

Several studies have been successfully completed in collaboration with several researchers who process voice signals using information technology computing. According to Garg et al. [7] there are various significant factors for processing human speech signal detection, including Cepstrum, Pitch, and Formant. One of the most important things in sound processing is estimating the pitch of the voice signal. Another study is Data Analysis of Speech Signals (Speech) Recording Indoors and Outdoors Open. In this study, we analysed the difference in sound signal when recorded indoors and outdoors in an open space. The recorded sound signal shows that the large amplitude in the open space will be greater than the signal peak. Voice signals are recorded in indoor and outdoor enclosed and open spaces. Based on the recording results, it can be seen that there is a difference in the sound signal in recording in a closed room and outdoors in an open space, greater than the peak of the sound signal recorded in enclosed spaces and outdoors [8].

In this audio signal processing, a digital sound processing system can be used to recognize the input sound and physically display the signal along with the frequency value generated by the related object source. This process is also known as speech recognition. Voice signal processing can be applied in various applications that can make it easier for humans. For example, for human speech recognition systems, detection of machine malfunctions, detection of heart rate abnormalities, and others. One other application is the acquisition of data using software. This aims to be able to assist humans in displaying, processing data sourced from a tool with its own characteristics automatically and structured so that when a physical quantity changes, we can display quantities that can be directly observed by the five human senses through visualization of digital signal graphs.

In this paper, we describe the process of data acquisition from audio and then the sound goes through several stages so that it can be known or provides information about the value of the resulting frequency. This paper discusses the results of the analysis of the output signal from the research of different voice signal data, namely the sound of rock songs, the sound of birds chirping, acoustic sounds and conversational sounds analysed using software. The stages of this research include designing the interface system on the software and collecting data samples in the form of audio data, namely the sound of rock songs, the sound of birds chirping, acoustic sounds and conversational sounds that are recorded and later acquired into sound signals. In this case, simulation software can be utilized to provide a user-friendly interface and environment for computing, simulation, evaluation, and optimization. It can also assist users (teachers, students, and engineers) in manipulating and presenting educational and professional work [9]. The success of this study will likely aid the creation of a more advanced automated environmental noise monitoring system capable of merging voice and visual data in a complimentary manner.

2. Signal acquisition

To communicate, one needs to generate sound signals in the form of barometric waves. The vibrations are then transmitted from the speaker's mouth through the listener's ear media. The sound comes from pressure waves in the neck, nose, and mouth [10]. Each sound source certainly has unique sound characteristics. A distinction can be made based on the value of the frequency and intensity of the sound emitted by the sound source. To determine the sound characteristics of the sound source, a recording device such as a microphone or underwater cheongmuugi is needed. The use of this recording device is called passive sounding. Voice signal processing requires a sophisticated and comprehensive application to be able to convert voice signals to provide signals. The signal processing step can be divided into three parts:

- a phase of pre-processing (acquisition, filtering, reaccentuation, segmentation),
- a phase of analysis (extraction and identification acoustic parameters: pitch, formants, MFCC),
- a phase of post-processing in order to improve the obtained results and to optimize the identified parameters.
2.1 Data acquisition
Data acquisition system is a system that functions to retrieve, collect and prepare data to process it to produce the desired data. The process by which physical phenomena from the real world are converted into electrical signals that are measured and converted into digital form for processing, analysis, and storage by a computer is referred to as data acquisition [11]. The types and methods chosen generally aim to simplify each step carried out in the entire process. The data acquisition system consists of a number of elements or components that are interconnected with each other formed in such a way that the system can function to retrieve, collect and store data quickly, real-time and accurately so that the data is then ready for further processing are measurement objects, Transducers, Amplifiers, Multiplexers, Data Acquisition Cards, computers and data acquisition software. Simulations of sound signal processing have been carried out where the recordings made are processed in such a way that they can provide information about the resulting frequency value, the sound signal graph and its spectrum [12].

Data collection systems are evolving rapidly with advancements in digital and information technologies. The data acquisition process requires a process of converting the physical amount of the data source into a digital signal format and processing it on a computer. Computerized processing and process control allow the software to perform data collection. This software offers the hope that the data collection process can be easily modified to meet your needs. The software provides hope that the data acquisition process can be varied easily according to needs.

2.2 Voice
Sound is a form of energy, like electricity and light. A sound is made when air molecules vibrate and move in a pattern called a wave, or sound wave. Sound has a certain frequency range and sound intensity that can or cannot be heard by humans. Alexander Graham Bell, an inventor of the telephone, discovered the unit for measuring the intensity of sound is the decibel (dB). While the unit of sound frequency is Hertz which was taken from Heinrich Rudolf Hertz to appreciate his contribution in the electromagnetic field. The frequency of sound that can be heard by humans is between 20 to 20,000 Hz. This range varies individually and generally depends on age. Later the saved sound will be converted to audio formats such as mp3, real audio, midi, wav, and others [13]. This study saves audio files in wav form because the software can only save sound formats in wav form. The parameters used in this study are amplitude and power. The magnitude and period parameters are less accurate in distinguishing between different sounds. If the amplitude of a sound is high, then the power is also stated to tend to be higher [14].

2.3 Waveform audio (.wav)
WAV is a Microsoft and IBM standard audio format for personal computers (PCs), typically using Pulse Code Modulation (PCM) encoding. Since WAV is uncompressed data, all audio samples are stored on the hard disk. For example, a piece of software that lets you create WAVs from analogue sound is Windows Sound Recorder. These audio files are relatively large in size and are rarely used on the Internet because WAV files have an upper limit of 2 GB [15].

2.4 Soundcard
A soundcard is a low-cost device found in all computers. It communicates via three jacks: line in, lineout, and microphone-in. Each of these jacks has multiple channels, each with its own ADC/DAC and input/output buffers. Since soundcards are used to process sound signals, their sampling rates are typically set to 44.1 kHz, though this can be changed such as through program up to a maximum of 96 kHz and a minimum of 5 kHz for commonly used PC motherboards in recent years. Because of its low cost and easy availability, using a data acquisition system developed with a sound card is advantageous. Furthermore, the operating voltage levels are low, allowing any hardware to be checked safely [16].
3. Design
In this study, sound is emitted through an electronic device. The sound is captured by the computer for processing. In this case there is a change from a physical signal to an electrical signal. This electrical signal is processed into a visible signal in the form of a graph so that it is easy to analyse. The process of this research is shown in Figure 1.

![Figure 1. Voice data acquisition block system](image)

3.1 Flowchart
This model is used to model the software created. Flowchart method is a method of depiction with a graph that aims to show the steps and sequence of procedures of the system created. Flowcharts make the process of the program flow described in a simple way so that it is easier to understand. The methodology used to achieve the objectives of this research is explained through the flow chart shown in Figure 2.

![Figure 2. Signal acquisition system flowchart](image)

3.2 System design
The GUI created for this study has three basic functions: (1) signal modelling and creation; (2) signal spectrum analysis; and (3) digital filter construction [17].
The menus designed in the software include record, sound, save, play, and reset as shown in figure 3. The components used in the software include:
a. Panels are usually used as a background or a place to design the interface and can also be grouped.
b. Push Button is usually used to run a function to be executed. When the GUI is run, the Push Button is clicked to perform certain functions
c. Edit Text is usually used to input data that is entered into the program.
d. Axes function to display graphics or images (image). Axes are not included in the GUI Control, but can be programmed so that users can interact with axes and graphical objects displayed through axes.
3.3 Figure and description

![Figure 3. Appearance and component on Graphical User Interface](image)

*Description:*

a. The Record Signal button is made of a push button component. The signal record button is used to record the signal.

b. When the interface will record a signal, this section will say start of recording automatically and when it will end the recording, this section will write end of recording automatically.

c. The sound button is made up of push button components to unmute. The mute button functions to sound the recorded signal.

d. The Save button is made of a push button component. The save button is used to save the recorded signal.

e. Button Load signal from the push button component. The signal load button is used to process the recorded signal.

f. Made from the edit text component. When the interface will record a signal, this section will say start of recording automatically and when it will end the recording, this section will write end of recording automatically.

g. The Unmute button is made up of push button components. The mute button functions to sound the recorded signal.

h. The reset button is made of push button components. The reset button is used to reset the signal to be recorded.

4. Results and discussion

4.1 System build

The next step after the program runs is to type the word "guide" at the command prompt menu. Then press Enter to display the GUI window, the quick start guide, consisting of the Create New GUI and Open Existing GUI tab menu. Choose to open the existing GUI that contains previously opened software, or choose the Browse button to open the application stored in the folder on your computer. To run the GUI application that has been created, press the Run button. After that, the Matlab application can be used.
The steps to perform audio signal acquisition are as follows:

a. To be able to perform voice signal acquisition, a GUI programming application is used. The first step is to create a GUI interface and code the commands for each pushbutton. with a display design like the one below:

![GUI display](image)

**Figure 4.** GUI display

b. Signal recording is done by pressing the record signal button

![Record voice using record signal button](image)

**Figure 5.** Record voice using record signal button

c. The following figure shows the situation when sound is being recorded.
d. Stages of sounding the recorded voice

Figure 6. Display while sound is being recorded

Figure 7. Unmute sound

e. Use the save button to save the generated sound signal
Figure 8. Save the voice signal

f. Signal load display

Figure 9. Signal load display

g. Press the reset button to reset and make a new recording
4.2 Signal analysis and results
In this case, the input data is speech signal data (speech), which is obtained by recording sound with condenser microphones, soundcards, laptop computers, and integrated software. The voice data in the recorded speech is the utterance of words/sentences in standard Indonesian as if speaking normally. The stages of this research are to build a system with a GUI and then record the sound samples that have been prepared. Then, as a final step, the results of the research as a data base voice signal (speech) were released in the form of an extension (format) called Waveform Audio File Format (WAV). The simulation is carried out by using real-time speech signal data stored in a WAV formatted database. And after that, the testing process is carried out to determine whether the trial process for implementation on speech signal data (speech) can run smoothly. The output form is a sinusoidal signal as a function of time (time domain).

The test results of four samples of sound signals are displayed as shown in the following figure. In Figure 11 below is a form of signal acquisition resulting from acoustic sound recording.
Figure 12 is a waveform of signal acquisition as a result of recording the sound of birds chirping.

![Waveform of birds chirping sound](image12)

**Figure 12.** Display signal waveform of birds chirping sound

Figure 13 is a form of signal acquisition resulting from a recorded voice conversation.

![Waveform of voice conversation](image13)

**Figure 13.** Display signal waveform of voice conversation

Figure 14 is a form of signal acquisition resulting from a rock song sound recording.
Figure 14. Display signal waveform of rock song sound

By looking at the results of the sound sample graph above, you can see the difference between the signal results of the four sound samples. It can be seen that rock songs have greater amplitude than other sound samples. The recording takes place indoors to ensure a good signal. The signal recorded in a closed room sounds very clear and clearer, according to Speech Signal Data Analysis (Speech) Recording Indoors and Outdoors Open, because there are no other sounds that enter when recording in the form of additional noise signals or interference that causes poor signal quality.

The result of the bird's song waveform appears to be more stable, and the amplitude of the dialogue voice waveform is smaller. A variable's amplitude is a measurement of how far and in which direction it differs from zero. As a result, signal amplitudes may be positive or negative. The sample value amplitudes of three different waveforms were given in the time-domain sequences. It's worth noting that some of the discrete amplitude values were positive while others were negative [18].

5. Conclusion
Audio signal acquisition is the process of converting physical phenomena in the real world into measurable electrical signals and then into a digital format for computer analysis, processing, and storage. A signal acquisition system has been successfully created that can visualize the process of converting audio signals to frequency signals. A Graphical User Interface has been successfully created that can visualize the process of changing voice signals into frequency signals. The results of the waveform of rock songs have a greater amplitude. While the sound of birdsong has a smaller amplitude so that the power of the sound of rock songs is higher than the sound of conversation. Whereas this design is aimed at beginners, it includes of the fundamental analytical operations, and its ability to read external signals from other devices or programs. Furthermore, it will also be very useful to be developed in the future as a means of environmental noise monitoring.

References
[1] Binanto I 2010 Multimedia Digital-Dasar Teori dan Pengembangannya (Yogyakarta: Andi Offset)
[2] Mutiara A D, Sutardi S and Ramadhan R 2016 semanTIK 2 29-37
[3] Yudihartono S and Suyuti S 2018 Logitech Teknik Elektro 1 36-9
[4] Furui S 1999 Proceedings of the 37th annual meeting of the Association for Computational Linguistics p 11-20
[5] Meltzner G S, Heaton J T, Deng Y, De Luca G, Roy S H and Kline J C 2017 *IEEE/ACM transactions on audio, speech, and language processing* **25** 2386-98 

[6] Kabari L G and Ogie K 2019 *International Journal Advance Research Computer Science* **10** 5-8 

[7] Garg D, Kaur S, Arora D 2012 *International Journal of Advances in Electrical and Electronics Engineering* **1** 278-283 

[8] Bouaffif L and Ouni K 2012 *6th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)* 788-91 

[9] Altim M Z, Hartono S Y and Ashad B A 2019 *Logitech Teknik Elektro* **2** 45-9 

[10] Khairunnisa K and Indrasary Y 2016 *Jurnal Simantec* **5** 75-84 

[11] Estrada R 2008 *J. Teknokrida*. **6** 

[12] Adler J, Azhar M and Supatmi S *Telekontran* **1** 16-23 

[13] Irtawaty A S, Ulfah M and Rukhyah S F 2019 *Jurnal Teknologi Terpadu* **7** 70-5 

[14] Rahandi A, Rachmawati D and Sembiring S 2012 *J Alkhawarizmi* **1** 

[15] Mitra S and Dey D 2012 *National Conference on Computing and Communication Systems* p 1-5 

[16] Huang Y S and Cai Y Q 2019 *J. Phys.: Conf. Ser.* **1314** 012018 

[17] Lyons R G 2004 *Understanding Digital Signal Processing* (New Jersey: Prentice Hall)