

MODUL 5

EKSTRAKSI CIRI SINYAL WICARA

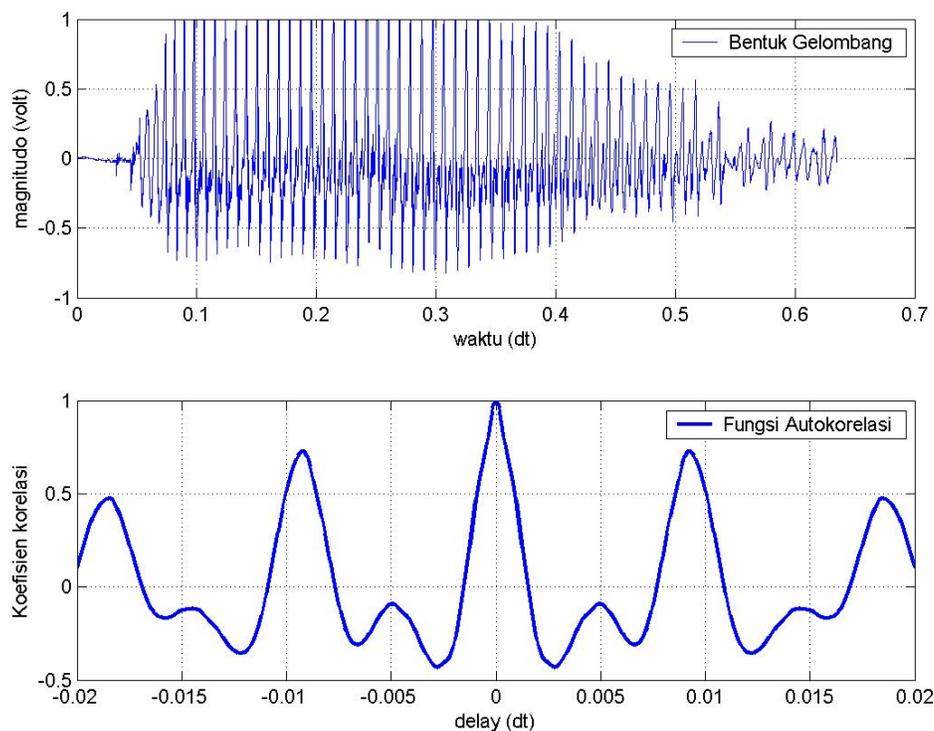
I. TUJUAN

- Mahasiswa mampu melakukan estimasi frekuensi fundamental sinyal wicara dari pengamatan spektrumnya dan bentuk gelombangnya
- Mahasiswa mampu menggambarkan bentuk *kuefrensi* dan *cepstrum* sinyal wicara

II. DASAR TEORI

2.1. Auto Korelasi Sinyal Wicara

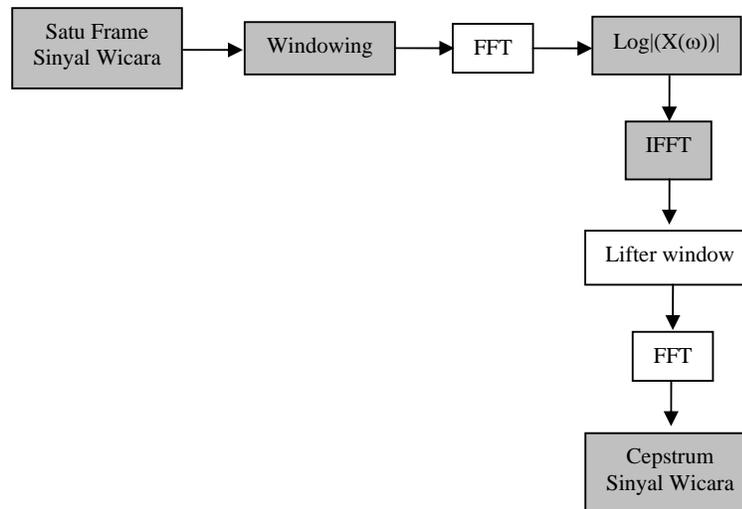
Tujuan melakukan estimasi dalam domain waktu adalah untuk mendapatkan nilai autokorelasi sinyal wicara. Nilai auto korelasi suatu sinyal wicara akan menunjukkan bagaimana bentuk gelombang itu membentuk suatu korelasi pada dirinya sendiri sebagai fungsi perubahan waktu ke waktu. Bentuk-bentuk yang sama atau mirip pada setiap *delay* waktu tertentu menunjukkan perulangan bentuk atau periodisitas pola sinyal wicara. Dengan demikian akan dapat kita lakukan estimasi nilai frekuensi fundamentalnya.



Gambar 1. Sinyal wicara dalam bentuk gelombang, dan fungsi auto korelasi

2.2. Analisa *Spectral* Sinyal Wicara

Proses ekstraksi ciri sinyal wicara didasarkan pada sebuah diagram blok yang cukup populer seperti berikut.



Gambar 2. Diagram blok ekstraksi ciri sinyal wicara.

Dengan mengikuti diagram blok diatas, kita akan mendapatkan langkah-demi langkah ekstraksi ciri. Pada Sub Bab ini pembicaraan berkisar pada proses melihat *short time fourier analysis* dari sinyal wicara, atau yang juga kita kenal sebagai *power spectral density* (PSD) sebuah sinyal wicara pada durasi atau frame tertentu. Dengan mengetahui bentuk PSD sinyal wicara kita akan mampu melakukan ekstraksi ciri sinyal wicara tersebut. Target akhir sub bab ini adalah mencari bentuk *power spectral density* (PSD).

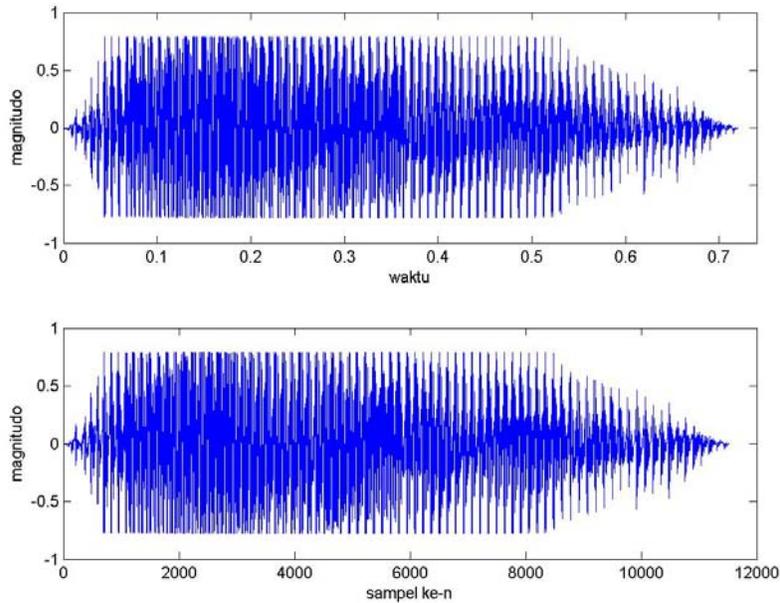
Sebelum proses pada Gambar diatas dilakukan ada baiknya kita melihat gambaran sebuah sinyal wicara yang telah kita simpan dalam bentuk file “*a.wav*”. Setelah kita dapatkan bentuknya dalam domain waktu seperti pada Gambar 3 bagian atas, selanjutnya kita coba melihatnya sebagai fungsi dari *sampling*. Dalam hal ini kita lihat bentuk sinyal wicara sesuai dengan urutan sampel yang ada. Seperti kita lihat bahwa untuk nilai sampel ke-700 sampai dengan sampel ke-8200, menunjukkan nilai magnitudo sinya yang relatif stabil.

Kita lanjutkan dengan melakukan pembentukan *frame* sebuah sinyal wicara seperti pada Gambar 4 bagian atas. Dengan melakukan *windowing* kita akan mendapatkan bentuk frame sinyal wicara terwindow seperti pada Gambar 4 bagian bawah. Sudah tentu kita paham untuk apa proses *windowing* dilakukan disini. Kalau kurang jelas anda pelajari lagi Modul 2 dan Modul 3. Dengan demikian tidak salah apabila kita mengambil satu

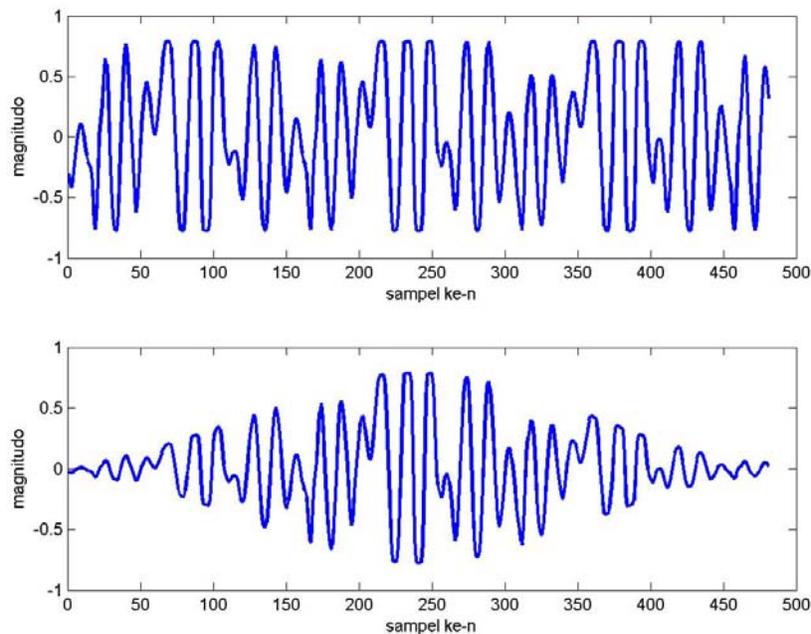
frame sinyal dari sampel ke-2000 sampai dengan sampel ke 2480. Karena dalam satu frame kita bentuk dari:

$$\begin{aligned} \text{Sampel/frame} &= (\text{sample/detik}) * (\text{detik/frame}) \\ &= 16000 * 0,06 = 480 \text{ sampel/frame} \end{aligned}$$

Hal ini dilakukan dengan menetapkan bahwa satu frame sinyal wicara sepanjang 50 ms.

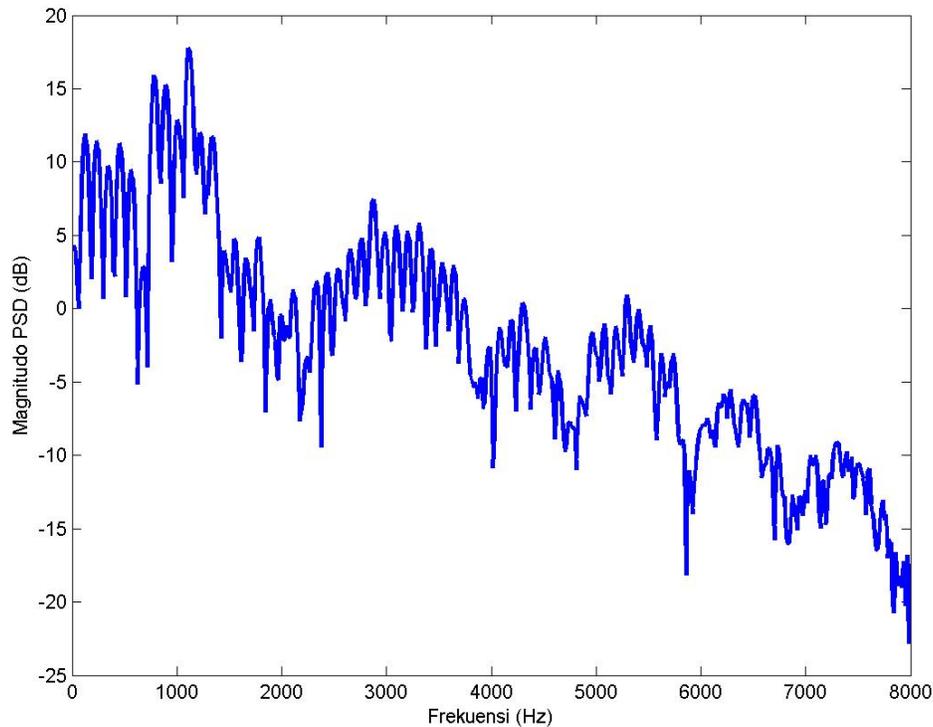


Gambar 3. Sinyal wicara dalam domain waktu dan sebagai fungsi sample ke-n



Gambar 4. Satu frame sinyal wicara dalam domain waktu

Proses dilanjutkan dengan melakukan transformasi sinyal ke dalam domain frekuensi. Dengan menggunakan fft dan proses logaritmik akan kita dapatkan nilai *power spectral density* (PSD) sinyal wicara seperti pada Gambar 5.



Gambar 5. Power spectral density sinyal wicara

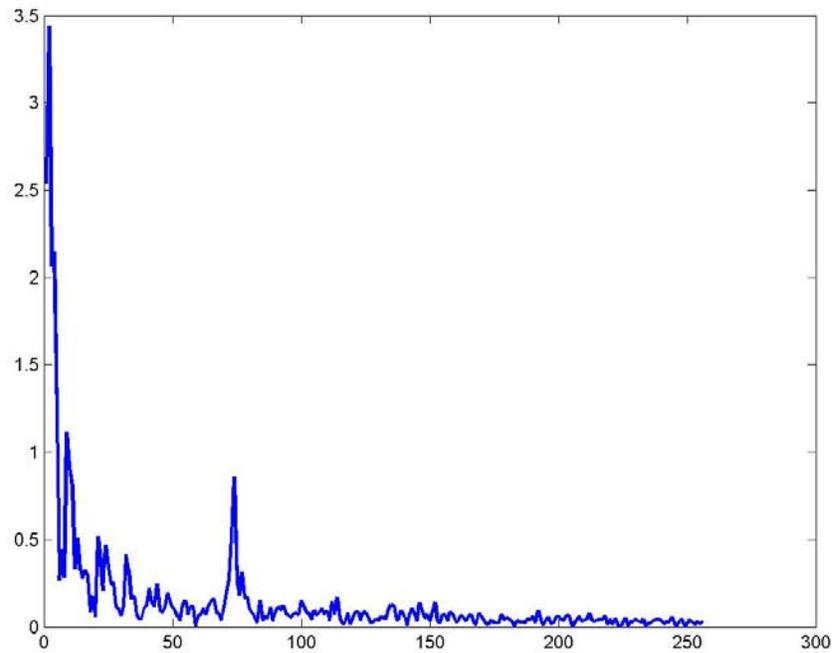
2.3. Cepstrum Sinyal Wicara

Cepstrum $c(\tau)$ didefinisikan sebagai *inverse* transformasi Fourier pada *short-time* nilai logaritmik spektrum amplitudo sebuah sinyal, $|X(\omega)|$. Jika log amplitudo spektrum tersusun dari banyak spasi harmonik yang teratur, maka analisis Fourier pada spektrum ini akan menunjukkan sebuah puncak yang berhubungan dengan jarak antar harmonisa tersebut, yang juga dikenal sebagai frekuensi fundamental.

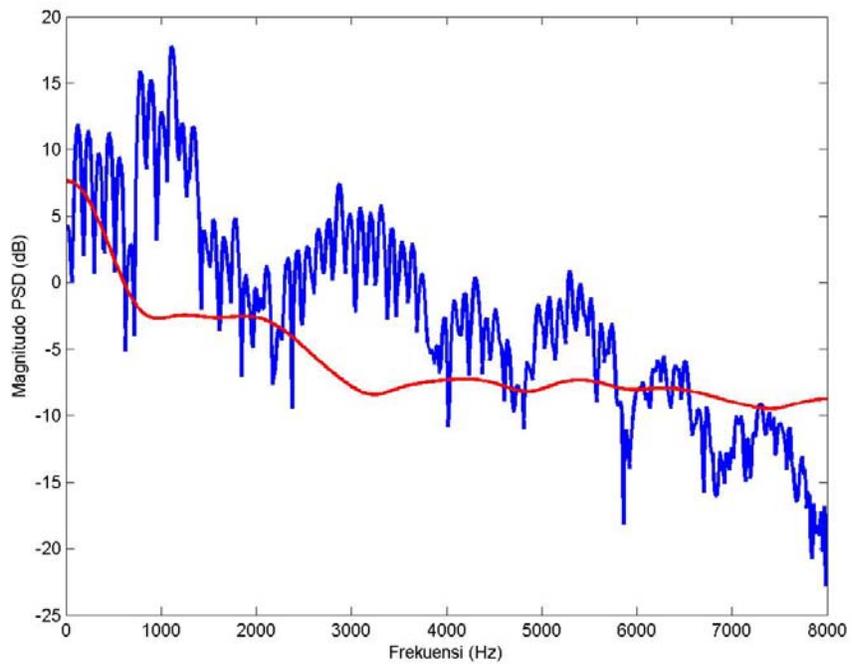
Dengan melakukan proses mengikuti diagram blok pada Gambar 2, maka proses berikutnya adalah melakukan **ifft** nilai PSD yang sudah diperoleh. Proses ini menghasilkan sebuah nilai *kuefreensi* dari sinyal wicara. Gambaran dari nilai *kuefreensi* seperti pada Gambar 6 berikut ini.

Langkah ini dilanjutkan dengan lifter window, yaitu proses pengambilan sebagian saja dari nilai *kuefreensi* sinyal wicara yang terdapat pada Gambar 6. Dalam hal ini kita bisa mengambil 16 sampai 20 nilai pertama yang muncul. Nilai ini sudah cukup representatif untuk mendapatkan ciri dari sinyal wicara.

Yang terakhir dari langkah kita adalah melakukan transformasi fourier pada hasil *lifter window* yang selanjutnya akan menghasilkan sebuah *cepstrum* dari sinyal wicara yang kita olah.



Gambar 6. Gambaran kuefrensi sinyal wicara "a.wav"



Gambar 7. Perbandingan nilai psd dan cepstrum sinyal wicara "a.wav"

Dari Gambar 7 diatas kita dapatkan gambaran bentuk cepstrum sinyal wicara menunjukkan pola yang mirip dengan pola PSD. Disini tampak bahwa bentuk cepstrum merupakan penghalusan dari bentuk PSD sebuah sinyal wicara.

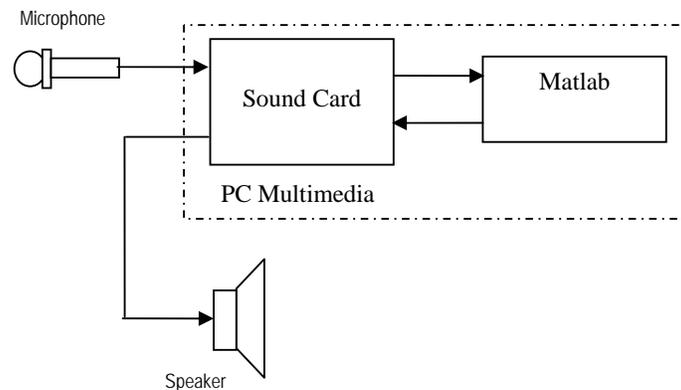
III. PERANGKAT YANG DIPERLUKAN

- 1 (satu) buah *PC Multimedia* lengkap *sound card* dan *microphone*
- Satu perangkat lunak *Matlab under windows*

IV. LANGKAH-LANGKAH PERCOBAAN

4.1. Penataan Perangkat

Sebelum melakukan percobaan harus dilakukan penataan seperti pada Gambar 8 berikut ini.



Gambar 8. Penataan perangkat percobaan pengukuran energi sinyal wicara

PC harus dilengkapi dengan peralatan *multimedia* seperti *sound card*, *speaker active* dan *microphone*. Untuk *microphone* dan *speaker active* bias juga digantikan dengan *head set* lengkap. Sebelum anda memulai praktikum, sebaiknya anda tes dulu, apakah seluruh perangkat *multimedia* anda sudah terintegrasi dengan PC.

Untuk membantu anda dalam menjalankan praktikum ini, disini diberikan beberapa contoh perintah dasar yang akan anda perlukan pada proses ekstraksi ciri sinyal wicara. Coba anda pelajari beberapa *library* yang ada di Matlab berikut ini.

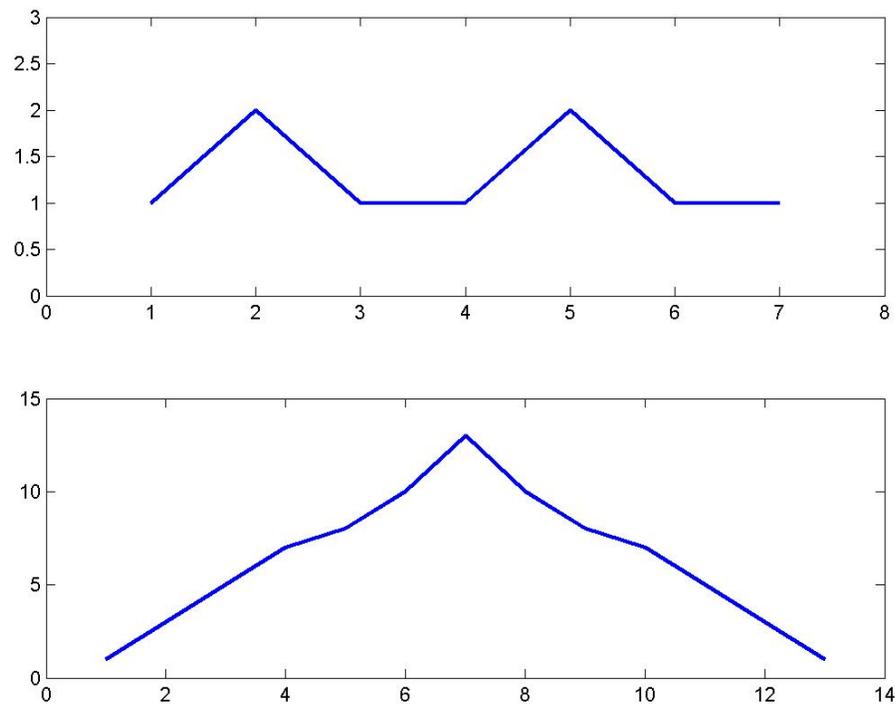
Korelasi

$C = \text{XCORR}(A,B)$, where A and B are length M vectors ($M > 1$), returns the length $2 \cdot M - 1$ cross-correlation sequence C. If A and B are of different length, the shortest one is zero-padded. C will be a row vector if A is a row vector, and a column vector if A is a column vector.

Contoh pemakaiannya seperti berikut:

```
clear all;
x=[1 2 1 1 2 1 1];
subplot(211)
plot(x)
axis([0 8 0 3])
xx=xcorr(x);
subplot(212)
plot(xx)
```

Hasilnya adalah seperti berikut ini.



Gambar 9. Contoh Proses Korelasi

Nilai sample ke-n

Membaca sebuah nilai pada sample tertentu

```
Fs=12000;
[x,Fs]=wavread('file_a');
x_240=x(2000:2240);
```

Dengan program ini akan didapatkan nilai x_240 sebagai nilai dari variabel x pada sampel 2000 sampai dengan sampel ke 2240.

Transformasi dari domain waktu ke frekuensi

[H,W] = FREQZ(B,A,N) returns the N-point complex frequency response vector H and the N-point frequency vector W in radians/sample of the filter:

Contoh pemakaiannya

```
X=freqz(x_wind);
```

Langkah ini akan mentransormasikan nilai x_wind menjadi bentuk domain frekuensi yang hasilnya disimpan pada variabel X. Kita harus tahu bahwa nilai yang disimpan dalam X merupakan bentuk kompleks.

Transformasi dari domain frekuensi ke domain waktu

IFFT(X) is the inverse discrete Fourier transform of X.

Contoh pemakaiannya

```
x_k = abs(iff(X_dB));
```

Langkah ini akan mentransformasikan dari domain frekuensi ke dalam domain waktu, operasinya kita kenal sebagai invers fast fourier transform.

4.3. Estimasi Frekuensi Fundamental dengan Domain Waktu

Pada percobaan ini kita coba untuk melakukan estimasi nilai frekuensi fundamental dengan menggunakan domain waktu. Dalam hal ini kita memanfaatkan fungsi auto korelasi suatu sinyal wicara. Langkah-langkahnya adalah sebagai berikut.

1. Buat program untuk memanggil file 'a.wav' .
2. Tampilkan hasilnya dalam domain waktuk sebagai *figure(1)*
3. Ambil beberapa sampel saja dari nilai a.wav yang sudah anda panggil. Dalam hal ini anda bisa menentukan sebanyak 200, 300, atau terserah anda.
4. Tambahkan program untuk mencari bentuk auto korelasinya.

5. Hitung jarak antar puncak satu dengan puncak berikutnya yang paling dekat. Coba aati juga jarak puncak ke 0, puncak ke 1, puncak ke 2. Arahkan pandangan anda geser ke kiri, perhatikan puncak ke-1, puncak ke-2, dst. Coba anda hitung jarak antar puncak tersebut. Tentu saja nilainya dalam satuan waktu, bisa mili detik atau yang lain.

4.3. Karakteristik Power Spectral Density

1. Panggil sinyal wicara "a.wav".
2. Tampilkan dalam domain waktu dan dalam bentuk sinyal sebagai fungsi sample ke-n.
3. Buat sebuah frame sinyal sepanjang 240 sampel, dalam hal ini anda bisa memilih bagian sinyal yang representatif, yang memiliki nilai cukup stabil selama durasi frame yang anda tetapkan panjangnya.
4. Transformasikan ke dalam domain frekuensi, anda bisa menggunakan perintah `fft` atau yang lain.
5. Rubah nilainya dalam parameter dB, $20 \log_{10}(\text{abs}(X))$

4.4 Mencari Bentuk Cepstral Sinyal Wicara

1. Lakukan hal yang sama seperti pada langkah percobaan 4.3, usahakan tampilan yang sudah anda peroleh anda *hold* supaya gambarnya tidak hilang.
2. Tambahkan program untuk melakukan invers `fft` anda.
3. Tampilkan bentuk kuefrensi yang telah anda peroleh, jika terjadi bentuk pencerminan pada sisi kiri dan sisi kanan, ambil sisi kiri saja dan tampilkan kembali dengan mengatur nilai sample tertentu saja yang anda olah.
4. Lakukan proses lifter window dengan cara ambil 16 titik nilai kuefrensi pertama saja. Jika anda ragu anda bisa mengambil 20 titik nilai kuefrenensi pertama.
5. Transformasikan kembali ke domain fekuensi, dan usahakan gambarnya menjadi satu dengan hasil tampilan PSD yang telah anda peroleh.

5. ANALISA DATA DAN TUGAS

1. Buat program GUI untuk program-program yang sudah anda kerjakan dari langkah 4.2 sampai 4.4 pada percobaan ini.