

# **Analisis Forman Frekuensi pada Suara Manusia dengan Menggunakan *Linear Prediction***

**IRMA AMELIA DEWI, MUHAMMAD ICHWAN, SALMA SILFIANA**

Program Studi Informatika Institut Teknologi Nasional Bandung  
Email: [irma\\_amelia@itenas.ac.id](mailto:irma_amelia@itenas.ac.id)

*Received* 10 Maret 2020 | *Revised* 25 April 2020 | *Accepted* 28 Mei 2020

## **ABSTRAK**

*Suara manusia merupakan natural language sebagai salah satu gaya interaksi dengan komputer. Manusia mempunyai ragam suara yang berbeda, dapat dilihat dari formant, pitch dan volume suara. Masukan perintah suara yang baik bagi komputer dibutuhkan proses pencarian kualitas suara berdasarkan forman frekuensi. Pada penelitian ini tahapan proses diawali dengan pre-processing yaitu preemphasis, frame blocking dan windowing kemudian dilanjutkan pencarian nilai forman menggunakan Linear Prediction. Hasil nilai forman yang didapatkan dicocokkan dengan nilai forman data latih yang berada pada database. Terdapat 2700 data suara uji dengan durasi perekaman suara dilakukan selama 1 detik. Berdasarkan hasil pengujian nilai forman yang diperoleh untuk F0 kisaran 0–423, nilai forman F1 kisaran 572-1678, nilai forman F2 kisaran 1536-2583, nilai forman F3 kisaran 2676-3384, nilai forman F4 kisaran 3519-4947.*

**Kata kunci:** *Forman, Frekuensi, Pitch, Linear Prediction*

## **ABSTRACT**

*Human voice is a natural language as a style of interaction with computers. Humans have a variety of sounds can be seen from the formant, pitch and volume. Entering voice commands that are good for the computer requires the search for sound quality based on formant frequencies. In this study, the process stages begin pre-emphasis, frame blocking and windowing for noise reduction and searching formant values using Linear Prediction. The formant value obtained is matched with the formant value of the training data in database. There are 2700 test sound data with recording duration is 1 second. Based on test results obtained formant values for F0 range of 0-423, value range 572-1678 formants F1, F2 formant values range from 1536 to 2583, the value of the range of 2676-3384 formant F3, F4 formant values of the range of 3519-4947.*

**Keywords:** *Formant, Frequency, Pitch, Linear prediction*

## 1. PENDAHULUAN

Suara manusia merupakan sinyal multidimensi yang membawa berbagai jenis informasi seperti frekuensi dasar (*Pitch*), energi dan durasi pengucapan. Sinyal suara juga dapat menyampaikan banyak informasi tentang pembicara seperti jenis kelamin, usia, identitas individu dan aspek lainnya (**Devi , Erwin, & Surbakti, 2017**). *Pitch* (F0) merupakan nilai perbedaan frekuensi suara yang disebabkan oleh membuka dan menutupnya pita suara (**Bhaskoro & D, 2012**).

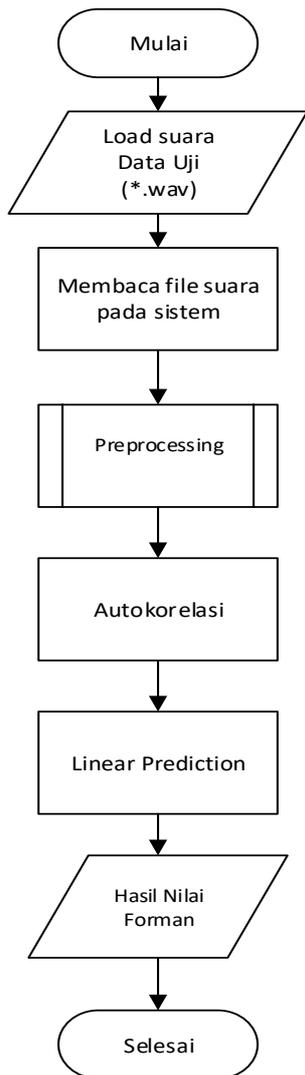
Setiap manusia memiliki keragaman jenis suara, walaupun terkadang terdapat kemiripan namun secara parameter dari bentuk suara tersebut berbeda-beda. Secara persepsi fisik manusia terhadap suara memiliki beberapa parameter, yaitu formant, jenis suara, *pitch*, *timbre* dan volume. Berdasarkan persepsi fisik, manusia dapat dengan mudah mengidentifikasi gender dari jenis suara tersebut. Salah satu parameter persepsi fisik yang cukup berpengaruh adalah Forman (**Bhaskoro & D, 2012**). Forman merupakan frekuensi resonansi alami yang dihasilkan dari pelbagai rongga mulut dan tenggorokan, output yang dihasilkan tergantung pada bentuk dan ukurannya. Sehingga forman menghasilkan 3 jenis formant yaitu F1, F2, dan F3.

Suara atau ucapan manusia dapat dijadikan salah satu cara untuk berinteraksi antara manusia dengan komputer. Setiap suara yang dimiliki manusia mempunyai keragaman bentuk yang berbeda-beda (**Permana, Nurhasanah, & Zulkarnain, 2018**) (**Nurhasanah, Zulkarnain, & Permatasari, 2017**). Perbedaan tersebut bisa dilihat dari jenis kelamin, *vocal range*, usia dan ragam suara yang berbeda. Ragam suara yang berbeda bisa disebabkan karena *formant*, *pitch* dan *volume* suara setiap orang yang berbeda (**Umar, Sunardi, & Gustafi, 2019**). Pada penelitian ini fokus bagaimana menganalisa kualitas frekuensi formant pada suara pembicara dalam pengucapan angka dengan menggunakan *Linear prediction* untuk mendapatkan nilai formant yang kemudian dicocokkan dengan formant yang berada pada data latih menggunakan metode KNN. Dengan hasil yang dikeluarkan oleh sistem adalah hasil kualitas suara.

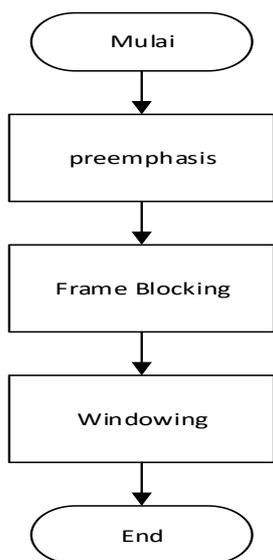
## 2. METODE PENELITIAN

Masukan pada sistem ini adalah nilai sinyal suara yang didapat dari masukan suara data latih ataupun data uji dengan format suara dalam bentuk *.wav* yang akan dilakukan proses *preprocessing* dan pencarian nilai formant pada suara yang sudah di-*inputkan* seperti pada Gambar 1. Kedepannya penelitian ini akan digunakan untuk menentukan kualitas suara manusia berdasarkan nilai forman sebagai inputan dari perintah suara ke sistem digital.

Pada sistem terdapat pencarian untuk mencari file data uji yang sudah dilakukan perekaman sebelumnya dengan melakukan proses *load* suara pada aplikasi yang sudah disiapkan. Sistem dapat melakukan proses menghilangkan gangguan suara yang dilakukan di bagian *Preprocessing* dan dilanjutkan dengan dilakukan pencarian nilai forman. Nilai forman yang didapatkan dari hasil proses *Linear Prediction*. Pada Gambar 2 merupakan tahapan pada proses pre-processing sinyal suara.



**Gambar 1. Flowchart Sistem Keseluruhan**



**Gambar 2. Tahap Pre-Processing Sinyal Suara**

## 2.1 Akuisi Data

Langkah awal yang dilakukan dalam penelitian ini adalah melakukan akuisisi data. Proses akuisisi ini suara dilakukan berdasarkan data sinyal suara yang diperoleh dengan cara merekam suara untuk data latih dan data uji menggunakan mikrofon yang memiliki spesifikasi dan dengan ruang tertentu untuk menekan adanya noise. **(Dewi, Zulkarnain, & Aprilia, 2018)**. Frekuensi sampling yang digunakan adalah frekuensi standar 16.000 Hz dengan format .wav, 16 bit/sample dan menggunakan channel mono. Durasi rekaman pada proses latih dan uji adalah 1 detik.

Persamaan (1) adalah rumus untuk mencari akuisisi data pada proses pengambilan *sampling*

$$X = FS \times dt(\text{detik}) \times \left(\frac{\text{bit}}{8}\right) \times j \quad (1)$$

Keterangan :

X :Data *sampling* sinyal

Fs :Frekuensi sampling

dt :Durasi rekaman (detik)

bit :Jumlah bit resolusi

j :1

Perhitungan pada proses akuisisi data untuk pengambilan sampling adalah :

$$X = 16000 \times 1 \times \left(\frac{16}{8}\right) \times 1 = 32000 \text{ byte}$$

Persamaan (2) adalah rumus untuk menghitung *sample rate*

$$\text{Sample rate} = \frac{Fs}{Ts} \quad (2)$$

Keterangan :

Fs : Frekuensi *Sampling*

Ts : Durasi

Perhitungan untuk mendapatkan sample rate adalah :

$$\text{Sample Rate} = \frac{32000}{1} = 32000 \text{ Hz}$$

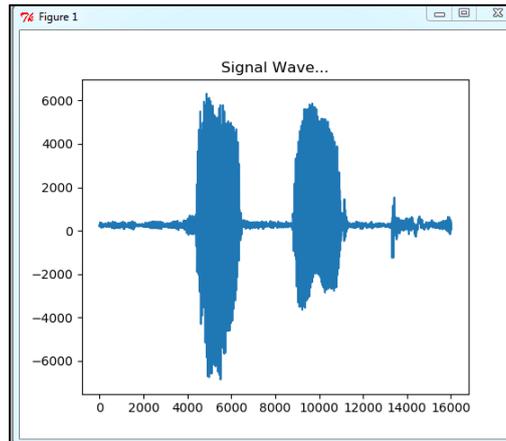
Persamaan (3) adalah rumus untuk menghitung *sample point*

$$\text{Sample point} = \text{sample rate} * \text{waktu ambil data} \quad (3)$$

Untuk mendapatkan *sample point* adalah :

$$\text{Sample point} = 32000 \times 0.01 = 320 \text{ sample point}$$

Untuk contoh perhitungan digunakan contoh data nilai suara sebanyak 8 angka dari salah satu *sample* suara. Data tersebut adalah (0.0057, 0.0033, 0.0040, 0.0039, 0.0016, -0.0014, -0.0059, 0.0055)



**Gambar 3. Sinyal Suara Asli**

Gambar 2. adalah gambar sinyal suara asli yang belum dilakukan pemfilteran dan perubahan sinyal suara.

## 2.2 *Pre-emphasize*

*Pre-emphasize* merupakan tahapan awal yang digunakan untuk meratakan sinyal *spectral* dan membuat sinyal rentan terhadap efek presisi dalam pemrosesan sinyal. Biasanya pada metode *preemphasize* digunakan *filter* yang bernama *High Pass Filter* berorde satu. *High Pass Filter* merupakan penapis yang meloloskan/meneruskan sinyal yang memiliki frekuensi tinggi dan menekan frekuensi yang lebih rendah daripada frekuensi *cutoff* (frekuensi kerja dari suatu sinyal)

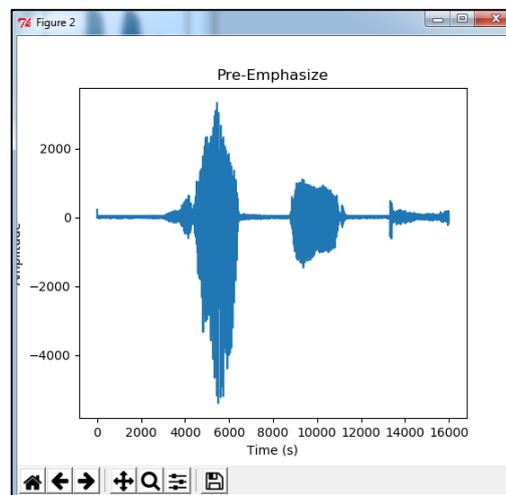
$$y[n] = s[n] - a s[n - 1] \quad (4)$$

Keterangan:

$y[n]$ : Sinyal hasil *pre-emphasis filter*

$s[n]$  : Sinyal sebelum *pre-emphasis filter*

$a$  : Merupakan konstanta *filter preemphasis*,  
biasanya bernilai 0,97



**Gambar 4. Sinyal Setelah Dilakukan Proses Preemphasis**

Hasil sinyal *pre-emphasize* pada data ke 0 (  $Y[0]$  ) sama dengan nilai sinyal sebelum proses *pre-emphasize* (  $S[0]$  ) karena belum terjadi proses *pre-emphasize* sehingga belum ada perubahan nilai. Diteruskan dengan perhitungan nilai yang selanjutnya dimana Y dan S akan mengitung data angka ke 1

$$Y1 = s[1] - \alpha s[1-1]$$

$$Y1 = 0,0033 - (0,97 * s[0] )$$

$$Y1 = 0,0033 - (0,97 * 0,0057)$$

$$Y1 = - 0,0022$$

Keterangan :

Y1 = sinyal hasil *preemphasis* yang pertama

s[1] = data yang ke 1 yaitu 0,0033 sesuai dengan data sample suara.

$\alpha$  = merupakan konstanta filter *preemphasis*, biasanya bernilai 0,97

s[1-1] = diambil dari s[n-1] yang dimana n adalah urutan data.

s[0] = data hasil *preemphasis* yang ke 0 (s[0] sama dengan y[0]) .

Berikutnya dicari nilai data sinyal baru yang diperoleh dari penambahan data sinyal sebelum proses *pre-emphasis* dengan data hasil *pre-emphasis* menggunakan persamaan berikut:

$$N_n = N_n + Y_n$$

$$N_0 = 0,0057 + 0,0057 = 0,0114$$

Keterangan :

$N_n$  = Data sinyal baru hasil *preemphasis*

$N_n$  = Nilai sinyal sebelum dilakukan *preemphasis*

$Y_n$  = Data hasil *preemphasis* sebelumnya

Begitu juga dengan ketujuh data yang lainnya, sehingga didapatkan nilai sinyal setelah *preemphasis* adalah : (0.0114, 0.0011, 0.0072 ,0.0077,0.0053, 0.0001, -0.0046, -0.0002)

### 2.3 Frame Blocking

Pada proses *frame blocking* sinyal suara dibagi menjadi beberapa *frame* untuk memudahkan dalam proses analisa suara. Pada pengolahan sinyal suara dibutuhkan sinyal yang memiliki parameter yang tetap (berubah dengan lambat) terhadap waktu (*time invariant*). Output dari proses *frame blocking* ini akan menghasilkan sinyal suara yang dianggap tidak berubah terhadap waktu (Dewi, Zulkarnain, & Aprilia, 2018) (Rabiner & Juang, 1993).

$$x_l n = \tilde{s} Ml + n, \quad n = 0, 1, \dots, N - 1, \\ l = 0, 1, \dots, L - 1$$

(5)

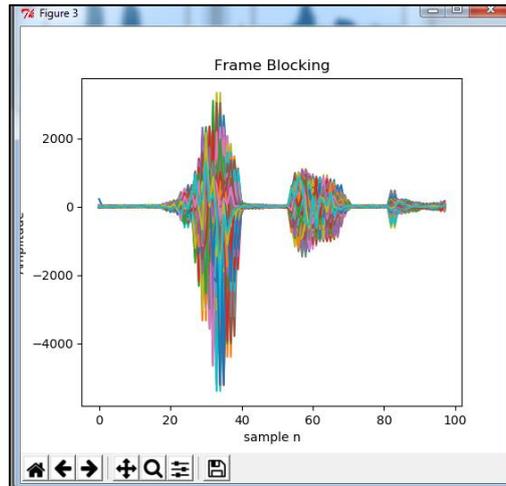
Keterangan :

$x_l n$  : Sinyal ke-n *frame blocking*

M : jarak antar frame

N : ukuran frame (*frame size*)

l : *Frame*



**Gambar 5. Sinyal Setelah Dilakukan Proses *Frame Blocking***

Diketahui nilai sinyal setelah preemphasis adalah : (0.0114, 0.0011, 0.0072 ,0.0077,0.0053, 0.0001, -0.0046, -0.0002)

$N = \text{Frekuensi Sampling} * \text{sample dalam satuan detik}$   
 $N = 16.000 * 0,02 = 320 \text{ sample}$

Dan akan didapatkan Overlappingframe (M) sebagai berikut  
 $M = \text{Frekuensi Sampling} * \text{sample dalam satuan detik}$   
 $M = 16.000 * 0,01 = 160$

$$x(n) = y(M+n)$$

Keterangan :

$x(n)$  = hasil frame blocking ke  $-n$   
 $y$  = nilai suara yang sudah dilakukan proses preemphasis  
 $M$  = Overlappingframe  
 $n$  = data ke- $n$

$$x(n) = y (M+n)$$

$$x(0) = y (M+0)$$

$$x(0) = 0.0114(M+0)$$

$$x(0) = 0.0114(160 + 0) = 1.824 \text{ Hz}$$

Sehingga didapatkan nilai sinyal setelah Frame Blocking adalah : (1.824 , 1.1771, 1.1664, 1.2551, 0.8692, 0.0165, -0,7363, -0.0334)

## 2.4 *Windowing*

Setelah proses *frame blocking* selesai dilakukan, tahapan berikutnya adalah proses *windowing*. *Windowing* adalah proses mengubah sinyal pada awal dan akhir setiap *frame* bernilai nol. Hal ini bertujuan untuk menghilangkan diskontinuitas sinyal pada awal dan akhir *frame*. *Hamming Window* merupakan model yang sering digunakan karena *Hamming Window* menghasilkan *noise* yang tidak terlalu besar (Dewi, Zulkarnain, & Aprilia, 2018). Untuk menghitung nilai hasil *windowing* digunakan Persamaan (6):

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right) \quad (6)$$

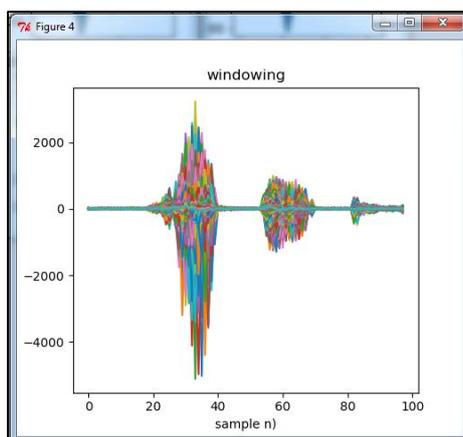
Keterangan :

$w(n)$  = hasil nilai *Windowing* yang ke- $n$

$\Pi = 3,14$

$N$  = Jumlah sample pada masing- masing *frame*

$n = 0,1,2,3, \dots .N-1$



**Gambar 6. Sinyal Setelah Dilakukan Proses *Windowing***

Diketahui nilai sinyal setelah *Frame Blocking* adalah : (1.824 , 1.1771, 1.1664, 1.2551, 0.8692, 0.0165, -0,7363, -0.0334) dengan nilai  $N = 320$  dari perhitungan  $N$  pada bagian *frame blocking* sebelumnya. Penyelesaian window  $w(n)$  :

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

$$w(0) = 0.54 - 0.46 \cos\left(\frac{2 \times 3.14 \times 0}{320 - 1}\right)$$

$$w(0) = 0.008 \cos \frac{0}{319}$$

$$w(0) = 0.008 \cos 0$$

$$w(0) = 0.008 \times 1 = 0.08$$

Untuk data *frame blocking* yang digunakan adalah : (1.824 , 1.1771, 1.1664, 1.2551, 0.8692, 0.0165, -0,7363, -0.0334)

sehingga :

$w(o) = \text{Nilai Frame Blocking} * \text{hasil } w(0)$

$w(o) = 1.824 * 0.08$

$w(o) = 0.14592$

Maka , Hasil dari *windowing* dari keseluruhan nilai yaitu (0.14592 , 0.094168, 0.093312 ,0.100408, 0.069536 , 0.00132, - 0.058936 , -0.002672)

## 2.5 *Linear Prediction*

Tahap untuk mencari nilai formant diambil data masukan berupa nilai yang dihasilkan setelah proses *windowing*. Tahap selanjutnya dari proses *windowing* adalah proses analisis autokorelasi (*autocorrelation analysis*).



**Gambar 7. Flowchart Linear Prediction**

Untuk mencari nilai *Autocorrelation* dapat menggunakan Persamaan 7

$$r_l(m) = \sum_{n=0}^{N-i-m} \tilde{x}_l(n) \cdot \tilde{x}_l(n+m) \quad , \quad m = 0, 1, \dots, p \quad (7)$$

Dimana :

$r_l(m)$  adalah Hasil sinyal autokorelasi

$\tilde{x}_l(n)$  adalah Hasil sinyal *windowing* ke-n

Jika data hasil *windowing* sampel data latih adalah (0.14592 , 0.094168, 0.093312 ,0.100408, 0.069536 , 0.00132, - 0.058936 , -0.002672) maka nilai autokorelasinya

$$\begin{aligned} m(1) &= \tilde{x}_l(1)\tilde{x}_l(1+1) \\ &= (0.14592)\tilde{x}_l(2) \\ &= (0.14592)(0.094168) \\ &= 1.374 \times 10^{-2} \end{aligned}$$

Maka , Hasil dari keseluruhan nilai yaitu ( $1.374 \times 10^{-2}$  ,  $0.946 \times 10^{-2}$ ,  $0.012 \times 10^{-2}$ ,  $0.027 \times 10^{-2}$ ) . sehingga untuk mendapatkan nilai autokorelasi dilakukan dengan menambahkan semua hasil perhitungan sebelumnya

$$\begin{aligned} r_l(m) &= 1.374 \times 10^{-2} + 0.946 \times 10^{-2} + 0.012 \times 10^{-2} + 0.027 \times 10^{-2} \\ &= 2.359 \times 10^{-2} \end{aligned}$$

Maka hasil dari nilai akhir dari nilai analisis autokorelasi adalah  $2.359 \times 10^{-2}$  .Langkah selanjutnya adalah melakukan perhitungan Levinson Durbin pada nilai autokorelasi yang didapat, seperti terlihat pada Persamaan 8 untuk mendapatkan koefisien lpc A(z).

Levinson Durbin :

$$\begin{bmatrix} R[0] & R[1] & R[2] & R[p] \\ R[1] & R[0] & R[1] & R[p-1] \\ R[2] & R[1] & R[0] & R[p-2] \\ R[p] & R[p-1] & R[p-2] & R[0] \end{bmatrix} \begin{bmatrix} a1 \\ a2 \\ a3 \\ ap \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ R(3) \\ R(p) \end{bmatrix} \quad (8)$$

Dengan solusi pemecahan terlihat pada Persamaan 9 sampai dengan Persamaan 15.

$$E^{(0)} = R(0) \quad (9)$$

$$k_{(i)} = R(i) - \sum_{j=1}^{i-1} a_j^{(i-1)} R(i-j) / E^{(i-1)} \quad (10)$$

$$a_i^{(i)} = k_{(i)} \quad (11)$$

$$a_j^{(i)} = a_j^{(i-1)} - k_{(i)} a_{i-j} \quad (12)$$

$$E^{(i)} = (1 - k_i^{(i)}) E^{(i-1)} \quad (13)$$

$$a_j = a_j^p \quad (14)$$

$$A(z) = 1 - \sum_{k=1}^p a_k z^{-k} \quad (15)$$

Keterangan :

$R(i)$  = nilai autokorelasi

$E(i)$  = error

$(i) k$  = koefisien pantulan

$(j) a$  = koefisien lpc

$A(z)$  = persamaan koefisien lpc

Setelah persamaan koefisien didapat, maka selanjutnya adalah mencari akar-akar persamaan koefisien  $A(z)$  dan kemudian dirubah ke dalam perubahan sudut untuk kemudian dihitung nilai formant (**Indra, 2016**).

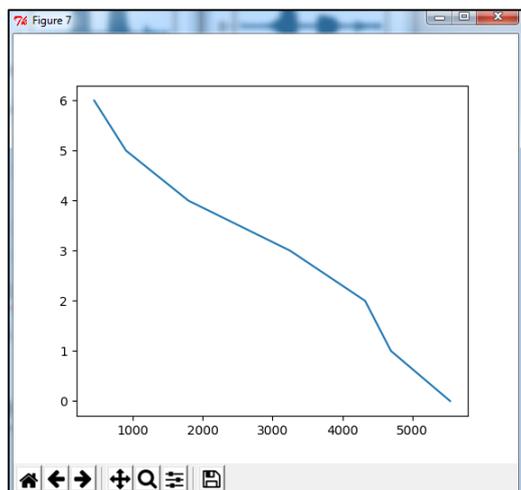
Karena koefisien LPC bernilai real, maka akan dilakukan proses pencarian akar yang dimana hanya bagian akar imaginary positif yang dipertahankan dan sudut yang ditentukan oleh persamaan seperti berikut :

$$\text{angz} = \sum \tan^{-1} \left( \frac{\text{Im} r_k}{\text{Re} r_k} \right) \quad (16)$$

Dalam persamaan diatas  $r_k$  adalah akar yang dipertahankan dari polynomial. Frekuensi sudut dalam rad / sampel diubah menjadi Hz dan lebar pita dari pembentuk dihitung oleh persamaan berikut

$$F_k = \left( \frac{F_s}{2\pi} \right) \tan^{-1} \left( \frac{\text{Im } r_k}{\text{Re } r_k} \right) \tag{17}$$

Nilai-nilai ini diurutkan dalam urutan menaik. Dan tiga nilai pertama di pilih untuk menjadi nilai 3 forman pertama (**Bhore & Shah, 2015**).



**Gambar 8. Sinyal *Linear Prediction***

### 2.6. Standar Suara Forman

Untuk membedakan vokal suara satu sama lain dapat dilihat melalui nilai forman frekuensi (**Russel, 2005**). untuk Batasan nilai forman terdapat sebagai berikut  
 Forman 0 atau pitch berada di kisaran 0Hz sampai kurang dari 500Hz  
 Forman 1 berada di kisaran 500Hz sampai kurang dari 1500Hz  
 Forman 2 berada di kisaran 1500Hz sampai kurang dari 2500Hz  
 Forman 3 berada di kisaran lebih dari 2500Hz sampai kurang dari 3500Hz  
 Forman 4 berada di kisaran lebih dari 3500Hz sampai kurang dari 4500Hz.

## 3. HASIL DAN PEMBAHASAN

Pada penelitian ini telah dilakukan pengujian suara oleh delapan orang. Untuk mendapatkan suara uji, dilakukan perekaman sebanyak sembilan puluh kali yang di mulai pengujian mengucapkan angka satu yang berjumlah sepuluh kali sampai pengucapan angka sembilan sepuluh kali. Jumlah dari suara uji terdapat 720 *file* wav yang akan dicari nilai formannya. Terdapat nilai forman minimum dan maksimumnya dalam satuan hertz(Hz) untuk suara perempuan dan pria seperti pada Tabel 1.

**Tabel 1. Nilai Forman Keseluruhan**

Pengujian Angka	F0		F1		F2		F3		F4	
	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX
Satu	0	487	572	1486	1536	2969	2676	3600	3519	4922
Dua	96	458	592	1448	1173	2890	2148	3888	3234	5224
Tiga	0	478	263	1887	1523	3072	2378	3871	2766	4946

Pengujian Angka	F0		F1		F2		F3		F4	
	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX
Empat	0	495	610	1477	1384	2456	2214	3371	3305	4682
Lima	87	465	587	2069	1597	2873	2386	3928	3412	4817
Enam	0	490	763	1722	1416	2195	2267	3502	3559	4242
Tujuh	149	490	549	2161	1590	3112	2533	4617	2985	4990
Delapan	0	423	515	1678	940	2583	2438	3384	3141	4947
Sembilan	0	491	504	2419	1607	3358	1753	3850	3376	4880

Pada Tabel 1. Terdapat nilai forman keseluruhan dari semua responden yaitu pria dan wanita untuk pengucapan angka satu sampai Sembilan. Untuk pengujian pengucapan angka satu terdapat nilai kisaran angka F0 yaitu dari 0 - 487Hz , F1 yaitu dari 572 - 1486Hz , F2 yaitu dari 1536Hz - 2969Hz , F3 yaitu dari 2767Hz - 3600 dan F4 yaitu 3519Hz -4922Hz.

Untuk pengujian pengucapan angka dua terdapat nilai kisaran angka F0 yaitu dari 96Hz - 458Hz , F1 yaitu dari 592Hz - 1448Hz , F2 yaitu dari 1173Hz - 2890Hz , F3 yaitu dari 2148Hz - 3888hZ dan F4 yaitu 3234Hz - 5224Hz.

Untuk pengujian pengucapan angka tiga terdapat nilai kisaran angka F0 yaitu dari 0Hz - 478Hz , F1 yaitu dari 263Hz - 1887Hz , F2 yaitu dari 1523Hz - 3072Hz , F3 yaitu dari 2378Hz - 3888hZ dan F4 yaitu 3234Hz - 5224Hz.

Pengucapan angka empat yaitu dari F0 yaitu dari 0Hz – 495Hz , F1 yaitu 610Hz – 1477Hz , F2 yaitu 1384Hz – 2456Hz, F3 yaitu 2214Hz – 3371Hz Dan F4 yaitu 3305Hz – 4682Hz.

Pengucapan angka lima yaitu dari F0 yaitu dari 87Hz – 465Hz , F1 yaitu 584H – 2069Hz , F2 yaitu 1597Hz – 2873 Hz , F3 yaitu 2386Hz – 3928Hz , F4 yaitu 3412Hz – 4817Hz.

Pengucapan angka enam yaitu dari F0 dari 0Hz – 490Hz , F1 yaitu 763Hz – 1722Hz , F2 yaitu 1416Hz – 2195Hz , F3 yaitu 2267Hz – 3502Hz , F4 yaitu 3559Hz – 4242Hz.

Pengucapan angka tujuh yaitu dari F0 yaitu dari 149Hz – 490Hz , F1 yaitu 549Hz – 2161Hz , F2 yaitu 1590Hz – 3112Hz , F3 yaitu 2533Hz – 4617Hz , F4 yaitu 2985Hz – 4990Hz.

Pengucapan angka delapan yaitu dari F0 yaitu dari 0Hz – 423Hz , F1 yaitu 515Hz – 1678Hz , F2 yaitu 940Hz – 2583Hz , F3 yaitu 2438Hz – 3384Hz , F4 yaitu 3141Hz – 4947Hz.

Pemgucapan angka sembilan yaitu dari F0 yaitu 0Hz – 491Hz , F1 yaitu 504Hz – 2419Hz , F2 yaitu 1607Hz – 3358Hz , F3 yaitu 1753Hz – 3850Hz , F4 yaitu 3376Hz – 4880Hz.

**Tabel 2. Nilai Forman Responden Wanita**

Pengujian Angka	F0		F1		F2		F3		F4	
	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX
Satu	185	495	527	1429	1516	2079	3265	3449	3540	4848
Dua	169	485	458	1434	1333	2890	2844	3888	3234	5224
Tiga	214	476	514	1460	1517	2486	2874	3377	3340	4629
Empat	207	378	671	1134	1530	1843	2390	3371	3328	4299
Lima	214	431	759	1494	1497	2873	2386	3482	3533	4576
Enam	231	415	763	1480	1550	2144	2267	3502	3606	4162
Tujuh	344	448	735	1509	1590	1463	2695	3190	2985	4920

Pengujian Angka	F0		F1		F2		F3		F4	
	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX
Delapan	125	415	675	1731	2532	1821	2637	3258	3141	4563
Sembilan	310	472	805	1474	1644	2491	2657	3533	3376	4691

Pada Tabel 2. Terdapat nilai forman responden wanita yang dimana dari nilai forman ke 0 hingga ke empat mempunyai nilai minimum dan maksimum yang berbeda-beda. Setelah dilakukan pencarian nilai forman dari masing-masing responden, setiap orang mempunyai nilai forman yang berbeda-beda karena karakter suara dan tinggi rendahnya suara seseorang mempengaruhi nilai forman pada orang tersebut.

Pengucapan angka satu yaitu dari F0 yaitu 185Hz – 495Hz , F1 yaitu 527Hz – 1429Hz , F2 yaitu 1516Hz – 2079Hz , F3 yaitu 3265Hz – 3449Hz , F4 yaitu 3540Hz – 4848Hz.

Pengucapan angka dua yaitu dari F0 yaitu 169Hz – 485Hz , F1 yaitu 458Hz – 1434Hz , F2 yaitu 1333Hz – 2890Hz , F3 yaitu 2844Hz – 3888Hz , F4 3234Hz – 5224Hz.

Pengucapan angka tiga yaitu dari F0 yaitu 214Hz – 476Hz , F1 yaitu 514Hz – 1460Hz , F2 yaitu 1517Hz – 2486Hz , F3 yaitu 2874Hz – 3377Hz , F4 yaitu 3340Hz – 4629Hz.

Pengucapan angka empat yaitu dari F0 yaitu 207Hz – 378Hz , F1 yaitu 671Hz – 1134Hz , F2 yaitu 1530Hz – 1843Hz , F3 yaitu 2390Hz – 3371Hz , F4 yaitu 3328Hz – 4299Hz.

Pengucapan angka lima yaitu dari F0 yaitu 214Hz – 431Hz , F1 yaitu 759Hz – 1494Hz , F2 yaitu 1497Hz – 2873Hz , F3 yaitu 2386Hz – 3482Hz , F4 yaitu 3533Hz – 4576Hz.

Pengucapan angka enam yaitu dari F0 yaitu 231Hz – 415Hz , F1 yaitu 763Hz – 1480Hz , F2 yaitu 1550Hz – 2144Hz , F3 yaitu 2267Hz – 3502Hz , F4 yaitu 3606Hz – 4162Hz.

Pengucapan angka tujuh yaitu dari F0 yaitu 344Hz – 448Hz , F1 yaitu 735Hz – 1509Hz , F2 yaitu 1590Hz – 1463Hz , F3 yaitu 2695Hz – 3190Hz , F4 yaitu 2985Hz – 4920Hz.

Pengucapan angka delapan yaitu dari F0 yaitu 125Hz – 415Hz , F1 yaitu 675Hz – 1731Hz , F2 yaitu 2532Hz – 1821Hz , F3 yaitu 2637Hz – 3258Hz , F4 yaitu 3141Hz – 4563Hz.

Pengucapan angka sembilan yaitu dari F0 yaitu 310Hz – 472Hz , F1 yaitu 805Hz – 1474Hz , F2 yaitu 1644Hz – 2491Hz , F3 yaitu 2657Hz – 3533Hz , F4 yaitu 3376Hz – 4691Hz.

**Tabel 3. Nilai Formant Responden Pria**

Pengujian Angka	F0		F1		F2		F3		F4	
	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX
Satu	0	459	527	1521	1327	2671	2676	3600	3519	4922
Dua	96	592	509	1448	1173	2578	2195	3447	3430	4235
Tiga	0	478	529	1887	1523	3072	2378	3871	2766	4946
Empat	0	495	610	1477	1384	2456	2214	3184	3305	4682
Lima	87	565	387	2069	1570	2582	2425	3928	3412	4817
Enam	0	490	768	1722	1416	2195	2557	3182	3559	4242
Tujuh	149	490	349	1941	1640	3112	2533	4617	3276	4990
Delapan	0	423	515	1678	1281	2583	2438	3384	3298	4947
Sembilan	0	491	504	2419	1607	3358	1753	3850	3393	4880

Pada Tabel 3. Terdapat nilai formant responden pria yang dimana dari nilai formant ke 0 hingga ke 4 mempunyai nilai minimum dan maksimum yang berbeda-beda. Setelah dilakukan pencarian nilai formant dari masing-masing responden, setiap orang mempunyai nilai formant yang berbeda-beda karena karakter suara dan tinggi rendahnya suara seseorang mempengaruhi nilai formant pada orang tersebut. Pada penelitian ini dilakukan pengujian terhadap pria yang berumur 17 tahun hingga 25 tahun. Dimana pada umur tersebut suara pria cenderung mempunyai karakteristik suara yang berat.

Pengucapan angka satu yaitu dari F0 yaitu 0Hz – 459Hz , F1 yaitu 527Hz – 1521Hz , F2 yaitu 1327Hz – 2671Hz , F3 yaitu 2676Hz – 3600Hz , F4 yaitu 3519Hz – 4922Hz.

Pengucapan angka dua yaitu dari F0 yaitu 96Hz – 592Hz , F1 yaitu 509Hz – 1448Hz , F2 yaitu 1173Hz – 2578Hz , F3 yaitu 2195Hz – 3447Hz , F4 yaitu 3430Hz – 4235Hz.

Pengucapan angka tiga yaitu dari F0 yaitu 0Hz – 478Hz , F1 yaitu 529Hz – 1887Hz , F2 yaitu 1523Hz – 3072Hz , F3 yaitu 2378Hz – 2871Hz , F4 yaitu 2766Hz – 4946Hz.

Pengucapan angka empat yaitu dari F0 yaitu 0Hz – 495Hz , F1 yaitu 610Hz – 1477Hz , F2 yaitu 1384Hz – 2456Hz , F3 yaitu 2214Hz – 3184Hz , F4 yaitu 2766Hz – 4946Hz.

Pengucapan angka lima yaitu dari F0 yaitu 87Hz – 565Hz , F1 yaitu 387Hz – 2069Hz , F2 yaitu 1570Hz – 2582Hz , F3 yaitu 2425 – 3928Hz , F4 yaitu 3412Hz – 4817Hz.

Pengucapan angka enam yaitu dari F0 yaitu 0Hz – 490Hz , F1 yaitu 768Hz – 1722Hz , F2 yaitu 1416Hz – 2195Hz , F3 yaitu 2557Hz – 3182Hz , F4 yaitu 3559Hz – 4242Hz.

Pengucapan angka tujuh yaitu dari F0 yaitu 149Hz – 490Hz , F1 yaitu 349Hz – 1941Hz , F2 yaitu 1640Hz – 3112Hz , F3 yaitu 2533Hz – 4617Hz , F4 yaitu 3276Hz – 4990Hz.

Pengucapan angka delapan yaitu dari F0 yaitu 0Hz – 423Hz , F1 yaitu 515Hz – 1678Hz , F2 yaitu 1281Hz – 2583Hz , F3 yaitu 2438Hz – 3384Hz , F4 yaitu 3298Hz – 4947Hz.

Pengucapan angka sembilan yaitu dari F0 yaitu 0Hz – 491Hz , F1 yaitu 504Hz – 2419Hz , F2 yaitu 1607Hz – 3358Hz , F3 yaitu 1753Hz – 3850Hz , F4 yaitu 3393Hz – 4880Hz.

#### **4. KESIMPULAN**

Berdasarkan penelitian mengenai sistem penentuan kualitas suara berdasarkan formant frekuensi yang telah dilakukan, diperoleh kesimpulan sebagai berikut :

1. Sistem yang dibangun dapat digunakan untuk mengenali suara berupa angka dengan pengumpulan data latih dengan kondisi kebisingan ruangan dari 30 dB sampai 90 dB.
2. Terdapat 2700 data suara uji dengan durasi perekaman suara dilakukan selama 1 detik. Berdasarkan hasil pengujian nilai formant yang diperoleh untuk F0 kisaran 0–423, nilai formant F1 kisaran 572-1678, nilai formant F2 kisaran 1536-2583, nilai formant F3 kisaran 2676-3384, nilai formant F4 kisaran 3519-4947.

## DAFTAR RUJUKAN

- Bhaskoro, S. B., & D, A. (2012). Aplikasi Pengenalan Gender Menggunakan Suara. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI 2012)*, (pp. 16-23).
- Bhore, S., & Shah, M. (2015). A Comparative Study of Formant Estimation. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, 4(12), 2879-2882.
- Devi , L., Erwin, & Surbakti, A. (2017). *Analisa Pengaruh Penyakit Flu Dan Batuk Terhadap Suara Penderita Dengan Menggunakan Komputer*. Retrieved from Repository University of Riau: repository.unri.ac.id
- Dewi, I. A., Zulkarnain, A., & Aprilia , A. (2018). Identifikasi Suara Tangisan Bayi menggunakan Metode LPC dan Euclidean Distance. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, 6(1), 153-164.
- Indra, T. P. (2016). *Pengindentifikasian Suara Tinggi Dan Rendah Pada Pria Dan Wanita Ditinjau Dari Pitch Dan Formant Menggunakan Metode Support Vector Machine*. Retrieved from Unikom Repository: <https://repository.unikom.ac.id/294/>
- Nurhasanah, Y. I., Zulkarnain, A., & Permatasari, D. (2017). Pengenalan Pembicara untuk Menentukan Gender Menggunakan Metode MFCC dan VQ. *MIND (Multimedia, Artificial Intelligence, Networking, Database) Journal*, 2(1), 34-47.
- Permana, I. S., Nurhasanah, I. Y., & Zulkarnain, A. (2018). Implementasi Metode Mfcc Dan Dtw Untuk Pengenalan Jenis Suara Pria Dan Wanita. *MIND (Multimedia, Artificial Intelligence, Networking, Database) Journal*, 3(1), 49 - 63.
- Rabiner, L., & Juang, B.-H. (1993). *Fundamentals of speech recognition*. New Jersey: Prentice-Hall, Inc.
- Russel, K. (2005). *General Phonetics*. Retrieved from home.cc.umanitoba.ca.
- Umar, R., Sunardi , S., & Gustafi, M. F. (2019). Analisis Statistik Manipulasi Pitch Suara Menggunakan Audio Forensik Untuk Bukti Digital. *Jurnal Mobile and Forensics (MF)*, 1(1), 1-12.