JITE, 2 (2) Januari 2019 ISSN 2549-6255 (Print) ISSN 2549-6247 (Online)

JITE
(Journal of Informatics and Telecommunication Engineering)

Available online http://ojs.uma.ac.id/index.php/jite

Analisis Fast Fourier Tansform untuk Pengenalan Voice Register Wanita dalam Teknik Bernyanyi

Fast Fourier Transform Analysis for Voice Recognition Registers Women in Techniques Singing

Muhathir[1)]*, Rizki Muliono [1)], Susilawati [1)]

1)Teknik Informatika, Fakultas Teknik, Universitas Medan Area, Indonesia

*Coresponding Email: muhathirbangdes@gmail.com

Abstrak

Automatic speech recognition merupakan kemampuan untuk menerima dan mengidentifikasi kata-kata yang diucapkan dengan mengubah sinyal analog ke digital, dan mengekstraksi karakteristik vokal unik seperti pitch, frekuensi, nada dan irama untuk membentuk model speaker atau sampel suara. Sampel suara yang digunakan yaitu voice register, voice register adalah pembagian wilayah suara manusia berdasarkan sumber suara, sensasi ruang resonansi, bentuk, warna, timbre suara, dan tinggi rendahnya nada yang dihasilkan. Fast Fourier Transform digunakan sebagai transformasi untuk mengolah sample suara yang akan diklasifikasi. FFT dalam mentransormasikan sinyal voice register hanya mampu mengklasifikasikan dengan rata-rata true positive rate 65.4%.

Kata Kunci: FFT, Voice Register, Automatic Speech Register.

Abstract

Automatic speech recognition is the ability to receive and identify spoken words by converting analog to digital signals, and extract unique vocal characteristics such as pitch, frequency, tone and rhythm to form speaker models or sound samples. Sound samples used are voice registers, voice registers are the division of the human voice region based on the source of sound, the sensation of space resonance, shape, color, sound timbre, and the high and low tone produced. Fast Fourier Transform is used as a transformation to process sound samples to be classified. FFT in transforming voice register signals is only able to classify with an average true positive rate of 65.4%.

Keywords: FFT, Voice Register, Automatic Speech Register.

How to Cite: Muhathir, M. Rizki, Susilawati. (2019). Analisis Fast Fourier Tansform untuk Pengenalan Voice Register Wanita dalam Teknik Bernyanyi. *JITE (Journal of Informatics and Telecommunication Engineering)*. 2 (2):92-98

PENDAHULUAN

Suara merupakan sarana utama komunikasi antara manusia yang sangat efektif (Fadlisyah & Muhathir, 2015). Selain efektif, manusia juga lebih familiar menggunakan suara dalam berkomunikasi. Pengolahan suara merupakan konsep paling utama untuk semua jenis sistem yang membutuhkan interaksi manusia dalam kegiatan sehari-hari (Goyal & Batra, 2017.). Pengolahan suara merupakan kemampuan untuk menerima dan mengidentifikasi kata-kata yang diucapkan (Swamy & K, 2013). Automatic speech recognition mengubah sinyal ucapan menjadi urutan kata dengan bantuan algoritma yang diimplemetasikan kedalam program komputer (Chauhan, Samal, & Ghoshal, 2015). Aspek perilaku suara manusia digunakan untuk identifikasi dengan mengubah frasa yang diucapkan dari format analog ke digital, dan mengekstraksi karakteristik vokal unik seperti pitch, frekuensi, nada dan irama untuk membentuk model speaker atau sampel suara (Parwinder & Rani, 2014).

Menurut (Fadlisyah & Muhathir, 2015) sistem pengolahan suara dapat dikategorikan menjadi empat jenis: *Isolated Words, Connected Words, Continuous Speech, Spontaneous Speech*. *Isolated Words* : merupakan pemrosesan suara dengan menerina satu kata pada satu waktu dimana sistem mendengar pembicara untuk menunggu antara ucapan-ucapan, *Connected Words :* system kata terhubung hampir sama dengan *Isolated Words*, tetapi memungkinkan terpisah ucapan-ucapan untuk menjadi giliran yang sama dengan jeda minimal antara ucapan, *Continuous Speech:* merupakan pengolahan suara berkelanjutan memungkinkan pengguna untuk berbicara hampir secara alami, sementara komputer menentukan konten (pada dasarnya ini dikte komputer). *Spontaneous Speech:* merupakan suara spontan yg dapat dikategorikan sebagai suara alami yang terdengar dan tidak terlatih. Sebuah sistem pengolahan suara yang spontan harus mampu menangani berbagai fitur suara yang alami seperti kata-kata “ums” dan “ahs” dan bahkan sedikit gagap.

Pengolahan suara pada biasanya hanya mengklasifikasikan atau mengidentifikasikan pemilik suara, dalam penelitian ini akan dianalisis register suara wanita dalam menyanyikan lagu menggunakan Fast Fourier Transform (FFT), dalam teknik bernyanyi biasanya wanita akan menggunakan register suara *Chest Voice*, *Head Voice*, *Falsetto*, dan

Vocalfry (Meiyanti, Subandi, Fuqara, Budiman , & A, 2017).

Voice Clasification

Berbagai metode yang digunakan dalam mengklasifikasi suara, seperti Jaringan Syaraf Tiruan (JST), DCT, Gabor Filter, Statistik, Wavelet, dan lain-lain. Dari ucapan manusia dapat diperoleh berbagai informasi antara lain: 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 (Meiyanti, Subandi, Fuqara, Budiman , & A, 2017).

Voice Register

Register suara adalah pembagian wilayah suara manusia berdasarkan sumber suara, sensasi ruang resonansi, bentuk, warna, timbre suara, dan tinggi rendahnya nada yang dihasilkan (Fric, Šram, & Jan, 2016) (M, F, & G). Beberapa register suara :

a. *Vocalfry* adalah suara manusia yang dihasilkan melalui pita suara atau getaran glottis dimana kedua belah pita suara menyatu sempurna dengan sedikit dorongan udara dari paru-paru.
b. *Chest voice* adalah suara yang dihasilkan apabila ruang resonansi terjadi di rongga mulut atau para ahli mengatakan di dada.
c. *Falsetto* adalah *head voice* yang resonansinya tidak mencapai rongga hidung atau kepala, Suara yang dihasilkan adalah berat dan sedikit sengau seperti gabus, serta nada yang dihasilkan tidak setinggi head voice.
d. *Head voice* adalah suara yang dihasilkan apabila ruangan resonansi terjadi di rongga hidung atau kepala. Sifat dari *head voice* adalah ringan, nyaring, lembut, lebih merdu dari *falsetto*, dan nada yang dihasilkan juga lebih tinggi dari *falsetto*.

Pre-Emphasized

Filter pre-emphasis merupakan salah satu model pemfilteran sinyal suara yang digunakan untuk mendapatkan bentuk spectral frekuensi sinyal suara yang lebih halus (Chauhan, Samal, & Ghoshal, 2015) (Parwinder & Rani, 2014).

Fast Fourier Transform

Fast Fourier Transform (FFT) merupakan pengembangan dari Fourier Transform (FT) yang ditemukan pada tahun 1965. Penemu FT adalah J. Fourier pada tahun 1822. FT membagi sebuah sinyal menjadi frekuensi yang berbeda-beda dalam fungsi eksponensial yang

kompleks(Sipasulta, Lumenta, & Sompie, 2014).

Fast fourier transformation merupakan proses lanjutan dari DFT. Transformasi *Fourier* ini dilakukan untuk mentransformasikan sinyal dari domain waktu ke domain frekuensi. Hal ini bertujuan agar sinyal dapat diproses dalam spektral substraksi. DFT dilakukan dengan mengimplementasikan sebuah transformasi, dengan panjang vektor *N* berdasarkan rumus (Hanggarsari, Fitriawan, & Yuniati, 2012) :

$$F(u) = \frac{1}{N} \sum_{x=0}^{x=N-1} f(x)\exp[-2j\pi ux/N$$

$$F(u) = \frac{1}{N} \sum_{x=0}^{x=N-1} fx(x)\left(\cos\left(\frac{2\pi ux}{N}\right) - sin\left(\frac{2\pi ux}{N}\right)\right)$$

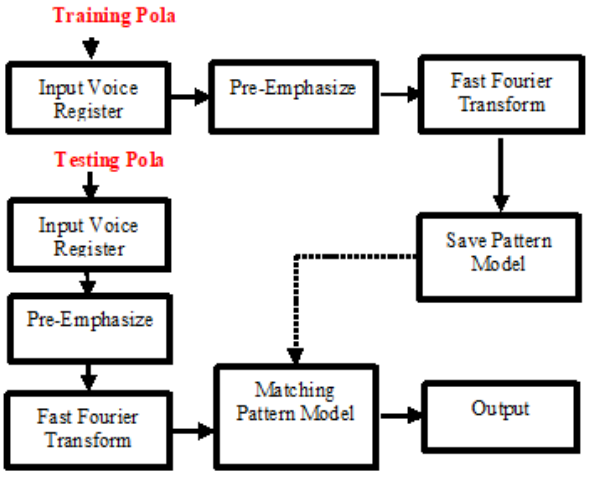
METODE PENELITIAN

Datasheet

Datasheet yang digunakan dalam penelitian ini yaitu file suara dengan register suara *Chest Voice*, *Head Voice*, *Falsetto*, dan *Vocalfry* yang diperoleh dari hasil rekaman suara wanita dengan bantuan software adobe audition 1.5.

Langkah Penelitian

Langkah penelitian secara umum yang dirancang pada penelitian meliputi dua buah proses, proses *training* data dan proses *testing* data, skema proses *training* dan *testing* dapat dilihat pada gambar 1.

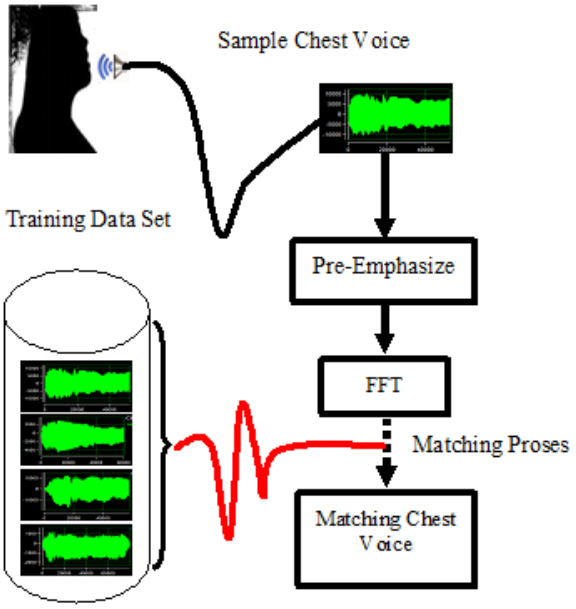


Gambar 1. Langkah Penelitian Secara Umum (Muhathir,2018) (Muhathir, Mawegkang, & Ramli,2017)

Pada Gambar 1 terdapat dua buah proses yaitu : proses *training* dan proses *testing*, pada proses *training* sinyal suara inputan dilakukan pre-processing dengan pre-emphasize untuk menghaluskan sinyal suara dan di ikuti dengan *Fast Fourier Transform* untuk mentransformasikan sinyal suara dari domain waktu ke domain frekuensi dan hasil transformasi disimpan sebagai model pola untuk proses pengujian, sedangkan pada proses pengujian sinyal suara inputan diproses dengan pre-emphasize dan dilanjutkan dengan memtransformasikan sinyal tersebut dengan menggunakan *Fast Fourier Transform,* kemudian masuk ketahap pencocokan model pola yang telah

disimpan pada tahap *training*, hasil klasifikasi akan menggolongkan register suara yang sangat mirip atau mendekati pola pada saat *training*.

Langkah penelitan secara keseluruhan pengenalan pola register suara yang dibangun dalam penelitian ini dapat diilustrasikan pada gambar 2.

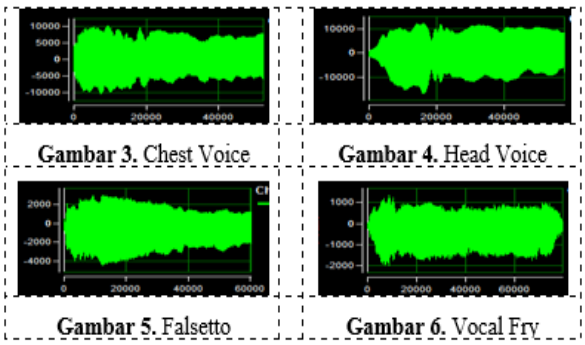


Gambar 2. Langkah Penelitian Secara Keseluruhan

HASIL DAN PEMBAHASAN

Sample of Voice Register

Sampel suara yang digunakan dalam penelitian ini berjumlah 134 sample yang mewakili register suara *Chest Voice*, *Head Voice*, *Falsetto*, dan *Vocalfry,* 60% dari 134 sample digunakan sebagai *training* data dan 40% digunakan sebagai *testing* data, gambar 3,4,5,6 menyajikan frequensi untuk masing-masing register suara



Gambar 3. Chest Voice

Gambar 4. Head Voice

Gambar 5. Falsetto

Gambar 6. Vocal Fry

Testing

Pengujian voice register dengan menggunakan Fast Fourier Transform (FFT) dilakukan 10 kali percobaan, hasil pengujian dapat dilihat pada tabel 1 sampai tabel 10, untuk memudahkan jenis register suara Chest Voice disingkat dgn CV, Head Voice disingkat dengan HV, Falsetto disingkat dengan FS dan Vocalfry disingkat dengan VF.

Tabel 1. Pengujian Pertama

		CV	*HV*	*FS*	*VF*
FFT	***CV***	**0.43**	0.21	0.07	0.29
	HV	0.17	**0.5**	0	0.33
	FS	0.08	0.25	**0.58**	0.08
	VF	0.12	0.06	0.06	**0.76**

Tabel 2. Pengujian Kedua

		CV	*HV*	*FS*	*VF*
FFT	***CV***	**0.31**	0.31	0.15	0.23
	HV	0.19	**0.75**	0	0.06
	FS	0.09	0.18	**0.64**	0.09
	VF	0.2	0	0.13	**0.67**

Tabel 3. Pengujian Ketiga

		CV	*HV*	*FS*	*VF*
FFT	***CV***	**0.43**	0.07	0.07	0.43
	HV	0.07	**0.6**	0.13	0.2
	FS	0.17	0.08	**0.67**	0.08
	VF	0.29	0.07	0.07	**0.57**

Tabel 4. Pengujian Keempat

		CV	*HV*	*FS*	*VF*

FFT	***CV***	**0.5**	0.08	0.17	0.25
	HV	0.14	**0.71**	0	0.14
	FS	0.23	0.08	**0.62**	0.08
	VF	0.19	0.06	0.06	**0.69**

Tabel 5. Pengujian Kelima

		CV	***HV***	***FS***	***VF***
FFT	***CV***	**0.73**	0.07	0.07	0.13
	HV	0.09	**0.73**	0.18	0
	FS	0.07	0	**0.93**	0
	VF	0.13	0	0.2	**0.67**

Tabel 6. Pengujian Keenam

		CV	***HV***	***FS***	***VF***
FFT	***CV***	**0.57**	0.29	0.07	0.07
	HV	0	**0.88**	0	0.12
	FS	0.18	0.18	**0.45**	0.18
	VF	0.07	0.07	0	**0.86**

Tabel 7. Pengujian Ketujuh

		CV	***HV***	***FS***	***VF***
FFT	***CV***	**0.79**	0.07	0	0.14
	HV	0.33	**0.58**	0.08	0
	FS	0	0.07	**0.93**	0
	VF	0.33	0	0.13	**0.53**

Tabel 8. Pengujian Kedelapan

		CV	***HV***	***FS***	***VF***
FFT	***CV***	**0.5**	0.17	0.08	0.25
	HV	0.08	**0.77**	0.15	0
	FS	0	0	**0.93**	0.07
	VF	0.17	0.2	0.33	**0.4**

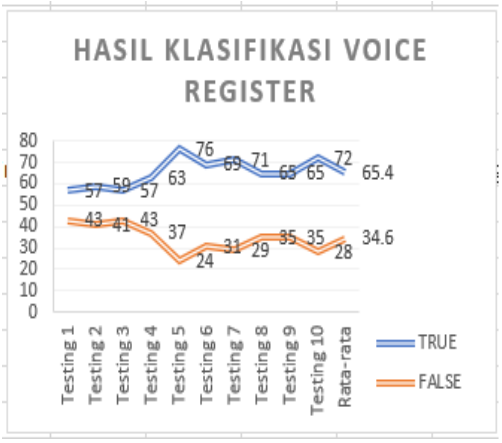
Tabel 9. Pengujian Kesembilan

		CV	***HV***	***FS***	***VF***
FFT	***CV***	**0.53**	0.07	0.2	0.2
	HV	0.07	**0.71**	0.14	0.07
	FS	0.11	0.11	**0.78**	0
	VF	0.24	0	0.18	**0.59**

Tabel 10. Pengujian Kesepuluh

		CV	***HV***	***FS***	***VF***
FFT	***CV***	**0.62**	0	0	0.38
	HV	0	**0.58**	0.08	0.33
	FS	0	0	**0.9**	0.1
	VF	0.12	0	0.12	**0.76**

Berdasakan hasil sepeuluh percobaan pengujian, hasil untuk masing-masing percobangan memiliki tingkat akurasi yang berbeda-beda. Pada hasil pengujian pertama FFT mampu mengenai pola voice register dengan akurasi 57%, hasil pengujian kedua FFT mampu mengenai pola voice register dengan akurasi 59%, hasil pengujian ketiga FFT mampu mengenai pola voice register dengan akurasi 57%, hasil pengujian keempat FFT mampu mengenai pola voice register dengan akurasi 63%, hasil pengujian kelima FFT mampu mengenai pola voice register dengan akurasi 76%, hasil pengujian keenam FFT mampu mengenai pola voice register dengan akurasi 69%, hasil pengujian ketujuh FFT mampu mengenai pola voice register dengan akurasi 71%, hasil pengujian kedelapan FFT mampu mengenai pola voice register dengan akurasi 65%, hasil pengujian kesembilan FFT mampu mengenai pola voice register dengan akurasi 65%, hasil pengujian kesepuluh FFT mampu mengenai pola voice register dengan akurasi 72%, dan rata-rata akurasi dari masing-masing pengujian yaitu 65.4%. untuk lebih jelasnya dapat dilihat pada gambar 7.



Gambar 7. Hasil pengujian

Pembahasan

Berdasarkan hasil penelitian yang telah dilakukan dengan menguji cobakan sepuluh kali percobaan belum memperoleh hasil yang maksimal dalam mengklasifikasikan voice register seperti yang ditunjukkan pada gambar 7, dari sepuluh kali percobaan hanya tiga percobaan yang memperoleh hasil akurasi diatas 70%.

SIMPULAN

Hasil penelitian menunjukkan FFT belum bekerja maksimal dalam mengklasifikasikan pola voice register.

Untuk meningkatkan kualitas FFT dalam mengklasifikasi, pada tahap pre-processing dapat menambahkan fitur filter seperti low pass filter high pass filter dan lain sebagainya.

DAFTAR PUSTAKA

Chauhan, H., Samal, S., & Ghoshal, A. (2015). VOICE RECOGNITION. International Journal of Computer Science and Mobile Computing (IJCSMC), 4(4).

Fadlisyah, & Muhathir. (2015). Perbandingan Unjuk Kerja Transformasi Wavelet, Mellin Dan Discrete Sine Transform (Dst) Untuk Pengenalanayat Al-Qur'an Pada Surat Yasiin 1-10 Melalui Suara. Techsi-Jurnal Teknik Informatika, 7(2).

Fric, M., Šram, F., & Jan, G. Š. (2006). Voice Registers, Vocal Folds Vibration Patterns And Their Presentation In Videokymography. in International Acoustical Conference.

Goyal, S., & Batra, N. (2017.). Issues and Challenges of Voice Recognition in Pervasive Environment. Indian Journal of Science and Technology, 10(30).

Hanggarsari, P. N., Fitriawan, H., & Yuniati, Y. (2012). SIMULASI SISTEM PENGACAKAN SINYAL SUARA SECARA REALTIME. ELECTRICAL Jurnal Rekayasa dan Teknologi Elektro.

M, M., F, F., & G, F. (n.d.). Voice register terminology and standard pitch. Quarterly Progress and Status Report STL-QPSR, 1963.

Meiyanti, R., Subandi, A., Fuqara, N., Budiman , M. A., & A, P. U. (2017). The Recognition Of Female Voice Based On Voice Registers In Singing Techniques In Real-Time Using Hankel Transform Method And Macdonald Function. in International Conference On Computing And Applied Informatics.

Muhathir. (2018). KLASIFIKASI EKSPRESI WAJAH MENGGUNAKAN BAG OF VISUAL WORDS. JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING (JITE), 1(2).

Muhathir, Mawengkang, H., & Ramli, M. (2017). KOMBINASI Z-FISHER TRANSFORM DAN BRAY CURTIS DISTANCE UNTUK PENGENALAN POLA HURUF JAR PADA CITRA AL-QURAN. Jurnal Bisman Info, 4(1).

Parwinder, P. S., & Rani, P. (2014). An Approach to Extract Feature using MFCC. IOSR Journal of Engineering (IOSRJEN), 4(8).

Sipasulta, R. Y., Lumenta, A. S., & Sompie, S. R. (2014). Simulasi Sistem Pengacak Sinyal Dengan Metode FFT (Fast Fourier Transform). E-journal Teknik Elektro dan Komputer.

Swamy, S., & K, V. R. (2013). An Efficient Speech Recognition System. Computer Science & Engineering: An International Journal (CSEIJ), 3(4).