

Multiband Spectral Subtraction dengan Menggunakan Gaussian Window untuk Meningkatkan Kualitas Sinyal Ucapan Berderau

Fitrilina¹, M. Idil Fitra¹, Dwi Harinitha²

¹Jurusan Teknik Elektro, Fakultas Teknik, Universitas Andalas

²Program Studi Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Padang

Informasi Artikel

Naskah Diterima : 26 Mei 2019

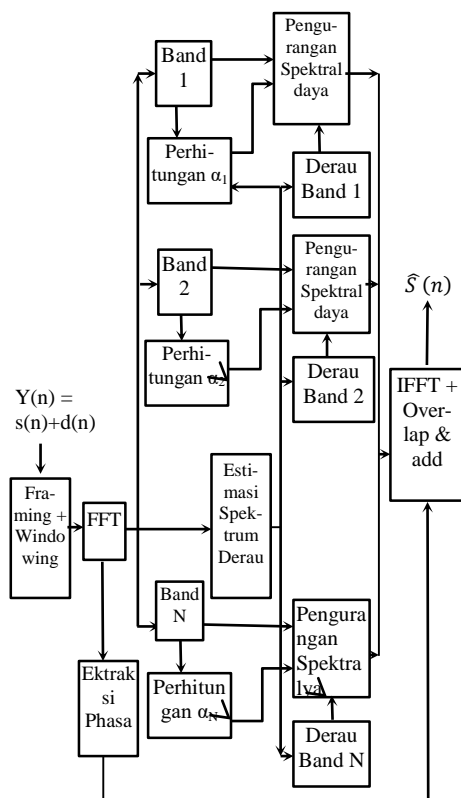
Direvisi : 29 Mei 2019

Disetujui : 14 Juni 2019

*Korespondensi Penulis :

fitrilina@eng.unand.ac.id

Graphical abstract



Abstract

The presence of noise damages the quality and intelligibility of speech signals, thereby reduces the performance of speech based applications. Noise is unavoidable and unpredictable. The most common type of noise found in the environment is non stationary noise. Non-stationary noise does not affect the speech signal uniformly over whole frequencies. Therefore conventional spectral subtraction methods are not effective. This paper, propose a Multiband spectral subtraction method which is divides the noisy speech signal into several frequency bands uniformly and reduce spectral noise in each band independently. The number of frequency bands tested are 4, 8 and 12 bands. The window type affects spectral estimates. Therefore this method is tested for three types of windows, hamming, sine multitapper with number of taper 4 and Gaussian. The performance of the proposed method is assessed based on the PESQ of the enhanced speech signal. The Gaussian window with side lobe attenuation 3.75 had the best PESQ in all number of frequency bands compared to hamming and sine multitaper. The best average PESQ is obtained using Multiband Spectral subtraction using a Gaussian window with side lobe attenuation 3.75 and number of bands 4.

Keywords: Multiband Spectral Subtraction, Gaussian Window dan PESQ

Abstrak

Keberadaan derau merusak kualitas dan kejelasan sinyal ucapan sehingga mengurangi kinerja aplikasi yang berbasis ucapan. Derau tidak dapat dihindari dan tidak terprediksi. Jenis derau yang paling umum ditemukan pada lingkungan adalah derau non stasioner. Derau non stasioner memberi pengaruh secara tidak seragam pada semua frekuensi sinyal ucapan. Oleh karena itu metoda pengurangan spektral yang konvensional menjadi tidak efektif. Pada penelitian ini diusulkan Multiband spectral subtraction. Metoda ini membagi sinyal menjadi beberapa band frekuensi secara seragam dan melakukan pengurangan spektral derau secara saling bebas pada setiap band. Jumlah band frekuensi yang diujikan yaitu 4, 8 dan 12 band. Jenis window mempengaruhi estimasi spektral. Oleh karena itu ini diujikan pada tiga jenis window yaitu hamming, sine multitapper dengan jumlah taper 4, dan gaussian window. Kinerja yang diusulkan dinilai berdasarkan nilai PESQ dari sinyal ucapan hasil perbaikan. Berdasarkan hasil pengujian didapatkan bahwa penggunaan gaussian window dengan redaman side lobe 3.75 memiliki nilai PESQ terbaik disemua jumlah band frekuensi dibandingkan hamming dan sine multitapper. Nilai rata-rata PESQ terbaik diperoleh dengan menggunakan Multiband Spectral subtraction menggunakan gaussian window dengan redaman side lobe 3.75 dan jumlah band 4.

Kata kunci: Multiband Spectral Subtraction, Gaussian Window and PESQ

© 2019 Penerbit Jurusan Teknik Elektro UNTIRTA Press. All rights reserved

1. PENDAHULUAN

Derau merupakan sinyal pengganggu yang tidak diinginkan karena dapat merusak sinyal ucapan. Keberadaan derau lingkungan tidak dapat diprediksi dan tidak dapat dihindari. Derau yang

bercampur dengan sinyal ucapan akan membuat kejelasan informasi dalam sinyal tersebut menjadi berkurang bahkan hilang. Oleh karena itu, dibutuhkan suatu metoda untuk menghilangkan atau mengurangi derau dalam sinyal ucapan sehingga kejelasan dan kualitas sinyal ucapan dapat diperbaiki. Perbaikan sinyal ucapan berderau sangat dibutuhkan untuk mendukung aplikasi seperti sistem pengenalan ucapan, sistem pengkodean ucapan, sistem telekonferensi, dan lain-lain. Pada penelitian ini digunakan *spectral subtraction* karena ini lebih sederhana, mudah diimplementasikan serta terbukti efektif mengurangi derau [1]. Prinsip dasar *spectral subtraction* adalah mengurangi *magnitude* atau daya sinyal derau dari sinyal ucapan berderau. Derau diasumsikan tidak berkorelasi dan bersifat aditif pada sinyal ucapan.

Permasalahan utama dalam *spectral subtraction* adalah bagaimana mengatasi derau yang bersifat non stasioner dan munculnya derau baru yang disebut derau musikal. *Spectral subtraction* yang diusulkan Boll [2] merupakan teknik yang cukup populer dalam mereduksi derau karena sederhana dan cukup mampu memperbaiki sinyal yang mendapatkan derau tambahan. Tetapi Boll belum mampu mengatasi munculnya derau musikal [3]. Penelitian [4] menggunakan Berauti *spectral subtraction* yang berusaha mengurangi derau musikal dengan menggunakan *factor oversubtraction* dan *spectral floor*. Berauti melakukan pengurangan *spectral* derau secara seragam pada semua frekuensi sehingga kurang efektif mengatasi derau yang bersifat non stasioner. Derau non stasioner merupakan derau yang memiliki spektral daya yang berubah-ubah sehingga akan memberikan pengaruh yang berbeda pada semua rentang frekuensi. Derau non stasioner merupakan jenis derau yang paling banyak ditemukan pada lingkungan tempat sistem bekerja.

Oleh karena sinyal derau tidak mempengaruhi sinyal ucapan secara seragam disemua rentang frekuensi, beberapa frekuensi bisa mendapat pengaruh yang lebih buruk dibandingkan frekuensi yang lain. Sebagai contoh, derau sinyal percakapan memiliki energi ucapan yang lebih besar pada frekuensi rendah, sehingga pada frekuensi rendah akan memberi pengaruh lebih kuat dibandingkan frekuensi tinggi. Oleh karena itu pada penelitian ini digunakan *Multiband Spectral Subtraction* (MBSS) yang melakukan pengurangan *spectral* derau secara tidak seragam pada setiap *band* frekuensi sinyal ucapan.

MBSS telah digunakan pada penelitian [3, 5]. Penelitian [3] menggunakan *hamming window* sedangkan penelitian [5] membandingkan penggunaan *hamming window* dengan *sine multitapper*. Penelitian [5] menyatakan bahwa *hamming window* merupakan jenis *window* yang banyak digunakan akan tetapi memiliki varians yang tinggi sehingga dapat menurunkan kualitas ucapan. Inilah yang menyebabkan penelitian [5] menggunakan *sine multitapper*. Tetapi hasil penelitian ini hanya menganalisa sinyal hasil perbaikan dalam bentuk spektrogram tanpa ada parameter yang terukur. Penelitian [6] menggunakan *Boll's Spectral Subtraction* dan membandingkan penggunaan *hamming window* dengan *Gaussian window*. Penelitian [6] mendapatkan bahwa penggunaan *Gaussian window* memberikan peningkatan nilai rasio sinyal terhadap derau (SNR) yang lebih baik dibandingkan dengan *hamming window*.

Berdasarkan penelitian seperti paparan di atas, maka pada penelitian ini dilakukan perbandingan tiga jenis *window* yaitu *Gaussian window*, *hamming window*, dan *sine multi tapper* dengan menggunakan *Multiband Spectral Subtraction* (MBSS). Kinerja dari metode *spectral subtraction* yang diusulkan dilihat berdasarkan nilai *Perceptual Evaluation of Speech Quality* (PESQ) dari sinyal hasil perbaikan.

2. METODE PENELITIAN

Metoda *Multiband Spectral Subtraction* yang diusulkan dalam penelitian ini digambarkan dengan diagram blok pada gambar 1 dan diimplementasikan dengan langkah-langkah sebagai berikut:

- 1) Pada sinyal suara berderau yang dapat dinyatakan seperti persamaan (1) dilakukan *Frame Blocking* dengan panjang *frame* 25 ms dan *overlap* 10 ms.

$$y(n) = s(n) + d(n), n \in (0, N - 1) \quad (1)$$

Dimana $y(n)$, $s(n)$ dan $d(n)$ secara berurutan adalah sinyal ucapan berderau, ucapan bersih dan derau. Sedangkan notasi n adalah indeks waktu diskrit dan N adalah jumlah sampel pada sinyal. Karena sinyal ucapan bersifat non stasioner dan agar dapat diproses dengan asumsi stasioner maka dilakukan proses *framing* yaitu membagi sinyal dalam durasi waktu yang singkat (*frame*) dan dilanjutkan dengan proses *windowing*.

- 2) *Windowing* dengan menggunakan 3 jenis *window* yaitu *hamming*, *sine multitaper* dengan nilai $k=4$, serta *Gaussian* dengan nilai redaman *side lode* 0.75, 1.5, 2.25, 3 dan 3.75. *Hamming Window*, *Sine Multitapper*, dan *Gaussian Window* dinyatakan dengan persamaan (2) – (4).

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right); 0 \leq n \leq N - 1 \quad (2)$$

$$\alpha_k(n) = \sin \frac{\pi k(n+1)}{N+1} \sqrt{\frac{2}{N+1}} \quad 0 \leq n \leq N - 1 \quad (3)$$

$$w_{gaussian}(n) = e^{-\frac{1}{2}\left(\alpha_{gw} \frac{n}{(N/2)}\right)^2}, 0 \leq |n| \leq \frac{N}{2} \quad (4)$$

Dimana N adalah jumlah sampel dalam satu *frame*, n adalah indeks sample, α_k adalah *sine tapper* yang ke- k , dan α_{gw} adalah redaman *side lobe Gaussian Window*.

- 3) Transformasi Fourier diskrit dengan menggunakan FFT mengubah sinyal domain waktu ke domain frekuensi, seperti pada persamaan (5).

$$Y(m, k) = S(m, k) + D(m, k) \quad (5)$$

Dimana $Y(m,k)$, $S(m,k)$ dan $D(m,k)$ adalah representasi $y(n)$, $s(n)$ dan $d(n)$ pada domain frekuensi saat *frame* m dan indeks frekuensi k . Persamaan (5) dapat dinyatakan dalam domain spektral daya seperti persamaan (6) yang menunjukkan bahwa jika spektrum derau dapat diestimasi, maka spektrum sinyal ucapan bersih juga akan dapat diestimasi dengan cara mengurangi spektrum derau dari spektrum sinyal ucapan berderau. Sehingga pada penelitian [2], *spectral subtraction* dapat dinyatakan seperti persamaan (7).

$$|Y(m, k)|^2 = |S(m, k)|^2 + |D(m, k)|^2 \quad (6)$$

$$|\hat{S}(m, k)|^2 = |Y(m, k)|^2 - |\hat{D}(m, k)|^2 \quad (7)$$

Dimana $|\hat{S}(m, k)|^2$ adalah estimasi spektrum daya sinyal bersih (sinyal yang telah diperbaiki) dan $|\hat{D}(m, k)|^2$ adalah estimasi spektrum daya derau.

- 4) Mengestimasi spektrum daya derau berdasarkan informasi *magnitude* sinyal berderau. Pada proses ini menggunakan *Voice Activity Detection* (VAD) untuk menentukan *frame* derau. Spektrum derau diestimasi dengan merata-ratakan spektrum daya pada *frame* derau.
- 5) Setiap *frame* dibagi menjadi N *sub band* secara seragam. Pada penelitian ini divariasikan 4, 8 dan 12 *sub band*. Hal ini dilakukan untuk mengetahui jumlah *sub band* yang terbaik.
- 6) Setiap *sub band* dihitung nilai *over subtraction* dengan menggunakan persamaan (8) – (9).

$$SNR_i(dB) = 10 \log_{10} \frac{\sum_{k=b_i}^{e_i} |Y_i(k)|^2}{\sum_{k=b_i}^{e_i} |\hat{D}_i(k)|^2} \quad (8)$$

$$\alpha_i = \begin{cases} 5 & ; SNR_i < -5 \text{ dB} \\ 4 - \frac{3}{20} (SNR_i) & ; -5 \leq SNR_i \leq 20 \text{ dB} \\ 1 & ; SNR_i > 20 \text{ dB} \end{cases} \quad (9)$$

Dimana b_i serta e_i adalah indeks frekuensi awal dan akhir pada *band* frekuensi ke- i , dan α_i adalah faktor *over subtraction* pada *band* ke- i .

- 7) Melakukan pengurangan Spektral untuk mendapatkan estimasi sinyal bersih (sinyal hasil perbaikan). Pengurangan Spektral dilakukan dengan menggunakan persamaan (10) – (12). Nilai *spectral floor* (β) yang digunakan adalah 0.003.

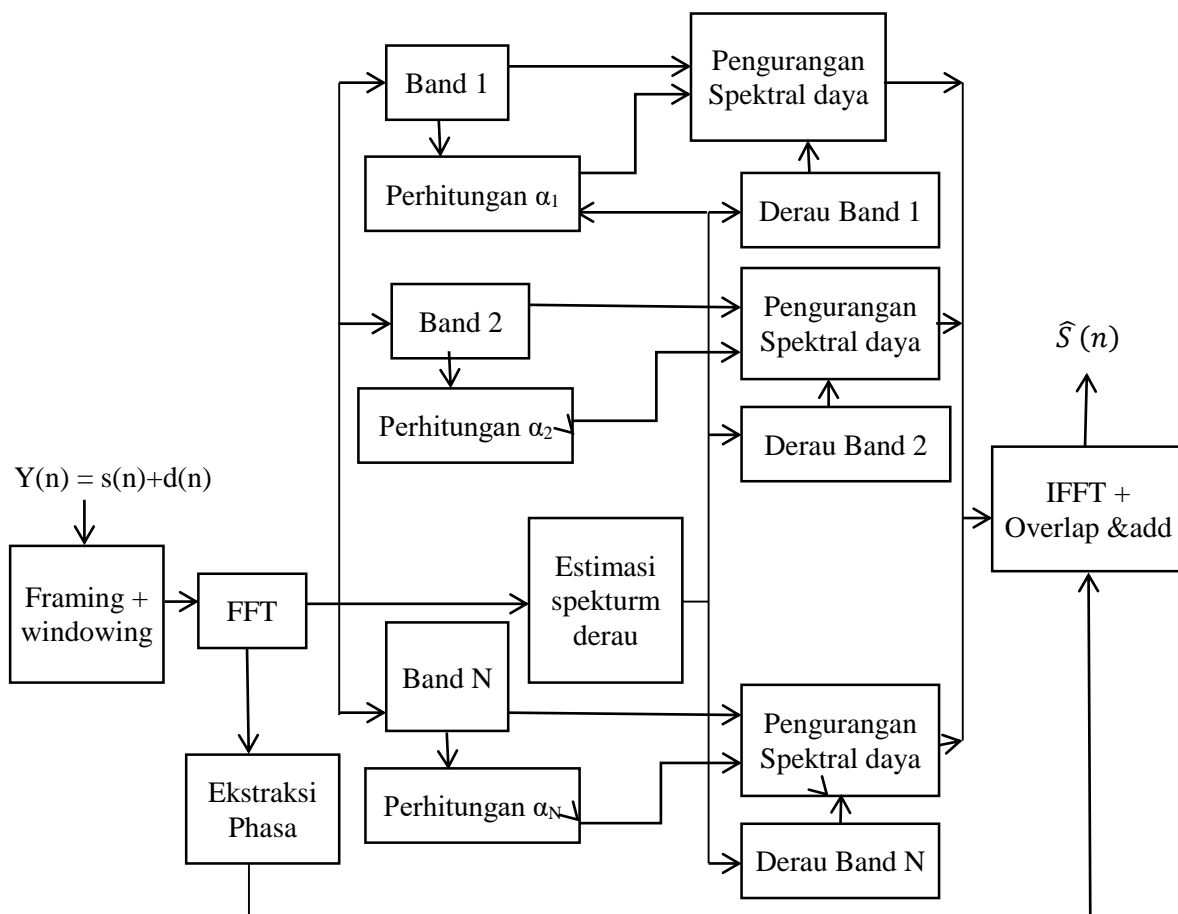
$$|\hat{S}_i(k)|^2 = |Y_i(k)|^2 - \alpha_i \delta_i |\hat{D}_i(k)|^2; b_i \leq k \leq e_i \quad (10)$$

$$|\hat{S}_i(k)|^2 = \begin{cases} |\hat{S}_i(k)|^2 & ; \text{jika } |\hat{S}_i(k)|^2 > 0 \\ \beta |Y_i(k)|^2 & ; \text{lainnya} \end{cases} \quad (11)$$

$$\delta_i = \begin{cases} 1 & ; f_i \leq 1\text{kHz} \\ 25 & ; 1\text{kHz} \leq f_i \leq \frac{f_s}{2} - 2\text{kHz} \\ 15 & ; f_i > \frac{f_s}{2} - 2\text{kHz} \end{cases} \quad (12)$$

Dimana δ_i untuk menyesuaikan pengurangan derau pada setiap *band*. Nilai δ_i secara empiris telah dinyatakan pada persamaan (12).

- 8) Mengubah kembali estimasi *spectrum* sinyal bersih ke dalam domain waktu dengan menggunakan IFFT.
- 9) Merekonstruksi estimasi sinyal bersih domain waktu pada setiap *frame* dengan melakukan *Overlap and Add* (OLA).



Gambar 1. Multiband Spectral Subtraction yang diusulkan

Adapun data ucapan yang digunakan pada penelitian ini bersumber dari *data base* NOIZEUS. Perincian data adalah sebagai berikut:

- 1) Sampel ucapan memiliki durasi 2 detik dengan frekuensi *sampling* 8 kHz.
- 2) Jumlah sampel ucapan berderau sebanyak 30 sampel untuk setiap jenis derau dan setiap level SNR.
- 3) Jenis derau yang diuji adalah AWGN, Suara Mobil, *Airport* dan Percakapan.
- 4) Level SNR yang diujikan 0, 5, 10 dan 15 dB.

- 5) Jumlah ucapan bersih 30 sampel yang akan digunakan untuk menghitung nilai PESQ sinyal hasil perbaikan.

Kinerja metoda MBSS yang diusulkan dianalisa berdasarkan nilai PESQ sinyal hasil perbaikan dan perbandingannya dengan sinyal berderau. PESQ merupakan rekomendasi P.862. ITU-T pada tahun 2000 [1]. Sejumlah penelitian menemukan bahwa PESQ memiliki korelasi sangat tinggi dengan hasil evaluasi menggunakan ukuran subjektif. PESQ memprediksi nilai *Mean Opinion Score* (MOS). Nilai MOS memiliki lima indikasi kualitas yaitu “sangat buruk”, “buruk”, “cukup”, “baik”, “sempurna” maka PESQ juga menghasilkan indikasi berupa nomor 1-5. Nomor 1 merepresentasikan kualitas sangat buruk dan nomor 5 merepresentasikan kualitas sempurna [9].

3. HASIL DAN PEMBAHASAN

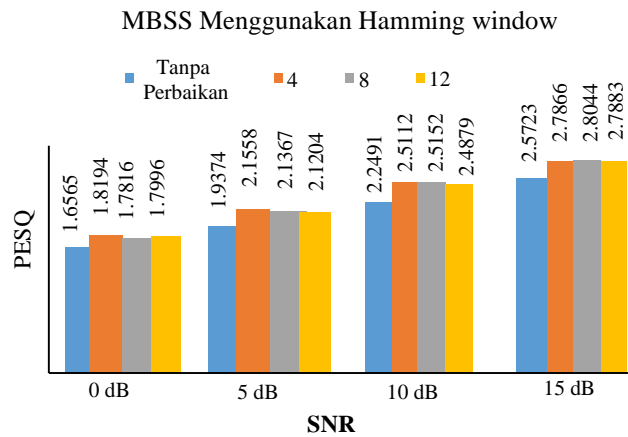
3.1 MBSS dengan Menggunakan Hamming Window

Tabel 1 merupakan nilai rata-rata PESQ dari MBSS menggunakan *Hamming Window* dengan memvariasikan jumlah *band*. Nilai ini didapat dengan merata-ratakan nilai PESQ dari 30 sampel ucapan pada masing-masing jenis derau dan level SNR. Pada tabel 1 terlihat bahwa nilai PESQ semakin tinggi seiring dengan kenaikan level SNR. Hal ini karena PESQ dan SNR merupakan dua parameter yang dapat digunakan untuk menunjukkan kualitas sinyal ucapan. Pada kondisi tanpa perbaikan untuk level SNR 0 dB, ucapan berderau AWGN memiliki nilai PESQ yang terendah dan ucapan berderau bandara yang tertinggi dibandingkan ucapan berderau lainnya pada SNR yang sama. Hal ini juga terjadi pada level SNR 5, 10 dan 15 dB. Setelah dilakukan perbaikan dengan menggunakan MBSS dengan *Hamming window*, maka dihasilkan peningkatan nilai PESQ pada semua jenis derau, semua level SNR dan semua jumlah *band* yang diujikan. Pada semua jumlah *band* yang diujikan, nilai PESQ sinyal hasil perbaikan yang tertinggi pada ucapan berderau percakapan dan terendah pada ucapan berderau AWGN.

Tabel 1. PESQ pada MBSS Menggunakan Hamming Window

Jumlah Band	SNR	Jenis Derau				Rata-rata
		Bandara	AWGN	Percakapan	Mobil	
Ucapan Berderau	0	1,7259	1,5614	1,7053	1,6337	1,6565
	5	2,0212	1,8311	2,006	1,8913	1,9374
	10	2,341	2,1339	2,3212	2,2006	2,2491
	15	2,6332	2,4714	2,6529	2,5318	2,5723
4	0	1,8647	1,682	1,8998	1,8311	1,8194
	5	2,2143	2,0456	2,2169	2,1464	2,1558
	10	2,5139	2,4956	2,512	2,5232	2,5112
	15	2,7735	2,7957	2,7774	2,7997	2,7866
8	0	1,8533	1,6496	1,8625	1,7611	1,7816
	5	2,218	2,001	2,2274	2,1006	2,1367
	10	2,5411	2,4663	2,5439	2,5094	2,5152
	15	2,8088	2,7827	2,8164	2,8097	2,8044
12	0	1,8616	1,6549	1,8863	1,7956	1,7996
	5	2,205	1,9753	2,2095	2,0918	2,1204
	10	2,5225	2,4281	2,5181	2,483	2,4879
	15	2,7984	2,7539	2,8066	2,7942	2,7883

Pada gambar 2, dari keseluruhan variasi jumlah *band* dan level SNR dapat dilihat bahwa MBSS menggunakan *Hamming Window* mampu meningkatkan kualitas sinyal ucapan berderau. Nilai rata-rata PESQ tertinggi untuk level SNR 0 dB dan 5 dB didapatkan pada jumlah *band* 4, sedangkan untuk level SNR 10 dan 15 dB pada jumlah *band* 8. Berdasarkan hasil tersebut, maka untuk *hamming window* lebih baik menggunakan jumlah *band* 4.



Gambar 2. Grafik Nilai Rata-rata PESQ dengan Variasi Jumlah Band pada MBSS Menggunakan Hamming Window

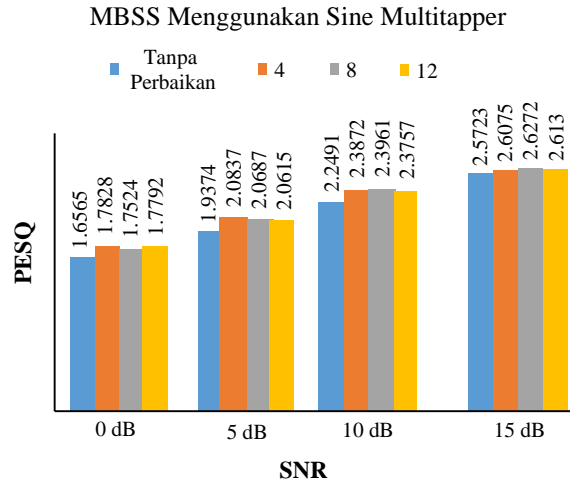
3.2 MBSS dengan Menggunakan Sine Multitapper

Tabel 2 menunjukkan kinerja MBSS menggunakan *sine multitapper* dengan jumlah *tapper* 4. Jika nilai PESQ ucapan berderau tanpa perbaikan dibandingkan dengan ucapan berderau hasil perbaikan, maka terlihat peningkatan nilai PESQ. Hal ini menunjukkan bahwa penggunaan *sine multitapper* pada metoda MBSS mampu memperbaiki kualitas ucapan berderau pada semua level SNR dan jumlah *band* yang diujikan. Ucapan hasil perbaikan dengan nilai PESQ tertinggi terdapat pada ucapan berderau percakapan dan terendah ucapan berderau AWGN. Hal ini terdapat pada semua pengujian jumlah *band*.

Tabel 2. PESQ pada MBSS Menggunakan Sine Multitapper (k=4)

Jumlah Band	SNR	Jenis Derau				Rata-rata
		Bandara	AWGN	Percakapan	Mobil	
Tanpa Perbaikan	0	1,7259	1,5614	1,7053	1,6337	1,6565
	5	2,0212	1,8311	2,006	1,8913	1,9374
	10	2,341	2,1339	2,3212	2,2006	2,2491
	15	2,6332	2,4714	2,6529	2,5318	2,5723
4	0	1,8396	1,6431	1,8454	1,8031	1,7828
	5	2,1455	1,9766	2,1453	2,0674	2,0837
	10	2,3998	2,3639	2,3959	2,3891	2,3872
	15	2,6018	2,608	2,6088	2,6115	2,6075
8	0	1,8321	1,6135	1,8278	1,7362	1,7524
	5	2,1479	1,9395	2,1552	2,0322	2,0687
	10	2,4247	2,3458	2,4271	2,3868	2,3961
	15	2,6315	2,6025	2,6463	2,6283	2,6272
12	0	1,8454	1,6296	1,8619	1,7799	1,7792
	5	2,1428	1,9199	2,1452	2,038	2,0615
	10	2,4103	2,315	2,409	2,3687	2,3757
	15	2,6203	2,5801	2,6391	2,6126	2,6130

Pada gambar 3 dapat dilihat bahwa peningkatan nilai rata-rata PESQ tertinggi untuk level SNR 0 dB dan 5 dB pada penggunaan jumlah *band* 4, sedangkan SNR 10 dB dan 15 dB pada jumlah *band* 8. Jika dirata-rata nilai PESQ pada seluruh level SNR disetiap jumlah *band* maka didapatkan nilai PESQ secara berturut-turut untuk *band* 4, 8 dan 12 adalah 2.2153, 2.2111, dan 2.2073. Oleh karena itu MBSS menggunakan *sine multitapper* dengan jumlah *tapper* 4 bekerja lebih baik pada *band* 4.



Gambar 3. Grafik Nilai Rata-rata PESQ dengan Variasi Jumlah Band pada MBSS Menggunakan Sine Multitapper

3.3 MBSS dengan Menggunakan Gaussian Window

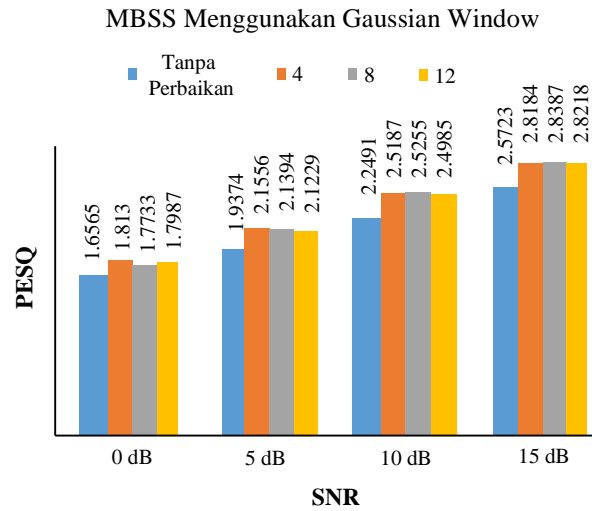
Kinerja MBSS menggunakan *Gaussian window* dengan nilai redaman *side lobe 0* ditunjukkan pada tabel 3. Ucapan berderau yang telah diperbaiki menggunakan MBSS dengan *Gaussian window* redaman *side lobe 0* memiliki nilai PESQ yang lebih tinggi dibandingkan dengan ucapan berderau tanpa perbaikan. Hal ini menunjukkan bahwa metoda MBSS menggunakan *Gaussian window side lobe 0* mampu memperbaiki kualitas ucapan berderau pada semua level SNR. Sama seperti penggunaan *hamming* dan *sine multitapper* pada MBSS, *Gaussian window* juga menghasilkan PESQ tertinggi pada ucapan hasil perbaikan berderau percakapan dan PESQ terendah pada ucapan hasil perbaikan berderau AWGN. Hal ini terjadi pada semua jumlah *band* yang diujikan.

Tabel 3. PESQ pada MBSS Menggunakan Gaussian Window dengan Redaman Side Lobe 0

Jumlah Band	SNR	Jenis Derau				Rata-rata
		Bandara	AWGN	Percakapan	Mobil	
Tanpa Perbaikan	0	1,7259	1,5614	1,7053	1,6337	1,657
	5	2,0212	1,8311	2,006	1,8913	1,937
	10	2,341	2,1339	2,3212	2,2006	2,249
	15	2,6332	2,4714	2,6529	2,5318	2,572
4	0	1,8557	1,6789	1,8925	1,8248	1,813
	5	2,2141	2,0539	2,2097	2,1447	2,156
	10	2,5203	2,5127	2,5176	2,524	2,519
	15	2,7971	2,8497	2,8057	2,8212	2,818
8	0	1,8398	1,6448	1,8628	1,7459	1,773
	5	2,2202	2,0108	2,2241	2,1024	2,139
	10	2,5515	2,4858	2,549	2,5157	2,526
	15	2,8331	2,8406	2,8445	2,8367	2,839
12	0	1,8508	1,6524	1,9025	1,7892	1,799
	5	2,2077	1,9824	2,2079	2,0936	2,123
	10	2,5313	2,4459	2,5266	2,4901	2,499
	15	2,8219	2,8093	2,8377	2,8184	2,822

Gambar 4 menunjukkan bahwa *Gaussian window* pada redaman *side lobe 0* mampu menaikkan nilai PESQ sinyal berderau, pada semua jumlah *band* yang diujikan. Sama seperti pada dua jenis

window sebelumnya bahwa untuk level SNR 0 dB dan 5 dB, PESQ tertinggi dengan menggunakan jumlah *band* 4 sedangkan untuk SNR 5 dB dan 10 dB PESQ tertinggi pada penggunaan jumlah *band* 8. Perbedaan nilai PESQ kedua *band* ini tidak signifikan sehingga dapat dikatakan jumlah *band* yang sebaiknya digunakan adalah 4.



Gambar 4. Grafik Nilai Rata-rata PESQ dengan Variasi Jumlah Band pada MBSS Menggunakan Gaussian Window dengan Redaman Side Lobe 0

3.4 Kinerja MBSS Berdasarkan Redaman Side Lobe Gaussian Window

Parameter redaman *side lobe* (α_{GW}) akan mempengaruhi bentuk *window* yang nantinya akan mempengaruhi estimasi *spectral* sinyal. Oleh Karena itu pada tabel 4, ditunjukkan kinerja MBSS menggunakan *gaussian window* dengan jumlah *band* 4 dan variasi redaman *side lobe*. Jika nilai PESQ ucapan tanpa perbaikan dibandingkan dengan ucapan hasil perbaikan menggunakan MBSS *Gaussian window*, maka terlihat bahwa dari semua nilai redaman *side lobe* Gaussian yang diujikan menghasilkan peningkatan nilai PESQ pada semua jenis derau yang diujikan. Nilai PESQ ucapan hasil perbaikan berderau percakapan memiliki PESQ terbaik sedangkan AWGN masih yang terendah untuk semua nilai redaman *side lobe* Gaussian.

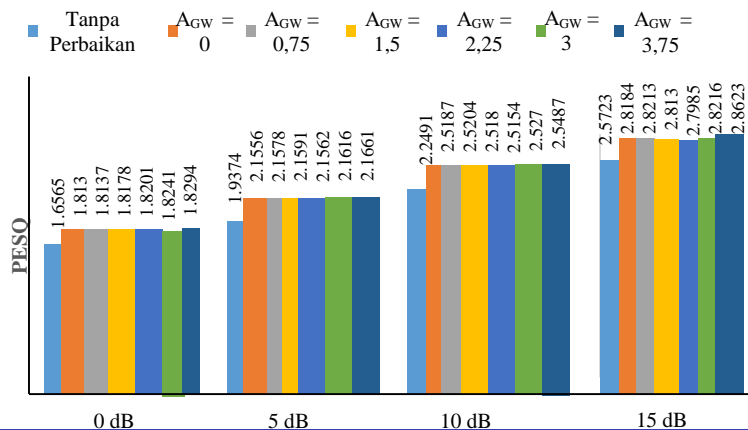
Tabel 4. PESQ pada MBSS Menggunakan Gaussian Window dengan Jumlah Band 4

Side Lobe	SNR	Jenis Derau				Rata-rata
		Bandara	AWGN	Percakapan	Mobil	
Tanpa Perbaikan	0	1,7259	1,5614	1,7053	1,6337	1,6565
	5	2,0212	1,8311	2,0060	1,8913	1,9374
	10	2,3410	2,1339	2,3212	2,2006	2,2491
	15	2,6332	2,4714	2,6529	2,5318	2,5723
0	0	1,8557	1,6789	1,8925	1,8248	1,8130
	5	2,2141	2,0539	2,2097	2,1447	2,1556
	10	2,5203	2,5127	2,5176	2,5240	2,5187
	15	2,7971	2,8497	2,8057	2,8212	2,8184
0,75	0	1,8567	1,6788	1,8929	1,8264	1,8137
	5	2,2163	2,0545	2,2122	2,1482	2,1578
	10	2,5235	2,5134	2,5187	2,5261	2,5204
	15	2,8003	2,8518	2,8088	2,8242	2,8213
1,5	0	1,8641	1,6807	1,8944	1,8321	1,8178
	5	2,2149	2,0570	2,2137	2,1508	2,1591
	10	2,5218	2,5095	2,5183	2,5223	2,5180
	15	2,7952	2,8355	2,8021	2,8190	2,8130
2,25	0	1,8640	1,6837	1,9006	1,8320	1,8201

	5	2,2159	2,0476	2,2131	2,1483	2,1562
	10	2,5189	2,5034	2,5179	2,5216	2,5154
	15	2,7839	2,8109	2,7884	2,8108	2,7985
3	0	1,8720	1,6819	1,9061	1,8363	1,8241
	5	2,2221	2,0549	2,2210	2,1483	2,1616
	10	2,5305	2,5189	2,5278	2,5308	2,5270
	15	2,8058	2,8311	2,8154	2,8340	2,8216
3,75	0	1,8836	1,6832	1,9084	1,8426	1,8294
	5	2,2301	2,064	2,2272	2,1431	2,1661
	10	2,5554	2,5410	2,5482	2,5502	2,5487
	15	2,8454	2,8772	2,8563	2,8703	2,8623

Perbandingan rata-rata PESQ pada semua variasi *side lobe Gaussian window* dapat dilihat pada gambar 5. Semua redaman *side lobe Gaussian* yang digunakan pada MBSS mampu memperbaiki sinyal ucapan berderau yang ditunjukkan dengan adanya peningkatan nilai PESQ pada semua level SNR yang diujikan. Semakin tinggi nilai redaman *side lobe* maka terdapat kecenderungan nilai PESQ semakin naik, tetapi kenaikannya tidak signifikan. Nilai redaman *side lobe* yang terbaik dari variasi yang dilakukan adalah nilai 3.75

MBSS Menggunakan Gaussian Window



Gambar 5. Rata-rata PESQ pada MBSS Menggunakan Gaussian Window dengan Jumlah Band 4, Variasi Redaman Side Lobe

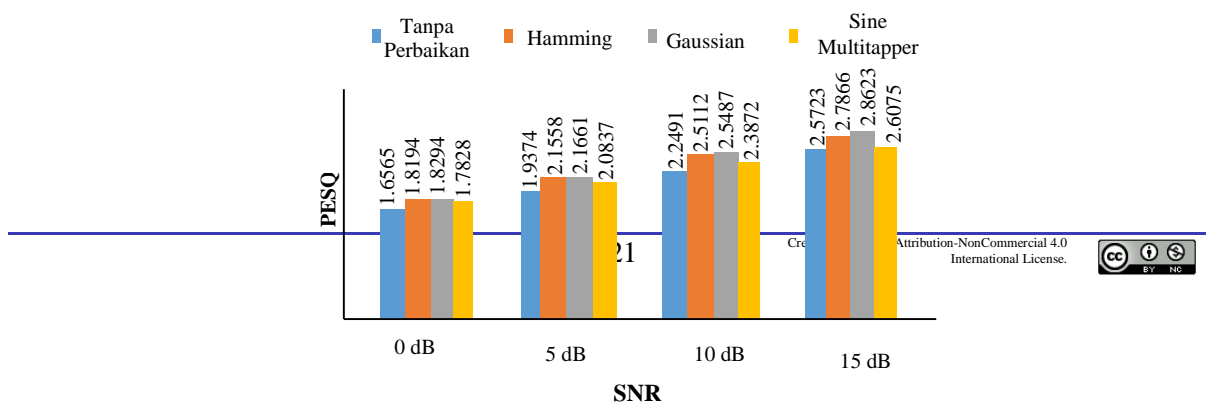
3.5 Perbandingan Jenis Window pada MBSS

Tabel 5 menunjukkan nilai rata-rata PESQ untuk 4 jenis derau dengan menggunakan MBSS dengan 3 jenis window yang berbeda. Jumlah band yang digunakan adalah 4 band. Pada Gaussian window menggunakan redaman side lobe 3.75. Berdasarkan tabel 5, penggunaan ketiga jenis window memberikan peningkatan PESQ yang lebih baik untuk jenis derau percakapan dan yang terendah untuk derau AWGN. Rata-rata nilai PESQ seperti pada gambar 6, menunjukkan bahwa semua jenis window mampu memperbaiki kualitas sinyal ucapan berderau untuk empat jenis derau di masing-masing level SNR, dengan adanya peningkatan PESQ dibandingkan kondisi tanpa perbaikan. Window yang memberikan peningkatan terbaik pada semua level SNR dan semua jenis derau adalah Gaussian window dengan redaman side lobe 3.75. Hamming window lebih baik dibandingkan Sine multitapper dengan jumlah tapper 4.

Tabel 5. PESQ MBSS dengan 4 Band

Jumlah Band	SNR	Jenis Derau				Rata-rata
		Bandara	AWGN	Percakapan	Mobil	
Tanpa Perbaikan	0	1,7259	1,5614	1,7053	1,6337	1,6565
	5	2,0212	1,8311	2,0060	1,8913	1,9374
	10	2,3410	2,1339	2,3212	2,2006	2,2491
	15	2,6332	2,4714	2,6529	2,5318	2,5723
Hamming	0	1,8647	1,6820	1,8998	1,8311	1,8194
	5	2,2143	2,0456	2,2169	2,1464	2,1558
	10	2,5139	2,4956	2,5120	2,5232	2,5112
	15	2,7735	2,7957	2,7774	2,7997	2,7866
Gaussian (3,75)	0	1,8836	1,6832	1,9084	1,8426	1,8294
	5	2,2301	2,064	2,2272	2,1431	2,1661
	10	2,5554	2,5410	2,5482	2,5502	2,5487
	15	2,8454	2,8772	2,8563	2,8703	2,8623
Sine Multi tapper	0	1,8396	1,6431	1,8454	1,8031	1,7828
	5	2,1455	1,9766	2,1453	2,0674	2,0837
	10	2,3998	2,3639	2,3959	2,3891	2,3872
	15	2,6018	2,6080	2,6088	2,6115	2,6075

Perbandingan Jenis Window Pada kinerja MBSS



Gambar 6. Rata-rata PESQ pada MBSS untuk Tiga Jenis Window dengan Jumlah Band 4

4. KESIMPULAN

Multiband spectra subtraction dengan menggunakan tiga jenis *window* dan tiga variasi jumlah *band* mampu memperbaiki kualitas sinyal ucapan berderau. Hal ini ditunjukkan dengan adanya peningkatan nilai PESQ sinyal hasil perbaikan dibandingkan dengan sinyal tanpa perbaikan. *Window* yang memberikan peningkatan PESQ terbaik adalah *Gaussian window*. Nilai redaman *side lobe* terbaik adalah 3.75 tetapi peningkatan nilai *side lobe* tidak memberikan peningkatan nilai PESQ yang signifikan. Jumlah *band* yang memberikan nilai PESQ yang terbaik adalah 4 *band*. Oleh karena itu *multiband spectra subtraction* yang diusulkan adalah MBSS menggunakan *Gaussian window* dengan jumlah *band* 4 dan redaman *side lobe* 3.75.

Ucapan Terima Kasih

Terima kasih disampaikan kepada Jurusan Teknik Elektro Fakultas Teknik Universitas Andalas yang telah memberikan dana penelitian DIPA Fakultas Teknik dengan nomor kontrak: 087/UN.16.09.D/PL/2018.

REFERENSI

- [1] Pardede., Hilman, F. *Nonlinear Spectral Subtraction Based on Tsallis Statistics for Speech Enhancement*. 2013. INKOM Vol. 7 No 1 Mei: artikel 228.
- [2] S. F. Boll. *Suppression of Acoustic Noise in Speech Using Spectral Subtraction*. 1979. IEEE Trans. Acoust: Speech Signal Process Vol. 27 No. 2: 113– 120.
- [3] Kamath, S., Loizou, P. *A Multiband Spectral Subtraction Methods for Enhancing Speech Corrupted by Colored Noise*. 2002. Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Processing.
- [4] M. Berouti., R. Schwartz., J. Makhoul. *Enhancement of Speech Corrupted by Acoustic Noise*. 1979. Proc. IEEE ICASSP: 208-211.
- [5] Supriya, P. Sarvade., Dr. Shridhar, K., Varun, P. Sarvade. *Multi-Band Spectral Subtraction for Speech Enhancement Using Sine Multitaper*. 2016. IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Vol 6 Issue 6 Ver. II Nov – Dec.
- [6] S. China, Venkateswarlu., A. Subba Rami Reddy, K. Satya Prasad. *Speech Enhancement Using Boll's Spectral Subtraction Method Based on Gaussian Window*. 2014. Global Journal of Researches in Engineering: F Electrical and Electronics Engineering Vol 14 Issue 6 Version 1.0.
- [7] Navneet, Upadhyay., Abhijit, Karmakar. *Single-Channel Speech Enhancement using Critical-Band Rate Scale Based Improved Multi-Band Spectral Subtraction*. 2013. Journal of Signal and Information Processing ISSN 2159-4465 Vol 04: 314.

- [8] Venkata Rami Reddy Datla. *Implementation and Evaluation of Spectral Subtraction (SS) with Minimum Statistics and Wiener Beamformer Combination*. 2013. Master Thesis of Electrical Engineering, School of Electrical Engineering, Blekinge Institute of Technology (BTH), Sweden.
- [9] Scoot, Pennock. *Accuracy of the Perceptual Evaluation of Speech Quality Algorithm*. 2017. Lucent Technologies.

