

Analisis Spektrogram Sinyal Suara Asli dan Suara Hasil Konversi Berbasis Derivative Gelombang Glotal

Sriwijanaka Yudi Hartono^{1,a)}, Saidah Suyuti^{2,b)}, Bambang Panji Asmara^{3,c)},
Bayu Adrian Ashad^{4,d)}, Nurul Khaerani Hamzidah^{5,e)}

^{1,2,4}Program Studi Teknik Elektro, Fakultas Teknik, Universitas Muslim Indonesia

³Program Studi Teknik Elektro dan Komputer, Fakultas Teknik, Universitas Negeri Gorontalo

⁵Jurusan Teknik Informatika dan Komputer, Politeknik Negeri Ujung Pandang

email: sriwijanaka.hartono@umi.ac.id^{a)}, saidahsuyuti06@gmail.com^{b)}, bambang@ung.ac.id^{c)}
bayuadrianashad@umi.ac.id^{d)}, nkhamzidah@poliupg.ac.id^{e)}



Abstrak

Ucapan manusia yang dihasilkan oleh suatu sistem produksi yang dibentuk oleh alat – alat ucap manusia. Proses ini dimulai dengan formulasi pesan dalam otak pembicara (penutur), pesan tersebut selanjutny akan diubah menjadi perintah-perintah yang diberikan kepada alat-alat ucap manusia sehingga dihasilkan ucapan yang sesuai dengan pesan yang ingin disampaikan. Transformasi sinyal suara (speech) berbasis derivative gelombang glottal digunakan dengan metode Liljenctranct fant. Pada metode ini parameter yang berperang penting adalah open Quotient (OQ), Speed Quotient (SQ), formant dan pitch. Keempat parameter tersebut sangat berpengaruh dalam menentukan kualitas konversi suara. Spektrogram adalah suatu representasi dari warna suara atau spektral yang bervariasi terhadap waktu yang menunjukkan tingkat insitas energi spektral (density Spectral). Berdasarkan dari data base sinyal suara yang diproses secara pengolahan sinyal digital diperoleh adanya perbedaan ketajaman spektral pada spektrogram antara sinyal asli dengan sinyal konversi, dimana tingkat ketajaman spektral warna hasil konversi lebih tinggi dari pada spektral suara asli. Hal ini dikarenakan adanya penguatan intensitas lebih dari spektral akibat transformasi derivative dari gelombang glottal akibat parameter OQ, SQ, formant dan pitch. Sinyal suara yang dihasilkan tidak berubah baik bentuk sinyal maupun durasinya, yang berubah hanyalah warna suaranya. Untuk membedakan suara hasil konversi dari suara pria lain dan suara wanita yaitu dengan merubah besarnya nilai keempat parameter tersebut secara tepat.

Kata Kunci: sinyal suara, konversi, spektrogram, *pitch*

A. Pendahuluan

Dalam hal melakukan wicara pastilah kita menginginkan informasi yang disampaikan akan dapat diterima dengan baik. Dalam hal ini merupakan suatu tantangan yang dihadapi oleh pendengar adalah mengartikan suara wicara ke dalam unit-unit bahasa yang memiliki suatu arti.

Dewasa ini telah banyak yang menarik untuk melakukan riset pengolahan sinyal digital pada sinyal suara, karena memiliki keunikan tersendiri dan menantang karena persepsi orang dalam menilai suara hanya berdasarkan alat indera pendengaran bukan sulitnya dalam pengolahannya.

Penelitian ini dilakukan dengan metode *Liljenctranct-Fant* dengan transformasi derivative berbasis gelombang glottal untuk konversi sinyal wicara ke berbagai warna

suara tanpa merubah bentuk dan durasi sinyal. Selanjutnya menganalisis bentuk spektrogram dari sinyal asli dan sinyal hasil konversi tersebut.

B. Teori Singkat

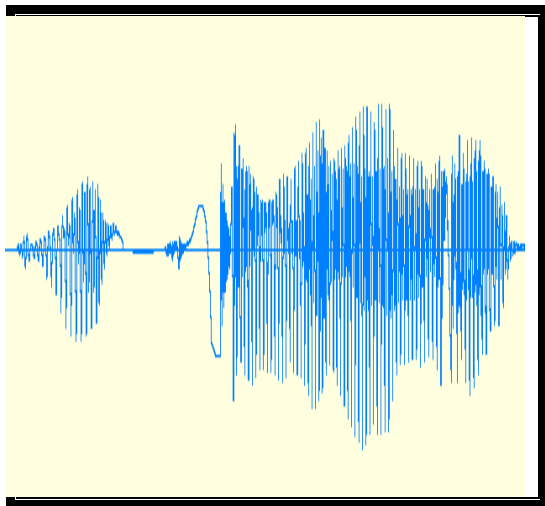
Pengolahan sinyal wicara deawas ini umumnya memanfaatkan kecangihan teknologi komputer yang semakin berkembang pesat. Salah satunya yang digunakan dalam kecanggihannya adalah pengolahan sinyal multimedia pada sinyal suara (*speech*)

a. Sinyal Wicara (*Speech*)

Sinyal wicara (suara) adalah sinyal yang dihasilkan oleh suara manusia dan biasanya mempunyai frekuensi kerja sekitar 0 Hz hingga 5 KHZ. Sinyal ini biasa juga disebut dengan sinyal ucapan yang dapat dianggap sebagai sinyal yang

berubah lambat pada durasi waktu yang sangat singkat yaitu pada durasi waktu sekitar 20 – 40 mS. Pada waktu tersebut karakteristik sinyal wicara dapat dianggap stasioner'

Karakteristik sinyal wicara adalah suatu bentuk gelombang bunyi yang kontinu dengan frekuensi dasar berkisar 100 – 400 Hz dimana didalamnya terjadi serangkaian perubahan harmonik pada fdkuensi dasar yang disebut *formant*. Formant pertama sering juga disebut sebagai pitch. Pitch manusia berbeda satu sama degan yang lainnya terutama pada jenis kelamin (Sriwijanaka, 2007). Berikut ini pada gambar 1 contoh bentuk sinyal wicara yang direkam dengan frekuensi sampling 1200 KHz dengan penuturan kata/ kalimat dalam bahasa indonesia.



Gambar 1. Bentuk sinyal wicara (speech)

b. Representasi Sinyal Wicara

Organ bicara dan organ pendengaran manusia telah didesain dan diciptakan dengan ukuran dan kemampuan tertentu. Berkaitan dengan hal tersebut, tipe komunikasi wicara manusia terbatas pada bandwidth sekitar 7-8 kHz. Sehingga eksitasi sinyal wicara di bagi atas:

- a. Elemental, yang terdiri atas:
 - Sura bergetar (voiced)
 - Suara tak bergetar (unvoiced)
- b. Kombinasi antara voiced dan unvoiced.

c. Derivated Gelombang Glottal metode LF

Proses suara sintesis terlebih dahulu diperkenalkan oleh Fant dan Lilljestrants adalah berbasis derivated gelombang glottal. Model LF memproses sinyal wicara dengan ketajaman bunyi aliran glottal adalah parameter R_k , R_g , R_a dan R_d yang akan menentukan nilai Open Quotient (OQ) dan nilai Speed Quotient (SQ). Parameter LF ini digunakan untuk membangkitkan formant glottal dari beberapa frekuensi formant, magnetudo dan bandwidth.

Nilai OQ dan SQ dalam metode LF ini sangat berpengaruh dalam menentukan kualitas konversi suara. Begitupula estimasi nilai koefisien parameter formant dan pitch.

Besarnya nilai SQ harus lebih besar dari pada nilai OQ dan estimasi nilai koefisien formant dan pitch. Parameter ini berkorelasi langsung dengan fiologikal sinyal wicara yang digunakan untuk sintesa atau konversi sinyal suara.

Parameter OQ adalah merupakan lingkaran vibrator glottal dan juga ditentukan oleh glottal pulse width. Besarnya nilai OQ untuk eksperimen konversi suara berada disekitar nilai range $0,4 < OQ < 1,0$. Sedangkan parameter SQ ditentukan oleh glottal pulse skewness atau sinyal yang tidak simetris yaitu perbandingan antara pening interval dengan closing interval. Besarnya nilai SQ dalam eksperimen $SQ > 0,7$ hal ini akibat adanya korelasi antara suara dengan tekanan phonation. Besarnya nilai SQ untuk eksprimen menurut Fant, nilainya harus berkorelasi dengan nilai OQ. Jika nilai OQ rendah maka nilai SQ harus besar.

Mengenai bentuk spektral dari derivative bentuk gelombang glottal dinyatakan berdasarkan spesifikasi dari frekuensi domain. Serta hubungannya dengan spektral aliran glottal yang dibedakan adanya menambahkan minus 6 dB/slope peroktafnya. Penambahan tersebut diusahakan agar agar dapat

membangkitkan suara yang bunyinya natural (Fant, 2004).

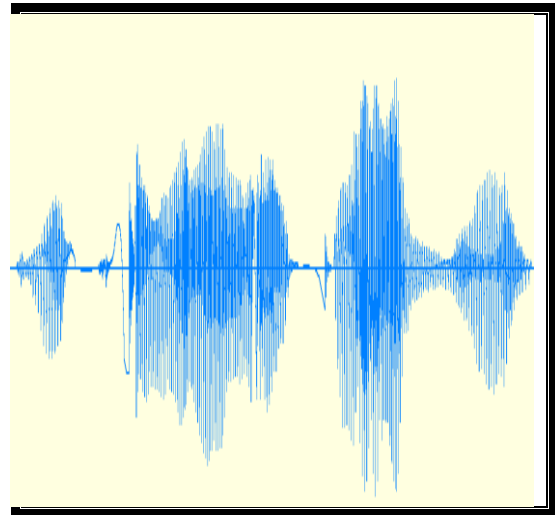
C. Metode Penelitian

Tahapan pertama yang dilakukan adalah pengambilan data base yaitu perekaman suara dengan menggunakan mikropon tipe dinamik dengan frekuensi sampling 12 kHz, 16 bit mono yang menghasilkan file berformat ".wav". data suara yang direkam di ruangan kedap suara adalah berupa rangkaian kata atau kalimat dalam bahasa Indonesia yang diucapkan oleh pengucap tunggal dengan intonasi normal seperti berbicara bias. Tahapan proses selanjutnya yaitu melakukan pengolahan sinyal digital berupa proses preemphasis dan frame serta windowing. Tahapan akhir adalah proses konversi sinyal asli hasil rekaman dengan menggunakan metode LF untuk mengkonversi ke suara pria lain karena suara yang direkam adalah pengucap tunggal pria. Dengan bantuan software *snack tcl* maka file sinyal wav tersebut di proses dirubah kebentuk file ".dat" dan ".par". selanjutnya di proses untuk konversi suara ke warna suara lainnya dengan metode LF yaitu dengan menentukan estimasi nilai OQ dan SQ serta parameter koefisien formant dan pitch. Hasil sinyal dari proses tersebut dapat dilihat dengan menggunakan aplikasi *wave surver*.

D. asil dan Pembahasan

a. Ptosos sinyal asli

Berdasarkan hasil perekamas suara yang dilakukan dengan format .wav. bahwa suatu sinyal memiliki suatu ciri pada suatu kawasan tertentu yakni suatu sinyal yang bergantung pada domain waktu. Akan tetapi pada suatu selang waktu tertentu seperti 20mS ia dapat merupakan suatu fungsi yang tidak bergantung pada waktu sehingga dapat memudahkan untuk analisis atau proses selanjutnya. Pada penelitian ini kata yang direkam adalah contohnya penuturan kata/ kalimat "/bukan lewat lagu/" yang menghasilkan file format ".Wav". seperti yang terlihat pada gambar 2 berikut ini :

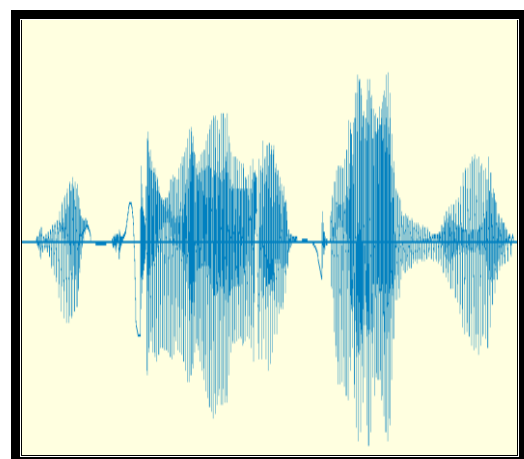


Gambar 2. Bentuk sinyal wicara pengucapan "/ bukan lewat lagu /"

b. Proses Sinyal Konversi

Pada proses konversi sinyal suara asli ke sinyal suara konversi ke suara lain dengan mengkonversi file ".wav" ke file ".dat" dan ".par" dengan bantuan software *snack tcl*.

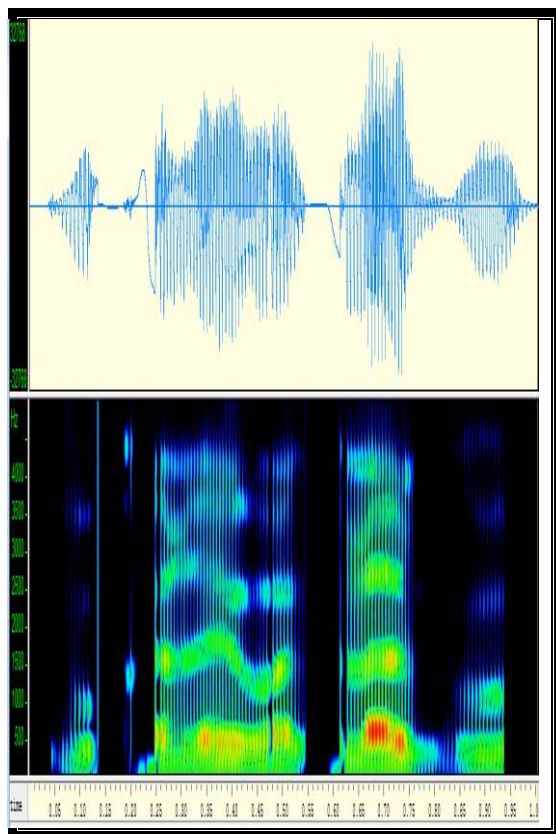
Dari *software snack tcl* untuk diproses dengan metode LF. Untuk konversi suara ke suara lain dilakukan estimasi nilai SQ dan OQ serta koefisien formant dan pitch. Berikut ini contoh hasil eksperimen dengan penuturan "/bukan lewat lagu/" dengan estimasi nilai SQ 1,9 , OQ = 1,25, koefisien formant = 1,4 dan koefisien pitch =1,15. Bentuk sinyal hasil konversi tersebut dapat di lihat pada gambar 3 berikut ini.



Gambar 3 bentuk sinyal hasil konversi dengsn metode LF
 Pada gambar 3 tersebut di atas adalah hasil konversi suara sinyal asli penutur suara pria ke suara wanita.

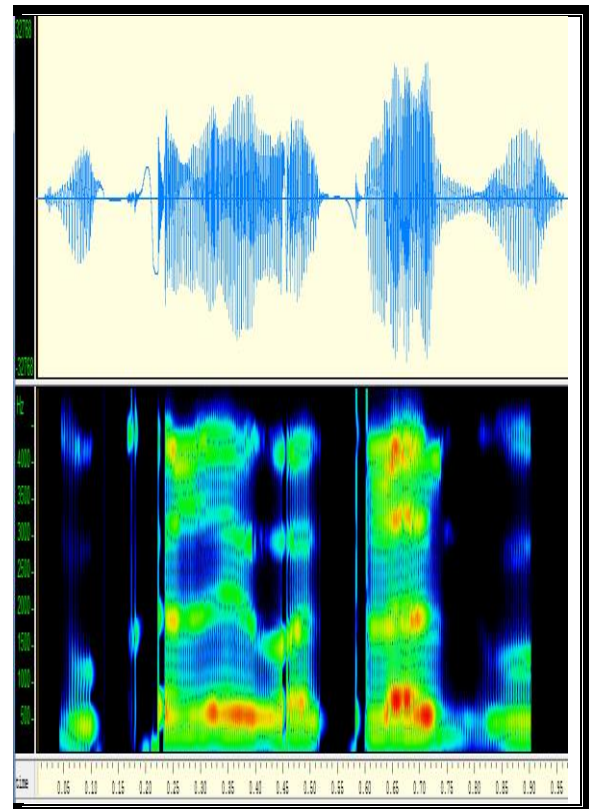
c. Proses Tampilan Spektogram sinyal

Spektogram adalah representasi bentuk spektral yang bervariasi terhadap waktu yang menunjukkan tingkat density spectral atau merupakan visualisasi dari berbagai nilai formant yang dilengkapi level energi yang bervariasi pada daerah domain waktu. Berikut ini akan ditampilkan spektogram suara asli pada gambar 4 dan spektogram suara konversi pada gambar 5.



Gambar 4. Bentuk visual spektogram sinyal suara asli penuturan kalimat “/bukan lewat lagu/”

Sedangkan pada gambar 5 di bawah ini adalah bentuk spektogram sinyal hasil konversi dari sinyal asli.



Gambar 5. Bentuk visualisasi sinyal suara hasil konversi dengan metode LF.

Tingkat kesesuaian nilai SQ dan Oq , koefisien formant dan pitch untuk melakukan konversi warna suara ke warna suara yang diinginkan adalah dengan cara melakukan eksperimen yang berulang-ulang sehingga diperoleh warna suara yang tepat dan tetap natural. Pada bentuk sinyal asli dan konversi terjadi perbedaan karena kibat peningkatan ketajaman warna suara dai keempat parameter dengan metode LF. Misalnya jika dikonversi ke suara wanita maka terjadi peningkatan frekuensi formant karena frekuensi suara pria lebih kecil dri frekuensi suara wanita.

Pada visualisasi spectogram terlihat bahwa menunjukkan pola khas berdasarkan kata/kalimat yang diucapkan dan pola khusu dari sebaran formatnt dari kata/kalimat yang diucapkan, spektogram mapu men visualisasikan pola formant, bandwidth pada pengucapan kata/kalimat tanpa terpengaruh pada tinggi rendahnya

frekuensi resonansi pada saat pengucapan kata-kata atau kalimat.

Terjadi perbedaan ketajaman level spektral antara sinyal asli dan sinyal konversi diakibatkan adanya pengaruh dari parameter metode LF yang berbasis transformasi derivative gelombang glotal dan pengaruh adanya penambahan -6 dB/slope peroktafnya. Spektogram dapat juga digunakan untuk analisa identifikasi suara seseorang.

E. Kesimpulan

Berdasarkan dari hasil penelitian yang dilakukan diperoleh kesimpulan:

1. Sinyal suara (wicara) terdapat sura yang bergetar dan bergetar yang bentuk suatu visual sinyal yang kontinyu terhadap waktu.
2. Konversi sinyal dengan metode LF harus memperhatikan 4 parameter yaitu OQ, SQ, parameter formant dan parameter pitch. Parameter OQ dan SQ merupakan faktor penentu kualitas konversi suara.
3. Terjadi perbedaan ketajaman atau level energi antara sinyal suara dengan sinyal hasil konversi karena pengaruh keempat dari parameter dalam metode LF dengan transformasi derivatibe berbasis gelombang glottal.dan akibat penambahan -6 dB/slope peroktafnya.

F. Saran

Dalam penelitian selanjutnya disarankan :

1. Mencari algoritma yang lebih simpel dan lebih mudah dari metode ini dengan melacak refrensi-referensi hasil riset dan mmembandingkan prosesnya.
2. Peneliti lebih menguasai GNU C jika menggunakan metode ini untuk melakukan sintesa sinyal wicara.
3. Mencoba membuat metode LF ini ke program lainnya seperti program Matlab.

G. Daftar Pustaka

- [1]. Childers Donald G, 2000, Speech processing and Synthesis Toolboxes, John Wiley And Sons, new York.
- [2]. Deller J.R. Proakis, 1993, Discrete Time processing of Speech, macmillan Publishing company New York.
- [3] Fant Gunnar, 2004, Speech Acoustic And Phonetic, Kluwer Academic London.
- [4]. Hartono Yudi Sriwijanaka, 2007, Transformasi Sinyal Wicara (Speech) Berbasis derivative gelombang Glottal, Tesis pascasarjana ITS Surabaya.