

Muhammad Raihan N (182019001)  
LITS PTU

- a. - Pembentukan ucapan dimulai dengan adanya hembusan udara yang menghasilkan oleh paru-paru.
- Cara kerjanya mirip seperti piston atau pompa yang ditekan untuk menghasilkan tekanan udara.
  - Ketika vocal cord berada dalam keadaan tegang, aliran udara akan menyebabkan terjadinya vibrasi pada vocal cord dan menghasilkan bunyi ucapan yang disebut voiced speech sound.
  - Ketika vocal cord berada dalam keadaan santai, aliran udara akan melalui daerah yang sempit pada vocal tract dan menyebabkan terjadinya turbulensi, sehingga menghasilkan suara yang dikenal sebagai unvoiced sound.
  - Ucapan dihasilkan sebagai rangkaian komponen-komponen bunyi-bunyi yang ~~berbeda~~ pembentukannya. Setiap komponen bunyi yang ~~berbeda~~ berbeda dibentuk oleh perbedaan posisi, bentuk, serta ukuran dari alat-alat ucap manusia yang berubah-ubah selama terjadinya proses produksi ucapan.
- b. - Silence (S) : keadaan diam, sunyi, tidak ada suara yang dihasilkan.
- Unvoiced (U) : keadaan pada saat pita suara tidak terdengar,  $\Rightarrow$  sinyal ucapan acak
  - Voiced (V) : keadaan pada saat pita suara dirangsang  $\Rightarrow$  sinyal ucapan quasi-periodik
- c. • Sebuah mic yang mengubah gelombang tekanan suara  $p(t)$  ke dalam sebuah sinyal ~~analog~~ listrik  $x_c(t)$ .
- Kemudian, sebuah pen-sampling pada interval waktu  $T_c$  ~~yang~~ dengan nilai-nilai tegang  $x_c(nT_c) = x_c(n)$ , dan terakhir sebuah pengubah analog ke digital (A/D) yang mengkuantitasikan tiap-tiap  $x_c(n)$ ,  $n=0,1,\dots,N-1$  ke dalam sebuah bilangan yang spesifik, biasanya 16 bit.
  - Sebuah pensampling dan pengubah A/D terdapat didalam computer audio cards ~~super~~



② a. FFT : Suatu sinyal direpresentasikan sebagai besaran fisik yang ~~mengubah~~ berubah-ubah menurut waktu, ruang, atau variabel lainnya.

- Power spectral : Energi pada sinyal suara ~~berasal~~ berasal dari udara yang ditimbulkan pada paru-paru dan diteruskan melewati kerongkongan.
- Pitch : merupakan periode ~~pengulangan~~ pengulangan pulsa yang disebabkan oleh membuka dan menutupnya pita suara.
- Formant : ~~Frekuensi~~ Frekuensi resonansi yang terjadi ditentukan oleh bentuk dan panjang vocal tract.

b. Dik:  $f = c/\lambda$

$c$  : ~~kecepatan~~ kecepatan suara ( 35.000 cm/sec )

suara dengan  $\lambda = 10^4$  cm memiliki frekuensi rendah  $f = 35$  Hz

suara dengan  $\lambda = 2$  cm memiliki frekuensi tinggi  $f = 17.500$  Hz

Maka:  $F_1 = c/\lambda_1 = c/(4L) = 35000/4 \times 17,5 = 500$  Hz

$F_2 = c/\lambda_2 = c/(4/3L) = 3c/4L = 3 \times 35000/4 \times 17,5 = 1500$  Hz

$F_3 = c/\lambda_3 = c/(4/5L) = 5c/4L = 5 \times 35000/4 \times 17,5 = 2500$  Hz

c.  $F_1 \approx 1000 - 1100$  (- adalah ~~kecepatan~~ diantara)

$F_2 \approx 2000 - 2100$

$F_3 \approx 2700 - 2800$

$F_4 \approx 3600 - 3700$

d.  $F_s = 11000$  Hz

Orsampling tiap frame = 25 ms

1 detik = 1000 ms

$25/1000 = 0,025$

Jawaban : jumlah sampel per frame =  $11000 \times 0,025 = 275$  sampel/frame

Jumlah frame =  $\frac{1000}{25} = 40$  frame/s

e. Dik: terdapat 25 gelombang

$5,030 - 4,970 = 0,06$

Jawab:  $\frac{25}{0,06} = 416,66$  Hz

### ③ a. - Cara kerja Blok diagram (LPC)

- Preemphasis : sinyal suara yang telah diubah menjadi sinyal digital,  $s(n)$ , dilewatkan pada sebuah filter yang berorde rendah.
- Frame Blocking : setelah di preemphasis, diblok menjadi beberapa bagian dg jumlah sampel  $N$ , dan tiap bagian dipisahkan dg sejumlah  $M$  sampel.
- Windowing : melakukan proses window pada setiap bagian sinyal yang telah dibuat sebelumnya.
- Analisa Autokorelasi : berfungsi untuk mendapatkan nilai koefisien auto korelasi dari setiap frame dari hasil windowing.
- Analisa LPC : semua nilai autokorelasi yang telah dihitung pada tahap sebelumnya akan diubah menjadi parameter LPC.
- Pengubahan parameter LPC : diturunkan dari koefisien LPC adalah koefisien cepstral (LPC,  $c(m)$ ).
- Cara kerja Blok diagram MFCC
  - Remove DC : untuk menghitung rata-rata dari data sampel suara, dan mengurangkan nilai setiap sampel suara dg nilai rata-rata.
  - Pre-emphasize filtering : mempertahankan frekuensi tinggi pada sebuah spectrum, yang umumnya tereliminasi pada saat proses produksi suara.
  - Frame blocking : sinyal ucapan yang terdiri dari  $S$  sampel ( $c(x)$ ) dibagi menjadi beberapa frame  $N$  dipisahkan oleh  $M$ .
  - Windowing :
  - FFT : Sebuah metode yang memungkinkan untuk melakukan analisa terhadap spectral properties dari sinyal yang diinputkan.
  - Mel Frequency Warping : menggunakan filterbank bertipe an untuk mengikuti aliran energi di frequency band tertentu.
  - DCTC : mengubah ke dalam domain waktu.
  - Cepstral Filtering : menghasilkan spektrum hasil dari domain prosesor sehingga dapat digunakan lebih baik untuk pattern matching.



## 6. LPC

### 1. Preemphasis

$$y(n) = s(n) - \alpha \cdot s(n-1)$$

$$y_0 = -0,085 - 0,95 \times 0 = -0,085$$

$$y_1 = (-0,067) - (0,95 \times (-0,085)) = 0,0137$$

$$y_2 = (0,043) - (0,95 \times (-0,067)) = 0,1066$$

$$y_3 = (0,029) - (0,95 \times (0,043)) = -0,0118$$

$$y_4 = (0,051) - (0,95 \times (0,029)) = 0,0234$$

$$s^1(0) = -0,085 + (-0,085) = -0,17$$

$$s^1(1) = -0,067 + 0,0137 = -0,0532$$

$$s^1(2) = 0,043 + 0,1066 = 0,1496$$

$$s^1(3) = 0,029 + (-0,0118) = 0,0171$$

$$s^1(4) = 0,051 + 0,0234 = 0,0744$$

### 2. Frame Blocking $N = 275$

### 3. Windowing

$$\tilde{x}(n) = x(n) \cdot w(n)$$

$$w_0 = 0,54 - 0,46 \cos \left( \frac{2 \cdot 3,14 \cdot 1}{275 - 1} \right) = 0,08$$

$$w_1 = 0,54 - 0,46 \cos \left( \frac{2 \cdot 3,14 \cdot 2}{275 - 1} \right) = 0,08$$

...

$$w_4 = 0,54 - 0,46 \cos \left( \frac{2 \cdot 3,14 \cdot 5}{275 - 1} \right) = 0,08$$

$$x_0 = -0,17 \times 0,08 = -0,0136$$

$$x_1 = -0,0532 \times 0,08 = -0,0042$$

$$x_2 = 0,1496 \times 0,08 = 0,0119$$

$$x_3 = 0,0171 \times 0,08 = 0,0013$$

$$x_4 = 0,0744 \times 0,08 = 0,0059$$

a. Analisis Autokorelasi

$$r(0) = (-0,0136 \times (-0,0136)) + (-0,0092 \times (-0,0092)) + (0,0119 \times (0,0119)) + (0,0013 \times (0,0013)) + (0,0059 \times (0,0059)) = 0,0393$$

c. Analisis LPC

$$k_0 = 0,0393$$

$$k_1 = (1 - 0,0393)^2$$