ANALISIS SPEKTRUM FREKUENSI NON-LINEAR SINYAL TUTUR DENGAN ALIH RAGAM FOURIER CEPAT

Salman Abd. Cadum[1], Prayoto[2], Adhi Susanto[3], Kirbani Sri Brotopuspito[4]
[1,3]Jurusan Teknik Elektro Fakultas Teknik Universitas Gadjah Mada
e-mail: salman_abd_2002@yahoo.com
[2]Jurusan Fisika Fakultas MIPA Universitas Gadjah Mada
[4]Jurusan Geofisika Fakultas MIPA Universitas Gadjah Mada

Abstrak

Pada penelitian ini akan diteliti analisis spektrum frekuensi non-linear sinyal tutur dengan menggunakan alih ragam Fourier cepat (Fast Fourier Transform, FFT). Hasil penelitian menunjukkan bahwa: suatu skala logaritmis akan memperluas daerah frekuensi yang rendah dari spektrum dan mempersempit daerah frekuensi yang tinggi pada tampilan, dibutuhkan suatu FFT yang jauh lebih besar guna mendapatkan resolusi frekuensi yang sangat tinggi pada frekuensi yang rendah, penerapan fungsi berbagai window terhadap data dapat membantu mengurangi efek kebocoran yang terjadi pada spektrum frekuensi, metode ini berjalan lebih cepat jika jumlah point data merupakan kelipatan dua (128, 256, 1024, 2048, atau 4096, dan seterusnya) dan memilih suatu resolusi frekuensi yang tepat serta resolusi waktu yang sesuai menjadi suatu kesesuaian antara kebutuhan untuk mengamati detail frekuensi yang baik dalam spektrum dengan kebutuhan untuk mengamati variasi waktu yang cepat dalam spektrum.

Kata kunci : *spektrum frekuensi non-linear, sinyal tutur, FFT*

1. PENDAHULUAN

Spektrum menggunakan suatu transform Fourier cepat (*Fast Fourier Transform*, FFT) matematis untuk melakukan analisis frekuensi. FFT biasanya dinyatakan dengan jumlah *point* data masukan yang digunakan dalam setiap perhitungan yang selalu berupa kelipatan dua (128, 256, 512, 1024, 2048, atau 4096). Resolusi frekuensi spektrum selalu merupakan nilai cuplikan *digital* sinyal *audio/*tutur yang dibagi dengan jumlah *point* data FFT. Semakin besar jumlah *point* data FFT, semakin baik resolusi frekuensi spektrum. Frekuensi maksimum yang dihitung oleh FFT dan batas frekuensi tertinggi spektrum adalah setengah nilai cuplikan *digital* [1].

Dalam setiap proses analisis spektrum, resolusi waktu dan resolusi frekuensi memiliki hubungan terbalik, resolusi frekuensi yang sangat bagus berkaitan dengan resolusi waktu yang buruk, sebaliknya resolusi waktu yang sangat bagus berkaitan dengan resolusi frekuensi yang buruk. Hubungan antara resolusi waktu (*Time Resolution*, TR) dalam detik dan resolusi frekuensi (*Frequency Resolution*, FR) dalam Hz adalah:

$$TR = 1/FR \tag{1}$$

Akibat logis prinsip ketidakpastian ini adalah bahwa memilih resolusi frekuensi yang sangat bagus akan menutupi *detail* waktu dalam spektrum sinyal *audio*/tutur. Memilih suatu resolusi frekuensi yang tepat dan resolusi waktu yang sesuai menjadi suatu kesesuaian antara kebutuhan untuk mengamati *detail* frekuensi yang baik dalam spektrum dengan kebutuhan untuk mengamati variasi waktu yang cepat dalam spektrum [2].

Teori Fourier menyatakan bahwa bentuk gelombang (*waveform*) periodis dapat dipecah ke dalam serangkaian sinusoid yang memiliki frekuensi, amplitude, dan fase yang berbeda. Dalam sintesis *wavetable*, amplitude parsialnya tetap, dan sementara hal ini umumnya cukup untuk memberikan kesan (*impression*) tipe warna nada (*timbre*) tertentu, keterbatasannya mencegahnya menerima hasil yang lebih tahan dalam kelangsungan hidup (*life-like*), hal ini terutama karena dua alasan: a) di alam, frekuensi parsial jarang merupakan perkalian bilangan bulat yang tepat (*exact*), dan b) amplitude parsial individu hampir selalu berbeda-beda relatif terhadap satu sama lain selama suatu nada (*tone*). Kedua efek ini mustahil dicapai dengan sintesis *wavetable*, dan sintesis aditif, oleh karena itu, dengan hubungannya yang jelas dengan analisis Fourier, maka Fourier menawarkan solusi masalah ini.

Penganalisa FFT mentransformasikan data dari domain waktu ke domain frekuensi dengan menghitung FFT. Hal ini didasarkan pada integral Fourier persamaan 2 berikut.

$$y = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft}dt \qquad (2)$$

Namun ini merupakan suatu bentuk yang dapat dihitung secara numeris. Integral ini mensyaratkan bahwa suatu sinyal kontinyu diintegralkan selama waktu yang tak terhingga, tentu saja, diinginkan hasil dalam waktu yang terhingga. Dan karena komputer berhubungan dengan angka, maka diperlukan *digitize* (men-*digital*-kan) bentuk gelombang, yang dapat membuat waktu bersifat diskrit. Kedua perubahan terhadap sinyal ini mengakibatkan kesalahan dalam spektrum frekuensi yang dihitung. Cuplikan sinyal pada waktu diskrit dapat menyebabkan *aliasing* (yang dapat terlihat sebagai sinyal bayangan (*phantom*) pada tampilan). Pengubahan batas integral panjang tak terhingga menjadi panjang terhingga dapat menyebabkan kesalahan yang disebut kebocoran (yang muncul sebagai energi dari titik tertentu dalam spektrum terbaur (*smeared*) naik dan turun melintasi spektrum). Karena ketidakmungkinan mengukur suatu sinyal untuk waktu yang tak terhingga, maka penganalisa mengubah batas integral ke panjang waktu yang dibutuhkan untuk mengumpulkan blok cuplikan. Blok cuplikan disebut *time record*. FFT mensyaratkan bahwa sinyal dalam *time record* diulang terus-menerus sepanjang waktu. Jika *time record* yang diulang secara aktual tampak seperti sinyal asli, maka tidak akan terjadi kebocoran. Jika, pada sisi lain, tak terlihat seperti sinyal aslinya, maka terjadi kebocoran. Penerapan fungsi *window* terhadap data yang ada akan dapat membantu mengurangi efek kebocoran dalam domain frekeunsi [3].

Guna menghitung spektrum *y* dari sinyal tutur *X* dapat digunakan FFT dengan menggunakan persamaan berikut:

$$y_m = \sum_{k=0}^{N-1} X_k . e^{2\pi ikm/N} \qquad (3)$$

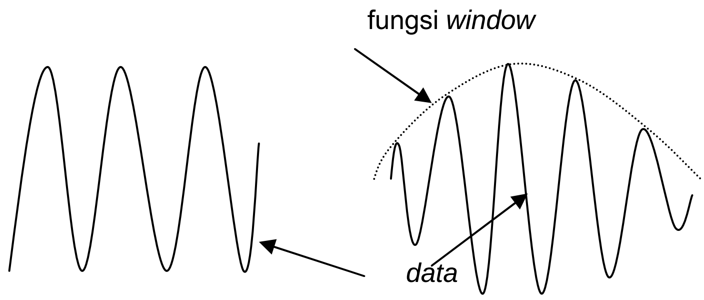
dengan $m = 0 \ldots N\text{-}1$, dan $k = 0 \ldots N\text{-}1$.

Demikian juga untuk menghitung sinyal *X* dari spektrum y dapat digunakan invers transform Fourier cepat (IFFT) dengan menggunakan persamaan berikut:

$$X_k = \frac{1}{N}\sum_{m=0}^{N-1} y_m . e^{-2\pi ikm/N} \qquad (4)$$

dengan $m = 0 \ldots N\text{-}1$, dan $k = 0 \ldots N\text{-}1$.

Pememilihan suatu *data window* biasanya dilakukan melalui suatu kompromi. Setiap *window* ‘mengotori’ (*smear*) spektrum frekuensi**.** Ini berarti bahwa suatu ‘puncak yang runcing’ yang berkaitan dengan suatu gelombang sinus dalam sinyal semakin melebar. Di sisi lain kebocoran palsu (*spurious*) ke dalam frekuensi di sekitarnya akan berkurang. Fungsi *window* disajikan pada Gambar 1.



Gambar 1. Gambar fungsi *window*

Analisis Frekuensi untuk suatu sinyal dengan tingkat tekanan suara yang konstan pada semua frekuensi seperti derau putih, suatu jumlah energi yang sama terkandung dalam setiap pita pengukuran. Suatu penganalisa spektrum yang dikalibrasikan secara tepat dengan masukan derau putih akan menunjukkan suatu spektrum, ketika menggunakan skala frekuensi logaritmis, lebar setiap pita pengukuran akan meningkat bersama dengan frekuensi. Tingkat tekanan suara yang diukur dalam setiap pita akan meningkat dengan nilai kenaikan sebesar 3dB per oktaf dalam frekuensi.

2. METODE PENELITIAN

Ukuran FFT dapat dihitung pada spektrum dengan menggunakan FFT pada *point* (128, 256, 512, 1024, 2048, atau 4096). Resolusi frekuensi paling tinggi pada spektrum adalah nilai cuplikan *digital* yang dibagi dengan ukuran FFT. Jika digunakan FFT yang lebih besar hanya untuk analisi resolusi yang tinggi atau dengan skala frekuensi logaritmis. FFT dengan resolusi yang lebih tinggi membutuh waktu lebih banyak untuk menghitung spektrum. Untuk alasan ini kadang-kadang sebaiknya mengurangi nilai cuplikan ketika merekam data *audio*/tutur, jika resolusi frekuensi yang tinggi dibutuhkan, dari pada menggunakan suatu FFT dengan resolusi yang lebih tinggi.

Algoritma dasar FFT berikut digunakan untuk menghitung spektrum *y* dari sinyal *X* sebagai berikut:

1. Mulai;
2. Konstanta
   sinyal ← Larik TFFTFloat;
3. Variabel:
   spektrum ← Larik TFFTComplex;
4. Isikanlah data dengan nilai nol hingga ke kelipatan 2 berikutnya dapat dijadikan suatu pilihan bila kecepatan dianggap penting.
5. Untuk m ← 0..N-1, dan k ← 0..N-1;
6. Hitunglah spektrum y sinyal x:

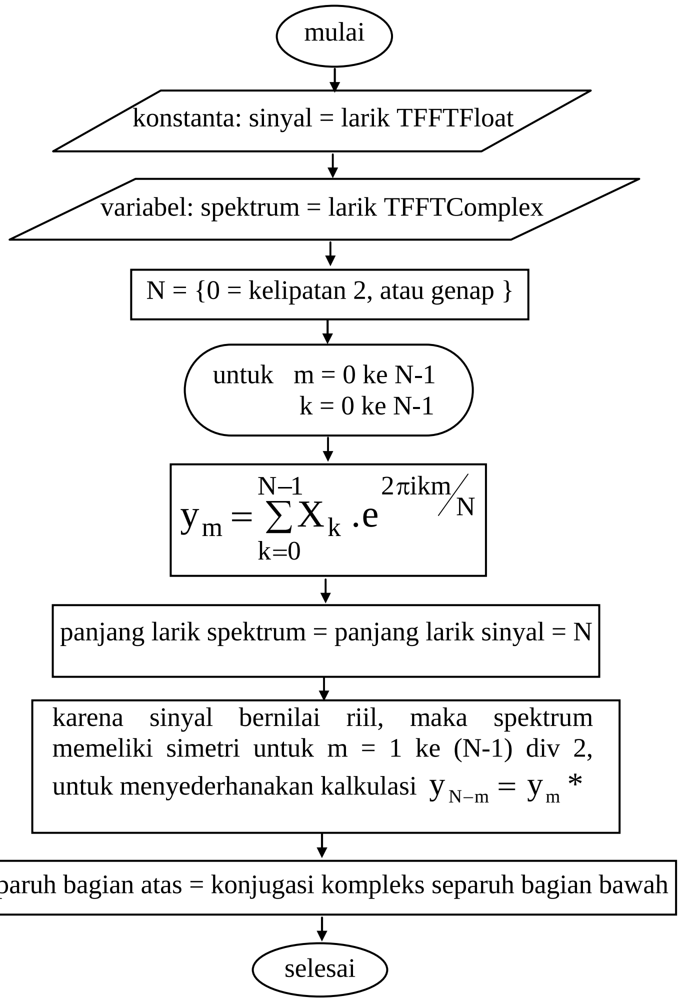
$$y_m = \sum_{k=0}^{N-1} X_k \,.\, e^{2\pi ikm/N}$$

7. Larik spektrum dan larik sinyal harus memiliki panjang yang sama N.
8. Algoritma ini berfungsi untuk setiap jumlah N, tetapi algoritma ini akan sangat cepat jika N adalah kelipatan 2, juga bekerja lebih cepat untuk N genap.
9. Karena sinyal masukan bernilai riil, spektrum memiliki simetri ini untuk m ← 1..(N-1) div 2, yang mungkin berguna untuk menyederhanakan beberapa kalkulasi:

$$y_{N-m} = y_m{}^*$$

10. Separuh bagian atas adalah konjugasi kompleks separuh bagian bawah**.**
11. Selesai.

Diagram alir algoritma FFT dapat dilihat pada Gambar 2.

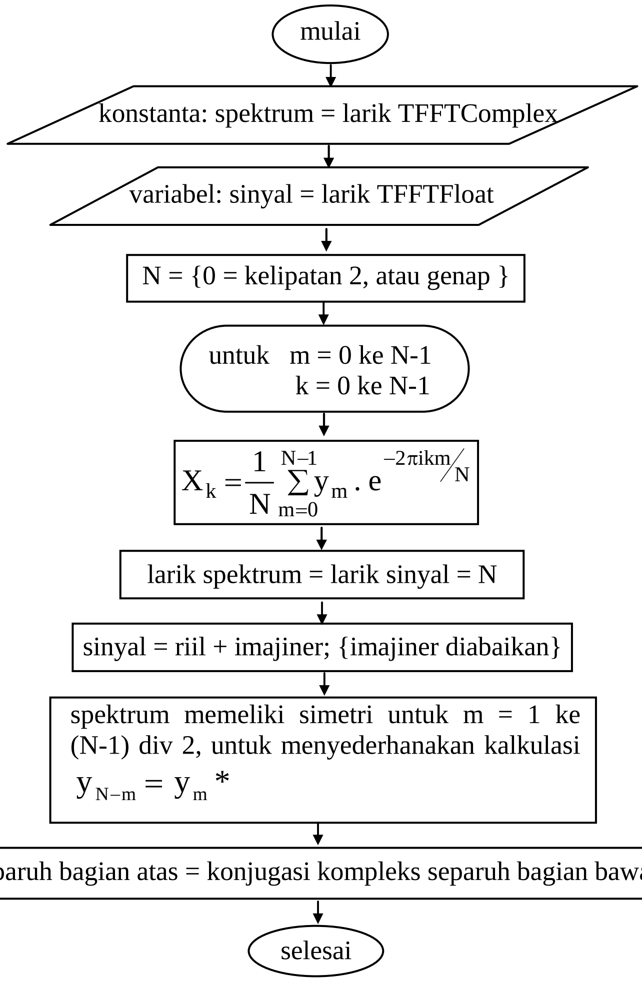


Gambar 2. Diagram alir algoritma dasar FFT

Algoritma dasar IFFT dapat menghitung sinyal *X* dari spektrum *y* sebagai berikut:

1. Mulai;
2. Konstanta;
3. spektrum ← larik TFFTComplex
4. Variabel:
5. sinyal ← Larik TFFTFloat;
6. Algoritma ini berfungsi untuk setiap jumlah N, namun sangat cepat jika N adalah kelipatan juga bekerja lebih cepat bahkan untuk N genap.
7. Untuk m ← 0..N-1, dan k ← 0..N-1,
8. Hitunglah sinyal x spektrum y:
   $$X_k = \frac{1}{N}\sum_{m=0}^{N-1} y_m \,.\, e^{-2\pi ikm/N}$$
9. Larik spektrum dan larik sinyal harus memiliki panjang yang sama N.
10. Algoritma dasar FFT menghasilkan suatu sinyal bernilai kompleks tetapi hanya bagian riil yang disalin ke keluaran larik sinyal, sementara bagian imajiner diabaikan.
11. Terdapat nol dimana spektrum memiliki simetri untuk m ← 1..(N-1) div 2: $y_{N-m} = y_m{}^*$
12. Separuh bagian atas adalah konjugasi kompleks separuh bagian bawah.
13. Selesai.

Diagram alir algoritma IFFT dapat disajikan pada Gambar 3.



Gambar 3. Diagram alir algoritma dasar IFFT

Skala frekuensi pada spektrum ini menggunakan skala frekuensi logaritmis untuk menghitung suatu spektrum. Suatu skala logaritmis akan memperluas daerah frekuensi rendah dari spektrum tersebut dan mempersempit daerah frekuensi yang tinggi pada tampilan.

Spektrum menyediakan suatu kapabilitas *logging* data otomatis untuk tujuan ini, baru kemudian ditentukan tingkat tekanan suaranya. Tingkat tekanan suara dinyatakan dalam desibel (dB) yang merujuk pada amplitude minimum suatu sinyal *audio*/tutur 16 *bit* yang dihasilkan oleh kartu suara. Jika amplitude suatu komponen frekuensi adalah A, maka tingkat sinyal dalam desibel ditetapkan sebagai S = 20*log (A). Hal ini menghasilkan suatu tingkat sinyal maksimum sebesar 90 dB yang berkaitan dengan nilai amplitude 16 bit maksimum sebesar 32767. Nilai ini sama dengan tingkat tekanan suara absolut pada setiap frekuensi plus suatu faktor desibel konstan yang sama dengan tingkat hasil (*gain*) yang digunakan oleh mikrofon, kartu suara komputer, dan penapis analisis spektrum yang dilakukan dalam perangkat lunak.

Tampilan spektrum menampilkan data dengan skala frekuensi logaritmis. Skala logaritma membagi sumbu frekuensi menjadi interval frekuensi logaritmis yang sama. Skala frekuensi logaritmis memberikan keunggulan lebih besar pada frekuensi yang rendah dengan memperluas ruang tampilan untuk frekuensi rendah dengan mengorbankan frekuensi yang tinggi. Tampilan bidang ($^1/_3$) oktaf adalah suatu kasus khusus pada suatu tampilan skala frekuensi logaritmis di mana spektrum tingkat tekanan suara pada pita frekuensi yang berdekatan dengan lebar sepertiga oktaf direncanakan.

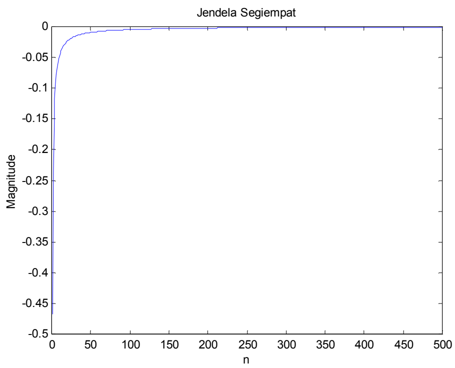
Skala logaritmis membutuhkan resolusi frekuensi yang sangat tinggi pada frekuensi yang rendah. Akibatnya adalah bahwa penggunaan suatu skala log membutuhkan suatu FFT yang jauh lebih besar guna mendapatkan resolusi frekuensi yang dibutuhkan. Mengingat muatan komputasional suatu FFT yang besar, skan terhadap berkas gelombang atau masukan *audio*/tutur dapat berjalan lebih lambat ketika menggunakan suatu skala frekuensi logaritmis.

Ketika menggunakan suatu skala frekuensi logaritmis, dipilih FFT terkecil yang memberikan resolusi frekuensi yang tepat pada spektrum atau bidang pada frekuensi yang rendah. Jika daerah frekuensi yang rendah pada suatu tampilan frekuensi logaritmis kelihatannya terhalang (*blocky*), dinaikkan resolusi frekuensi rendah dengan dipilihnya suatu ukuran FFT yang lebih besar.

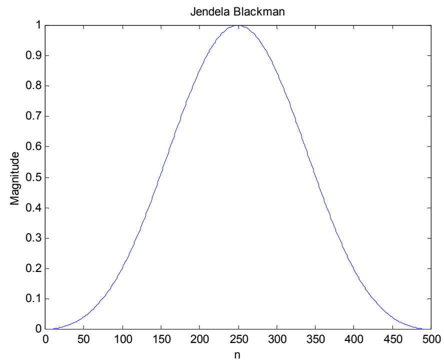
Penting untuk memahami bagaimana pilihan terhadap suatu skala frekuensi logaritmis mempengaruhi tampilan spektrum tingkat tekanan suara. Spektrum mengukur tingkat tekanan suara pada pita frekuensi sempit diskrit di sepanjang spektrum. Suatu sinyal akan menghasilkan ketajaman, karena rutin FFT menganggap masukan sebagai suatu sinyal periodis yang dapat mencacat (*distort*) spektrum. Untuk mengurangi masalah tersebut suatu *window* data pertama-tama dapat diterapkan pada sinyal.

Tabel 1. Persamaan *window* yang diterapkan pada spektrum

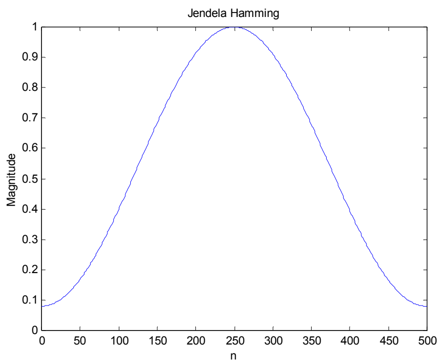
No	Tipe	Nilai cuplikan	Kisaran k
1	Segiempat	$w_k = \begin{cases} 1 \\ 0 \end{cases}$	$0 \le k \le N-1$ else
2	Segitiga Bartlett , atau Parzen	$w_k = 1 - \left\| 1 - \frac{2k}{N-1} \right\|$	$0 \le k \le N-1$
3	Hanning	$w_k = \frac{1}{2}\left(1 - \cos\frac{2k\pi}{N-1}\right)$	$0 \le k \le N-1$
4	Hamming	$w_k = 0.54 - 0.46\cos\frac{2k\pi}{N-1}$	$0 \le k \le N-1$
5	Blackman	$w_k = 0.42 + 0.5\cos\frac{2k\pi}{N-1} + 0.08\cos\frac{4k\pi}{N-1}$	$0 \le k \le N-1$
6	Welch	$\text{Wlech}(x,\tau) = \begin{cases} 1 - \left(\frac{x}{\tau}\right)^2 \\ 0 \end{cases}$	$\|x\| \prec \tau$ else
7	FlatTop	$w_k = 0.2810638602 - 0.520897135 * \cos(x) +$ $0.1980389663 * \cos(2x)$ $x = 2\pi k / (N-1)$	$0 \le k \le N-1$
8	Nuttall	$w_k = 0.375 - 0.5 * \cos(x/(n-1)) + 0.125 * \cos\left(2x/(n-1)\right)$ $x = 2\pi k / (N-1)$	$0 \le k \le N-1$



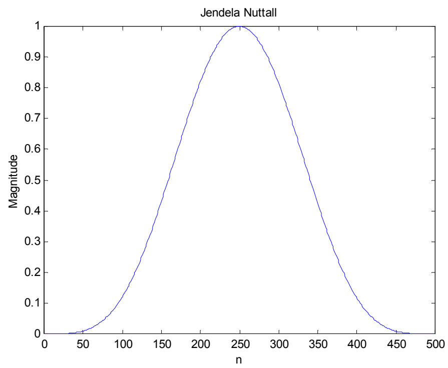
(a) Jendela Segiempat



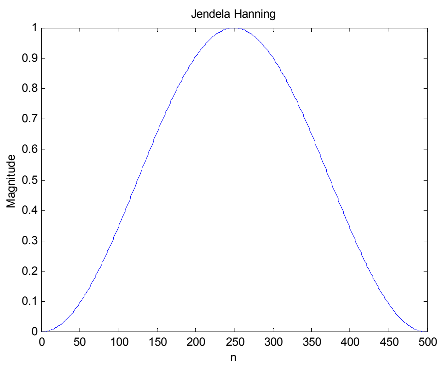
(b) Jendela Blackman



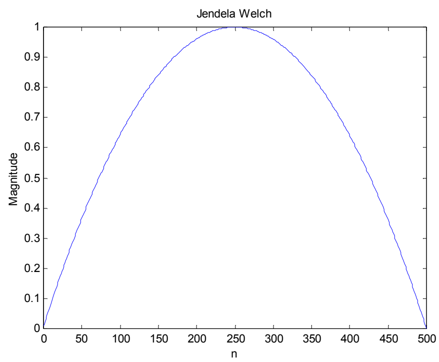
(c) Jendela Hamming



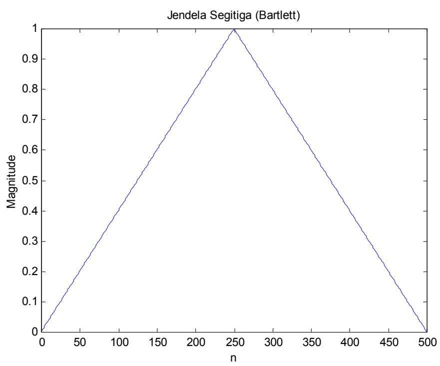
(d) Jendela Nuttall



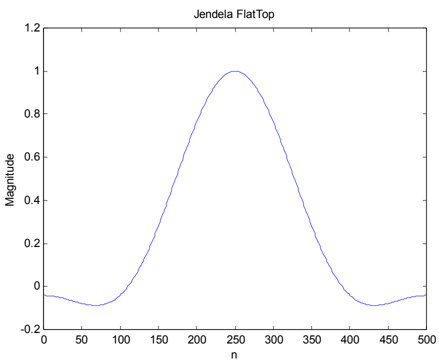
(e) Jendela Hanning



(f) Jendela Welch



(g) Jendela Segitiga (Bartlett)

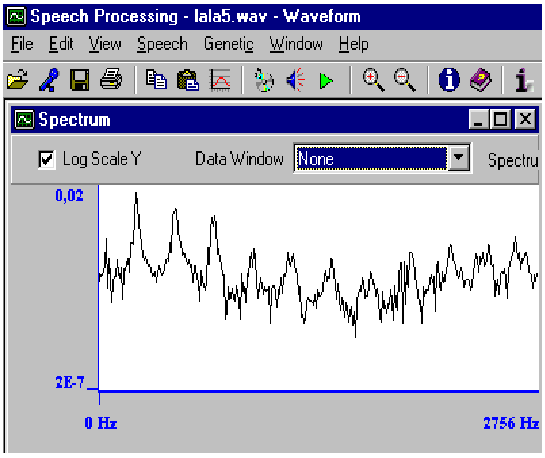


(h) Jendela FlatTop

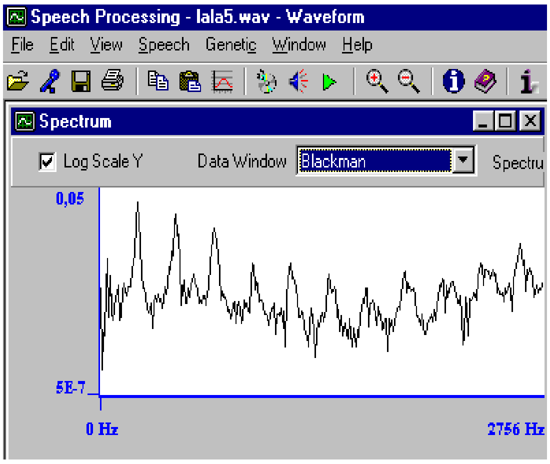
Gambar 4. Berbagai jendela yang diterapkan pada analisis spektrum

3. HASIL DAN PEMBAHASAN

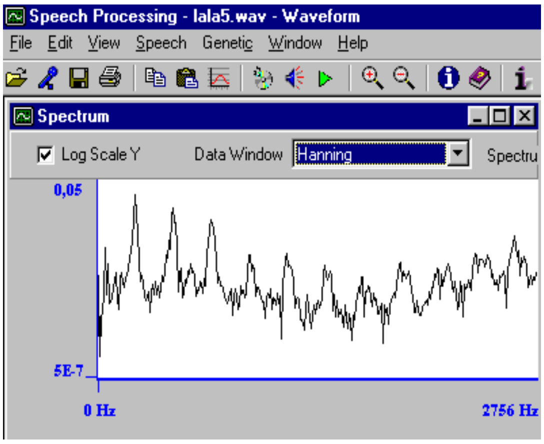
Spektrum dengan skala logaritmis dapat disajikan pada Gambar 5



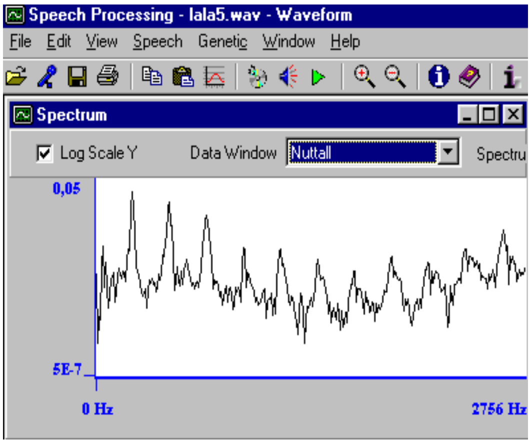
(a) Spektrum skala logaritmis dengan *non-window*



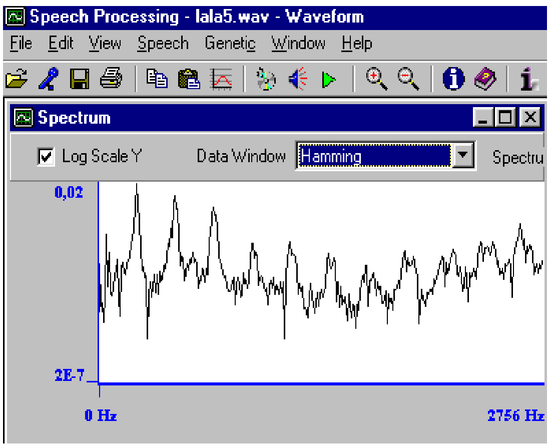
(b) Spektrum skala logaritmis dengan *window* Blackman



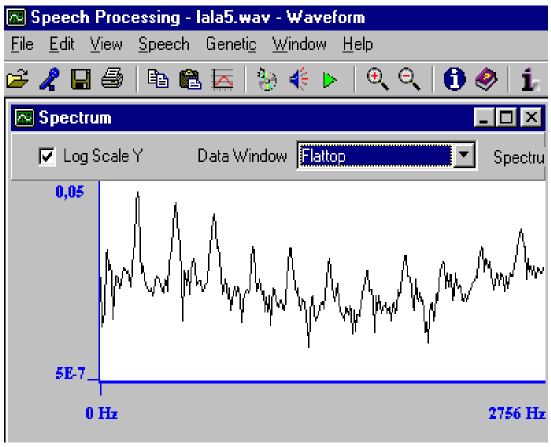
(c) Spektrum skala logaritmis dengan *window* Hanning



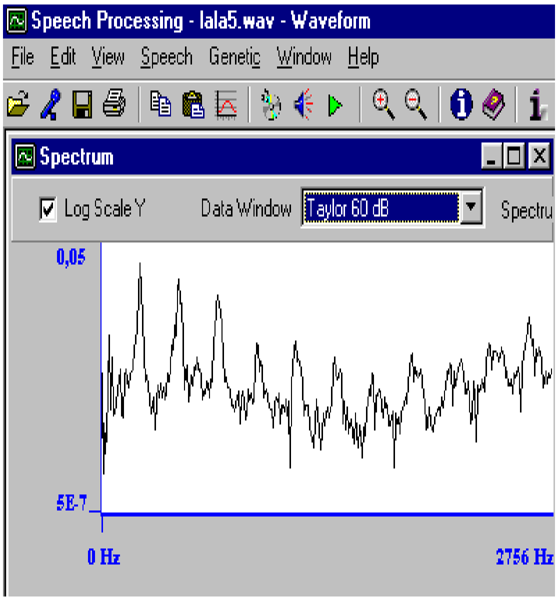
(d) Spektrum skala logaritmis dengan *window* Nuttal



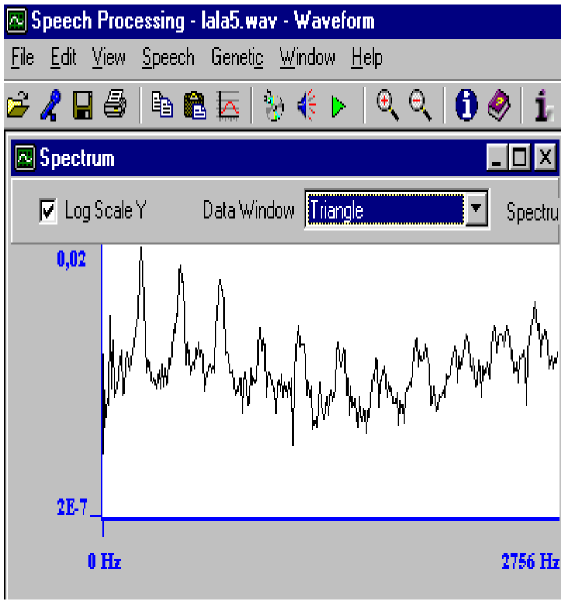
(e) Spektrum skala logaritmis dengan *window* Hamming



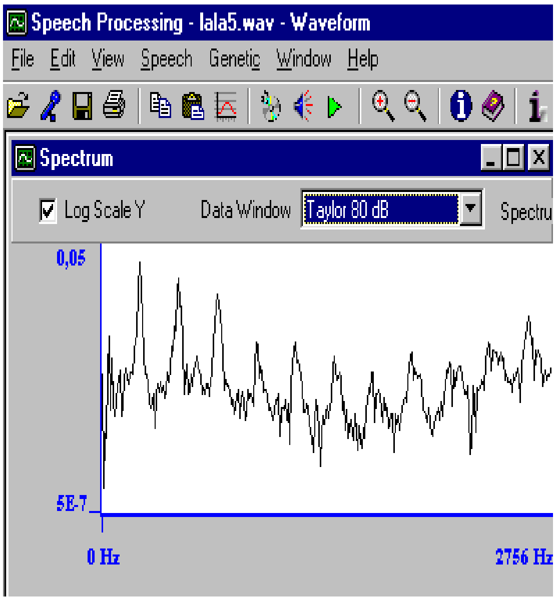
(f) Spektrum skala logaritmis dengan *window* Flattop



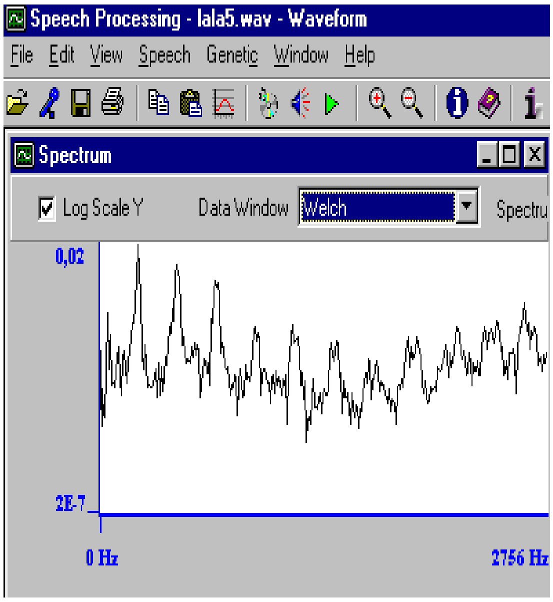
(g) Spektrum skala logaritmis dengan *window* Taylor 60 dB



(h) Spektrum skala logaritmis dengan *window* Triangle



(i) Spektrum skala logaritmis dengan *window* Taylor 80 dB



(j) Spektrum skala logaritmis dengan *window* Welch

Gambar 5. Tampilan spektrum skala logaritmis dengan menggunakan berbagai *window*.

Pada Tabel 2 berikut ditunjukkan perbandingan efek pemilihan *window* terhadap spektrum sinyal tutur.

Tabel 2. Perbandingan pemilihan *window* terhadap analisis spektrum frekuensi non-linear

No.	Nama window	Tingkat sidelobe tertinggi	Sidelobe fall-off	Keterangan
1.	(tanpa window)	13.3dB	6dB/Oct	Lobe utama yang mungkin yang paling sempit, tetapi terjadi kebocoran yang cukup besar.
2.	Bartlett (Triangle)	26.6dB	12dB/Oct	Welch memiliki suatu lobe utama yang sedikit lebih sempit dibanding Bartlett, tetapi sidelobe-nya agak lebih tinggi.
3.	Welch (parabola)	?	12dB/Oct	Welch memiliki suatu lobe utama yang sedikit lebih sempit dibanding Bartlett, tetapi sidelobe-nya agak lebih tinggi.
4.	Hanning	31.5dB	18dB/Oct	Lobe utama yang agak sempit, reduksi kebocoran jauh asimptotik yang bagus.
5.	Hamming	43dB	6dB/Oct	Lobe utama yang agak sempit, reduksi kebocoran dekat yang bagus.
6.	Blackman	58.11dB	18dB/Oct	Lobe utama yang lebih luas dibanding Hanning dan Hamming, reduksi kebocoran dekat dan jauh yang bagus.
7.	Nuttall	93.3dB	18dB/Oct	Reduksi kebocoran dekat yang lebih bagus daripada window Blackman tetapi lobe utama yang lebih luas.
8.	FlatTop	43.2dB	6dB/Oct	Window Flattop digunakan ketika menghitung amplitude spektrum suatu puncak dengan energi spektrum mendekati aslinya dan menyediakan akurasi amplitude yang bagus. Window Flattop memiliki suatu lobe utama yang luas dengan suatu 'flat top' melewati dua garis spektrum. Ini berarti bahwa setiap frekuensi sinyal yang berada di antara dua garis spektrum akan terlihat pada garis spektrum yang terdekat tanpa kehilangan amplitude. Window FlatTop digunakan ketika pengukuran komponen spektrum amplitude dari frekuensi arbitrer dianggap penting.

4. SIMPULAN

Berdasarkan hasil dan pembahasan dapat ditarik kesimpulan sebagai berikut:

1. Suatu skala logaritmis akan memperluas daerah frekuensi yang rendah dari spektrum, dan mempersempit daerah frekuensi yang tinggi pada tampilan.
2. Skala frekuensi logaritmis membutuhkan resolusi frekuensi yang sangat tinggi pada frekuensi yang rendah, oleh karena itu dibutuhkan suatu FFT yang jauh lebih besar guna mendapatkan resolusi frekuensi yang dibutuhkan.
3. Penerapan fungsi berbagai *window* terhadap data dapat membantu mengurangi efek kebocoran yang terjadi pada spektrum frekuensi.
4. Metode ini berjalan lebih cepat jika *point* data merupakan kelipatan dua (128, 256, 1024, 2048, atau 4096, dan seterusnya).
5. Memilih suatu resolusi frekuensi yang tepat dan resolusi waktu yang sesuai menjadi suatu kesesuaian antara kebutuhan untuk mengamati *detail* frekuensi yang baik dalam spektrum dengan kebutuhan untuk mengamati variasi waktu yang cepat dalam spektrum.

DAFTAR PUSTAKA

[1]. Rabiner, L., R., and R. W. Schafer, **"Digital Processing of Speech Signals**". Prentice Hall, 1978.

[2]. Sugamura, N., and F. Itakura, "**Speech Analysis and Synthesis Methodes Developed at ECL in NTT- from LPC to LSP**", Speech Communication, 5:199-215, 1986.

[3]. Fallside, F., and W.A. Woods, "**Computer Speech Processing**", Cambridge University Engineering Department, Printice Hall, 1985.