TEKNIK PEMROSESAN SINYAL MULTIMIKROFON MENGGUNAKAN OVERLAP AND ADD SHORT TIME FOURIER TRANSFORM UNTUK MENGURANGI DAMPAK REVERBERASI REKAMAN SINYAL SUARA MANUSIA YANG TERDEGRADASI OLEH REVERBERASI RUANGAN

Era Khaterin Panjaitan[1], Bambang Hidayat[2], Koredianto Usman[3]

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

Abstrak

Reverberasi (multiple echo) adalah sinyal noise akustik yang muncul dalam suatu ruang tertutup akibat adanya superposisi dari suatu pantulan multiple dan difraksi dari dinding dan objek-objek yang ada di dalam ruangan tersebut. Sinyal yang tereverberasi merupakan masalah kompleks yang sulit dicari solusinya, karena sinyal suara asli berkorelasi dengan efek reverberasi.
Efek reverberasi akan menyebabkan terdegradasinya sinyal suara asli dan juga karakteristik spektrumnya. Hal tersebut akan sangat mengganggu persepsi pendengar terhadap informasi yang disampaikan, sehingga dibutuhkan suatu alat bantu berupa sistem yang mampu mengurangi efek reverberasi tersebut.
Dalam Tugas Akhir ini akan dilaksanakan penelitian proses dereverberasi menggunakan metode analisis sintesis Overlap and Add Short Time Fourier Transform (STFT) dengan menggunakan sistem multimicrophone (dua buah mikrofon) dalam berbagai ukuran ruangan (kecil, sedang, dan besar). Sistem ini menggunakan multimicrophone karena sesuai untuk aplikasi hands free telephony dan audio conferencing.
Sistem multimicrophone akan menghasilkan dua suara yang mempunyai karakteristik spectral yang saling berkorelasi. Prinsip kerja sistem adalah membandingkan dua sinyal tereverberasi. Karakteristik spektral dari dua sinyal tereverberasi mempunyai sifat koheren yaitu mempunyai korelasi tinggi untuk sinyal asli dan akan mempunyai korelasi rendah untuk sinyal reverberasi.
Dari hasil penelitian didapatkan bahwa sistem yang menggunakan nilai gain terkecil dan overlap 75% memberikan keluaran sinyal dereverberasi yang paling optimal pada ruangan kecil dan besar dimana pada kondisi ini, sistem memberikan nilai Mean Square Error dan Reverberation Time paling kecil.

Kata Kunci : spectral suara, koheren, reverberasi, korelasi, multimicrophone, STFT

Abstract

Reverberation (multiple echo) is the acoustic noise signals that appear in an enclosed space due to a superposition of multiple reflections and diffraction from walls and objects in the room. Reverberation signal is a complex problem that hard to find a solution, because the original voice signal is correlated with reverberation effects.
Reverberation effect will cause the original voice signal degradation and also the characteristics of the spectrum. This will greatly disturb the audience's perception of the information presented, so it is required a system that can reduce the effects of these reverberation.
In this Final Project will be carried out the research process dereverberation, that use analytical synthesis method Overlap and Add Short Time Fourier Transform (STFT) using multi microphone system (two microphones) in a variety of room sizes (small, medium, and large). This system uses multi microphone as appropriate for applications hands-free telephony and audio conferencing. Multi microphone system will produce two sounds that have spectral characteristics that correlate to each other. Working principle of this system is to compare the two signals that have been given reverberation. Spectral characteristics of the two signals, that have been given reverberation, have coherent characteristic, which have a high correlation to the original signal and will have a low correlation to the reverberation signal.
From this research are expected that system which uses small gain and overlap 75% give optimum dereverberation signal output in small and large room where in this state, system can get small value of Mean Square Error and Reverberation Time.

Keywords : spectral of sound, coherent, reverberation, correlation, multi microfon, STFT

BAB I
PENDAHULUAN

1.1. Latar Belakang

Reverberasi (*multiple echo*) adalah sinyal noise akustik yang muncul dalam suatu ruang tertutup akibat adanya superposisi dari suara pantulan *multiple* dan difraksi dari dinding dan objek-objek yang ada di dalam ruangan tersebut. Ketika ada seseorang berbicara dalam satu ruangan, efek reverberasi tersebut akan berdampak pada terdegradasinya sinyal suara asli dan juga karakteristik spektrumnya. Suara yang terdegradasi karena reverberasi ruangan akan sangat mengganggu persepsi dari pendengar terhadap informasi yang disampaikan. Efek ini dapat dikurangi dengan penggunaan sebuah *microphone* yang diletakkan dengan jarak yang dekat dengan sumber suara. Namun, cara ini tidak sesuai untuk aplikasi *hand free telephony* dan *audio conferencing*. Teknik dereverberasi sangat dibutuhkan untuk mendapatkan kejelasan dan kejernihan suara dari sumber suara yang tereverberasi.

Efek penting yang harus diperhatikan dari reverberasi terhadap suara adalah adanya superposisi *spectral* antara sumber suara asli dengan *spectral* sinyal reverberasi . Energi *spectral* dari sinyal yang tercampur akan mengaburkan *spectral* dari sinyal suara tersebut. Akibatnya suara yang diterima pendengar menjadi kabur, tidak jelas, dan sulit dimengerti.

Sinyal yang tereverberasi merupakan masalah kompleks yang sulit dicari penyelesaiannya karena sinyal suara asli berkorelasi dengan efek reverberasi. Melalui *multimicrophone* (dua buah mikrofon) akan diperoleh dua suara yang mempunyai karakteristik *spectral* yang saling berkorelasi. Koefisien korelasi *spectral* suara akan lebih tinggi dibandingkan dengan level korelasi dari sinyal reverberasinya.

Dalam Tugas Akhir ini akan dilaksanakan penelitian proses dereverberasi menggunakan metode analisis sintesis *Overlap and Add Short Time Fourier Transform* dengan menggunakan faktor koreksi yang diturunkan dari besaran korelasi *spectral* dari dua sumber suara yang berasal dari sistem *multimicrophone*. Sistem ini diharapkan mampu meminimalisasi efek reverberasi ruangan pada sinyal suara, sehingga pendengar mampu memperoleh informasi yang tepat dari sumber.

1.2. Rumusan Masalah

Dari latar belakang yang telah disampaikan sebelumnya, maka dapat dijabarkan beberapa rumusan masalah yang dibahas pada Tugas Akhir ini, yaitu:

1. Bagaimana memperbaiki kualitas suara yang terdegradasi karena adanya dampak reverberasi ruangan (*multiple echo*)?
2. Bagaimana pengaruh dari setiap ruangan (yang mempunyai respons impuls berbeda) yang digunakan sebagai model penelitian dengan ukuran yang berbeda-beda dan materi-materi yang ada di dalamnya akan memberikan efek reverberasi pada sinyal suara asli?
3. Bagaimana karakteristik *Short Time Spectral* sinyal suara terdegradasi dan karakteristik *Short Time Spectral* sinyal suara yang diinginkan?
4. Bagaimana penerapan proses *Multimicrophone Overlap and Add STFT* untuk mendapatkan sinyal suara yang lebih jelas?

1.3. Tujuan

Hasil yang ingin dicapai melalui penelitian yang dilakukan pada Tugas Akhir ini adalah untuk memperoleh sinyal suara yang lebih jelas dari sinyal suara yang terdegradasi akibat reverberasi ruangan melalui metode *Overlap and Add Short Time Fourier Transform (STFT)* dari sinyal suara *multimicrophone*.

1.4. Batasan Masalah

Beberapa batasan masalah pada penelitian Tugas Akhir ini adalah:

1. Untuk pemrosesan, digunakan input rekaman sinyal suara manusia yang melakukan jenis pembicaraan: menyanyi, berpidato, dan melakukan percakapan yang terdiri atas 5 orang laki-laki dan 5 orang perempuan berdurasi 10 detik dan 5 detik, serta direkam dengan menggunakan *sound recorder*.
2. Metode yang digunakan adalah *Multimicrophone Overlap and Add Short Time Fourier Transform.*
3. Format penyimpanan file nada adalah *wave format* (*.wav) dan *mono*.
4. Frekuensi *sampling* yang digunakan adalah 8 kHz.
5. Jumlah mikrofon yang digunakan adalah sebanyak 2 buah yang setara.

6. Sinyal suara yang diproses tidak mengandung noise, hanya reverberasi.
7. Impuls respon ruangan h(n) yang digunakan adalah impuls respon ruangan h(n) yang disimulasikan dengan ukuran ruangan, materi, dan waktu reverberasi yang berbeda.
8. Simulasi h(n) yang digunakan untuk penelitian ada tiga jenis, yaitu ruangan berukuran kecil, sedang, dan besar.
9. Menggunakan *window* dengan tipe *Fixed Length Window*, yaitu: *Hamming window.*
10. Simulasi menggunakan *toolbox* Matlab versi 7.6.

1.5. Metode Penelitian

Beberapa langkah penelitian yang dilakukan untuk mendapatkan hasil yang diharapkan sesuai dengan Tugas Akhir ini adalah:

1. Membuat blok simulasi analisis dan sintesis sinyal suara, untuk mengetahui karakteristik *speech signal* (sinyal suara manusia).
2. Membuat simulasi impuls respons ruangan dengan ukuran dan materi yang berbeda, yang bisa memberikan dampak reverberasi yang signifikan.
3. Menggunakan *speech signal* yang memiliki frekuensi yang dapat didengar oleh manusia, yang akan dikonvolusikan dengan respons impuls ruangan. Keluaran yang diperoleh berupa sinyal yang tereverberasi.
4. Membuat blok simulasi dengan input sinyal suara yang telah tereverberasi akibat impuls respons ruangan yang akan dianalisis dan diproses dengan *Multimicrophone Overlap and Add STFT* untuk menghilangkan dampak reverberasinya yang dilanjutkan dengan proses sintesis STFT untuk mendapatkan sinyal output yang telah dikurangi reverberasinya sampai dengan standar persepsi normal yang telah ditetapkan.
5. Proses evaluasi dari hasil simulasi dengan cara membandingkan sinyal suara asli, sinyal reverberasi, dan sinyal yang diproses dengan perhitungan parameter objektif, yaitu: MSE (*Mean Square Error*) dan RT (*Reverberation Time*), serta evaluasi parameter subjektif, DCR (*Degradation Category Rating*).
6. Pengambilan kesimpulan dan penyusunan Laporan Tugas Akhir.

1.6. Sistematika Penulisan

Secara umum keseluruhan Tugas Akhir ini dibagi menjadi lima bab bahasan. Penjelasannya adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, metode penelitian, rencana kerja, dan sistematika penulisan.

BAB II DASAR TEORI

Bab ini membahas teori mengenai karakteristik sinyal suara, akustik ruangan, reverberasi, STFT, korelasi, dan *Magnitude Squared Coherence* (MSC).

BAB III PERANCANGAN DAN REALISASI SISTEM

Bab ini membahas proses perancangan sistem dereverberasi sinyal suara dengan sistem *multimicrophone* menggunakan metode analisis dan sintesis *Overlap and Add STFT*.

BAB IV HASIL ANALISIS IMPLEMENTASI

Bab ini berisi hasil dari penelitian dan menguraikan analisis dari metode dereverberasi sinyal suara yang digunakan pada Tugas Akhir ini, yaitu *multimicrophone* dengan metode *Overlap and Add STFT* dan menguraikan hasil pengukuran parameter objektif dan penilaian parameter subjektif.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil Tugas Akhir dan saran untuk pengembangan-pengembangan lebih lanjut.

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil percobaan dan analisis yang diperoleh, maka dari Tugas Akhir yang berjudul Teknik Pemrosesan Sinyal Multimikrofon Menggunakan *Overlap and Add Short Time Fourier Transform* untuk mengurangi dampak reverberasi rekaman sinyal suara manusia yang terdegradasi oleh reverberasi ruangan ini dapat disimpulkan beberapa hal, antara lain:

1. Berdasarkan hasil analisis nilai rata-rata MSE pada tabel 4.1 dan nilai rata-rata RT pada table 4.2, sistem yang dirancang akan memberikan keluaran yang paling baik untuk jenis masukan:
   - suara laki-laki menyanyi dan suara perempuan berpidato saat diimplementasikan pada ruangan berukuran kecil.
   - suara laki-laki berpidato dan percakapan saat diimplementasikan pada ruangan berukuran besar.
2. Dalam analisis pengaruh overlap terhadap kualitas sinyal hasil dereverberasi, semakin besar *overlap* yang diberikan maka semakin baik kualitas dereverberasi yang diperoleh dengan rentang overlap yang efektif adalah 50%-87,5%. *Overlap paling optimal adalah 87,5%, tetapi* untuk memperoleh *trade-off* antara kualitas dan lamanya waktu yang diperlukan untuk menjalankan program, maka *overlap* yang digunakan pada sistem ini adalah *overlap* sebesar 75%.
3. Sinyal suara manusia akan memberikan performansi semakin baik jika gain yang digunakan semakin kecil karena gain hasil modifikasi dari *coherence* melakukan koreksi terhadap efek reverberasi yang terdapat pada sinyal suara.
5. Berdasarkan hasil pengukuran subjektif pada bab 4, sistem bekerja paling baik pada ruangan berukuran sedang dengan nilai rata-rata DCR adalah 3,4 yang artinya suara hasil dereverberasi masih dapat didengar dengan jelas tetapi mengalami degradasi yang sedikit mengganggu. Menurut 30 orang koresponden, suara hasil dereverberasi yang paling baik adalah pada sistem yang menggunakan *overlap* 75% karena pada sistem yang menggunakan *overlap* 25% dan 50% terjadi aliasing sehingga suara output lebih cepat daripada suara aslinya dengan nilai DCR adalah 4. Sistem dengan variasi gain juga memberikan hasil yang optimal pada gain terkecil dengan nilai DCR adalah: 4.

5.2 Saran

Ada beberapa hal yang dapat dijadikan saran untuk perkembangan Tugas Akhir ini, di antaranya:

1. Proses perekaman dilakukan dengan sebaik mungkin sehingga dihasilkan performansi yang terbaik. Baik di sini maksudnya adalah proses perekaman menggunakan kabel, dilakukan di tempat diruang *unechoic* (sepi dan kedap suara).
2. Efek ruangan (RIR) yang diberikan dapat dikembangkan menjadi efek ruangan yang sebenarnya (bukan buatan) yang bisa diperoleh melalui hasil pengukuran langsung.
3. Metode analisis dan sintesis yang digunakan dapat dikembangkan menjadi metode yang *adaptive* terhadap pitch suara manusia.
4. Gain yang diberikan dapat dikembangkan menjadi gain yang dilengkapi dengan *voice non voice silence detection* dan memiliki threshold tertentu.
5. Metode yang digunakan dapat dikembangkan menjadi metode transformasi lain, seperti: wavelet. Metode wavelet lebih baik daripada metode yang digunakan dalam tugas akhir ini (STFT) karena pada metode wavelet, window yang digunakan bersifat variabel dan *adaptive*. Pada metode STFT, window yang digunakan adalah *fixed length window*.
6. Agar *processing* lebih cepat, sistem diimplementasikan dengan *hardware* yang *processing*-nya lebih cepat (*array processor*) atau menggunakan Xeon.

DAFTAR PUSTAKA

[1] Bishnu S. Atal, Vladimir Cuperman, and Geisha A., 1993. *Speech and Audio Coding for Wireless and Network Applications.* Kluwer academic publisher: Boston/Dordrecth/London

[2] Jae S. Lim and Oppenheim Allan, 1988, *Advanced Topic in Signal Processing.* Englewood Cliffs, New Jersey.
A. Akbari Azirani, R. Le Bouquin, and G. Faucon, "Enhancement of Speech Degraded by Coherent and Incoherent Noise Using A Cross-Spectral Estimator." *IEEE Trans. Speech, and Audio Processing,* vol. 5, no. 5, pp. 484-487, 1997.

[3] Leigh D. Alsteris, Paliwal, K. K, 2006. *Short Time Phase Spectrum in Speech Processing : A Review and Some Experimental Result.* School of Microelectronic Engineering, Griffith University, Brisbane, QLD 4111, Australia

[4] Sieben, G.W., Gold, M.A., 2000. *Ten Ways to Provide a High Quality Acoustical Environtment in School. Journal Acoustic,* Vol. 31. P. 376-384.

[5] Ternstr ¨om, S, 1989. *Long-time Average Spectrum Characteristics of Different Choirs in Different Rooms*. *Journal STL-QPSR,* Vol.30. P. 015-031.