Tugas Pengganti Ujian Tengah Semester

Filter Wiener untuk Pengurangan Noise dan Speech Enhancement

Group 1

Muhammad Aidan Daffa Junaidi- 1906300800

Fatma Putri Ramadhani - 1906381735

Dept. of Electrical Engineering, Computer Engineering Study Program

Faculty of Engineering, Universitas Indonesia

Kampus Baru UI, Depok 16424

Email:

[muhammad.aidan@ui.ac.id](mailto:muhammad.aidan@ui.ac.id)

[fatma.putri@ui.ac.id](mailto:fatma.putri@ui.ac.id)

***Abstract*— Speech enhancement merupakan hal yang krusial di masa sekarang. karena di masa sekarang banyak bentuk komunikasi dan persebaran informasi yang dilakukan secara online. Namun, tidak semua orang memiliki ruang pribadi untuk berkomunikasi digital. Oleh karena itu, kebisingan latar belakang dari kondisi dalam ruangan dapat mengganggu komunikasi selama perekaman. Denoising speech memiliki banyak manfaat misalnya dalam komunikasi suara atau pengenalan suara di mana diperlukan proses denoising yang cepat. Makalah ini akan membahas penggunaan Wiener filter pada speech compression dan speech enhancement berdasarkan paper yang diambil dari IEEE.**

***Kata kunci*—Makalah, Wiener, filter, speech compression, speech enhancement.**

# Pendahuluan

Data transmisi secara digital menggunakan kecepatan yang tinggi, bahkan hingga 2.5 Tb/s. Meskipun begitu, kebutuhan transmisi suara secara digital masih tetap tinggi. Karena kebutuhan yang tinggi, maka akan terjadi kekurangan kapasitas *bandwidth* yang tinggi sehingga menyebabkan distorsi sinyal. Oleh sebab itu, peningkatan sinyal suara masih sangat diperlukan.[1]

Untuk mengontrol *noise*, dapat menggunakan pendekatan dengan *time-domain* Wiener Filter. Metode yang digunakan adalah dengan mengontrol jumlah sisa *noise* dan juga dengan menggabungkan komponen, seperti *noise reduction* dan *echo cancellation*.[2]

Metode lain yang diajukan pada tahun 2006 adalah *Two-Step Noise Reduction* (TSNR), yaitu dengan menekan dan menghapus efek gema pada suara dan menjaga suara yang diperlukan.

# Studi Literatur

Seperti yang telah dijelaskan pada bagian sebelumnya, ada beberapa artikel dan tes yang menjadi referensi dan juga memungkinkan Anda untuk mengikuti tes ini. Sebelum itu, Anda perlu mengetahui apa sebenarnya kompresi audio atau pengkodean audio itu. Kompresi audio adalah proses kompres atau kompresi ukuran file yang dibuat dengan merekam suara manusia dengan sejumlah besar data. Kompresi/encoding audio diperlukan untuk mengurangi ukuran file dengan menghapus redundansi dari audio. Cara paling umum untuk menghilangkan redundansi bahasa adalah pengkodean transformasi. Saat ini dimungkinkan untuk menggunakan beberapa teknik konversi untuk mengompres sinyal audio.

1. Speech enhancement using long short term memory with trained speech features and adaptive wiener filter

Speech enhancement adalah proses meningkatkan kejelasan sinyal speech yang terdegradasi oleh kebisingan latar belakang. Dalam penelitian ini, model peningkatan sinyal suara baru diperkenalkan menggunakan deep learning. Model yang diusulkan dibagi menjadi dua fase: (i) training (ii) testing. Pada fase training, spektrum noise dan spektrum sinyal diperkirakan menggunakan faktor matriks non-negatif (NMF) dari sinyal input noise. Kemudian fitur Empiris Mean Decomposition (EMD) diekstraksi dari filter Wiener. Sinyal denoise diperoleh dari EMD, frekuensi shell diperkirakan dan fitur delta fraksional AMS diekstraksi. Kontribusi utama dari penelitian ini adalah penggunaan model Long Short Term Memory (LSTM) untuk memperkirakan secara akurat koefisien tuning filter Wiener η untuk semua sinyal input. Extracted Features (EMD) melatih LSTM untuk mendekode sinyal spektral, dan keluaran EMD adalah sinyal ucapan yang ditingkatkan noise.[3]

Penelitian ini bertujuan untuk meningkatkan sinyal speech dengan berbagai sumber kebisingan. Bagian hasil mengevaluasi model yang diusulkan dengan sumber kebisingan yang berbeda seperti "kebisingan bandara, kebisingan pameran, kebisingan restoran, kebisingan stasiun, dan kebisingan jalanan". Berbagai sumber kebisingan dianalisis dengan nilai SNR yang berbeda dalam hal pengukuran kualitas ucapan. Dengan memanfaatkan filter wiener yang dimodifikasi dan model LSTM berbantuan fitur yang diekstraksi, sinyal denoise speech diperoleh. Dibandingkan dengan model yang ada, metode yang diusulkan lebih bagus performa nya dibanding SDR, PESQ, CORR, ESTOI, STOI, dan SNR, serta nilai RMSE yang lebih rendah. Selain itu, model yang diusulkan mengatasi kekurangan seperti pengurangan kejelasan ucapan, PESQ yang lebih rendah, ketahanan yang lebih rendah, tidak cocok untuk lingkungan kebisingan yang kompleks, kualitas ucapan yang lebih rendah, dan SNR yang rendah. Namun, metode yang diusulkan ini memiliki kekurangan pada beberapa sumber kebisingan dan tidak menentukan magnitudo spektral dan estimasi fase spektral.[3]

1. Enhancement of single channel speech quality and intelligibility in multiple noise conditions using wiener filter and deep CNN

Deep Neural Network telah menjadi metode yang sangat baik untuk meningkatkan sinyal speech karena memberikan hasil yang baik dibandingkan dengan metode tradisional. Dalam penelitian ini menjelaskan transformasi dalam sinyal speech yang ditingkatkan dengan menerapkan Deep Convolutional Neural Network (Deep CNN) untuk memodelkan hubungan non-linier dan membandingkannya dengan metode filter Wiener, yang merupakan teknik terbaik untuk peningkatan speech diantara metode tradisional. Denoising dilakukan dalam domain frekuensi dan dikonversi kembali ke domain waktu untuk menganalisis metrik kinerja seperti kualitas ucapan dan kejelasan ucapan. Kualitas ucapan dianalisis menggunakan signal-to-noise ratio (SNR) dan perceptual assessment of speech quality (PESQ). Kejelasan speech dianalisis menggunakan short-term objective intelligibility (STOI). Kedua metode mengevaluasi denoised speech, dan analisis hasil menunjukkan bahwa SNR metode filter Wiener tradisional jauh lebih baik dibandingkan dengan Deep CNN. Namun, bahasa berbasis PESQ dan STOI Deep CNN melakukan prosedur penyaringan Wiener. [4]

Evaluasi dilakukan dalam kondisi kebisingan yang berbeda untuk menganalisis kinerja sistem peningkatan speech, dan hasilnya menunjukkan bahwa sistem berbasis Deep CNN mengungguli teknik filter Wiener tradisional dengan kinerja terbaik dalam hal kualitas dan kejelasan. Kualitas sinyal denoise speech berdasarkan SNR menunjukkan peningkatan yang kuat dengan sinyal denoise speech yang di filter Wiener. Namun, Deep CNN memberikan hasil yang sangat baik dalam hal kualitas dan kejelasan yang dianalisis menggunakan skor PESQ dan STOI. Oleh karena itu, perlu dicatat bahwa kinerja Deep CNN mengungguli teknik filter Wiener tradisional.[4]

# Metodologi

1. Wiener Filter

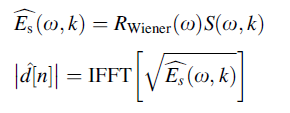
Wiener Filter bertujuan untuk menghitung statistik dari sinyal yang tidak diketahui dengan mengambil sinyal yang mirip sebagai *input* dan melakukan *filter* untuk membuat perkiraan sebagai *output*-nya. Wiener Filter didasarkan pada perhitungan dan pemisahan sinyal dari suara yang jelas dari sinyal dengan *noise* yang terdistorsi.

Karena Wiener Filter melakukan *noise reduction*, hal ini dapat berpengaruh pada integritas sinyal suara. Maka dari itu, perlu dilakukan perubahan pada Wiener Filter sehingga sinyal suara dapat dibedakan dengan jelas.

C(⍵) = spektrum sinyal

B(⍵) = spektrum *noise power*

S(⍵) = spektrum dari sinyal suara yang memiliki *noise*



Es = Estimasi sinyal yang sudah ditingkatkan

Dengan menggabungkan besarnya data spektrum dari suara yang jelas dengan fase dari suara yang masih memiliki *noise,* maka estimasi dari suara yang sudah ditingkatkan dapat dicapai.



ď = Estimasi suara yang sudah ditingkatkan

1. Two Step Noise Reduction

Two Step Noise Reduction merupakan teknik yang memiliki kemampuan untuk segera melacak ketidakstasioneran sinyal speech tanpa menimbulkan efek denoised. Selain itu, dalam aplikasi pengenalan speech otomatis, algoritma TSNR menunjukkan pengurangan kesalahan substitusi dan penyisipan yang signifikan yang mengarah ke peningkatan kinerja pengenalan relatif yang substansial. [5]

1. Harmonic Regeneration Noise Reduction

Ketika SNR terlalu kecil, maka masih terdapat distorsi pada harmoni suara. Maka dari itu, dibuatlah *Harmonic Regeneration* dimana sinyal yang terdistorsi akan dibuat sinyal harmoniknya menggunakan fungsi non-linear. Kemudian, sinyal ini digunakan untuk menghitung supresi sehingga dapat menjaga harmoni suara.

1. Arsitektur Sistem

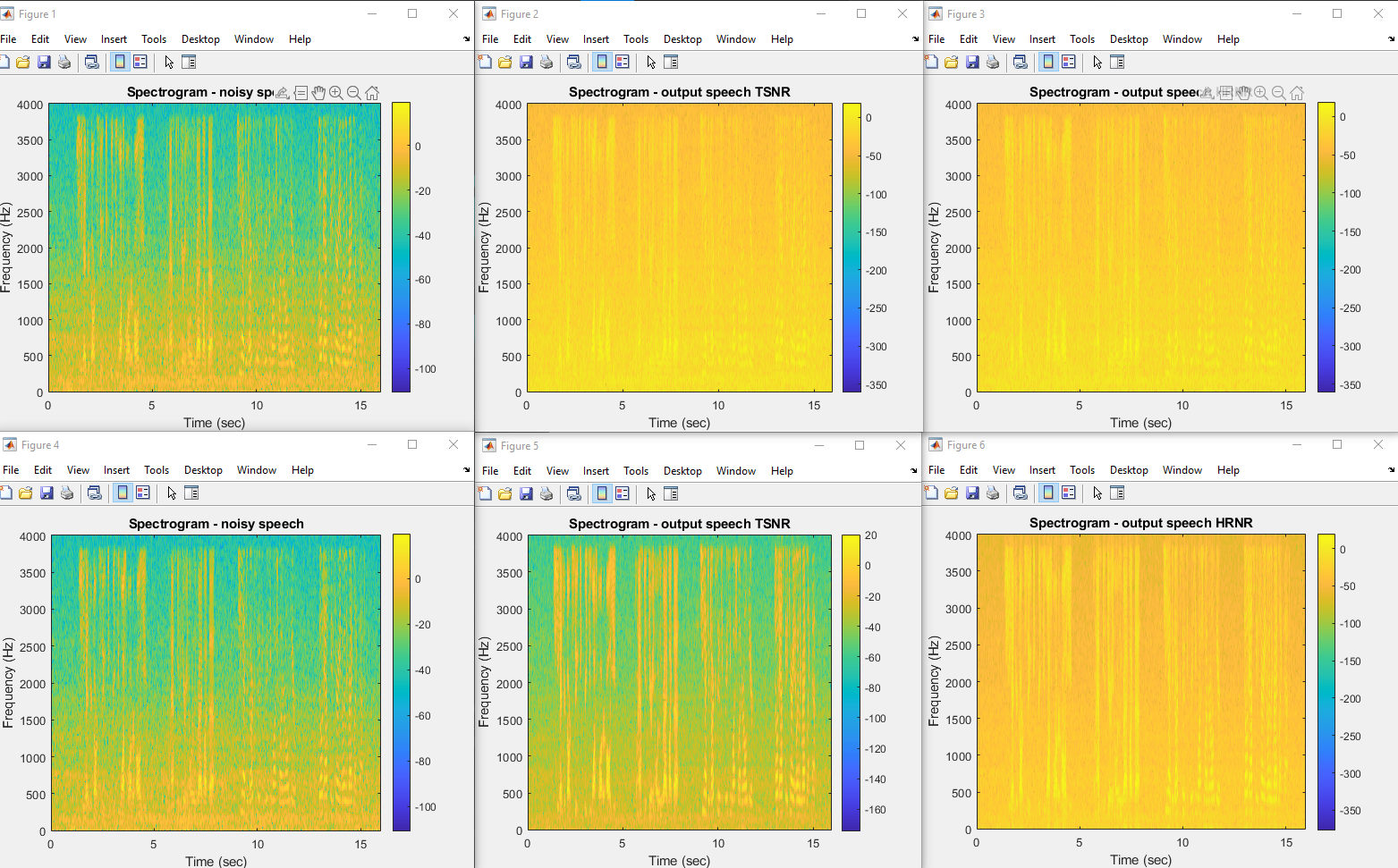
Pada penelitian kali ini, *sample* suara yang digunakan adalah suara seseorang berbicara dengan *sampling rate* 4kHz. Pada *background* suara, terdapat suara mobil sebagai *noise*-nya.

Penelitian dilakukan dengan menerapkan Wiener Filter dengan metode Two Step Noise Reduction (TSNR) dan Harmonic Regeneration Noise Reduction (HRNR). Parameter yang menjadi *input* dari fungsi adalah ns, yaitu suara yang masih memiliki *noise*, fs, yaitu *sampling frequency* dalam Hz, dan IS, yaitu banyaknya *sampel* dari periode aktivitas yang sedang tidak berbicara. IS yang digunakan pada penelitian ini adalah 100 dan 10.000 untuk membandingkan seberapa baik *noise reduction* yang dihasilkan berdasarkan perbedaan IS.

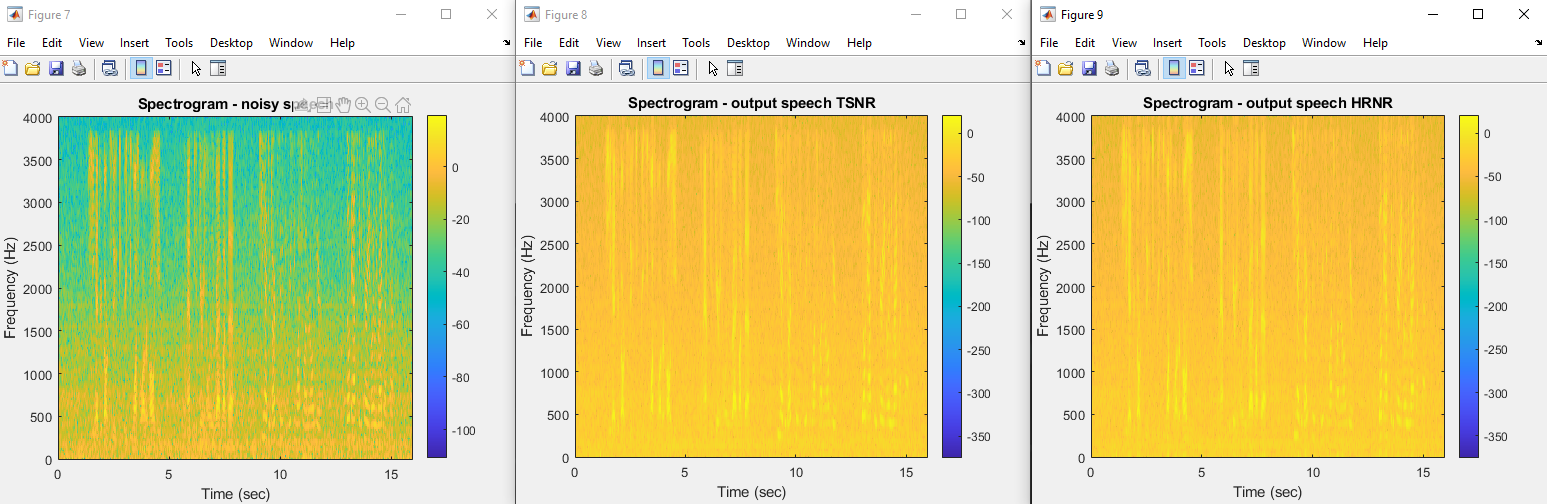
Output dari fungsi ini adalah esTSNR, yaitu suara yang sudah ditingkatkan dengan metode Two Step Noise Reduction, dan esHRNR, yaitu suara yang sudah ditingkatkan dengan metode Harmonic Regeneration Noise Reduction Method.

# Pembahasan

Percobaan pertama adalah melihat hasil spektrum dari Wiener Filter dengan metode Two Step Noise Reduction dan Harmonic Regeneration Noise Reduction dengan IS 100.



Percobaan kedua adalah melihat hasil spektrum dari Wiener Filter dengan metode Two Step Noise Reduction dan Harmonic Regeneration Noise Reduction dengan IS 100.000.



Berdasarkan hasil yang didapatkan, Wiener Filter dengan IS 100 masih memiliki suara *noise* yang dapat terdengar. Sementara Wiener Filter dengan IS 100.000, hasil suara menjadi lebih jelas dan suara *noise* hampir tidak terdengar.

# kesimpulan

Pertemuan secara digital dan komunikasi multimedia telah banyak digunakan oleh banyak orang. namun masih banyak platform yang kurang membantu menghapus suara noise dari sinyal speech yg digunakan. makalah ini menggunakan wiener filtering dalam usaha untuk menghapus noise pada sebuah suara. dari paper yang diacu pada referensi nomor 1 didapat bahwa baik TSNR dan HRNR menunjukkan hasil yang baik dalam pengurangan noise. Kedua teknik tersebut berhasil meredam kebisingan pada saat tingkat kebisingan sangat tinggi. Namun, untuk semua tingkat kebisingan, TSNR menunjukkan hasil pengurangan kebisingan yang sedikit lebih baik daripada HRNR. Berdasarkan kecenderungan di atas, untuk tingkat kebisingan rendah atau sangat rendah, TSNR dan HRNR tampaknya tidak efektif, karena pengurangan kebisingan tidak signifikan dibandingkan dengan tingkat kebisingan itu sendiri.

# References

[1] Hilal H. Nuha, Ahmad Abo Absa, “Noise Reduction and Speech Enhancement Using Wiener Filter”, in *International Conference on Data Science and Its Applications, ICoDSA 2022* Bandung, Indonesia, 2022.

[2] Sebastian Braun, Konrad Kowalczyk, Emanuël A. P. Habets, “Residual noise control using a parametric multichannel Wiener filter”, 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, QLD, Australia, 2015.

[3] Anil Garg, “Speech enhancement using long short term memory with trained speech features and adaptive wiener filter”,doi.org/10.1007/s11042-022-13302-3, 2022.

[4] D. Hepsiba, Judith Justin, “Enhancement of single channel speech quality and intelligibility n multiple noise conditions using wiener filter and deep CNN”, doi.org/10.1007/s00500-021-06291-2 ,Germany, 2021.

[5] Cyril Plapous, Claude Marro, Laurent Mauuary, Pascal Scalart,”A TWO-STEP NOISE REDUCTION TECHNIQUE”, Conference Paper in Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on · June 2004

[6] Liu Zengyuan, Dong Anming, “A Speech Denoising Algorithm Based on Harmonic Regeneration ”, IOP Conference Series: Earth and Environmental Science, doi:10.1088/1755-1315/332/2/022042, 2019.

# 