ANALISIS DETEKSI TIPE SUARA PADA PRIA DAN WANITA MENGGUNAKAN METODE MEL-FREQUENCY CEPSTRAL COEFFICIENT DAN KLASIFIKASI SUPPORT VECTOR MACHINE MULTI-KELAS ONE-AGAINST-ALL

Reni Dwifebrianti[1], Bambang Hidayat[2], Inung Wijayanto[3]

[1]Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

Abstrak

Dalam dunia musik, suara manusia merupakan bagian yang penting selain instrument musik. Manusia memiliki tipe suara yang berbeda-beda yang dapat dibedakan dalam dua golongan besar yaitu tipe suara pada pria dan wanita. Untuk tipe suara pada pria dibagi lagi menjadi tenor, bariton, dan bass, sedangkan pada wanita dibagi menjadi sopran, mezzo sopran, dan alto. Dalam menentukan tipe suara biasanya dilakukan dengan cara manual yang masih memiliki ketergantungan pada pelatih vokal dan bantuan alat musik (umumnya piano), biasanya dilakukan dalam penentuan kelompok paduan suara. Pada penelitian sebelumnya telah dilakukan percobaan pendeteksian tipe suara, namun hasil akurasinya masih sangat rendah.

Pada tugas akhir ini, untuk membuat sistem pendeteksi tipe suara pada pria dan wanita pertama-tama dilakukan ekstraksi ciri menggunakan metode Mel-Frequency Cepstral Coefficient yang merupakan representasi terbaik dari analisis timbre. Dari hasil ekstraksi ini maka diperolehlah vektor ciri yang membedakan antara tipe suara yang satu dengan yang lain. Selain itu, hasil dari MFCC ini juga dapat memberikan informasi mengenai perbedaan gender dari masing-masing tipe suara yang juga akan digunakan sebagai info untuk membantu proses klasifikasi. Info pitch juga digunakan dalam membantu proses klasifikasi karena data suara yang digunakan merupakan rekaman pada range nada tertentu. Untuk mengklasifikasikan hasil ciri yang diperoleh, dicarilah hyperplane terbaik menggunakan metode Support Vector Machine yang kemudian digunakan sebagai pemisah ciri antar kelas.

Hasil dari pengujian yang telah dilakukan pada tugas akhir ini menunjukkan bahwa sistem dapat mendeteksi tipe suara pria dan wanita dengan akurasi terbaik 80,08%. Sistem dapat dikatakan cukup baik karena sudah mencapai akurasi yang diharapkan, namun masih perlu dikembangkan lagi untuk memperoleh hasil yang maksimal.

Kata Kunci : tipe suara, MFCC, SVM, timbre

Abstract

In musical world, human voice is the important part beside the musical instrument. Human have different voice types that can be differentiated into two big classes, there are man and woman voice type. Man voice type can be divided into tenor, bariton, and bass. Woman voice type can be divided into sopran, mezzo-sopran, and alto. Generally, determination of human voice type done manually, it still depends on vocal coach and musical instrument (for choir group determination, generally). In previuos research, experiment has be done to detect voice type, but the result give bad accuracy.

Making of detector system of man and woman voice type in this final project, using Mel-Frequency Cepstral Coefficient as feature extraction method that represent timbre analysis best. Result of this extraction give the characteristics that differentiate each voice type. Besides that, MFCC result can give information about gender difference from each voice type which is also used as info for classification process. Pitch info also used to support classification process because human voice data are recorded in certain pitch. To classify feature extraction result, the best hyperplane is searched by using Support Vector Machine method which is used as feature separator between classes.

Testing result in this final project show that system can detect man and woman voice type with the best accuracy 80,08%. It can be stated that this system work good enuogh because can reach the expected accuracy, however it still need developed to get maximum result.

Keywords : voice type, MFCC, SVM, timbre

BAB 1
PENDAHULUAN

1.1 Latar Belakang

Musik sudah menjadi bagian yang tak terpisahkan dari kehidupan manusia. Banyak hobi yang berkaitan dengan musik, mulai dari hobi bermain instrumen musik sampai dengan hobi berolah suara atau menyanyi. Dalam teori musik, manusia memiliki tipe suara yang berbeda-beda baik pria maupun wanita. Tipe suara pada pria dibagi menjadi tenor, bariton, dan bass, sedangkan pada wanita dibagi menjadi sopran, mezzo sopran, dan alto, namun sebagian besar orang terutama orang awam mengalami kesulitan dalam menentukan tipe suara mereka. Dalam menentukan tipe suara biasanya dilakukan dengan cara manual menggunakan bantuan alat musik (umumnya piano) oleh seorang ahli atau pelatih vokal, biasanya dilakukan dalam penentuan kelompok paduan suara. Namun penentuan dengan cara ini memiliki ketergantungan pada pelatih vokal dan alat musik. Untuk itu, dibuatlah sistem pendeteksi untuk mempermudah penentuan tipe suara pada pria dan wanita yang sekaligus sebagai pendorong pengembangan teknologi dalam bidang olah vokal.

Penelitian ini bertujuan untuk membuat sistem yang dapat mendeteksi tipe suara pada pria dan wanita berdasarkan *feature* yang didapat. Pada penelitian sebelumnya, Indrajit Prawira Yudha Telah mensimulasikan "Identifikasi Jenis Suara Manusia Berdasarkan Jangkauan Vokal Menggunakan Jaringan Saraf Tiruan Backpropagation" dengan akurasi maksimal 69.41 %[14].

Dalam tugas akhir ini dilakukan deteksi tipe suara pria dan wanita dengan menggunakan *Mel-Frequency Cepstral Coefficient* (MFCC) untuk mengekstraksi ciri. MFCC merupakan salah satu metode yang banyak digunakan dalam bidang *speech technology*, baik *speaker recognition* maupun *speech recognition.* Selanjutnya *Support Vector Machine* digunakan sebagai metode klasifikasi ciri. SVM merupakan salah satu metode yang handal yang digunakan untuk pengenalan pola (*pattern recognition*) dengan mencari *hyperplane*/ pemisah terbaik antara kedua kelas[7].

1.2 Tujuan

Tujuan dilakukannya penelitian pada tugas akhir ini adalah:

1. Membuat sistem yang dapat mendeteksi tipe suara pada pria dan wanita,

2. Menerapkan MFCC sebagai metode ekstraksi ciri dan menganalisis kinerjanya,
3. Menerapkan SVM sebagai metode klasifikasi ciri dan menganalisis kinerjanya,
4. Menganalisis kualitas performansi sistem pendeteksi dilihat dari nilai akurasinya.

1.3 Rumusan Masalah

Permasalahan yang akan akan dibahas pada tugas akhir ini adalah:

1. Bagaimana membuat sistem yang dapat mendeteksi tipe suara pada pria dan wanita menggunakan MFCC dan SVM,
2. Bagaimana menerapkan dan menganalisis kinerja MFCC untuk ekstraksi ciri,
3. Bagaimana menerapkan dan menganalisis kinerja SVM untuk klasifikasi ciri,
4. Bagaimana analisis kualitas performansi sistem dilihat dari nilai akurasinya.

1.4 Batasan Masalah

Dalam pembahasannya, penelitian tugas akhir ini dibatasi dengan hal-hal berikut:

1. Format data suara yang diolah adalah *wave* (*.wav).
2. Proses ekstraksi ciri menggunakan *Mel-Frequency Cepstral Coefficient* dan proses klasifikasi menggunakan *Support Vector Machine.*
3. Data suara yang digunakan hanya pada pria dan wanita kisaran usia 18-23 tahun.
4. Tipe suara yang dideteksi adalah sopran dan alto pada wanita, sedangkan pada pria adalah tenor dan bass.
5. Sampel suara yang digunakan berupa ucapan ”A” pada *range* nada tertentu, yaitu C4-A5 untuk Sopran, F3-F5 untuk Alto, C3-C5 untuk Tenor, dan E2-E4 untuk Bass, dengan frekuensi *sampling* 8000 Hz.
6. Simulasi dilakukan dengan menggunakan *software* MATLAB R2010a secara *non-realtime.*

1.5 Metodologi Penelitian

Metodologi yang digunakan dalam penelitian tugas akhir ini adalah sebagai berikut:

1. Studi Literatur

   Bertujuan untuk mempelajari dasar teori mengenai klasifikasi tipe suara manusia, *Mel-Frequency Cepstral Coefficient* dan *Support Vector Machine*.

2. Pengumpulan Data

   Bertujuan untuk memperoleh sampel dari masing-masing variasi tipe suara pada pria dan wanita sesuai dengan batasan masalah yang akan dijadikan sebagai *database* ciri.

3. Implementasi Perangkat Lunak

   Bertujuan untuk melakukan implementasi pada perangkat lunak sesuai dengan perancangan yang telah dilakukan.

4. Pengujian Sistem

   Bertujuan untuk melihat tingkat akurasi sistem dalam melakukan pendeteksian tipe suara pria dan wanita.

5. Analisis Performansi

   Bertujuan untuk menganalisis performansi sistem yang telah dibuat dilihat dari tingkat akurasi.

6. Penyusunan Laporan

   Bertujuan untuk menyusun keseluruhan hasil penelitian dan menarik kesimpulan dari penelitian yang telah dilakukan.

1.6 Sistematika Penulisan

Penelitian ini disusun menjadi 5 bab dengan rincian sebagai berikut:

BAB I PENDAHULUAN

Berisi tentang latar belakang, tujuan, rumusan masalah, batasan masalah, metodologi dan sistematika penulisan.

BAB II DASAR TEORI

Berisi tentang dasar teori yang mendukung dan menjadi literatur dalam tugas akhir ini.

BAB III PERANCANGAN SISTEM

Berisi diagram alir serta tahap-tahap yang menjelaskan proses kerja/ perancangan sistem.

BAB IV PENGUJIAN SISTEM DAN ANALISIS

Berisi skenario pengujian dan analisis terhadap hasil pengujian serta performansi sistem.

BAB V PENUTUP

Berisi kesimpulan dan saran dari hasil Tugas Akhir untuk pengembangan lebih lanjut.

BAB 5
PENUTUP

5.1 Kesimpulan

Setelah dilakukan pengujian dan analisis terhadap sistem deteksi tipe suara pria dan wanita pada Tugas Akhir ini, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Sistem dapat mendeteksi *gender* dengan akurasi optimal 96,27%.
2. Sistem dapat mendeteksi tipe suara pria dan wanita dengan akurasi optimal 80,08%.
3. Spesifikasi nilai parameter yang memberikan akurasi paling optimal pada deteksi tipe suara pria dan wanita adalah panjang *frame* 256, *overlap* 50%, jumlah *mel bank filter* 31, jumlah koefisien MFCC 20, orde filter *daubechies* db4, dan SVM dengan kernel RBF 3.
4. Berdasarkan hasil skenario pengujian terhadap beberapa parameter, parameter yang mempengaruhi hasil akurasi secara signifikan adalah jumlah koefisien MFCC, orde filter *daubechies* dan *kernel function* pada SVM.
5. Berdasarkan hasil pengujian terhadap sinyal suara dengan penambahan derau, menunjukan bahwa sistem rentan terhadap derau, dengan akurasi terendah terjadi pada sinyal dengan penambahan derau suara manusia.

5.2 Saran

Tugas akhir ini masih dapat dikembangkan agar diperoleh hasil yang lebih baik. Adapun saran yang dapat diberikan untuk pengembangan untuk tugas akhir ini adalah :

1. Menggunakan jenis rekaman suara yang lain, dapat berupa rekaman suara nyanyian ataupun rekaman lainnya pada nada tertentu.
2. Mengembangkan sistem menjadi sistem *real-time*.
3. Mensimulasikan sistem dengan menggunakan bahasa pemrograman selain Matlab.
4. Pada dasarnya penggunaan MFCC sudah memberikan hasil yang baik, namun selain menggunakan MFCC dapat juga dilakukan percobaan menggunakan metode lain seperti *spectral centroid*, *spectral roll-off* dan *spectral flux* yang juga merupakan representasi ciri timbre.
5. Mengeksplor penggunaan SVM lebih dalam lagi, misalnya dicoba dibandingkan dengan metode SVM multi kelas *one-against-one*. Atau dapat dilakukan pengujian

menggunakan metode klasifikasi selain SVM untuk mencari hasil yang paling optimal.

DAFTAR PUSTAKA

[1] Ben-Hur, Asa and Jason Weston. *A User's Guide to Support Vector Machines*. Colardo State University and Princenton, NJ 08540 USA.

[2] Cernocky, Jan dan Valentina Hubeika. *Fundamental Frequency Detection*. DCGN FIT BUT Brno.

[3] Gerhard, David. 2003. *Pitch Extraction and Fundamental Frequency: History and Current Techniques*. Departement of Computer Science University of Regina, Canada.

[4] Hasan, Rashidul Md., Mustafa Jamil, Md. Golam Rabbani, dan Md. Saifur Rahman. 2004. *Speaker Identification Using Mel Frequency Cepstral Coefficients*. Bangladesh University of Engineering and Technology, Dhaka.

[5] Kura, Vijay B. 2003. *Novel Pitch Detection Algorithm With Application to Speech Coding*. B. Tech Jawaharlal Institute of Technological University.

[6] Mustofa, Ali. 2007. *Sistem Pengenalan Penutur dengan Metode Mel-frequency Wrapping*. Jurnal Teknik Elektro. Vol. 7. No. 2: 88 – 96.

[7] Nugroho, Anto Satriyo, dkk. 2003. *Support Vector Machine-Teori dan Aplikasinya dalam Bioinformatika*. Kuliah Umum Ilmu Komputer.com.

[8] Prahallad, Kishore. *Speech Technology : A Practical Introduction, topic : Spectrogram, Cepstrum and Mel-Frequency Analysis*. Carnegie Mellon University & International Institute of Information Technology Hyderabad. Slide.

[9] Pramesthi, Dhita Khairunnisa. 2011. *Analisis Dialek Suara Teleponi dengan Mel-Frequency Cepstral Coefficient berbasis Pengolahan Sinyal Digital*. Bandung: Institut Teknologi Telkom.

[10] Purwanto, Kristiawan dan Tutug Dhanardono. *Simulasi Reduksi Derau Sinyal Suara Pada Gedung Kebun Raya Purwodadi Dengan Metode DWT*. ITS, Surabaya.

[11] Santosa, Budi. 2007. *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.

[12] Simanungkalit, Nortier. 2008. *Teknik Vokal Paduan Suara*. Jakarta: Gramedia Pustaka Utama.

[13] Yen, Joe. *Wavelet for Acoustics*, Technical Report R98942097 (citation for B3).

[14] Yudha, Indrajit Prawira. 2012. *Sistem Identifikasi Jenis Suara Manusia Berdasarkan Jangkauan Vokal Menggunakan Jaringan Saraf Tiruan Backpropagation*. Bandung: Institut Teknologi Telkom.