

# Classification Of Women's Voices Using Fast Fourier Transform (FFT) Method

Made Sri Ayu Apsari<sup>a1</sup>, I Made Widiartha<sup>a2</sup>

<sup>a</sup>Informatics Department, Udayana University  
Bali, Indonesia

<sup>1</sup>ayuapsaarii25@gmail.com

<sup>2</sup>madewidiartha@unud.ac.id

## Abstract

*Everyone has a different kind of voice. Based on gender, voice type is divided into six parts, namely soprano, mezzo soprano, and alto for women; and tenor, baritone, and bass in men. Each type of sound has a different range and with different frequencies. This study classified the type of voice in women using the Fast Fourier Transform (FFT) method by recording the voices of each user which would then be processed using the FFT method to obtain the appropriate sound range. This research got results with an accuracy of up to 80%. The results obtained from this study are quite appropriate and it is proven that the FFT method can be used in digital signal processing.*

**Keywords:** Fast Fourier Transform, Type of Voice, Women, Frequency, Vocal Range

## 1. Introduction

Sound processing is an activity of producing a toned sound, either with lyrics or not. The human voice is produced by the vocal cords. Each person has a different type of voice, which is divided into six types based on gender. The voice in men is divided into three, there are tenor, baritone, and bass; while the voice in women is divided into three, there are soprano, mezzo soprano, and alto [6]. However, not everyone knows what kind of voice they have. The type of human voice can be determined based on the vocal range it has. Vocal range can be determined by synchronizing the vocals that come out of the human voice with the cord notes on certain musical instruments [5].

In previous studies, a system was designed to classify both male and female voices using the FFT method and the MATLAB software. The type of sound is obtained based on the vocal range. The design and testing of this system get appropriate results where the FFT method can classify sounds based on frequency ranges and vocal ranges[1].

Classification of children's voices was also carried out using the FFT method. In this case, a study was conducted that could identify children's voices and classify these voices into types of human voices based on frequency using the FFT method where signals from the time domain are converted into signals in the frequency domain. In this study, it is equipped with a GUI that makes it easy for users[3].

In the research "character recognition of aceh male voice using the FFT method"[4] the author created a system that can recognize a person's voice character. This study carried out the recognition of the voice character of Acehnese men spoken by children, adolescents, and adults. The results obtained from testing in this study show an accuracy rate of 81.2% using the FFT (Fast Fourier Transform) method.

Rut Arini also conducted other research, which raised the topic of identifying human voice signals using the FFT method. In this study, a system was designed to identify the human voice by matching the voice entered with the existing database. Where the recognition rate gets better if the frame rate is higher, and in this research the best recognition rate is 96% at 256 frame

blocking. However, the voice recognition rate drops when the frame blocking value is below 128, namely 16, 32 and 64[7].

Based on the case studies previously described, it is evident that the FFT (Fast Fourier Transform) method can be used to perform digital signal processing, and get accurate results.

## 2. Research Methods

### 2.1 Type of Human Voice

According to Alen (2011), there are six types of voices that are owned by humans, which are differentiated by gender. Types of voices that are owned by women are soprano, mezzo soprano, and alto. Soprano voice is the highest voice for women. This type of sound is in the range C4 (middle C) to C6 (high C). Not a few people who have this type of voice can sing the notes G or A over C6. The mezzo-soprano voice is a type of voice in women that is between the soprano voice and the alto (contralto) voice. This type of sound is in the range A3 (tone A below middle C) to A5 (note A second above middle C). Alto (contralto) is the lowest type of voice in women. This voice is in the range of F3 (F note below middle C) to E5 (second E note above middle C).

|              |       |                   |
|--------------|-------|-------------------|
| Sopran       | C4-C6 | 261.626 - 1046.50 |
| Mezzo-Sopran | A3-A5 | 220.000 - 880.000 |
| Alto         | F3-F5 | 174.614 - 698.456 |

**Figure 1.** Woman type of voice

### 2.2 Data

In this study using data in the form of voices from 10 women who of course have different types of voices, which are recorded and stored in a file with the .wav format. The voice input file will be processed using the FFT method so that the type of sound can be determined.

### 2.3 Method

#### 2.3.1 Fourier Transform

Fourier transform is a transformation model that converts the time domain into the frequency domain. analyzes in the frequency domain are widely used such as filtering. By using the Fourier transform, the signal can be seen as an object in the frequency domain. Fourier transform can be defined by equation (1) as follows:

$$\begin{aligned}
 X(f) &= \int_{-\infty}^{\infty} x(t) e^{-i2\pi ft} dt \\
 &= \int x(t) \cos(2\pi ft) dt - i \int x(t) \sin(2\pi ft) dt
 \end{aligned} \tag{1}$$

Can also be written as :

$$\begin{aligned}
 \int_{-\infty}^{\infty} x(t) \cos(2\pi ft) dt &\rightarrow \sum_n x(n\Delta t) \cos(2\pi f n \Delta t) \Delta t \\
 &= \sum_n x(n\Delta t) \cos(2\pi n m \Delta t \Delta f) \Delta t \\
 &= \sum x(n\Delta t) \cos(2\pi n m N n) \Delta t
 \end{aligned}$$

The time domain of a signal period is expressed as  $T = N\Delta t$ , while in the frequency domain  $\Delta f = f_s / N$  where  $\Delta f$  represents the interval between the frequency and  $f_s = 1 / \Delta t = N / \Delta t$ . Thus, in equation (4)  $\Delta t \Delta f = 1 / N$ , which is a link between the time domain and the frequency domain. In accordance with the Nyquist equation, where the minimum sampling frequency ( $F_s$ ) is 2 times the frequency of the analog signal to be converted ( $F_{in\ max}$ ) to avoid frequency aliasing effects. Aliasing is the appearance of the same frequency from the

transformation results which cannot distinguish between the original frequency and the image frequency.

### 2.3.1.1 Discrete Fourier Transform (DFT)

Discrete Fourier Transform is a defined series in the discrete frequency domain that represents the Fourier transform of a finite series. DFT is a way of converting something or a signal from the time domain to the frequency domain. Discrete signals can be used by using another method, namely discrete Fourier transform. Discrete Fourier transform is used to convert discrete data from time domain to frequency domain. The discrete signal is generated from the analog signal defined in equation (2).

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi f t} dt \quad (2)$$

### 2.3.1.2 Fast Fourier Transform (FFT)

Fast Fourier Transform (FFT) is a Fourier transform algorithm developed from the Discrete Fourier Transform (DFT) algorithm. The Fast Fourier Transform algorithm is very efficient in calculating the DFT coefficient and can reduce the enormous computational complexity. FFT is a method that converts signals from time domain to frequency domain. By using this FFT method, the computation rate of the Fourier transform calculation can be increased. The formula of the FFT method can be defined as follows in equation (3).

$$X[k] = \sum_{n=-11}^{N-1} x(n)W_N^{kn} \quad (3)$$

## 2.4 Flowchart

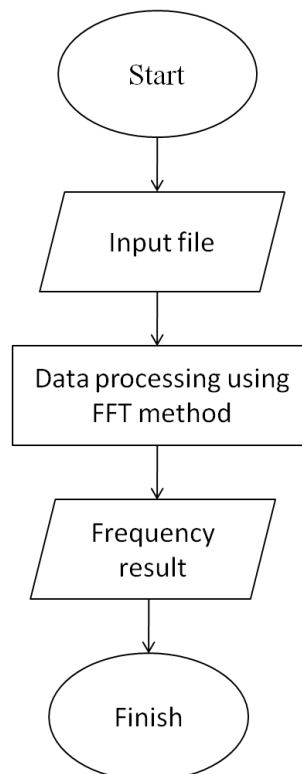


Figure 2. Flowchart

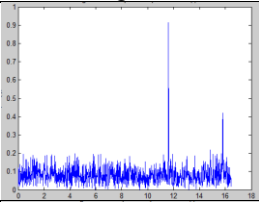
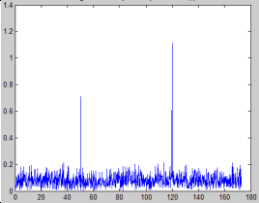
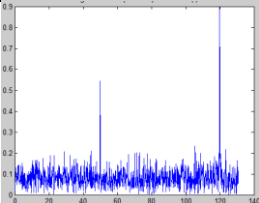
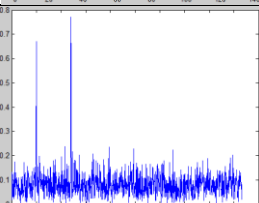
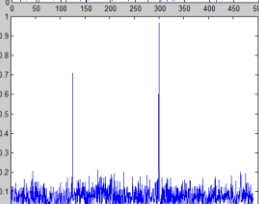
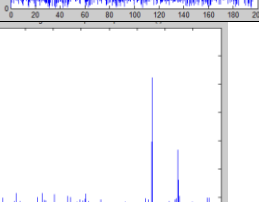
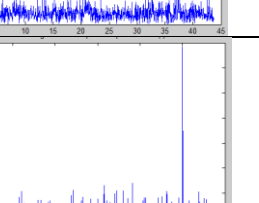
The picture above is a picture of the flowchart in this study, this research was conducted to identify the type of voice that a woman has, whether she has a soprano, mezzo soprano, or alto voice. The file that is entered at the beginning is in the form of a voice recording file from 10 women which is recorded within 5 seconds, then the recording file will be processed using the FFT method to search for signals in the frequency domain. There are three types of sound that will be used as a reference in identification, namely soprano, mezzo soprano, and alto. Soprano is the highest type of voice in women, the frequency of this type of sound is

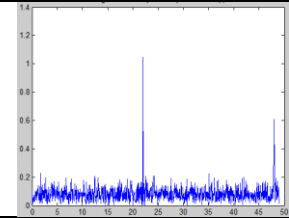
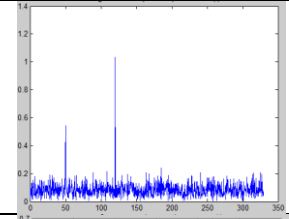
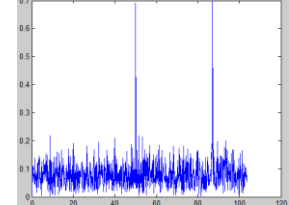
261,626 - 1046.50. Mezzo soprano is a type of voice in women that has a frequency between 220,000 - 880,000. And the last is the type of alto voice, which is the voice of women with a frequency between 174,614 - 698,456.

### 3. Result and Discussion

Voice recording is done once for each person, after recording the sound will be processed to get a frequency that matches the type of sound it has. Sounds recorded for 5 seconds each.

**Table 1.** Research result

| No. | Name   | Type of voice | Classificataion result | Figure  |
|-----|--------|---------------|------------------------|---|
| 1.  | User 1 | Sopran        | Sopran                 |    |
| 2.  | User 2 | Mezzo sopran  | Mezzo sopran           |    |
| 3.  | User 3 | Mezzo sopran  | Mezzo sopran           |   |
| 4.  | User 4 | Alto          | Alto                   |  |
| 5.  | User 5 | Mezzo sopran  | Mezzo sopran           |  |
| 6.  | User 6 | Sopran        | Sopran                 |  |
| 7.  | User 7 | Sopran        | Sopran, Alto           |  |

|     |         |        |                         |  |
|-----|---------|--------|-------------------------|--|
| 8.  | User 8  | Sopran | Mezzo sopran,<br>sopran |  |
| 9.  | User 9  | Alto   | Alto                    |  |
| 10. | User 10 | Sopran | Sopran                  |  |

The table above shows that each woman who was asked to record her voice had a different type of voice. User 1 has a soprano voice, which means he has a voice range between C4 - C6 and with frequencies ranging from 261,626 - 1046.50. Users 2 and 3 both have the mezzo soprano voice that is in the A3 - A5 voice range with a frequency between 220,000 - 880,000. User 4 has an alto voice with a voice range at F3 - F5 with a frequency between 174,614 - 698,456. User 5 has the same mezzo soprano voice type as user 2 and user 3. User 6 has a soprano voice that is in the range C4 - C6 and is on a frequency between 261,626 - 1046.50. User 7 is identified as having two types of voice, namely soprano and alto, which are on the frequency 174,614 - 1046.50. User 8 is also identified as having two types of voices, namely soprano and mezzo soprano with a frequency between 220,000 - 1046.50. User 9 has an alto voice with a vocal range of F3 - F5 and a frequency between 174,614 - 698,456. And user 10 has a soprano voice with a vocal range of C4 - C6. This voice recording shows quite appropriate results because it shows compatibility with existing data.

#### 4. Conclusion

This research on identifying the type of female voice using the FFT method shows quite appropriate results. However, there are several drawbacks that cause the results of this study to be less than optimal. This research got results with an accuracy of up to 80%. In this study, we have not used features that directly provide the results of identifying the type of sound. In the future, it is hoped that this research can be refined with the addition of other features in order to get maximum and accurate results.

#### References

- [1] E. Sidabutar, and E.P.Laksana, "Pengklasifikasian Suara Menggunakan Metode FFT pada Software Matlab untuk Mengetahui Tipe Suara Manusia", *Maestro*, vol.1, no.2, p.357-364, 2018.
- [2] H.M. Arkaan, I. Fauzi, L.W.A.Rosyid, and A.Junaidi, "Klasifikasi Ciri Suara Manusia Berbasis Matlab Menggunakan Metode Fast Fourier Transform", *J.of INIST*, vol.2, no.1, p.001-006, 2019.
- [3] R.N.Annisa, Suprayogi, and H.Bethaningtyas, "KLASIFIKASI SUARA ANAK-ANAK DENGAN MENGGUNAKAN METODE FAST FOURIER TRANSFORM", *e-Proceesing of Engineering*, vol.6, no.1, p.1141-1148, 2019.

- [4] Mursyidah, Jamilah, and Zayya, "Pengenalan Karakter Suara Laki-Laki Aceh Menggunakan Metode FFT (Fast Fourier Transform)", *Jurnal Infomedia*, vol.2, no.1, p.21-24, 2017.
- [5] Simanungkalit, N, "Musik," in *Teknik Vokal Paduan Suara*, Jakarta : Gramedia Pustaka Utama, 2013, ch.1, pp.1-6.
- [6] Phillips, Pamela S., "Determining your voice type," in *Singing For Dummies*, 2<sup>nd</sup> Ed. Canada : John Wiley & Sons, 2010, ch.2, pp.17-25.
- [7] R.A.L.Sibarani, "IDENTIFIKASI SINYAL SUARA MENGGUNAKAN METODE FAST FOURIER TRANSFORM (FFT) BERBASIS MATLAB", Universitas Sumatera Utara, 2018.