

KLASIFIKASI SUARA LOVEBIRD DENGAN METODE MEL FREQUENCY CEPSTRAL COEFFICIENT (MFCC) DAN FUZZY LOGIC

Warble Of Lovebird Classification Using Mel Frequency Cepstral Coefficient (MFCC) and Fuzzy Logic

Wulandary Ika Hanesia¹Dr. Ir. Bambang Hidayat, DEA²Unang Sunarya, S. T., M. T.³

^{1,2,3}Fakultas Teknik Elektro – Universitas Telkom
Jl. Telekomunikasi, Dayeuh Kolot, Bandung 40257 Indonesia

¹wlndrika@yahoo.co.id²bhidayat@telkomuniversity.ac.id³unangsunarya@telkomuniversity.ac.id

ABSTRAK

Perkembangan teknologi yang semakin pesat dapat memudahkan kehidupan masyarakat. Tidak hanya untuk pendidikan, pemerintahan atau fasilitas kesehatan, Tetapi juga berkembang sebagai bagian dari gaya hidup, salah satunya dalam menyalurkan hobi khususnya untuk pencinta kicau burung. Pada Tugas Akhir ini akan dibuat suatu sistem aplikasi yg dapat mengklasifikasikan jenis suara kicau lovebird yang bagus dan tidak bagus. Sistem ini menerapkan *metode Mel Frequency Cepstral Coefficients (MFCC)* sebagai ciri untuk membedakan jenis kualitas kicau burung dan algoritma *Fuzzy Logic* untuk pengklasifikasian kicau burung. Pada hasil penelitian ini adalah didapatkan nilai akurasi sebesar 91,67% dengan waktu komputasi 190.229 detik. Diharapkan dengan kemampuan sistem ini ,dapat membantu para pecinta kicau burung mengetahui kualitas suara burung yang bagus, sedang, dan tidak bagus.

Kata kunci: Lovebird, Kicau burung, MFCC, *Fuzzy Logic*

ABSTRACT

Technological growth makes people life easier in many fields. Not only for education, government or health facilities, it also grows on lifestyle field. For example, many people spend their time for hobby, especially for bird lovers, chirping lovebird contest existence increasingly go up recently. This final project will build up an application system which can classify type and quality of Lovebird's chirping. This system applied Mel Frequency Cepstral Coefficients (MFCC) Method as a characteristic to differentiate quality bird's chirping and Fuzzy Logic to classify type bird's chirping. This research gain 91.667% of accuracy with computing time 190.229 second.

With this system ability, expected can help and attract bird lovers to know sounds quality bird's chirping whether its good, fair, or bad.

Keywords : Lovebird, Warble, MFCC, *Fuzzy Logic*

1. Pendahuluan

Perkembangan teknologi yang semakin pesat memberikan banyak keuntungan pada masyarakat. Salah satu keuntungannya yaitu memberikan kemudahan dalam memenuhi kebutuhan, begitu juga dalam menyalurkan hobi yang khususnya untuk komunitas pecinta burung kicau. Kemudahan yang diperlukan oleh komunitas tersebut adalah adanya sistem yang dapat mendeteksi suara kicau burung, sehingga mereka mudah dalam menentukan suara kicau burung yang bagus pada setiap perlombaan burung kicau. Eksistensi perlombaan lovebird ternyata tidak diikuti dengan pengetahuan tentang seni suara lovebird. Akibatnya banyak kicau mania yang tertipu dalam memilih lovebird yang berkualitas.

Bedasarkan urain masalah diatas maka penulis membuat 'Klasifikasi Suara Lovebird dengan *Metode Mel Frequency Cepstral Coefficient (MFCC)* dan *Fuzzy Logic*' yaitu suatu aplikasi Matlab yang dapat memudahkan para pecinta burung mengetahui suara burung lovebird yang berkualitas dan dapat bersaing di perlombaan agar pemilihan lovebird memiliki objektivitas yang tinggi dan terukur secara ilmiah.

2. DASAR TEORI

2.1 MFCC (Mel Frequency Cepstrum Coefficients)

MFCC merupakan salah satu metode yang digunakan untuk melakukan *feature extraction* pada sebuah sinyal. Metode ini adalah metode standar untuk melakukan ekstraksi ciri dimana metode ini mengadopsi cara kerja dari pendengaran manusia (Davis dan Mermelstein, 1980). Sinyal suara akan di *filter* secara linear untuk frekuensi dibawah 1000 Hz dan secara logaritmik untuk frekuensi diatas 1000 Hz[4].

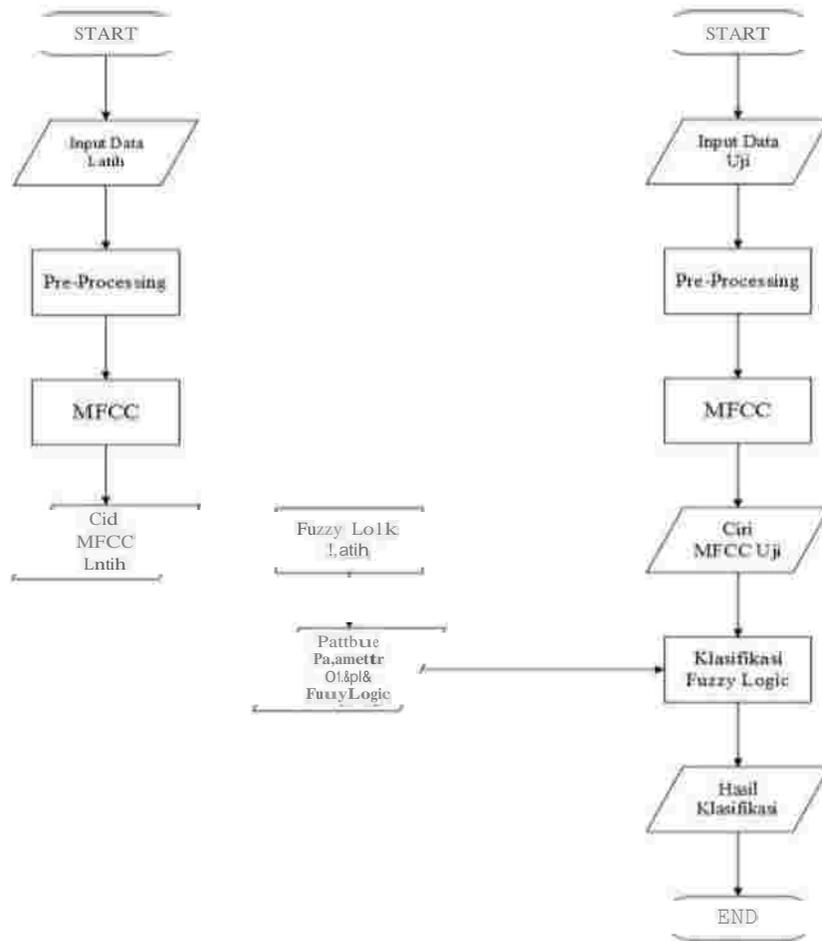
2.2 Fuzzy Logic

Logika fuzzy merupakan sebuah teori yang pertama kali diperkenalkan pada tahun 1965 oleh Prof. Lotfi A. Zadeh. Dasar logika fuzzy adalah teori himpunan fuzzy, dimana peranan derajat keanggotaan sebagai penentu keberadaan elemen dalam suatu himpunan sangatlah penting[3]. Pengklasifikasian suara dengan aturan berbasis *fuzzy* terdiri dari tiga komponen utama yaitu *Fuzzification, Inference, dan Defuzzification*.

3. Model dan Perancangan Sistem

3.1 Perancangan Sistem

Berikut adalah alur perancangan sistem klasifikasi kualitas suara burung:



Gambar 1 Diagram Alir Perancangan Sistem

3.2 Baca Audio

Akuisisi data merupakan tahap perekaman suara. Proses perekaman suara dilakukan dengan menggunakan media telepon genggam jenis iphone 5S. Frekuensi sampling yang digunakan adalah 44100 Hz. Format data suara disimpan dalam format *Waveform (*.wav)*.

3.3 Preprocessing

3.3.1 Inisialisasi Daerah Silent

Proses ini merupakan proses melakukan pencarian daerah *silent* berdasarkan nilai standar deviasi dari masing-masing sinyal suara. Proses inisialisasi ini hanya dilakukan pada daerah *silent* awal sinyal, bukan pada bagian tengah maupun akhir sinyal

3.3.2 Normalisasi

Pada proses ini dicari nilai maksimum dari data hasil *autocropping*. Keseluruhan nilai data *autocropping* kemudian dibagi dengan nilai maksimum data yang telah ditemukan. Hal ini bertujuan agar amplitudo memasuki rentang -1 sampai 1.

3.3.3 Silent Removal

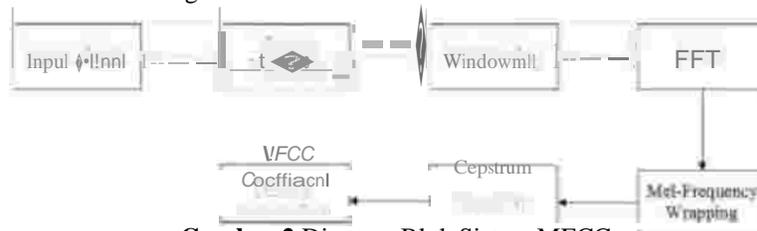
Proses ini merupakan proses untuk mencari dan menghilangkan daerah *silent* berdasarkan nilai standar deviasi dari masing-masing sinyal suara. Sedangkan untuk rumus standar deviasi sendiri adalah [5]:

$$std = \left(\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{1/2} \tag{1}$$

Setelah mendapatkan daerah silent, lalu dicari indeks maksimum dan minimum dari daerah silent tersebut dengan tujuan menghilangkan daerah *silent* dari indeks maksimum dan minimum tersebut sehingga daerah *silent* yang dihilangkan adalah daerah *silent* awal dan daerah *silent* akhir dari sinyal suara.

3.4 Ekstraksi Ciri

Ekstraksi ciri merupakan cara yang dilakukan untuk mendapatkan ciri dari sebuah suara. Proses ini merupakan tahap yang penting dalam mendeteksi suara bagus, suara sedang, dan suara tidak bagus. Metode yang digunakan dalam ekstraksi ciri adalah MFCC. Berikut diagram blok sistem MFCC :



Gambar 2 Diagram Blok Sistem MFCC

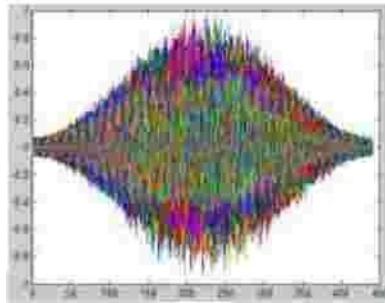
3.4.1 Frame Blocking

Dalam penelitian ini, ukuran *frame* yang digunakan adalah 0.01 dan 0.005 dalam sekon. Selain membagi bagi sinyal kedalam beberapa *frame*, dilakukan juga proses *overlap* antar *frame* dimana panjang *overlap* nya setengah dari panjang *frame*. Tujuan dari proses ini ialah agar tidak ada sedikitpun sinyal yang hilang.

3.4.2 Windowing

Proses *windowing* digunakan untuk mengurangi kebocoran *spectral* dan mengurangi efek diskontinuitas di awal dan akhir masing-masing *frame* yang dapat terjadi akibat proses *frame blocking*. Pada penelitian ini menggunakan *window Hamming*. Berikut persamaan window *Hamming* [5]:

$$w(n) = \begin{cases} 0.54 - 0.46 \cos \left(\frac{2\pi n}{N-1} \right) & 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

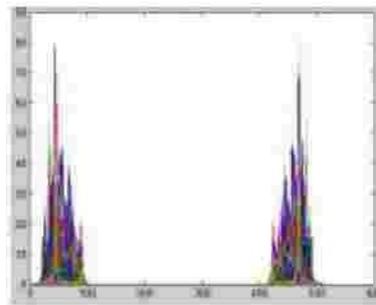


Gambar 3 Hasil Simulasi Windowing

3.4.3 FFT

Fast Fourier Transform merupakan proses realisasi dari DFT (*Diskrit Fourier Transform*) yaitu mentransformaikan sinyal dari domain waktu ke domain frekuensi. Berikut adalah persamaan dari FFT [5]:

$$f(n) = \sum_{k=0}^{N-1} y_k e^{-2\pi jkn/N}, n = 0,1,2,\dots,N - 1 \tag{3}$$



Gambar 4 Hasil Simulasi FFT

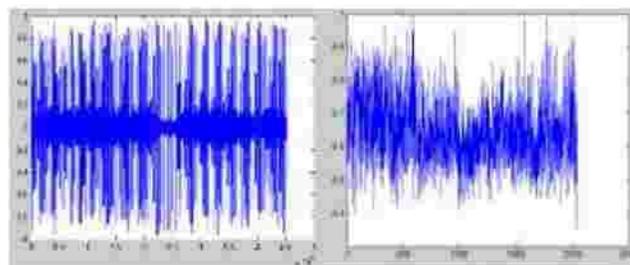
3.4.4 Mel Frequency Warping

Mel Frequency Warping umumnya dilakukan menggunakan *filterbank*. Filter ini merupakan bagian dari MFCC yang menerapkan cara kerja dari organ pendengaran manusia. Pada penelitian ini parameter filterbank yang digunakan adalah 20. Satu relasi frekuensi suara (dalam Hz) dengan skala *mel* adalah [5]

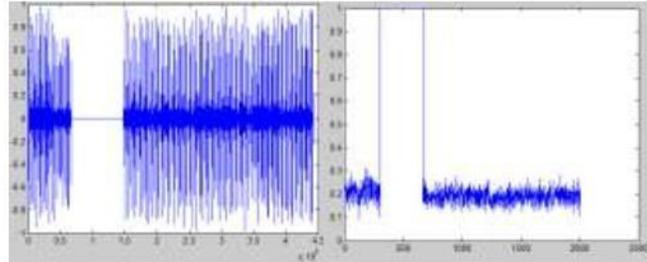
$$f_{mel} = 2592 * \log_{10} \left(1 + \frac{f}{700} \right) \tag{2-3}$$

3.4.5 Cepstrum

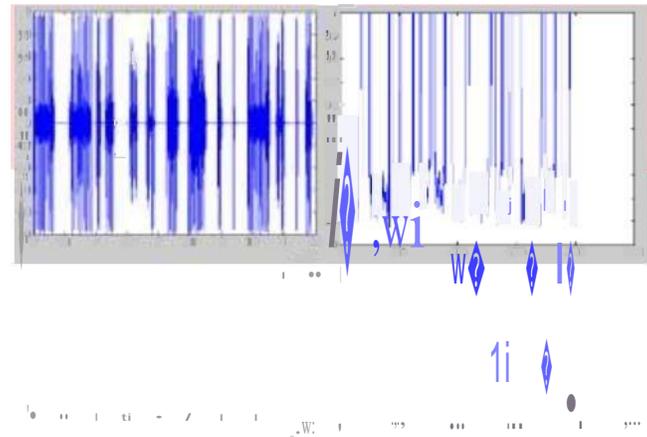
Cepstrum ialah tahap akhir dari proses ini, yaitu mengubah spektrum *log-mel* menjadi domain waktu. Hasil dari proses *Cepstrum* inilah yang disebut sebagai *Mel Frequency Cepstral Coeficient* (MFCC).



Gambar 3.4 Contoh Sinyal Inputan dan Hasil Ekstraksi Ciri MFCC Suara Bagus



Gambar 3.5 Contoh Sinyal Inputan dan Hasil Ekstrasi Ciri MFCC Suara Sedang



Gambar 3.6 Contoh Sinyal Inputan dan Hasil Ekstrasi Ciri MFCC Suara Tidak Bagus

4. Pengujian Sistem dan Analisis

4.1 Analisis Pengujian Akurasi dan Waktu Komputasi pada Koefisien MFCC dan Jumlah Frame

Pada bagian ini akan dilihat pengaruh perubahan parameter jumlah koefisien MFCC terhadap akurasi yang diperoleh dan waktu komputasinya. Pada ekstraksi ciri MFCC parameter yang digunakan adalah tetap, yaitu ukuran frekuensi *cutoff* = 0.50, range *frekuensi cutoff* = [150 4000], *number of filterbank channel* = 20, *cepstral sine lifter parameter* = 22, dan parameter yang diubah adalah ukuran frame = 0.01 dan 0.005 sekon. Sedangkan pada klasifikasi *Fuzzy Logic*, parameter yang tetap adalah nilai *epoch* = 5.

Tabel 4.1 Hasil Simulasi MFCC dan Ukuran Frame Terhadap Nilai Akurasi dan Waktu Komputasi

FRAME	MFCC											
	40		41		42		43		44		45	
	A	T	A	T	A	T	A	T	A	T	A	T
0.01	70.833	189.313	91.667	189.938	79.167	190.98	83.333	190.003	79.167	190.618	79.167	189.976
0.005	58.333	507.808	62.5	446.669	58.333	490.025	58.333	487.157	58.333	467.016	58.333	448.98

Dari hasil pengujian tersebut diperoleh hasil akurasi maksimal adalah 91,1667% pada saat jumlah koefisien nilai MFCC = 41, dengan nilai ukuran frame = 0.01sekon. Waktu komputasi ekstrasi ciri rata-rata yang diperoleh dari jumlah koefisien nilai MFCC = 41 dan nilai ukuran frame = 0.01 sekon adalah 189.938 detik untuk keseluruhan proses ekstrasi ciri pengujian suara

4.2 Analisis Pengujian Akurasi dan Waktu Komputasi pada Epoch dan Jumlah Frame

Pada skenario 2, setiap perubahan nilai *epoch* dan ukuran frame akan berpengaruh pada perhitungan akurasi dan waktu komputasi saat klasifikasi. Interval yang digunakan dalam pengujian ini adalah 5. Berdasarkan pengujian yang telah dilakukan suara uji diamati dengan jumlah *epoch* pada saat 1, 5, 10, 15, dan 20.

Tabel 4.2 Hasil Simulasi Epoch dan Ukuran Frame Terhadap Nilai Akurasi dan Waktu Komputasi

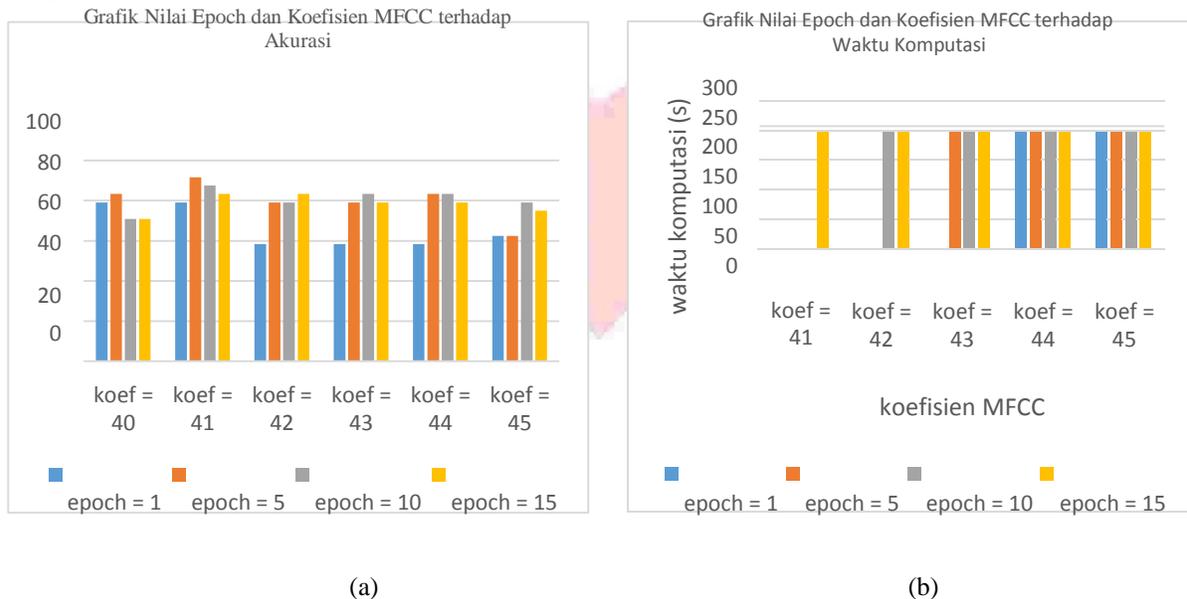
FRAME	Epoch							
	1		5		10		15	
	A	T	A	T	A	T	A	T

0.01	79.167	0.315	91.667	0.229	87.5	0.293	83.333	0.303
0.005	66.667	0.283	62.5	0.291	66.667	0.33	66.667	0.347

Dari hasil pengujian epoch dan ukuran frame (dalam sekon) diperoleh hasil akurasi maksimal adalah 91,667% dan waktu komputasi 0.229 detik pada saat nilai *epoch* = 5 untuk keseluruhan proses klasifikasi suara lovebird.

4.3 Analisis Pengujian Akurasi dan Waktu Komputasi pada Koefisien MFCC dan *Epoch*

Pada ekstraksi ciri MFCC parameter yang digunakan adalah tetap, yaitu ukuran frekuensi *cutoff* = 0.50, ukuran lebar *frame* = 0.01, range *frekuensi cutoff* = [150 4000], *number of filterbank channel* = 20, *cepstral sine lifter parameter* = 22, dan parameter yang diubah adalah nilai koefisien MFCC dan nilai *epoch*.



Gambar 4.1(a),(b) Grafik Hasil Akurasi dan Waktu Komputasi

Dari hasil pengujian epoch dan koefisien MFCC diperoleh hasil akurasi maksimal adalah 91,667% dan waktu komputasi 190.224 detik pada saat nilai *epoch* = 5 untuk keseluruhan proses klasifikasi suara lovebird.

5. Kesimpulan

Berdasarkan pengujian dan analisis sistem yang telah dilakukan terhadap proses klasifikasi suara batuk dapat ditarik kesimpulan yaitu:

1. Metode *Mel Frequency Cepstral Coefficient* (MFCC) sebagai ekstraksi ciri dan *Fuzzy Logic* sebagai klasifikasi dapat digunakan untuk mengklasifikasikan suara lovebird berkualitas bagus, sedang, dan tidak bagus.
2. Ukuran lebar *frame*, banyaknya koefisien MFCC, dan nilai iterasi (*epoch*) berpengaruh terhadap akurasi sistem.
3. Perancangan sistem pengklasifikasian kualitas suara bagus, sedang dan tidak bagus sesuai standar PBI menghasilkan nilai akurasi maksimal sebesar 91.667% dengan waktu komputasi keseluruhan saat ekstraksi ciri sampai klasifikasi selama 190.224 *second* dimana kondisi maksimal ini diperoleh dari ukuran lebar *frame* sebanyak 0.01 dalam sekon atau 441 sample, koefisien MFCC sebanyak 41 buah koefisien dan nilai iterasi (*epoch*) sebanyak 5 kali dengan 15 data latih dan 24 data uji.
4. Pada sistem pengklasifikasian ini, akurasi sistem berbanding lurus dengan waktu komputasi

Adapun beberapa saran untuk penelitian-penelitian selanjutnya adalah mengubah metoda ekstraksi ciri dan klasifikasi untuk meningkatkan nilai akurasi. Selain itu diharapkan sistem selanjutnya dapat dikembangkan pada aplikasi android dan dapat digunakan secara *real time*.

Daftar Pustaka:

- [1] Handono Biantono, Setiawan, Rusli. 2013. *Sukses Menangkarkan dan Memelihara Lovebird*. Jakarta: Penebar Swadaya.
- [2] Susanto Dian Adijaya. 2013. *Rahasia Burung Juara*. Jakarta: PT Trubus Swadaya.
- [3] G. J. Klir dan B. Yuan, *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems*, Singapore: World Scientific, 1996.
- [4] Muda, L, Mumtaj, B., & Elamvazuthi, L, 2010, *Voice Recognition Algorithm using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DWT) Techniques*, Journal of Computing Volume 2, Issue 3, March 2010.
- [5] N. K. R. Intan, *Analisis Dan Simulasi Sistem Penerjemah Kata Berbahasa Bali Ke Bahasa Inggris Berbasis Speech To Text Secara Real Time Menggunakan Metode Klasifikasi Hidden Markov Model*, Tugas Akhir, Bandung: Universitas Telkom, 2014.

