

## ANALISIS DEREVERBERASI SINYAL SUARA MANUSIA DENGAN MENGUNAKAN METODE SPECTRAL SUBTRACTION

Aditya Pratama<sup>1</sup>, Bambang Hidayat<sup>2</sup>, Koredianto Usman<sup>3</sup>

<sup>1</sup>Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

<sup>1</sup>[myname@example.com](mailto:myname@example.com)

---

### Abstrak

Reverberasi adalah fenomena bercampurnya sinyal audio asli dengan sinyal-sinyal noise yang timbul karena efek pantulan. Karakteristik reverberasi yang dialami oleh suatu sinyal bersesuaian dengan fungsi transfer ruangan (Room Impulse Response) yang menyebabkan sinyal hasil pantulan memiliki delay yang bervariasi tergantung ukuran ruangan. Sinyal reverberasi dalam kasus ini sulit dihilangkan karena sinyal reverberasi berkorelasi dengan sinyal asli, sehingga metode untuk mengurangi sinyal reverberasi ini berbeda dengan metode pengurangan noise biasa. Metode yang digunakan untuk pengurangan sinyal reverberasi disebut dereverberasi.

Dereverberasi adalah metode untuk mengurangi sinyal reverberasi sehingga kualitas sinyal suara menjadi lebih baik. Salah satu metode dereverberasi yang dibahas dalam tugas akhir ini adalah metode spectral subtraction. Input yang digunakan dalam analisis adalah sinyal suara manusia tereverberasi tanpa noise, sedangkan output yang diharapkan adalah berkurangnya sinyal reverberasi sehingga terjadi peningkatan kualitas sinyal suara.

Dari hasil simulasi menggunakan metode spectral subtraction didapatkan hasil yang cukup baik dari segi performansi nilai Mean Square Error (MSE), Energy to Late Reverberation Energy Ratio (ELR), Early to Total Reverberation Energy Ratio (ETR), dan Mean Opinion Score (MOS). Untuk jenis ruangan small room didapat performansi MSE maksimal sebesar 0,0944%, performansi ELR maksimal sebesar 0,0022%, performansi ETR maksimal sebesar 0,0018%, dan nilai MOS maksimal sebesar 4,0333.

Untuk jenis ruangan medium room didapat performansi MSE maksimal sebesar 11,0213%, performansi ELR maksimal sebesar 2,4028%, performansi ETR maksimal sebesar 2,3858%, dan nilai MOS maksimal sebesar 3,7333. Sedangkan untuk jenis ruangan large room didapat performansi MSE maksimal sebesar 48,5678%, performansi ELR maksimal sebesar 14,1619%, performansi ETR maksimal sebesar 14,0263%, dan nilai MOS maksimal sebesar 3,8333. Nilai performansi maksimal di atas semuanya didapat pada sampel suara perempuan berpidato.

Kata Kunci : sinyal suara manusia, reverberasi, dereverberasi, spectral subtraction.

---

Telkom  
University

### Abstract

Reverberation is a phenomenon in which characteristics of original audio signal mixed with noise signal that is emerged due to reflection effect. Reverberation characteristic of a signal is corresponding room impulse response that cause reflected signal to have certain delay times related to room dimension. This reverberated can be hardly cleaned since the reverberation signal is correlated to the original signal, so the method used to clean this reverberated signal is different from usual noise reduction method. Method used for cleaning this reverberated signal is called dereverberation..

Dereverberation is a method to reduce the reverberant signal so that the signal quality improved. One kind of dereverberation method that is discussed in this final assignment is spectral subtraction method. Input that is used in analysis is noiseless speech signal, and the expected output is signal with less reverberant so that the signal quality can be improved.

From simulation process using spectral subtraction method some good performance value are obtained. The performance consists of several parameters which is Mean Square Error (MSE), Energy to Late Reverberation Energy Ratio (ELR), Early to Total Reverberation Energy Ratio (ETR), and Mean Opinion Score (MOS). For small room simulation maximum MSE performance is 0,0943%, maximum ELR performance is 0,0022%, maximum ETR performance is 0,0018%, and maximum MOS is 4,0333.

For medium room simulation maximum MSE performance is 11,0213%, maximum ELR performance is 2,402829383%, maximum ETR performance is 2,3858%, and maximum MOS is 3,7333. For large room simulation maximum MSE performance is 48,5678%, maximum ELR performance is 14,1619%, maximum ETR performance is 14,0263%, and maximum MOS is 3,8333. Those maximum performances are all obtained in woman's speech voice sample.

Keywords : speech signal, reverberation, dereverberation, spectral subtraction

---

## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Reverberasi adalah *noise* akustik yang muncul pada ruangan tertutup berupa kumpulan pantulan dan difraksi suara oleh dinding dan obyek yang terdapat dalam suatu ruangan. Ketika seorang pembicara berbicara dalam sebuah ruangan, kumpulan *echo* tercampur dalam suara asli dan mengubah karakteristik *spectral* dan *temporal* suara asli. Efek ini dapat dikurangi dengan penggunaan *microphone* di dekat sumber suara, namun hal ini tidak efektif untuk aplikasi komunikasi langsung antara manusia dengan mesin yang bersifat *hands free*, seperti *Automatic Speech Recognition*, dan *Automatic Speaker Verification*.

Salah satu efek penting yang timbul karena reverberasi adalah *overlap-masking*: energi dari *phoneme* melebar sesuai waktu dan mengalami *overlap* dengan energi *phoneme* setelahnya. Hal ini berakibat adanya pengaburan (*blur*) dan *masking* pada karakteristik *spectral* dari *phoneme*.

Penyebab dari pelebaran energi ini adalah kumpulan refleksi dan difusi sinyal suara oleh batas- batas ruangan dan obyek penghalang dalam suatu ruangan. Fenomena yang terjadi adalah energi sinyal tereverberasi melemah secara eksponensial sesuai dengan konstanta waktu yang dipengaruhi oleh karakteristik ruangan. Hal ini terjadi karena pengaruh absorpsi energi oleh udara dan reflektor pada ruangan.

Secara intuisi, perubahan energi *reverberant tail* dari sebuah *phoneme* akan mengalami pelemahan secara eksponensial sesuai dengan *Room Impulse Response* (RIR). Repartisi sepanjang frekuensi dari energi *reverberant* akan tergantung pada repartisi sepanjang frekuensi *phoneme*. Oleh karena itu, maka pelebaran energi sinyal suara menjadi *reverberation tail* dapat secara kasar dimodelkan berdasarkan karakteristik *phoneme* awal dan waktu reverberasi ruangan. Model ini akan dapat digunakan untuk memperkirakan dan mengurangi energi *reverberant* dari sinyal *reverberant*.

Dalam tugas ini akan digunakan metode *Spectral Subtraction* untuk mengurangi efek reverberasi.

## 1.2 Tujuan Penulisan

1. Mensimulasikan efek reverberasi untuk ruangan kecil, sedang, dan besar.
2. Melakukan proses dereverberasi untuk ruangan-ruangan tersebut.
3. Mengukur performansi sistem melalui parameter *Mean Square Error* (MSE), *Early to Late Reverberation Energy Ratio* (ELR), *Early To Total Reverberation Energy Ratio* (ETR), dan *Mean Opinion Score* (MOS).
4. Mengetahui pengaruh koefisien *moving average* terhadap performansi dereverberasi dengan metode *spectral subtraction* untuk jenis suara menyanyi, pidato, dan percakapan

## 1.3 Perumusan masalah

1. Bagaimana memperbaiki kualitas suara yang terdegradasi karena adanya reverberasi (*multiple echoes*)?
2. Bagaimana pengaruh respon impuls dari setiap ruangan yang mempunyai ukuran dan waktu reverberasi yang berbeda?
3. Bagaimana menerapkan metode *spectral subtraction* untuk mendapatkan kualitas sinyal suara tereverberasi yang berkurang efek reverberasinya?

## 1.4 Batasan Masalah

1. Sinyal suara yang diproses adalah sinyal suara manusia yang melakukan jenis pembicaraan: menyanyi, berpidato, dan melakukan percakapan yang direkam dengan menggunakan *sound recorder*.
2. Sinyal suara tereverberasi yang diproses tidak mengandung *noise*.
3. Respon impuls ruangan  $h(n)$  yang digunakan adalah respon impuls ruangan  $h(n)$  yang dibuat simulasinya dengan parameter waktu reverberasi yang disesuaikan dengan referensi.
4. Jenis *window* yg digunakan adalah *window Hamming* dengan panjang 256 poin dengan FFT 256 poin dan *overlap* 75%.
5. Frekuensi sampling yang digunakan adalah 8 kHz.

6. Simulasi  $h(n)$  yang digunakan untuk penelitian ada 3 jenis, yaitu ruangan berukuran kecil, sedang, dan besar yang telah didefinisikan sesuai hasil simulasi
7. Suara asli yang akan diproses didapat dari rekaman.
8. Simulasi menggunakan MATLAB versi R2008a.

### 1.5 Manfaat dan Aplikasi Penelitian

1. Perancangan studio rekaman musik yang baik dengan reverberasi minimal.
2. Memaksimalkan kinerja alat pendengaran tuna rungu dengan aplikasi dereverberasi guna membantu tuna rungu untuk dapat mendengarkan musik yang lebih jelas dan berkurang efek reverberasinya.

### 1.6 Metodologi Penelitian

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

#### 1. Studi Literatur

Studi literatur bertujuan untuk mengumpulkan konsep-konsep yang berhubungan dengan pengolahan sinyal akustik musik dan analisis sinyal akustik musik. Adapun sumber yang didapat mencakup buku referensi, internet dan diskusi.

#### 2. Pengumpulan Data

Pengumpulan data bertujuan untuk mendapatkan beberapa data sinyal musik piano tunggal bersih dari noise yang akan digunakan sebagai masukan dari sistem. Pengumpulan data didapat dari internet.

#### 3. Perancangan Sistem

Perancangan sistem bertujuan untuk menentukan metodologi pengembangan sistem dengan memodelkan sistem yang akan diimplementasikan.

#### 4. Implementasi Sistem

Implementasi system bertujuan untuk melakukan implementasi metode pada sistem sesuai dengan analisis perancangan yang telah dilakukan dengan simulasi yang menggunakan *software* MATLAB R2008a.

#### 5. Analisis Performansi

Analisis performansi meupakan uji coba pada sistem yang telah dibuat dan akan dilakukan analisis data hasil implementasi yang telah dilakukan.

#### 6. Pengambilan Kesimpulan

Setelah data dari hasil implementasi dianalisis maka langkah terakhir adalah mengambil kesimpulan dari hasil analisis.

### 1.7 Sistematika Penulisan

Tugas akhir ini disusun dengan sistematika penulisan yang terdiri dari lima bab yang disusun sebagai berikut :

#### **BAB I       Pendahuluan**

Bab ini terdiri dari latar belakang masalah, tujuan penulisan, perumusan masalah, batasan masalah, manfaat dan aplikasi penelitian, metodologi penelitian dan sistematika penulisan.

#### **BAB II       Dasar Teori**

Bab ini berisi tentang dasar teori yang mendukung dan mendasari penulisan tugas akhir ini, yaitu membahas konsep yang berkaitan dengan sinyal akustik musik.

#### **BAB III     Perancangan dan Simulasi Sistem**

Bab ini menguraikan proses perancangan dalam blok diagram sistem. Disamping itu juga dimasukkan skenario implementasi untuk berbagai kondisi, Pembuatan sistem ini dilakukan dengan menggunakan *software* MATLAB R2008a.

#### **BAB IV Pengujian Sistem dan Analisis Hasil Simulasi**

Bab ini menjelaskan tentang analisis masalah dengan bantuan implementasi sistem dan juga sumber-sumber yang ada. Selain itu juga membahas dan melakukan analisis hasil pengujian sistem tersebut dari perancangan sistem dan simulasi.

#### **BAB V Kesimpulan dan Saran**

Bab ini merumuskan kesimpulan dari perancangan dan analisis yang telah dilakukan dan saran untuk pengembangan penelitian yang lebih lanjut.



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Dari pengujian dan analisis sistem yang telah dilakukan untuk melakukan proses dereverberasi sinyal suara manusia menggunakan metode *spectral subtraction*, dapat diambil kesimpulan sebagai berikut :

1. Simulasi reverberasi dan dereverberasi sinyal suara manusia dengan metode *spectral subtraction* telah berhasil dilakukan untuk ruangan kecil, sedang, dan besar.
2. Pada simulasi *small room*, nilai MSE, ELR, dan ETR mengalami perbaikan, namun tidak terjadi perubahan yang signifikan seiring perubahan nilai  $\beta$ . Perbaikan MSE, ELR, dan ETR paling tinggi masing-masing sebesar 0,0944%, 0,0022%, dan 0,0018% untuk sampel suara perempuan berpidato.
3. Pada simulasi *medium room*, nilai MSE, ELR, dan ETR mengalami fluktuasi nilai perbaikan sesuai dengan nilai  $\beta$ . Perbaikan MSE maksimal terjadi pada sampel suara perempuan berpidato sebesar 11,0213% dengan  $\beta=0,1$ . Perbaikan ELR maksimal terjadi pada sampel suara perempuan berpidato sebesar 2,4028% dengan  $\beta=0,7$ . Perbaikan ETR maksimal terjadi pada sampel suara perempuan berpidato sebesar 2,3858% dengan  $\beta=0,7$ .
4. Pada simulasi *large room*, nilai MSE, ELR, dan ETR mengalami fluktuasi nilai perbaikan sesuai dengan nilai  $\beta$ . Perbaikan MSE maksimal terjadi pada sampel suara perempuan berpidato sebesar 48,5678% dengan  $\beta=0,1$ . Perbaikan ELR maksimal terjadi pada sampel suara perempuan berpidato sebesar 14,1619% dengan  $\beta=0,8$ . Perbaikan ETR maksimal terjadi pada sampel suara perempuan berpidato sebesar 14,0263% dengan  $\beta=0,8$ .
5. Nilai MSE awal sinyal tereverberasi untuk simulasi *medium room* lebih besar 0.5dB bila dibandingkan dengan nilai MSE *medium room*. Sedangkan nilai MSE awal untuk *large room* mengalami peningkatan sebesar 0.1 dB jika dibandingkan dengan nilai MSE awal pada simulasi *medium room*.

## 5.2 Saran

Untuk pengembangan tugas akhir ini selanjutnya, dapat dilakukan dengan cara:

1. Data sampel sinyal suara manusia diolah *frame by frame* sehingga dalam pengolahannya dapat digunakan secara *real-time*.
2. Menggunakan lebih banyak jenis sampel suara.
3. Menambah reverberasi pada sinyal suara manusia untuk menguji ketahanan sistem.



## DAFTAR PUSTAKA

- [1] Allen, Jont dan David Bekley. “*Image Method for Efficiently Simulating Small Room Acoustics*”. Journal of Acoustics Society of America. April 1979.
- [2] Furui, Sadaoki. “*Digital Speech Processing, Synthesis, and Recognition*”. Tokai University Press. 1985.
- [3] Habets, Emanuel Anco Peter. “*Single and Multi Microphone Speech Dereverberation using Spectral Enhancement*”. Eindhoven University Press.2007.
- [4] Savioja, Lauri. “*Modelling Techniques for Virtual Acoustics*”. Helsinki University of Technology. 2000.

