

ANALISIS PERFORMANSI DAN SIMULASI VOICE ACTIVITY DETECTION (VAD) SECARA REAL TIME MENGGUNAKAN METODE LINEAR DISCRIMINANT ANALYSIS (LDA)

Ahmad Sirojuddin Zarkasyi¹, Bambang Hidayat², Gelar Budiman³

¹Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

Abstrak

Dalam bidang teknologi telekomunikasi, khususnya dalam hal voice, kita mengenal istilah Voice Activity Detection (VAD). Secara umum, Voice Activity Detection (VAD) didefinisikan sebagai sebuah proses pengidentifikasian keadaan voice atau silent dalam sebuah sinyal suara. Penggunaan VAD umumnya digunakan pada speech recognition dan speech coding. Dalam speech recognition, VAD digunakan sebagai indikator suatu keadaan voice ataupun silent dari suatu sinyal suara sedangkan didalam speech coding / kompresi, VAD digunakan untuk memfasilitasi skema Discontinuous Transmission (DTX). Adapun DTX itu sendiri merupakan suatu model transmisi yang memiliki cara kerja hanya akan mentransmisikan sedikit bit - bit ketika keadaan suatu speech dalam kondisi silent maupun dengan background noise dengan tujuan mengefisienkan pengkodean. Semakin akurat metode VAD yang digunakan maka akan semakin bagus performansinya dan akan mengefisienkan pengiriman bit - bitnya.

Dalam mengidentifikasi keadaan voice maupun silent, VAD dikatakan akurat tergantung dari algoritma yang dibuat. Dalam Tugas Akhir ini metode VAD yang digunakan adalah metode Linear Discriminant Analysis (LDA). Sistem akan mengekstraksi sinyal suara yang dipengaruhi noise dengan SNR yang berbeda - beda menjadi dua macam ciri yaitu Zero Crossing dan Energi bit. Dua macam ciri tersebut akan dijadikan input dalam metode LDA dan diharapkan mampu memberikan output yang memiliki performansi SDER, NDER dan OVER yang bagus.

Dalam simulasi VAD menggunakan metode Linear Discriminant Analysis (LDA) ini didapatkan performansi SDER, NDER dan OVER dibawah 12% dan dapat disimpulkan bahwa Linear Discriminant Analysis (LDA) dapat digunakan sebagai salah satu metode Voice Activity Detection karena performansinya cukup bagus hingga mencapai 88%.

Kata Kunci : DTX, Voice Activity Detection (VAD), real time, Linear Discriminant Analysis (LDA) , SDER, NDER, OVER.

Telkom
University

Abstract

In the field of telecommunication, especially in terms of voice, we recognize the term Voice Activity Detection (VAD). In general, Voice Activity Detection (VAD) is defined as a process of identifying the circumstances in a voice or silent sound signals. VAD is commonly used in speech recognition and speech coding. In speech recognition, VAD is used as an indicator of a silent state or a voice state of a voice signal, while in the speech coding / compression, VAD is used to facilitate the Discontinuous Transmission (DTX) scheme. Discontinuous Transmission (DTX) is a transmission method which will work only transmit a few bits of silent or noise background condition in order of coding efficiency. The more accurate the VAD method the better will be the performance and will make the data streamline more efficient. Linear Discriminant Analysis (LDA) is used to VAD method of this final project.

VAD accuracy of identification voice or silent condition is depend on the algorithm. Linear Discriminant Analysis (LDA) is used to VAD method of this final project. System will make two extractions from mixed between speech and noise of signal which have different SNR. Two of extractions are Zero Crossing and bit energy. This parameter will be input of LDA method and it hopes can give a good SDER, NDER, and OVER performance from the output.

VAD using LDA method simulation has 12 % SDER, NDER, and OVER performance and it can conclude that LDA can be use as one of VAD method because it has a good performance with more than 88%.

Keywords : DTX, Voice Activity Detection, real time, Linear Discriminant Analysis (LDA), SDER, NDER, OVER.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam bidang teknologi telekomunikasi, khususnya dalam hal voice, kita mengenal istilah *Voice Activity Detection* (VAD). Secara umum, *Voice Activity Detection* (VAD) didefinisikan sebagai sebuah proses pengidentifikasian keadaan *voice* atau *silent* dalam sebuah sinyal suara. Penggunaan VAD umumnya digunakan pada *speech recognition* dan *speech coding*. Dalam *speech recognition*, VAD digunakan sebagai indikasi suatu keadaan *voice* ataupun *silent* dari suatu sinyal suara sedangkan didalam *speech coding* / kompresi, VAD digunakan untuk memfasilitasi skema *Discontinuous Transmission* (DTX). DTX itu sendiri didefinisikan sebagai sebuah mode transmisi yang hanya akan menstranisikan sedikit saja bit – bit ketika keadaan dalam kondisi silent maupun dengan *background noise* sehingga ini akan mengefisienkan pengkodeannya.

Dalam mengidentifikasi kondisi *voice* ataupun *silent* ini sangatlah penting karena jika salah dalam mengidentifikasi keadaan *voice* ataupun *silent* tersebut maka akan banyak bit yang akan ditransmisikan sehingga skema *Discontinuous Transmission* (DTX) tidak bisa digunakan. Permasalahan tersebut timbul akibat kurang akuratnya algoritma yang digunakan untuk mengidentifikasi keadaan *voice* ataupun *silent* dalam *Voice Activity Detection* (VAD) ini. Dalam kehidupan sehari – hari *noise – noise* yang mempengaruhi performansi VAD ini sangatlah banyak kita jumpai. *Noise – noise* tersebut antara lain *Vehicle Noise* yaitu noise yang berasal dari mesin kendaraan bermotor. *Babble Noise* yaitu noise yang ditimbulkan akibat banyaknya orang yang berbicara dalam satu waktu, hal ini dapat kita jumpai di gedung A atau B IT Telkom, pasar, restoran, mall dll. Selain itu masih banyak contoh – contoh noise yang mempengaruhi performansi VAD ini dan *noise – noise* tersebut memiliki tingkat SNR yang berbeda – beda.

Untuk mengatasi permasalahan diatas mengenai tingkat keakuratan performansi pengidentifikasian, maka dalam tugas akhir ini akan digunakan salah satu algoritma pengidentifikasi yaitu menggunakan metode *Linear Discriminant Analysis* (LDA) dan dalam sistem tersebut akan dilatih dan diuji menggunakan SNR yang berbeda – beda. Diharapkan metode *Linear Discriminant Analysis* (LDA) juga akan memberi tingkat akurasi performansi SDER, NDER dan OVER dibawah 15%. Serta juga dapat

BAB I PENDAHULUAN

memperbaiki performansi TA sebelumnya^[1], sehingga dalam Tugas Akhir ini performansi VAD dianalisis secara *real time*.

1.2 Perumusan Masalah

Perumusan masalah yang menjadi acuan dalam pembuatan tugas akhir ini adalah:

1. Bagaimana suatu sistem dapat mendeteksi keadaan *voice* atau *silent* dengan *input* yang memiliki kondisi *voice* tercampur dengan *vehicle noise* atau dengan *babble noise* yang memiliki SNR 25 dB, 20 dB, 15 dB, 10 dB, 5 dB.
2. Bagaimana algoritma *Linear Discriminant Analysis* (LDA) dapat mendeteksi pola dengan baik.
3. Bagaimana menentukan parameter – parameter yang digunakan di dalam algoritma *Linear Discriminant Analysis* (LDA) dan di dalam sistem agar dapat mendeteksi dengan akurat.
4. Menentukan persentase keakuratan algoritma *Linear Discriminant Analysis* (LDA) untuk algoritma VAD yaitu SDER, NDER dan OVER.
5. Bagaimana performansi antara VAD secara *real time* dengan menggunakan algoritma *Linear Discriminant Analysis* (LDA) serta VAD secara tidak *real time* menggunakan *JST Backpropagation*.

1.3 Tujuan

Adapun tujuan dari penelitian untuk Tugas Akhir ini adalah :

1. Mensimulasikan dan menganalisis kinerja *Voice Activity Detection* (VAD) secara *real time* menggunakan metode *Linear Discriminant Analysis* (LDA).
2. Membuktikan apakah algoritma *Linear Discriminant Analysis* (LDA) cocok digunakan sebagai algoritma VAD.
3. Menganalisis performansi (SDER, NDER, OVER) dari *output* algoritma *Linear Discriminant Analysis* (LDA) dengan *input noise* yang memiliki SNR (*Signal to Noise Ratio*) yang berbeda – beda.
4. Mengetahui keakuratan algoritma *Linear Discriminant Analysis* (LDA) untuk VAD dengan *background vehicle noise* dan *babble noise*.
5. Dapat memperbaiki performansi TA sebelumnya^[1] yang secara tidak *real time*.

1.4 Batasan Masalah

Dalam pembahasannya, Tugas Akhir ini dibatasi pada hal-hal sebagai berikut:

1. Menggunakan *Microphone* jenis MIC-010 untuk merekam *speech* dan *noise*.
2. *Voice Activity Detection* (VAD) ini dibuat berupa simulasi.
3. Menggunakan *software* Matlab R2009a untuk simulasi VAD.
4. Metode yang digunakan adalah *Linear Discriminant Analysis* (LDA) .
5. Filter yang digunakan yaitu High Pass Filter (HPF). Karena pada perancangan Tugas Akhir ini ingin meredam *noise* yang tercampur dengan *speech*, dimana *speech* memiliki frekuensi tinggi sedangkan *noise* sendiri memiliki frekuensi rendah.
6. Menggunakan *software* Cool edit Pro v2.0.
7. *Object* yang diteliti adalah *output* sistem dengan *input* yang dipengaruhi *Vehicle Noise* dan *Babble Noise* yang memiliki SNR berbeda – beda (sesuai dengan rumusan masalah).

1.5 Metode Penelitian

Metode penelitian yang digunakan dalam Tugas Akhir ini adalah :

1. Study Literatur
Bertujuan untuk mempelajari dasar teori tentang segala literatur mengenai *Voice Activity Detection* (VAD), *real time*, metode LDA dan *software* matlab.
2. Observasi lapangan
Bertujuan untuk mendapatkan *recording* data tentang background noise yaitu berupa *vehicle noise* dan *babble noise*.
3. Study pengembangan aplikasi
Bertujuan untuk menentukan rancangan suatu sistem dan melakukan analisa terhadap sistem tersebut.
4. Implementasi Perangkat Lunak
Bertujuan untuk melakukan implementasi metode pada perangkat lunak sesuai dengan analisa perancangan yang telah dilakukan.
5. Analisis Performansi
Bertujuan untuk melakukan analisa performansi yang dapat dicapai oleh sistem.

6. Pengambilan Kesimpulan

Bertujuan untuk mengambil kesimpulan berdasar analisis yang sudah didapatkan.

1.6 Sistematika Penulisan

Penelitian ini disusun menjadi 5 BAB, dengan rincian sebagai berikut :

BAB I PENDAHULUAN

Berisi tentang latar belakang masalah, perumusan masalah, batasan masalah, tujuan, metode penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI (VOICE ACTIVITY DETECTION (VAD) DAN JST LINEAR DISCRIMINANT ANALYSIS (LDA))

Berisi tentang dasar – dasar teori yang diperlukan serta literatur literatur yang mendukung dalam *Voice Activity Detection (VAD)*, *real time* dan metode *Linear Discriminant Analysis (LDA)*.

BAB III PERANCANGAN SISTEM DAN SIMULASI

Membahas tentang sistem *Voice Activity Detection (VAD)* menggunakan algoritma *Linear Discriminant Analysis (LDA)*, Ekstraksi Ciri, parameter performansi (SDER, NDER dan OVER).

BAB IV ANALISA HASIL SIMULASI SISTEM

Menjelaskan serta menganalisis hasil keluaran dari *Voice Activity Detection (VAD)* menggunakan metode *Linear Discriminant Analysis (LDA)* setelah dilatih dan diuji menggunakan noise yang memiliki tingkat SNR yang berbeda beda. Menganalisa performansi *Voice Activity Detection (VAD)* secara *real time* menggunakan metode *Linear Discriminant Analysis (LDA)*.

BAB V KESIMPULAN DAN SARAN

Berisi kesimpulan akhir dan saran pengembangan lebih lanjut.

BAB V

PENUTUP

5.1 Kesimpulan

Dari hasil analisa pengujian sistem *Voice Activity Detection* (VAD) menggunakan metode *Linear Discriminant Analysis* (LDA) ini dapat diambil kesimpulan sebagai berikut:

1. Pada Tugas Akhir ini telah dianalisis kinerja *Voice Activity Detection* (VAD) secara *real time* menggunakan metode *Linear Discriminant Analysis* (LDA) untuk mengidentifikasi sinyal suara dalam keadaan *speech* maupun *noise* dengan performansi sekitar 88%.
2. Dalam pengujian sistem menggunakan *Vehicle Noise* yang telah ditraining sebelumnya dengan input SNR yang berbeda – beda (5dB, 10dB, 15dB, 20dB, dan 25dB) didapat bahwa SNR 25dB memiliki *error* sistem yang kecil, dimana nilai *Speech Detection Error Rate* (SDER) mencapai dibawah 11%. Sedangkan nilai *Noise Detection Error Rate* (NDER) mencapai kurang dari 1% dan untuk nilai *Overall Detection Error Rate* (OVER) mencapai 12% ini dikarenakan nilai SDER dan NDER yang cukup kecil mencapai dibawah 11% dan 1%.
3. Dalam pengujian sistem menggunakan *Vehicle Noise* yang telah ditraining sebelumnya dengan input SNR yang berbeda – beda (5dB, 10dB, 15dB, 20dB, dan 25dB) didapat bahwa SNR 25dB memiliki *error* sistem yang kecil, dimana nilai *Speech Detection Error Rate* (SDER) mencapai dibawah 11%. Sedangkan nilai *Noise Detection Error Rate* (NDER) mencapai kurang dari 1% dan untuk nilai *Overall Detection Error Rate* (OVER) mencapai 12% ini dikarenakan nilai SDER dan NDER yang cukup kecil mencapai dibawah 11% dan 1%.
4. Performansi TA sebelumnya mengenai VAD secara tidak *real time* menggunakan Jaringan Saraf Tiruan *Backpropagation* (JST BP) mencapai 100%, yang artinya metode LDA dapat dikatakan kurang bagus dikarenakan sistem VAD kali ini diuji secara *real time*, dimana sistem secara *real time* memiliki beberapa kelemahan, seperti input tercampur oleh noise yang seharusnya input tersebut bersih dari noise dan juga *microphone* yang dipakai kualitasnya kurang begitu bagus.

5.2 Saran

Pengembangan yang dapat dilakukan dalam Tugas Akhir ini adalah :

1. Menggunakan noise yang bervariasi
2. Menggunakan metode lain seperti jaringan saraf tiruan untuk memperbaiki performansi Voice Activity Detection.
3. Menggunakan *microphone* yang kualitasnya bagus sehingga tidak mempengaruhi terhadap performansi sistem.
4. Tugas Akhir ini juga dapat dikembangkan menggunakan TMS untuk pengembangan sistem.



DAFTAR PUSTAKA

- [1] Adhy Wibowo, Suryo. *Voice Activity Detection (VAD) Menggunakan Jaringan Saraf Tiruan Backpropagation Dengan Background Vehicle Noise Dan Babble Noise*. Tugas Akhir IT Telkom. 2007.
- [2] Kausari, Mulki. 2008. *Komputerisasi Iridologi untuk Mendeteksi Kondisi Ginjal Menggunakan Principal Komponen Analysis (PCA) dan K-Nearest Neighbors (KNN)*. Sekolah Tinggi Teknologi Telkom, Bandung.
- [3] *Modul Praktikum Pengolahan Sinyal Digital*. Laboratorium Pengolahan Sinyal Digital. IT Telkom, Bandung. 2010.
- [4] Prijono, Agus., Wijaya, Marfin Candra. 2007. *Pengolahan Citra Digital Menggunakan Matlab*. Bandung: Informatika.
- [5] Restuti, Sinta. 2010. *Analisis dan Implementasi Metode Linear Discriminant Analysis dan Support Vector Machine pada Pengenalan Iris Mata*. Institut Teknologi Telkom, Bandung.
- [6] Sugiharto, Aris. 2006. *Pemrograman GUI dengan MATLAB*. Yogyakarta : Andi.
- [7] Tyas Saksono, Hanung. *Pendeteksian Kanker Paru-paru dengan Menggunakan Transformasi Wavelet dan Linear Discriminant Analysis (LDA)*. 2010. Tugas Akhir IT Telkom. 2010.
- [8] Source code LDA, <http://ilmu-komputer.net/lda-matlab-source-code/>, diakses tanggal 17 Januari 2011.
- [9] Voice Activity Detection, http://en.wikipedia.org/wiki/Voice_activity_detection, diakses tanggal 18 Januari 2011.
- [10] Mikrofon, <http://id.wikipedia.org/wiki/Mikrofon>, diakses tanggal 18 Januari 2011
- [11] LDA, http://en.wikipedia.org/wiki/Linear_discriminant_analysis, diakses tanggal 18 Januari 2011.
- [12] Voice Detection, http://www.acm.org/crossroads/xrds13-4/voice_detection.html, diakses tanggal 17 Januari 2011.