

*MAKALAH*  
**ANALISIS CEPSTRUM SINYAL SUARA**



**Disusun Oleh:**  
**NENI ARYANI**  
*L2F 300 543*

**JURUSAN TEKNIK ELEKTRO**  
**FAKULTAS TEKNIK UNIVERSITAS DIPONEGORO**  
**S E M A R A N G**  
**2 0 0 2**

## DAFTAR ISI

JUDUL .....	1
ABSTRAK .....	1
1. Pendahuluan .....	1
2. Pemodelan Sinyal Suara .....	1
3. Alihragam Fourier Diskrit .....	2
4. Analisis Cepstrum .....	3
5. Pembingkai dan Penjendelaan .....	5
6. Kesimpulan .....	6
DAFTAR PUSTAKA	

# ANALISIS CEPSTRUM SINYAL SUARA

## Abstrak

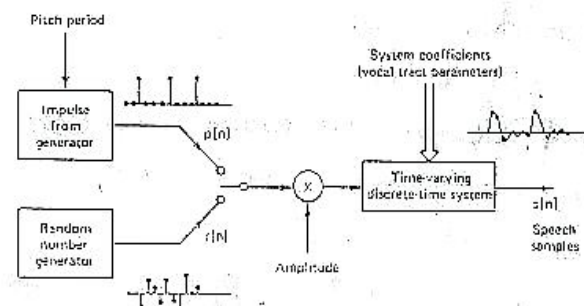
*Cepstrum merupakan transformasi Fourier dari logaritma spektrum suatu sinyal. Dengan operasi logaritma, analisis cepstrum digunakan untuk menganalisis sinyal suara yang diasumsikan sebagai hasil konvolusi dari jalur vokal dengan impuls periodik (untuk vokal) yang dibangkitkan oleh generator. Hasil dari cepstrum dikenal dengan sebutan pitch. Periode pitch digunakan untuk mengatur selang waktu proses pembangkitan sinyal suara.*

## 1. Pendahuluan

Sinyal ucapan atau suara yang dihasilkan oleh manusia berubah-ubah sesuai dengan perubahan tekanan udara, dengan demikian sinyal ucapan ini dapat dianggap berubah secara alami dan bentuk sinyal yang dihasilkan adalah suatu bentuk sinyal yang berbeda-beda untuk setiap manusia. Ini adalah suatu bentuk sifat yang unik dari sinyal ucapan, yang memungkinkan kita untuk mengenali suara masing-masing manusia.

Sebagai tahap awal penelitian sinyal ucapan adalah dengan menganalisis sinyal ucapan sebagai proses pengenalan parameter dan karakteristik sinyal ucapan. Salah satu parameter yang ada pada sinyal ucapan adalah *pitch*. Penentuan *pitch* sinyal ucapan suara manusia menggunakan metode analisis *cepstrum*.

## 2. Pemodelan Sinyal Ucapan



Gambar 1 Model Waktu Diskrit Produksi Sinyal Ucapan

Gambar 1 adalah penggambaran bentuk gelombang sinyal ucapan pada waktu diskrit. Pada model ini, contoh sinyal ucapan dianggap bentuk output konvolusi dari variasi waktu sistem waktu diskrit dan resonansi sistem jalur vokal. Bentuk permodelan saklar antara impulse periodic dan random noise tergantung pada tipe suara yang dihasilkan. Sinyal ucapan diasumsikan sebagai hasil dari konvolusi 2 buah sinyal.

$$s(n) = v(n) * p(n) \quad \text{untuk } 0 \leq n \leq L-1 \quad (2.1)$$

dengan  $v(n)$  adalah respon impuls jalur vokal dan  $p(n)$  adalah periodic (untuk sinyal suara ucapan) atau random noise ( untuk sinyal ucapan tak bersuara).

Pada kenyataannya, model dari persamaan (2.1) terdapat efek diskontinuitas pada awal dan akhir sinyal , sehingga diperlukan penjedeleaan untuk memperhalus awal dan akhir sinyal. Sehingga input sistem homomorphic dekonvolusi adalah :

$$x(n) = w(n)s(n)$$

jika  $w(n)$  berubah lambat terhadap perubahan  $v(n)$ , maka analisis dapat dituliskan,

$$x(n) = v(n) * p_w(n)$$

dimana

$$p_w(n) = w(n) p(n)$$

### 3. Alihragam Fourier Diskret ( *Discrete Fourier Transform* )

Agar suatu sinyal dalam kawasan waktu mudah untuk dianalisis, maka sinyal tersebut dapat dialihragamkan terlebih dahulu ke dalam kawasan frekuensi. Untuk itu dapat digunakan Alihragam Fourier (*Fourier Transform*).

Apabila suatu sinyal diskret dinyatakan dengan runtun  $X(n)$ , maka Alihragam Fourier-nya adalah  $X(e^{j\omega})$ . Besar  $X(e^{j\omega})$  didapat dari persamaan (2.2), sedangkan Alihragam Fourier Balik-nya (*Inverse Fourier Transform*) didapat dari persamaan (2.3) di bawah ini :

$$X(e^{j\omega}) = \sum_{-\infty}^{\infty} x(n) \cdot e^{-j\omega n} \quad (2.2)$$

$$X(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) \cdot e^{j\omega n} d\omega \quad (2.3)$$

Agar alihragam tersebut dapat dihitung dengan komputer maka jumlah runtun  $x(n)$  yang akan diolah harus terbatas. Dalam hal ini dapat digunakan persamaan Alihragam Fourier Diskret (*Discrete Fourier Transform/ DFT*) yang menerapkan perhitungan pada  $x(n)$  dengan selang tertentu.

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot W^{kn} \quad ,k = 0, 1, 2, 3 \dots N-1 \quad (2.4)$$

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \cdot W^{-kn} \quad ,k = 0, 1, 2, 3 \dots N-1 \quad (2.5)$$

dimana  $W = e^{-j2\pi / N}$

Persamaan (2.4) menunjukkan DFT dan persamaan (2.5) menunjukkan Alihragam Fourier Diskret-Balik (*Inverse Discrete Fourier Transform / IDFT*) untuk N buah data. Apabila data yang diolah semakin banyak (N makin besar), maka akan membutuhkan waktu yang semakin lama pula.

#### 4. Analisis Cepstrum

Konsep dari analisis *cepstrum* adalah perhitungan melalui metode *Real Cepstrum* (RC) [3], RC pada sinyal suara  $s(n)$ , didefinisikan sebagai :

$$c_s(n) = \text{IDFT} \{ \log | \text{DFT} \{ s(n) \} | \} \quad (2.6)$$

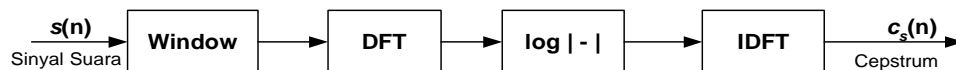
dengan menganggap bahwa DFT dari  $s(n)$  adalah  $S(\omega)$ , maka persamaan menjadi :

$$c_s(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log | S(\omega) | e^{j\omega n} d\omega \quad (2.7)$$

dimana  $n$  = data cuplik untuk deret genap.

$C_s(\omega) = \log | S(\omega) |$  adalah real dan genap.

Untuk proses perhitungan dari RC dapat dilihat pada Gambar 2.



Gambar 2. Blok analisis perhitungan Real *Cepstrum* [3]

Sinyal suara dalam bentuk sinyal diskret  $s(n)$  digunakan sebagai masukan, dengan data yang ada dilakukan penjendelaan (*window*), kemudian data di-DFT-kan.

$$s(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k)$$

$$S(e^{j\omega}) = \sum_{n=-\infty}^{\infty} s(n)e^{-j\omega n} \quad (2.8)$$

Tujuan DFT pada sinyal masukan, adalah agar sinyal masukan yang dianggap sebagai hasil operasi konvolusi (permodelan sintesis suara Gambar 1) dapat diubah ke dalam bentuk perkalian. Dengan menerapkan definisi konvolusi pada sinyal masukan yang dianggap sebagai keluaran, maka

$$S(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \left[ \sum_{k=-\infty}^{\infty} x(k)h(n-k) \right] e^{-j\omega n}$$

$$S(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \left[ \sum_{k=-\infty}^{\infty} x(k)h(n-k) \right] e^{-j\omega n}$$

$$S(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \left[ x(k)e^{-j\omega k} h(n-k)e^{-j\omega(n-k)} \right]$$

$$S(e^{j\omega}) = \sum_{k=-\infty}^{\infty} \left[ x(k)e^{-j\omega k} \sum_{n=-\infty}^{\infty} h(n-k)e^{-j\omega(n-k)} \right]$$

$$S(e^{j\omega}) = X(e^{j\omega}) \cdot H(e^{j\omega}) \quad (2.9)$$

Dan menggunakan fungsi logaritma, dimana logaritma suatu perkalian 2 buah komponen, sama dengan penjumlahan logaritma masing-masing komponen tersebut.

$$S(e^{j\omega}) = X(e^{j\omega}) \cdot H(e^{j\omega})$$

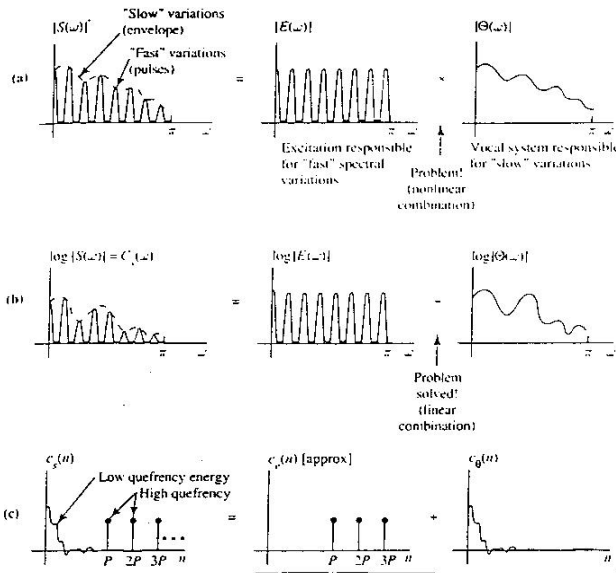
$$\text{Log}S(e^{j\omega}) = \text{Log}[X(e^{j\omega}) \cdot H(e^{j\omega})]$$

$$\text{Log}S(e^{j\omega}) = \text{Log}[X(e^{j\omega})] + \text{Log}[H(e^{j\omega})] \quad (2.10)$$

Maka sinyal suara dapat terpisahkan menjadi komponen-komponen sinyal yang membentuknya. Untuk sinyal *cepstrum* dipilih sinyal real dan deret genap. Fungsi IDFT digunakan untuk mengembalikan sinyal ke ranah waktu, sehingga sinyal *cepstrum*  $c_s(n)$  akan diperoleh. Dengan memperoleh sinyal

*cepstrum*, maka periode pitch akan diperoleh dengan mendeteksi periode munculnya puncak sinyal *cepstrum* tertinggi.

Menurut *Noll* cepstrum akan dapat dimengerti lebih jelas apabila dapat melihat proses perhitungan analisa cepstrum secara satu persatu melalui tampilan komponen sinyal pada gambar .... dengan menganggap persamaan  $S(\omega) = E(\omega)\Theta(\omega)$ , sebagai acuan sampai mencari nilai  $c_s(n)$ . Dimana antara komponen sinyal variasi lambat (*envelope*) dan komponen sinyal variasi cepat (*pulses*) akan terlihat terpisah pada akhir perhitungan, dengan menganggap komponen *envelope* adalah  $|\Theta(\omega)|$  dan komponen *pulses* adalah  $|E(\omega)|$  maka proses tersebut akan terlihat sebagai berikut:



Gambar 3. Dasar Proses Penggambaran Analisis Cepstrum

## 5. Pembungkaihan dan Penjendelaan

Salah satu sifat dari sinyal ucapan adalah berubah terhadap waktu (*non stationary*), sehingga untuk menganalisa sinyal ucapan diperlukan pembatasan waktu sinyal ucapan yaitu pembungkaihan dan penjendelaan.

Pembungkaihan adalah pembatasan sinyal suara yang digunakan untuk pergeseran penjendelaan (*windowing*).

Pada tahap ini tiap sinyal didalam masing-masing bingkai diatas dijenelakan satu per satu. Kegunaan dari penjenelakan ini adalah untuk mengurangi kesenjangan (*discontinuitas*) sinyal pada awal dan akhir bingkai (*frame*). Dengan penjenelakan ini sinyal akan meruncing menuju nol pada awal dan akhir bingkai. Adapun persamaan penjenelakan adalah:

$$x(n) = s(n) \cdot w(n), \quad 0 \leq n \leq N - 1 \quad (2.11)$$

Untuk metode analisis cepstrum pada tugas akhir ini digunakan jendela Hamming sebagai berikut:

$$w(n) = 0,54 - 0,46 \cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \leq n \leq N - 1 \quad (2.12)$$

Jendela hamming dipilih, karena jendela ini sering dipakai pada analisa sinyal untuk sistem pengenalan pola (*pattern recognition*).

## 6. Kesimpulan

1. Analisis cepstrum bisa digunakan untuk menganalisis sinyal suara.
2. Hasil dari analisis cepstrum adalah pitch.

## DAFTAR PUSTAKA

1. Abraham Peled dan Bede Liu, "*Digital Signal Processing – Teory, Design and Implementation*", John Wiley and Sons, 1996.
2. Alan V. Oppenheim, "*Applications of Digital Signal Processing*", Prentice – Hall Inc, 1978.
3. Alan V. Oppenheim, "*Speech Analysis – Synthesis System Based on Homomorphic Filtering*", J. Acoust. Soc. Am., vol 45, 1969.
4. Alan V. Oppenheim dan Ronald W. Schaffer, "*Digital Signal Processing*", Prentice-Hall of India, 1998.
5. Alan V. Oppenheim dan Alan S. Willsky dengan S. Hamid Nawab, "*Sinyal & Sistem*", Jilid I, Erlangga, 1997.
6. Douglas O'Shaughnessy, "*Speech Communication, Human and Machine*", Addison – Wesley Publishing Company, 1987.
7. Gordon E. Peaton, "*Voice Processing*", Mc. Graw Hill, 1992.



