

Perbandingan Prediksi Linear Menggunakan Metoda Autokorelasi Dan Metoda Burg

Djadjat Sudarajat¹, Andi Rosano²

¹Universitas Bina Sarana Informatika
e-mail: djadjat.dsj@bsi.ac.id

²Universitas Bina Sarana Informatika
e-mail: andi.aox@bsi.ac.id

Abstrak - Pada sistem pengenalan ucapan manusia digunakan metoda pengenalan parameter-parameter yang dihasilkan oleh ucapan manusia untuk mengenalinya. Parameter-parameter ucapan yang dihasilkan bersifat unik yaitu tidak ada yang sama pada tiap orang seperti yang terjadi pada sidik jari manusia tidak ada yang sama, sehingga parameter-parameter rongga pengubah getaran ucapan manusia dapat digunakan untuk mengenali ucapan seseorang. Beberapa metoda tiruan rongga pengubah getaran ucapan manusia diantaranya menggunakan metoda Autokorelasi dan metoda Burg. Pada penelitian ini akan diamati perbedaan kinerja metoda Autokorelasi dan metoda Burg dengan mengamati perbedaan spektrum pada tiap panjang frame yang berbeda-beda. Dari hasil pengukuran dapat dilihat bahwa pada perubahan panjang frame yang berbeda-beda pada metoda Burg tidak banyak perubahannya dibandingkan dengan metoda Autokorelasi. Pada metoda Autokorelasi terjadi perubahan spektrum ketika panjang frame makin pendek, sedangkan pada metoda Burg hampir tidak mengalami perubahan yang banyak. Sedangkan kesalahan rekonstruksi sinyal sering terjadi pada frame yang panjang karena kemungkinan terdapat dua jenis sinyal voiced/unvoiced lebih besar dari pada menggunakan frame yang pendek. Dengan demikian untuk menghindari kesalahan rekonstruksi sinyal lebih baik menggunakan frame yang pendek, dan menggunakan metoda Burg dengan frame yang pendek menunjukkan kinerja spektral yang lebih baik dibandingkan dengan metoda Autokorelasi.

Kata Kunci: LPC, Speech Recognition, Speaker Recognition, DSP.

Abstract - In the human speech recognition system, the method of recognizing the parameters generated by human speech is used to recognize it. The resulting speech parameters are unique, that is, nothing is the same for each person, as is the case with human fingerprints, nothing is the same, so the cavity parameters for changing human speech vibrations can be used to recognize someone's speech. Several methods of artificial cavity modifying human speech vibrations include the Autocorrelation method and the Burg method. In this study, the differences in the performance of the autocorrelation method and the Burg method will be observed by observing the difference in spectrum at different lengths of the frame. From the measurement results, it can be seen that there are not many changes in the different frame length changes in the Burg method compared to the Autocorrelation method. In the autocorrelation method, there is a change in the spectrum when the length of the frame is getting shorter, while in the Burg method there is almost no change. Meanwhile, signal reconstruction errors often occur in long frames because the possibility that there are two types of voiced / unvoiced signals is greater than using short frames. Thus, to avoid the error signal reconstruction is better to use a short frame, and using the Burg method with a short frame shows a better spectral performance compared to the autocorrelation method.

Keywords: LPC, Speech Recognition, Speaker Recognition, DSP

PENDAHULUAN

Sejak pertama kali diperkenalkan oleh N. Wiener tahun 1966, teknik prediksi linier menjadi sangat terkenal hingga saat ini digunakan di banyak bidang aplikasi. Terutama dibidang pengolahan suara yang pertama kali digunakan untuk sistem analisis dan sintesis suara oleh Itakura & Saito tahun 1968 dan Atal & Schroeder tahun 1968. Keistimewaan dari

teknik prediksi linier tersebut terletak pada cara pembentukan kembali (rekonstruksi) bentuk gelombang suara dan karakteristik spektrum suara dapat diperoleh secara efisien dan akurat dengan hanya menggunakan beberapa parameter yang dapat diperoleh relatif cukup sederhana (Furui, 1989). Ekuivalensi persamaan dikawasan waktu diskrit dari persamaan model analisis berikut :

$$E(z) = 1/G.S(Z).A(Z) \dots\dots\dots(1)$$



Dengan koefisien filter $a(i)$ yang didefinisikan di persamaan berikut :

$$A(Z) = \sum_{i=0}^p a(i) z^{-i} \dots\dots\dots(2)$$

$a(i)$ = koefisien filter, $a(0) = 1$

maka : $e(n) = \sum_{i=0}^p a(i) s(n - i) \dots\dots\dots(3)$

atau $e(n) = s(n) + \sum_{i=1}^p a(i) s(n - i) \dots\dots\dots(4)$

atau $e(n) = s(n) - \hat{s}(n) \dots\dots\dots(5)$

dengan $\hat{s}(n) = - \sum_{i=1}^p a(i) s(n - i) \dots\dots\dots(6)$

$\hat{s}(n)$ dinamakan besaran prediksi (tanda topi menunjukkan estimasi).

Fungsi $e(n)$ dapat diintegrasikan sebagai error prediksi antara besaran cuplikan $s(n)$ dan besaran prediksi $\hat{s}(n)$. Koefisien $-a(i)$, $i = 1, 2, \dots, p$, didefinisikan sebagai koefisien prediksi dengan tanda negatif dipilih untuk menyederhanakan perhitungan selanjutnya, dan pada umumnya untuk pengkodean sinyal suara manusia jumlah koefisien prediksi dipilih antara $p=8$ sampai $p=12$ (Markel J.D, 1976). Koefisien prediksi $a(i)$ dihitung dengan kriteria harga rata-rata kesalahan kuadrat $E[e^2(n)]$ dibuat sekecil mungkin (minimal) agar besaran prediksi $\hat{s}(n)$ sedekat mungkin dengan besaran cuplikan $s(n)$.

Ide dasar dari prediksi linear adalah memprediksi besar cuplikan $s(n)$ dengan mempergunakan kombinasi linear dari beberapa cuplikan sebelumnya $s(n-1)$, $s(n-2)$, ..., $s(n-p)$. Transformasi-z dari persamaan (6) dapat ditulis :

$$\hat{S}(z) = F(z) S(z) \dots\dots\dots(7)$$

Dengan :

$$F(z) = - \sum_{i=1}^p a(i) z^{-i} \dots\dots\dots(8)$$

Didefinisikan sebagai transformasi-z dari filter prediktor linier. Transformasi-z dari persamaan (5) dan (6), model prediksi linier dapat ditulis :

$$\begin{aligned} E(z) &= S(z) - \hat{S}(z) = S(z) - F(z)S(z) \\ E(z) &= S(z)[1-F(z)] \dots\dots\dots(9) \end{aligned}$$

Atau :

$$E(z) = S(z)A(z)$$

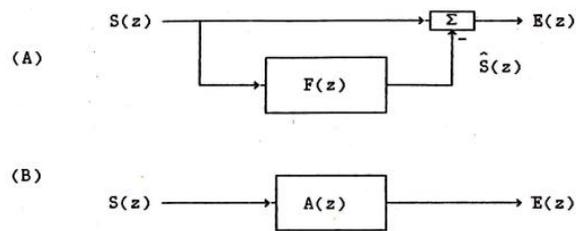
Dengan :

$$A(z) = 1 - F(z)$$

Atau :

$$A(z) = 1 + \sum_{i=1}^p a(i) z^{-i} \dots\dots\dots(10)$$

Blok diagram model prediksi linier diperlihatkan di gambar (1.A) dan (1.B) secara berturut-turut dalam bentuk filter prediktor $F(z)$ dan filter invers $A(z)$.



Sumber: (Sudaradjat, 1993)

Gambar 1. Model prediksi linier.

(A) dalam bentuk filter prediktor $F(z)$

(B) dalam bentuk filter invers $A(z)$ dimana $A(z)=1-F(z)$

1. Metoda Autokorelasi.

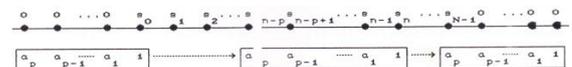
Pada metoda ini kriteria meminimalkan harga rata-rata kesalahan kuadrat (mean square error) didefinisikan sebagai persamaan :

$$E_p = \sum_{n=0}^{N+p-1} (e_p(n))^2 = \min \dots\dots\dots(11)$$

Dengan $e_p(n)$ adalah error prediksi yang diperoleh melalui hubungan di persamaan (3) :

$$E_p(n) = \sum_{i=0}^p a(i) s(n - i)$$

Deretan $e_p(n)$ diperoleh dari konvolusi antara deretan koefisien prediksi $a(i)$ dengan blok data sepanjang N dan $s(n)=0$ untuk $n<0$ dan $n>N-1$ seperti terlihat di gambar (2.A).



Sumber: (Sudaradjat, 1993)

Gambar 2.A Operasi konvolusi pada metoda Autokorelasi

dari gambar (2.A) diatas terlihat bahwa konvolusi tersebut menghasilkan deretan $e_p(n)$ sepanjang $N+p$, yang merupakan panjang daerah batas operasi di persamaan (11).

2. Metoda Burg

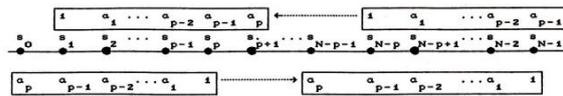
Metoda ini sedikit berbeda dengan metoda Autokorelasi, dimana harga $s(n)$ tidak dibuat nol untuk $s(n)$ diluar blok data operasi konvolusi seperti terlihat di gambar (2.B) dengan operasi konvolusinya dilakukan kedepan dan kebelakang, sehingga batas daerah jumlah total kuadrat error minimalnya mulai dari p sampai $N-1$. Hal itu dapat dinyatakan dalam persamaan (12) :

$$E_p = \sum_{n=p}^{N-1} [(e_p^+(n))^2 + (e_p^-(n+1))^2] = \min \dots\dots\dots(12)$$

Dimana : $e_p^+(n)$ adalah error prediksi kedepan (*forward prediction error*).

$e_p^-(n+1)$ adalah error prediksi kebelakang

(backward prediction error).

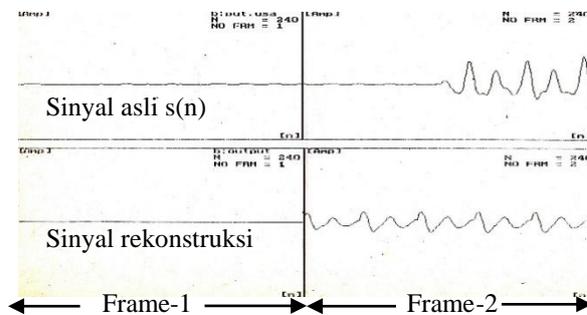


Sumber: (Sudaradjat, 1993)

Gambar 2.B Operasi konvolusi pada metoda Burg.

3. Kesalahan rekonstruksi sinyal.

Pada gambar (3) terlihat sinyal asli $s(n)$ dibandingkan dengan sinyal rekonstruksi hasil pengenalan ucapan "usai" dari seorang laki-laki berumur 31 tahun pada frame 1, dan 2. Jumlah frame untuk ucapan "usai" ini berjumlah 33 frame tetapi tidak semuanya ditunjukkan di gambar ini, panjang tiap frame 240 sample atau 30 mdetik dengan frekuensi cuplik 8 KHz, dan sinyal rekonstruksi dengan bit-rate 2,4 KBit/det.



Sumber: (Djadjat Sudaradjat, 2020)

Gambar 3. Perbandingan sinyal asli $s(n)$ dengan sinyal rekonstruksi pada frame ke-1, dan ke-2.

Pada frame ke-2 terjadi salah rekonstruksi yang seharusnya kombinasi *voiced* dan *un-voiced*, tetapi sinyal rekonstruksi menterjemahkan sebagai *voiced* semuanya. Untuk memperkecil salah rekonstruksi tersebut tentu dapat diatasi bila panjang framenya diperkecil dari 240 sampel menjadi 30 sample.

METODOLOGI PENELITIAN

Prosedur menghitung koefisien prediksi dengan metoda Autokorelasi atau Burg menggunakan algoritma yang telah dibahas di bab PENDAHULUAN.

1. Algoritma Autokorelasi.

Diagram alir algoritma Autokorelasi dapat dilihat di gambar (4) dengan langkah-langkah prosedur sbb :

1. Hitung autokorelasi data cuplikan sesuai dengan persamaan (Hernando Castaneda Mari, 2018) :

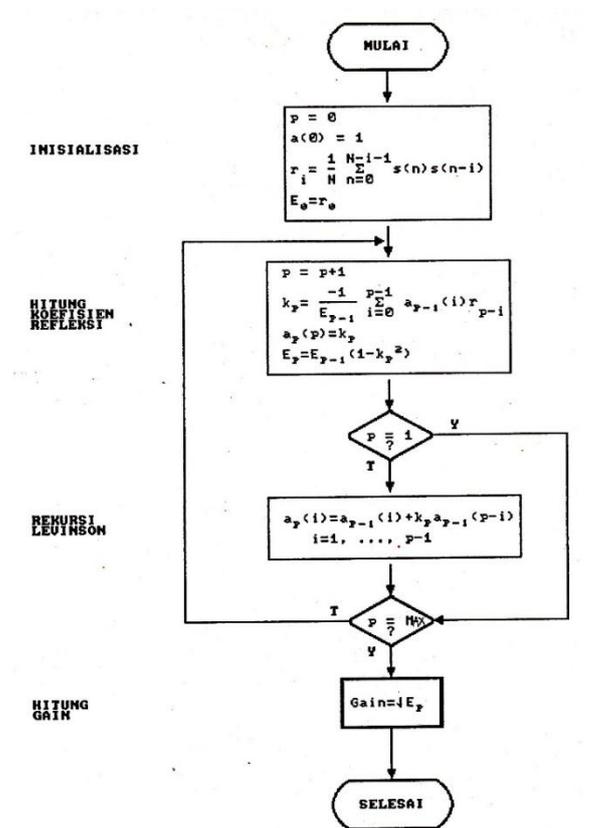
$$r_i = (1/N) \sum_{n=0}^{N-1-i} s(n) s(n-i),$$

$i=0, 1, 2, \dots, p$

dimana : p adalah jumlah orde prediksi,
 r_i adalah sample autokorelasi ke- i
 N adalah jumlah data cuplikan dalam satu frame.

- $s(n)$ adalah data cuplikan saat ke- n .
 $s(n-i)$ adalah data cuplikan saat ke- $(n-i)$.
2. Untuk tahap awal $p=0$, dipilih koefisien prediksi $s_{p=0}(0)=1$, karena itu diperoleh : $E_0=r_0$.
3. Untuk tahap p selanjutnya ($1 \leq p \leq \text{orde mak}$) :
 - a. dipilih $a_p(0)=1$.
 - b. hitung koefisien refleksi :

$$k_p = -1/E_{p-1} \cdot \sum_{i=0}^{p-1} a_{p-i}(i) r_{p-i}$$
 - c. koefisien prediksi $a_p(p) = k_p$
 - d. hitung energi error prediksi : $E_p = E_{p-1}(1-k_p^2)$
4. Koefisien prediksi dihitung dari persamaan : $a_p(i) = a_{p-i}(i) + k_p a_{p-i}(p-i)$
5. Hitung gain $G = \sqrt{E_p}$



Sumber: (Djadjat Sudaradjat, 2020)

Gambar 4. Diagram alir Algoritma Autokorelasi.

2. Algoritma Burg.

Diagram alir algoritma Burg dapat dilihat di gambar (5) dengan langkah-langkah sbb :

1. Untuk tahap awal $p=0$, dipilih :
 - a. $a_p(0) = 1$
 - b. $E_0 = r_0$
 - c. Error prediksi kedepan : $e_0^+(n) = s(n)$
Error prediksi kebelakang : $e_0^-(n) = s(n-1)$
2. Untuk tahap p selanjutnya ($1 \leq p \leq \text{orde mak}$) :
 - a. Hitung koefisien refleksi (Hernando Castaneda Mari, 2018):

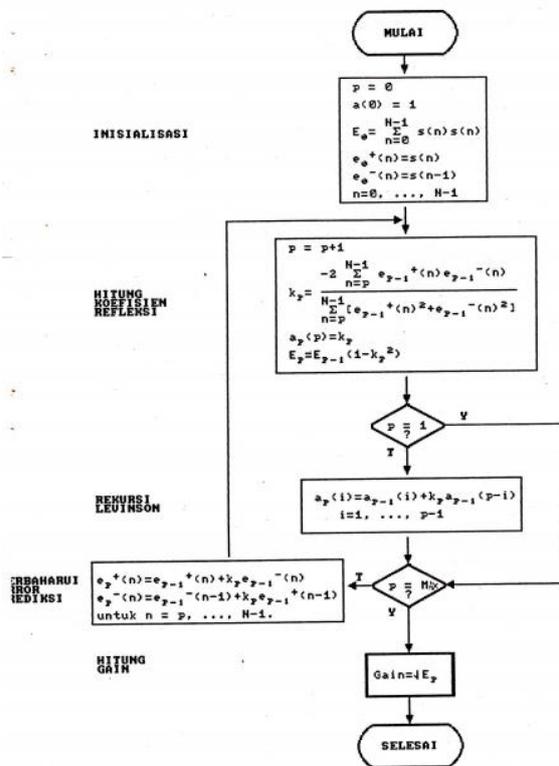
$$k_p = \frac{-2 \sum_{n=p}^{N-1} e_{p-1}^+(n) e_{p-1}^-(n)}{\sum_{n=p}^{N-1} [(e_{p-1}^+(n))^2 + (e_{p-1}^-(n))^2]}$$

- b. $a_p(p) = k_p$
- c. Hitung energi error prediksi : $E_p = E_{p-1}(1 - k_p^2)$
- 3. Koefisien prediksi dihitung dari persamaan :
 $a_p(i) = a_{p-1}(i) + k_p a_{p-1}(p-i)$
- 4. Untuk $p \leq n \leq N-1$, error prediksi dihitung dari persamaan :

$$e_p^+(n) = e_{p-1}^+(n) + k_p e_{p-1}^-(n)$$

$$e_p^-(n) = e_{p-1}^-(n-1) + k_p e_{p-1}^+(n-1)$$

- 5. Hitung gain $G = \sqrt{E_p}$



Sumber: (Sudaradjat, 1993)

Gambar 5. Diagram alir Algoritma Burg.

HASIL DAN PEMBAHASAN

Tujuan dari uji spektral ini adalah untuk membandingkan metoda analisis Autokorelasi dengan metoda analisis Burg dilihat dari perubahan estimasi spektral model *Autoregressive* apabila panjang frame analisis berubah. Kombinasi sistem yang dipilih dalam penelitian ini ditunjukkan pada tabel 1 dan jumlah koefisien prediksi yang digunakan adalah $p=10$.

Tabel 1. Kombinasi sistem pada uji spektral

Sistem	Panjang Frame (N)	Metoda analisis
1	240	Autokorelasi
2	240	Burg
3	160	Autokorelasi
4	160	Burg
5	80	Autokorelasi
6	80	Burg
7	30	Autokorelasi
8	30	Burg

Sumber: (Sudaradjat, 1993)

Pada pengujian ini akan diamati perubahan bentuk spektral beberapa fonem dari beberapa bunyi seperti diperlihatkan pada tabel 2.

Tabel 2. Uji bentuk spektral pada beberapa contoh fonem dari beberapa bunyi.

Fonem	Bunyi
/u/	usai
/a/	kata
/o/	obat
/i/	ikan

Sumber: (Sudaradjat, 1993)

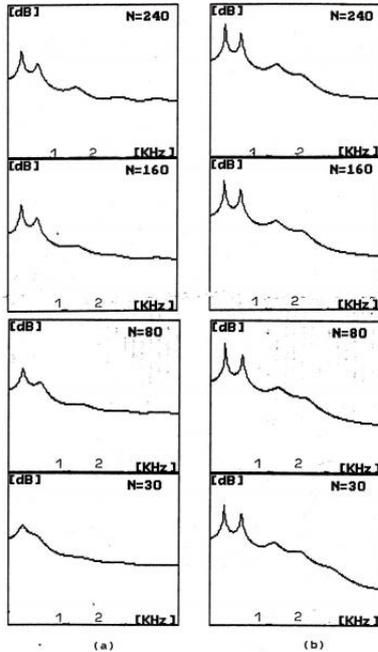
Kemudian dilihat karakteristik spektralnya, diamati, dan dibandingkan diantara kedelapan sistem tersebut, seperti terlihat di gambar (6) memperlihatkan satu contoh gambar spektral fonem /u/ dalam kata "usai" dari kedelapan sistem tersebut di tabel 1. Contoh gambar spektral untuk fonem /a/, /o/, dan /i/ berturut-turut dalam kata "kata", "obat", dan "ikan" diperlihatkan di gambar 7, 8, dan 9.

Dari hasil coba spektral digambar 6, 7, 8 dan 9 memperlihatkan bentuk spektral yang berubah untuk panjang frame (N) yang memendek, dan perubahan lebih banyak terjadi pada sistem dengan metoda Autokorelasi dibandingkan dengan metoda Burg. Keadaan tersebut menunjukkan bahwa pada frame yang pendek, perubahan spektral lebih banyak terjadi pada metoda Autokorelasi dari pada metoda Burg. Seperti dibahas di bab PENDAHULUAN, pada metoda Autokorelasi dilakukan perhitungan konvolusi dengan asumsi bahwa harga cuplikan diluar frame yang dianalisa diasumsikan berharga nol (untuk mengurangi efek waktu tak terbatas), sehingga dengan asumsi tersebut pada hakekatnya akan ada informasi yang hilang pada frame yang pendek yang mengakibatkan terjadinya perubahan spektral seperti yang terlihat di gambar 6a, 7a, 8a, dan 9a. Penampakan gambar 6, 7, 8, dan 9 diperoleh dari dari pemrograman diagram alir Algoritma Autokorelasi (Gambar 4) dan pemrograman diagram alir Algoritma Burg (Gambar 5) menggunakan bahasa pemrograman Turbo Pascal (Sudaradjat, 1993). Dengan demikian pada frame yang pendek, estimasi spektral dengan metoda Burg menunjukkan kinerja yang lebih baik dari pada metoda Autokorelasi, ada beberapa keuntungan dengan menggunakan panjang frame yang lebih pendek diantaranya lebih efisien dalam perhitungan, dan kecil

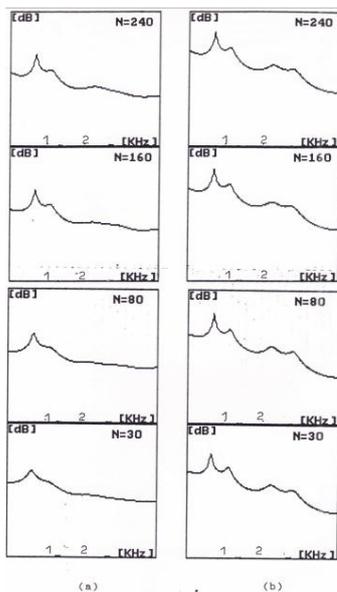
kemungkinan salah rekonstruksi yang disebabkan oleh kemungkinan terdapat dua jenis sinyal *voice* dan *unvoiced* pada satu frame sinyal asli sehingga pada sinyal rekonstruksi terjadi kesalahan rekonstruksi yang diterjemahkan hanya sebagai salah satu jenis sinyal *voiced* atau *unvoiced*..

Sumber: (Sudaradjat, 1993)

Gambar 6. Perubahan bentuk spektral fonem /u/



bunyi *usai*. (a) Metoda Autokorelasi, (b) Metoda Burg.

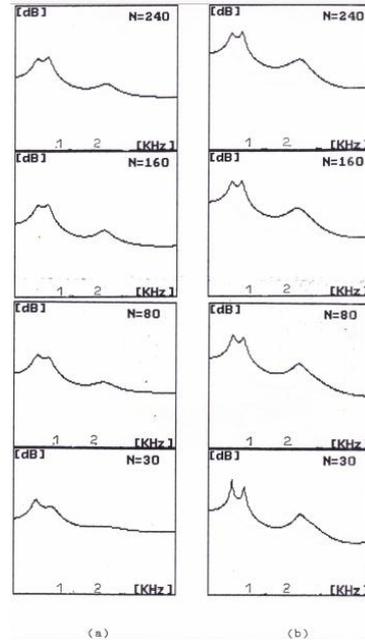


Sumber: (Sudaradjat, 1993)

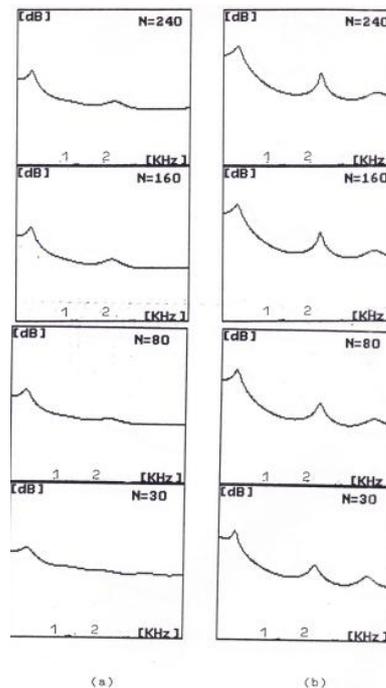
Gambar 7. Perubahan bentuk spektral fonem /a/ bunyi *kata*. (a) Metoda Autokorelasi, (b) Metoda Burg.

Sumber: (Sudaradjat, 1993)

Gambar 8. Perubahan bentuk spektral fonem /o/



bunyi *obat*. (a) Metoda Autokorelasi, (b) Metoda Burg.



Sumber: (Sudaradjat, 1993)

Gambar 9. Perubahan bentuk spektral fonem /i/ bunyi *ikan*. (a) Metoda Autokorelasi, (b) Metoda Burg.

KESIMPULAN

Dari hasil pengamatan perubahan bentuk spektral pada metoda Autokorelasi dan metoda Burg terlihat bahwa untuk perubahan panjang frame N = 240, 160, 80, dan 30 mengalami perubahan bentuk spektral pada metoda Autokorelasi, sedangkan pada metoda

Burg tidak mengalami banyak perubahan,

Dengan demikian pada frame yang pendek, estimasi spektral dengan metoda Burg menunjukkan kinerja yang lebih baik dari pada metoda Autokorelasi, ada beberapa keuntungan dengan menggunakan panjang frame yang lebih pendek diantaranya lebih efisien dalam perhitungan, dan kecil kemungkinan salah rekonstruksi. Karena pada satu frame yang lebih panjang ada kemungkinan terdapat dua jenis sinyal *voiced* dan *unvoiced* pada satu frame sinyal aslinya sedangkan pada sinyal rekonstruksi diterjemahkan sebagai salah satu jenis sinyal *voiced* atau *unvoiced* sehingga akan terjadi kesalahan rekonstruksi yang seharusnya diterjemahkan sebagai sinyal *voiced* dan *unvoiced* seperti sinyal aslinya tetapi sinyal rekonstruksi hanya diterjemahkan sebagai salah satu jenis sinyal *voiced* atau *unvoiced*.

REFERENSI

- Djadjat Sudaradjat, S. A. (2020). Aplikasi Pengolahan Sinyal Suara pada Teknologi Kecerdasan Buatan. *INSANtek – Jurnal Inovasi dan Sains Teknik Elektro*, 88-95.
- Furui, S. (1989). *Digital Speech Processing, Synthesis, and Recognition*. New York: Marcel Dekker.
- Hernando Castaneda Mari, M. J. (2018). Theory , Algorithms , Implementation and Practice of Power Density Signal by Autocorrelation Modeling. <https://www.researchgate.net/publication/325019198>, 1-20.
- Jurafsky. (2018). *Speech and Language Processing*. Third Edition Draft.
- Kala, A. &. (2015). Speech Analysis and Synthesis using Vocoder. *International Journal For Trends In Engineering & Technology*.
- Markel J.D, G. J. (1976). *Liner Prediction of Speech*. New York: Springer Verlag.
- Modi, D. (2015). Speech Cpmression using LPC. *Adaptive Signal Processing Term Paper*, 1-4.
- Prayank, S. &. (2013). Speech Processing. *International Journal of Engineering, Sciences & Emerging Technologies*, 83-87.
- Rabiner, L. R. (1977). On the Use of Autocorrelation Analysis for Pitch Detection. *IEEE Trans. On Acoustics, Speech, and Signal Processing, Vol. ASSP-25, No.1 Feb*.
- Rabiner, L. R. (2007). *Introduction to Digital Speech Processing*. Foundations and Trends in Signal Processing.
- Sudaradjat, D. (1993). *Pemrosesan Sinyal Suara Dengan Metoda LPC*. Bandung: ITB.

PROFIL PENULIS

Ir. Djadjat Sudaradjat, MT. Lulus S-1 tahun 1987, Jurusan Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Bandung. Lulus S-2 tahun 1993, Sub Program Studi Sistim Informasi Listrik, Program Magister Elektroteknik, Program Pasca Sarjana, Institut Teknologi Bandung. Sekarang menjadi Dosen di Universitas Bina Sarana Informatika, NIDN : 0302036201. No. Reg. Pendidik : 20103106503481. <https://sudaradjat.blogspot.com/>