The recognition of female voice based on voice registers in singing techniques in real-time using hankel transform method and macdonald function

R Meiyanti¹, A Subandi¹, N Fuqara², M A Budiman¹ and A P U Siahaan³,⁴

¹ Universitas Sumatera Utara, Medan, Indonesia
² Universitas Malikussaleh, Aceh Utara, Indonesia
³ Universitas Pembangunan Panca Budi, Medan, Indonesia
⁴ School of Communication and Computer Engineering, Universiti Malaysia Perlis

rinimeiyanti522@gmail.com, aminsubandi@yahoo.com, aransq92@gmail.com, mandrib@usu.ac.id and andiesiahaan@gmail.com

Abstract. A singer doesn’t just recite the lyrics of a song, but also with the use of particular sound techniques to make it more beautiful. In the singing technique, more female have a diverse sound registers than male. There are so many registers of the human voice, but the voice registers used while singing, among others, Chest Voice, Head Voice, Falsetto, and Vocal fry. Research of speech recognition based on the female's voice registers in singing technique is built using Borland Delphi 7.0. Speech recognition process performed by the input recorded voice samples and also in real time. Voice input will result in weight energy values based on calculations using Hankel Transformation method and Macdonald Functions. The results showed that the accuracy of the system depends on the accuracy of sound engineering that trained and tested, and obtained an average percentage of the successful introduction of the voice registers record reached 48.75 percent, while the average percentage of the successful introduction of the voice registers in real time to reach 57 percent.

1. Introduction

Voice recognition technique consists of two main stages, training and recognition. In the training stage, an identified voice is inserted into a database. In the recognition stage, an undefined voice signal will be identified. Voice recognition method is used for both verifications as well as identification. [1] The voice register is the division of the human voice area by sound source, the sensation of resonance space, shape, color, sound timbre, and the resulting low tone. Along with audio signal separation, audio classification is important due to following reasons:

- Different audio types should be processed differently;
- The searching space after classification is reduced into a particular subclass during the retrieval process. [2]

Maybe we are often confused and wondering about terms in the distribution of voice registers like falsetto, head voice, chest voice, vocal fry and some other terms. But simply we can tell the difference. In this system, the author uses the Hankel Transformation which aims to recognize the input sound as a sample and can normalize sound waves. In addition, the author also uses Macdonald's function in
this voice recognition system as a combination to produce better and more accurate system performance in voice recognition process.

2. Theories

2.1. Voice Classification
There are various approaches that can be used such as Artificial Neural Network (JST), DCT, Gabor Filter, Statistics, Wavelet, and others, to classify the sound. Various information that can be obtained from speech sounds coming from the mouth of humans, among others: Speech Identity, Expression, Dialect/Slog/Tribe, Gender Type, Distance of sound sources, Voice rate, Age, Word Recognition, Loud level, Saturation rate, Health condition and language quality.

2.2. WAV
WAV is Microsoft's standard audio format and IBM for personal computer (PC), usually using PCM coding (Pulse Code Modulation). WAV is uncompressed data so that all audio samples are stored all on the hard disk. Software that can create WAV from analog sound, for example, is Windows Sound Recorder.

2.3. Hankel Transform
The Hankel Transform is an integral transform developed by Hermann Hankel. Further computational investigation reveals a number of instances in which one member of a pair of sequences with the same Hankel transform is the Binomial or Invert transform of the other. [3] General Equation of Hankel Transform:

\[ F_v(k) = \sum_{r=0}^{N-1} f(r) \times r \times K_v(kr) \]  

For \( r = 0, 1, 2, 3, ..., N-1 \)

2.4. Macdonald Function
Certain problems of mathematical physics arising in spheroidal or cylindrical domains, e.g. Laplace's equation, have solutions that involve Macdonald's functions. [4] The Macdonald function is a modification of the Bessel function and the Cylinder function, which is made by H. M. Macdonald. The general equation of the Macdonald functions are:

\[ K_v(kr) = \frac{\pi j_v(kr) - i_v(kr)}{2 \sin \nu \pi} \]  

Where \( \nu \) is any non-integral real number

\[ j_v(kr) = \sum_{m=0}^{\infty} \frac{(kr)^{\nu + 2m}}{m! \Gamma(\nu + m + 1)} \]  

\[ i_v(kr) = \sum_{m=0}^{\infty} \frac{(kr)^{\nu + 2m}}{m! \Gamma(-\nu + m + 1)} \]
2.5. Gamma Function
In mathematics, the gamma function (presented by the Greek capital letter $\Gamma$) is an extension of the factorial function, with its argument shifted down by 1, to real and complex numbers. That is, if $n$ is a positive integer, then:

$$\Gamma(n) = (n - 1)!$$

Gamma function is defined for all complex numbers, except for negative integers and zeros. For complex numbers whose real part is positive, the gamma function is defined by an incorrectly converging integral.

In the previous research, the calculations using Hankel transform and Macdonald function can recognize the types of sounds such as soprano, alto, tenor and bass with different accuracy for both records sound and the type of voice in real time, where the average for recorded sounds reaches accuracy up to 66.87 %, while for real-time voice has accuracy up to 72%. [5]

3. The Scheme of System
The scheme of this study is divided into several stages, which are described in the following:

3.1. Overall Research Scheme for Macdonald Function and Hankel Transformation
The scheme of the overall method developed in this study is illustrated in Figure 1:

![Figure 1. Overall system scheme.](image)

3.2. Voice coaching scheme using Macdonald function and Hankel’s transform
Flowchart for voice coaching processes are developed based on Figure 2:
3.3. Voice Testing scheme using Macdonald function and Hankel’s transform

Flowchart for voice testing processes are developed based on Figure 3:

**Figure 2.** Voice coaching scheme using Macdonald function and Hankel’s transform.

**Figure 3.** Voice testing scheme using Macdonald function and Hankel’s transform.
4. User Interface

4.1. View of The Main Form
Before the operation, the program will display a page like this:

![Main Form Image]

Figure 4. View of the main form.

4.2. The Test Result
This form is intended for testing of Head Voice register in real-time, here is a view of test results:

![Test Result Image]

Figure 5. Sample testing process of head voice registers in real-time.

The picture shows that when the user inputs one of the voice register samples, the program will make an introduction by calculating the energy value in the sound sample and then get the end result of the introduction process, namely the voice register type.
5. Conclusions
The conclusions of this study are as follows:

- This application has a weakness that is only able to recognize the sound based on the energy value ratio between the training process and the testing process. The program can not directly detect the origin of the sound source being pronounced because the real voice register is not only seen from the high tone generated but also seen from the sound source, the sensation of resonance space and the color of the sound;

- Based on system performance, voice recognition process in real-time is more effective than voice recognition process by a record;

- The method applied in this application achieves the average percentage of the successful voice record recognition record reached 48.75%, while the average percentage of the successful introduction of voice registers in real time reached 57%.

References
[1] Ayad R R 2013 Security of Iris Recognition and Voice Recognition Techniques Ibn AL-Haitham Jour. for Pure & Appl. Sci. 26 326-57.
[2] Kumar, Choudhary, Tembhare, Pote 2014 Audio Signal Separation and Classification: A Review Paper International Journal of Innovative Research in Computer and Communication Engineering 2 Issue 11.
[3] John W L 2001 The Hankel Transform and Some of its Properties Journal of Integer Sequences 4 0115.
[4] Passian A, Simpson H, Kouchebian S, and Yakubovich S B 2009 On the orthogonality of the MacDonald’s functions. Journal of Mathematical Analysis and Applications, 360(2), 380–390.
[5] Fuadi W, Ula M., and Sadli M 2015 The Introduction Types of Vocal Sound in Choir in Realtime Using Hankel Transformation and Macdonald Function, 6(January), 1–14.
[6] Eric S W 2015 Applications of Incomplete Gamma Functions to the Incomplete Normal Distribution Open Journal of Applied & Theoretical Mathematics (OJATM) 1 No. 1 pp. 01-06.