

SIMULASI PENGKODE SUARA PITA LEBAR MENGUNAKAN METODE RPE-LTP

Taufan Sandra, Ikhwana Elfitri
Jurusan Teknik Elektro, Fakultas Teknik UNAND

ABSTRAK

Untuk meningkatkan kualitas dan kenyamanan dalam berkomunikasi, telah berkembang suatu teknik pengkodean suara pita lebar (*wideband speech coding*). Dalam pengkodean suara pita lebar (*wideband speech*), pita dibatasi pada 50 – 7000 Hz dan sinyal disampling dengan frekuensi 16 kHz, yang dapat menghasilkan kualitas suara mendekati kualitas berkomunikasi *face-to-face*. Dibandingkan dengan suara telepon *narrowband*, batas bawah rentang frekuensi diturunkan dari 300 Hz menjadi 50 Hz akan memberikan peningkatan kealamian, kenyataan dan kenyamanan. Batas atas rentang frekuensi dinaikkan dari 3400 Hz menjadi 7000 Hz akan memberikan perbedaan bunyi yang berdesis dan karena itu kejelasan suara menjadi lebih baik. Suara dengan *bandwidth* 50 – 7000 Hz tidak hanya memperbaiki kejelasan dan kealamian suara, tetapi juga menambah kejernihan suara dan terdengar lebih nyata. Dalam penelitian ini dilakukan perancangan program simulasi pengkodean suara pita lebar dengan menggunakan metode RPE-LTP. Penggunaan filter LPC orde 16 dengan panjang setiap frame 20 ms dan setiap frame berisi 320 sampel, mampu mempertahankan kejernihan dan kealamian sinyal suara pita lebar yang dikodekan serta menghasilkan suara sintesis hasil pengkodean yang terdengar lebih nyata.

1. PENDAHULUAN

Umumnya sistem pengkodean suara yang digunakan didasari pada *bandwidth* suara telepon, yaitu *narrowband* (pita sempit), yang secara nominal dibatasi sekitar 300 – 3400 Hz dan disampling dengan frekuensi 8 kHz. Sifat keterbatasan *bandwidth* pada *Public Switched Telephone Network* (PSTN) konvensional menimbulkan batasan pada kualitas komunikasi. Peningkatan *end-to-end* jaringan digital, seperti sistem *wireless* generasi kedua dan ketiga, ISDN dan suara pada jaringan paket, memungkinkan penggunaan *bandwidth* suara yang lebih lebar yang akan memberikan suatu kualitas komunikasi yang jauh lebih baik dari PSTN dan memberikan sensasi berkomunikasi *face-to-face* (berhadap-hadapan) [1].

Pengkodean suara (*speech coding*) merupakan proses penting dalam sistem komunikasi suara. Fungsinya adalah merepresentasikan sinyal suara dalam suatu deretan bit (*bit stream*) sehingga dapat ditransmisikan secara digital melalui berbagai sistem, seperti telepon kabel (*wire telephone*), telepon seluler (*mobile phone*), dan telepon via internet (*voice over internet protocol*). Sampai saat ini, berbagai penelitian telah dilakukan untuk menemukan metode yang efisien dalam merepresentasikan sinyal suara, dan beberapa diantaranya telah direkomendasikan oleh International Telecommunication Union (ITU) dan European Telecommunication Standard Institute (ETSI) sebagai standar dalam berbagai aplikasi komunikasi suara. Untuk *narrowband speech*, ITU-T mengeluarkan standar pengkode suara seperti MP-MLQ (G.723.1), LD-CELP (G.728), CS-ACELP (G.729). Sedangkan ETSI mengeluarkan standar RPE-LTP (ETSI full-Rate GSM 6.10) [2].

Kebanyakan frekuensi sinyal suara berada di bawah 7 kHz, walaupun dapat meningkat ke frekuensi yang lebih tinggi terutama sekali pada bunyi yang tidak disuarakan. Untuk meningkatkan kualitas dan kenyamanan dalam berkomunikasi, telah berkembang suatu teknik pengkodean suara pita lebar (*wideband speech coding*). Dalam pengkodean suara pita lebar, pita dibatasi pada 50 – 7000 Hz dan sinyal disampling dengan frekuensi 16 kHz, yang dapat menghasilkan kualitas suara mendekati kualitas berkomunikasi *face-to-face*. Dibandingkan dengan suara telepon *narrowband*, batas bawah rentang frekuensi diturunkan dari 300 Hz menjadi 50 Hz akan memberikan peningkatan kealamian, kenyataan dan kenyamanan. Batas atas rentang frekuensi dinaikkan dari 3400 Hz menjadi 7000 Hz akan memberikan perbedaan bunyi yang berdesis dan karena itu kejelasan suara menjadi lebih baik. Suara dengan *bandwidth* 50 – 7000 Hz tidak hanya memperbaiki kejelasan dan kealamian suara, tetapi juga menambah kejernihan berkomunikasi, kenyamanan pengguna dan suara berkomunikasi yang lebih nyata [1].

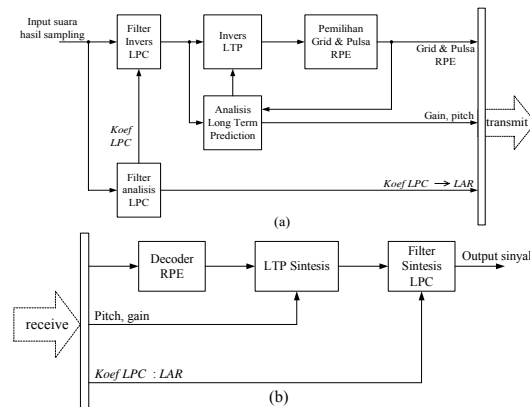
Peningkatan kebutuhan penggunaan *bandwidth* dalam melayani berbagai aplikasi komunikasi 3G tetap menuntut adanya penghematan ukuran data yang ditransmisikan. Karena itu, pengembangan teknologi pengkodean suara pita lebar juga harus memperhatikan ukuran sinyal suara. Dengan menerapkan proses pengkodean sinyal sumber untuk pita lebar, diharapkan ukuran sinyal suara dapat ditekan sekecil mungkin dengan tetap mempertahankan kualitas, kealamian serta kejernihan berkomunikasi. Dengan demikian, sinyal suara dapat ditransmisikan dengan *bit rate* transmisi serendah mungkin, sehingga dapat menghemat penggunaan *bandwidth* frekuensi. Sasaran akhir dari sistem pengkodean suara pita lebar ini adalah untuk

mendapatkan kualitas suara yang secara perceptual tidak berbeda dengan suara aslinya [2].

Berbagai penelitian untuk pengkodean suara pita lebar semakin berkembang. Untuk jaringan telepon, ITU-T telah mengeluarkan beberapa standar pengkodean suara pita lebar, yang pertama adalah Adaptive Differential Pulse Code Modulation (ADPCM) dengan bit rate 48, 56 dan 64 kbps. Berikutnya adalah Modulated Lapped Transform (MLT) yang bekerja pada bit rate 24 dan 32 kbps. Standar terbaru adalah Adaptive Multirate Wideband (AMR-WB) dengan bit rate 6,6 – 23,85 kbps. Untuk sistem komunikasi bergerak, AMR-WB juga telah dijadikan standar untuk sistem GSM/WCDMA oleh 3GPP. Sedangkan 3GPP2 mengeluarkan standar pengkodean suara pita lebar Variable-Rate Multimode Wideband (VMR-WB) untuk sistem CDMA2000 dengan bit rate rata-rata 5,77 kbps untuk mode 4 [2].

2. TEKNIK PENGKODE SUARA RPE-LTP

Pengkode suara jenis RPE-LTP (*Regular Pulse Excitation Long Term Prediction*) terbagi atas tiga bagian utama pada sisi pengirim, begitu juga pada sisi penerima. Gambar 2.1 adalah blok diagram pengkode suara RPE-LTP [8]. Pada sisi pengirim (*encoder*) terdapat blok LPC (analisis LPC dan filter invers LPC), blok LTP dan blok RPE. Pada sisi penerima (*decoder*) terdapat blok sintesis RPE, blok sintesis LTP dan blok sintesis LPC.



Gambar-2.1. Blok Diagram Pengkode Suara RPE-LTP. (a) *encoder*, (b) *decoder*.

2.1. Analisis Prediksi Linear

LPC analisis dilakukan untuk mengekstrak sampel-sampel sinyal membentuk koefisien prediksi Linear (*LPC Coefficient*). Dasar teknik prediksi ini adalah bahwa diantara sampel-sampel sinyal suara yang berdekatan mempunyai korelasi yang tinggi, sehingga beberapa nilai sampel dapat didekati dan direpresentasikan dengan suatu sampel atau koefisien prediksi [11].

Suatu prediktor linear dengan koefisien-koefisien prediksi a_k dan cuplikan sinyal prediksi dilambangkan dengan $\tilde{s}(n)$, sehingga cuplikan yang terjadi dapat dituliskan dalam persamaan sebagai berikut [4],

$$\tilde{s}(n) = \sum_{k=1}^p a_k s(n-k) \tag{2.1}$$

Dalam melakukan prediksi terjadi perbedaan antara sinyal prediksi dengan sinyal aslinya. Perbedaan yang terjadi ini disebut sebagai error (kesalahan) prediksi ($e(n)$). Dan dapat dinyatakan sebagai [4] :

$$e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{k=1}^p a_k s(n-k) \tag{2.2}$$

Nilai error sinyal ($e(n)$) biasa disebut dengan sinyal residu atau sinyal eksitasi. Masalah utama yang timbul dalam analisa prediksi ini adalah menentukan koefisien prediktor $\{a_k\}$ langsung dari sinyal suara, sehingga sifat-sifat spektral filter digital sesuai dengan spektral gelombang suara.

Agar koefisien prediktor $\{a_k\}$ benar-benar menunjukkan spektrum tiap *frame*, maka kesalahan prediksi $e(n)$ untuk keseluruhan sinyal prediksi harus sekecil mungkin [9]. Besarnya kesalahan prediksi biasanya dinyatakan sebagai energi sinyal residu (E). Energi sinyal residu tersebut didefinisikan sebagai [6] :

$$E = \sum_{n=-\infty}^{\infty} e^2(n) \tag{2.3}$$

$$= \sum_{n=-\infty}^{\infty} \left[s(n) - \sum_{k=1}^p a_k s(n-k) \right]^2$$

Karena besarnya nilai E dipengaruhi oleh faktor $\{a_k\}$, maka minimalisasi E bisa dinyatakan berdasarkan turunan parsial E terhadap $\{a_k\}$, yaitu [6] :

$$\frac{\partial E}{\partial a_k} = 0 \tag{2.4}$$

Karena sifat spektral gelombang suara berubah setiap saat maka koefisien prediktor pada waktu n harus diperkirakan untuk jangka waktu yang sangat pendek di sekitar waktu n sinyal percakapan. Berdasarkan hal inilah dilakukan penghitungan koefisien LPC dengan berbagai macam metode. Jadi diasumsikan bahwa segmen sinyal suara nilainya adalah 0 (nol) di luar interval pengamatan [9]. Ada tiga metode umum untuk menghitung nilai koefisien LPC, yaitu metode autokorelasi, metode kovarian dan metode lattice.

➤ **Metode Autocorrelation**

Pada dasarnya metode *autocorrelation* merupakan suatu bentuk prediksi *forward linear*. Didasarkan pada persamaan 2.3 dan 2.4, maka didapatkan persamaan [6] :

$$\sum_{k=1}^p a_k \sum_{n=-\infty}^{\infty} s(n-i)s(n-k) = \sum_{n=-\infty}^{\infty} s(n-i)s(n), \quad (2.5)$$

$1 \leq i \leq p, 1 \leq k \leq p$

Fungsi *autocorrelation* $R(i)$ bisa didefinisikan sebagai persamaan berikut [6] :

$$R(i) = \sum_{n=i}^{N-1} s(n-1)s(n), \quad (2.6)$$

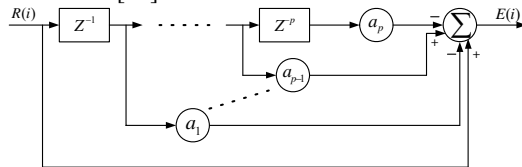
$1 \leq i \leq p$ (asumsi $s(0) = 0$)

Substitusi persamaan (2.6) ke dalam persamaan (2.5), didapatkan persamaan :

$$\sum_{k=1}^p R(i-k)a_k = R(i), \quad \text{atau} \quad (2.7)$$

$$R(i) - \sum_{k=1}^p R(i-k)a_k = 0, \quad 1 \leq i \leq p$$

Persamaan (2.7) tersebutlah yang merupakan suatu bentuk prediksi *forward*. Prediksi *forward* adalah prediksi sampel suara pada suatu saat melalui kombinasi sampel sebelumnya [10]. Hasil dari prediksi *forward* inilah yang merupakan sinyal residu (pada persamaan (2.7) dianggap 0). $R(i)$ adalah nilai koefisien autokorelasi [9]. Untuk lebih jelasnya bisa dilihat pada blok diagram pada Gambar 2.5 berikut [10] :



Gambar 2.2. Blok Diagram *autocorrelation*.

Untuk mendapatkan koefisien a_k , maka bisa digunakan algoritma *levinson durbin*, seperti persamaan berikut [6] :

$$k_i = \frac{R(i) - \sum_{k=1}^{i-1} a_k(i-1)R(i-k)}{E(i-1)} \quad (2.8)$$

$$a_k(i) = a_k(i-1) - k(i)a_{i-k}(i-1), \quad 1 \leq k < i \quad (2.9)$$

$$E(i) = (1 - k(i)^2)E(i-1) \quad (2.10)$$

Dengan asumsi bahwa $a_k = a_k(i)$ pada saat iterasi ke- i (jumlah iterasi maksimum adalah jumlah orde filter) dengan nilai inisial $E(0) = R(0)$ dan $a(0) = 0$ (E adalah sinyal residu analisis).

➤ **Metode Covariance**

Secara umum metode ini sangat mirip dengan metode *autocorrelation*, hanya saja *covariance* dan *autocorrelation* berbeda dalam hal penempatan analisis *window* pada saat menentukan koefisien refleksinya [6]. Pada metode *covariance* energi sinyal residunya menjadi :

$$E = \sum_{n=-\infty}^{\infty} e^2(n)w(n) \quad (2.11)$$

Sebagai catatan disini bahwa nilai $e(n)$ adalah energi sinyal tanpa *windowing* (pada autocorr, nilai $e(n)$ merupakan nilai setelah di-*windowing*), $w(n)$ merupakan analisis *window* (proses *windowing*).

Minimalisasi E bisa dinyatakan berdasarkan turunan parsial terhadap $\{a_k\}$, yaitu [6] :

$$\frac{\partial E}{\partial a_k} = 0 \quad (2.12)$$

Didapatkan suatu turunan rumus sebagai berikut :

$$\sum_{k=1}^p \phi(i,k)a_k = \phi(i,0), \quad 1 \leq i \leq p, \quad (2.13)$$

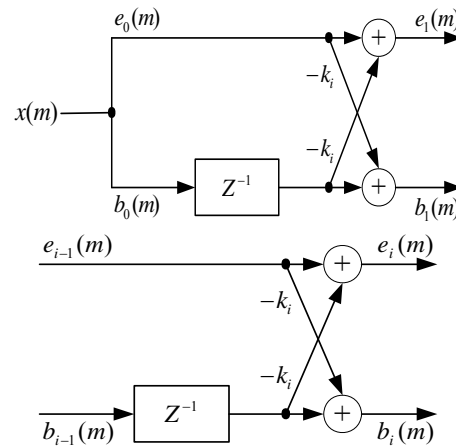
dimana $\phi(i,k)$ adalah fungsi *covariance* dari $s(n)$ yang didefinisikan sebagai :

$$\phi(i,k) = \sum_{n=-\infty}^{\infty} w(n)s(n-i)s(n-k) \quad (2.14)$$

➤ **Metode Lattice**

Metode *lattice* merupakan pengembangan dari metode *autocorrelation*. Pada metode *lattice*, koefisien refleksinya didapatkan berdasarkan *forward* dan *backward predictor* sinyal input (suara)

secara langsung [9]. Koefisien k_i inilah yang dinamakan *PARCOR* (*Partial Correlation*). Secara umum koefisien refleksi merupakan interpretasi negatif *PARCOR* antara dua sampel sinyal. Struktur *lattice* bisa dilihat seperti blok diagram pada Gambar 2.3 [9].



Gambar 2.3. Struktur *Lattice*.

Dalam hal ini $e_i(m)$ adalah *forward error* prediksi, sedangkan $b_i(m)$ adalah *backward error* prediksi dalam i orde filter. Metode *lattice* bisa dirumuskan sebagai berikut [3,9] :

$$k_i = \frac{2 \sum_{m=0}^{N-1} e^{(i-1)}(m) \cdot b^{(i-1)}(m-1)}{\sum_{m=0}^{N-1} [e^{(i-1)}(m)]^2 + \sum_{m=0}^{N-1} [b^{(i-1)}(m-1)]^2} \quad (2.15)$$

Untuk memilih metode yang akan dipakai, perlu diketahui terlebih dahulu faktor performansi, efisiensi dan stabilitas masing-masing metode [11]. Kestabilan pada proses ini adalah suatu hal yang penting karena pada saat dilakukan pengkodean nantinya, bisa terjadi error yang menyebabkan *decoder* menggunakan kode yang berbeda dari yang dikirimkan. Bila ditinjau dari segi performansinya, metode *lattice* dapat menghasilkan kualitas suara yang relatif lebih baik jika dibandingkan dengan metode *autocorrelation* dan *covariance*. Hal tersebut

dikarenakan pada metode *lattice* memiliki tingkat distorsi yang paling kecil [9].

Akan tetapi koefisien refleksi atau partial correlation, memiliki sensitifitas terhadap error kuantisasi. Sehingga koefisien tersebut kurang sesuai untuk dilakukan pentransmisian. Karena inilah, maka diperlukan suatu parameter yang dapat merepresentasikan koefisien refleksi tersebut agar error kuantisasi tidak terlalu mempengaruhi terjadinya perubahan nilai koefisien refleksi. Salah satunya adalah dengan menggunakan transformasi LAR (log area ratio). Seperti persamaan berikut [9] :

$$LAR(i) = \log_{10} \frac{1+k}{1-k} \quad (2.16)$$

Koefisien LAR inilah yang merupakan koefisien prediksi yang akan dikuantisasi dan hasilnya ditransmisikan. Selain ditransmisikan, LAR kembali dikonversi menjadi PARCOR yang kemudian digunakan pada filter invers LPC. Pada filter, dibentuk sinyal suara prediksi dengan mensintesis nilai PARCOR, dan selanjutnya sinyal prediksi dibandingkan dengan sinyal masukan. Selisih yang diperoleh dari perbandingan tersebut dinamakan *LPC residual*.

2.2. Analisis Pitch

Sinyal residu filter LPC memiliki spektrum yang unik dan spesifik yaitu terdapatnya spektrum yang bisa diasumsikan muncul secara periodik [11]. Besar amplitudo spektrum dinamakan *pitch gain* dan periode waktu munculnya spektrum periodik disebut *pitch period* dan *delay*.

Nilai *gain* dan *delay* dapat diperoleh melalui analisis *pitch* pada blok *Long Term Prediction* (LTP). Untuk analisis *pitch*, setiap segmen dibagi menjadi empat subsegmen dan penganalisaan dilakukan di setiap subsegmen tersebut. Tujuannya adalah untuk mengefisienkan proses analisis pada proses eksitasi berikutnya [3]. Hasil filter *pitch* disebut *pitch residual* yang memiliki amplitudo spektrum yang cukup kecil [3].

2.3. Pemilihan Grid dan Pulsa RPE

Pada blok ini kemudian dianalisa calon sinyal eksitasi berdasarkan panjangnya pada setiap subframe. Dalam hal pengeksitasiannya didapatkan sinyal dengan amplitudo bervariasi tetapi dengan jarak antara impuls adalah sama. Sehingga untuk penganalisaannya hanya perlu diketahui posisi impuls pertama. Pada blok RPE nantinya akan dihasilkan jumlah posisi *grid* RPE, nilai eksitasi maksimum dan nilai amplitudo impuls yang nantinya akan dikuantisasi secara skalar.

Pada *decoder*, pulsa-pulsa RPE, *Gain* dan *Delay* digabungkan lagi pada filter sintesis yang identik dengan LPC residual. Kemudian koefisien LAR juga diinterpolasi, dan dikonversi menjadi PARCOR, selanjutnya digabungkan dengan sinyal sintesis keluaran filter sintesis *pitch* untuk membentuk sinyal suara diskrit.

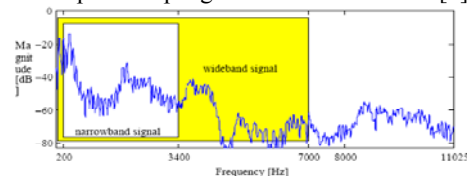
2.4. Suara Pita Lebar (Wideband Speech)

Lebar pita frekuensi yang biasanya digunakan untuk percakapan telepon adalah pada rentang frekuensi 300 – 3400 Hz, yang biasa dikenal dengan *narrowband speech*. Walaupun suara dengan rentang frekuensi 300 – 3400 Hz mempunyai kualitas yang dapat diterima (*toll quality*), ada motivasi yang kuat untuk meningkatkan kualitas suara, terutama dengan adanya teknologi 3G, dengan memperlebar pita frekuensi menjadi 50 – 7000 Hz. Suara pada rentang frekuensi ini biasanya disebut dengan suara pita lebar (*wideband speech*)[2].

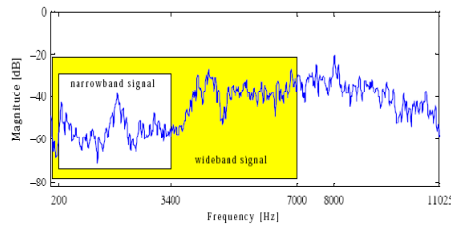
JP Adoul dalam [13] mengatakan ada beberapa kekurangan *narrowband speech*, yaitu : suara terdengar tidak alami, orang yang berbicara terasa berada pada jarak yang jauh dan susah membedakan antara huruf-huruf yang terdengar berdesis (*fricative sounds*), seperti “s” dan “f”. Semua kekurangan ini akan menyebabkan orang yang mendengarkan *narrowband speech* menjadi kelelahan (*listener fatigue*). Karena itulah muncul motivasi yang kuat untuk menerapkan suara pita lebar pada berbagai aplikasi sistem komunikasi suara.

Dengan menurunkan batas bawah rentang frekuensi (*frequency cut-off*) dari 300 Hz menjadi 50 Hz, suara akan terdengar lebih alami dan berarti. Sedangkan menaikkan batas atas rentang frekuensi dari 3400 Hz menjadi 7000 Hz, akan memudahkan pendengar untuk membedakan bunyi-bunyi yang berdesis (*fricative sounds*). Karl Hellwig dalam [14] menegaskan bahwa lebar pita (*bandwidth*) yang lebih besar akan meningkatkan kualitas suara sehingga dapat mengurangi tekanan (*stress*) selama melakukan percakapan telepon. C. Quinquis dalam [15] membuktikan dengan eksperimen bahwa pengkode suara pita lebar (*wideband speech coder*) menghasilkan suara dengan kualitas yang lebih baik dibandingkan *narrowband speech coder*.

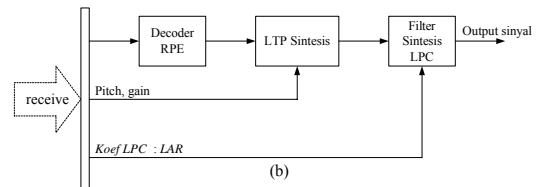
Gambar 2.4 menunjukkan spektrum energi dari suatu sinyal suara yang disuarakan untuk *narrowband* dan *wideband*. Sebuah spektrum energi yang mirip dari suatu sinyal suara yang tidak disuarakan ditunjukkan pada Gambar 2.5. Pada periode suara yang tidak disuarakan, energi penting yang berada diatas 4 kHz difilter untuk kasus suara *narrowband*. Hal ini mempengaruhi kejelasan, seperti perbedaan antara “s” dan “f”. Pada periode suara yang disuarakan, sebagian besar energi berada pada frekuensi rendah, dan menfilter energi dibawah 300 Hz dapat mempengaruhi kealamian suara [1].



Gambar-2.4 Contoh spektrum energi dari sebuah sinyal suara yang disuarakan.



Gambar- 2.5 Contoh spektrum energi dari sebuah sinyal suara yang tidak disuarakan.



Gambar-3.1 Blok Diagram Perancangan Pengkode Suara RPE-LTP. (a) encoder, (b) decoder.

2.5. LPC Pada Pengkode Suara Pita Lebar

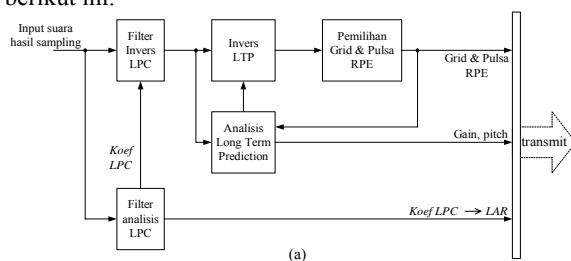
Pengkode suara AMR-WB (*Adaptive Multirate Wideband*) dikembangkan dari pengkode suara ACELP (*Algebraic Code-Excited Linear Prediction*) [16]. Walaupun bekerja dengan multi bitrate, setiap bitrate menggunakan *bitstream* yang sama. *Bitstream* disusun sedemikian rupa sehingga mode dengan bitrate yang lebih rendah dapat ditempatkan (*embedded*) pada mode dengan bitrate yang tinggi. Dengan demikian, *receiver/decoder* selalu dapat melakukan rekonstruksi sinyal suara walaupun sebagian dari bit-bit yang berasal dari bitrate yang lebih tinggi tidak dikirimkan [17].

Seperti beberapa pengkode suara lainnya, ACELP menggunakan metode prediksi linear yang dikenal dengan *Linear Predictive Coding* (LPC). LPC digunakan untuk melakukan estimasi selubung spektrum suara (*spectral envelope estimation*). Salah satu keuntungan menggunakan LPC adalah kemampuannya untuk menurunkan entropy sinyal suara sehingga proses kuantisasi lebih efisien. Hal ini diperoleh dengan cara mendeteksi dan menghilangkan informasi yang dapat diprediksi (*predictable*) dan mengandung perulangan (*redundant*). Dengan demikian, sinyal suara dapat direpresentasikan dengan jumlah bit yang lebih kecil. Pada kebanyakan pengkode suara, hasil estimasi berupa koefisien LPC dikonversi ke dalam parameter *Line Spectral Frequency* (LSF), kemudian dikuantisasi dan ditransmisikan secara digital [2].

3. PERANCANGAN PENKODE SUARA

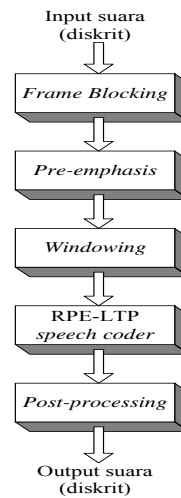
WIDEBAND RPE-LTP

Blok diagram perancangan pengkode RPE-LTP secara umum adalah seperti pada Gambar 3.1 berikut ini.



Sinyal suara yang diinputkan ke sistem pengkode suara merupakan sinyal diskrit yang dihasilkan dari proses sampling. Berdasarkan standar pengkodean suara pita lebar (*wideband speech coding*), maka besarnya frekuensi sampling adalah 16 kHz. Berarti terdapat 16.000 *sample* yang dikirimkan setiap detik. Adapun sinyal yang dihasilkan oleh *decoder* pada sistem pengkode suara juga merupakan sinyal diskrit, atau biasa disebut dengan sinyal suara sintesis.

Proses perancangan sistemnya perlu melalui beberapa tahapan. Tahapan tersebut dapat diilustrasikan seperti blok diagram pada Gambar 3.2 berikut ini [9].



Gambar-3.2 Tahapan perancangan sistem pengkode suara *Wideband RPE-LTP*.

3.1. Frame Blocking

Pada pengolahan suara, model produksi sinyal secara umum digunakan sebagai sebuah analisis. Hal ini akan menghasilkan suatu nilai yang keliru selama durasi sinyal yang panjang. Karena itu dalam penganalisaannya, suara dibagi ke dalam bentuk durasi yang lebih singkat, biasanya disebut *frame* [3]. Sinyal suara diasumsikan *stationer* pada setiap *frame* dan hal ini akan sangat membantu pada proses selanjutnya. Pada sistem ini sinyal hasil cuplikan dibagi ke dalam *frame-frame*, yang setiap *frame* berisi 320 cuplikan. Sehingga lamanya waktu pada setiap *frame* adalah 20 ms.

3.2. Pre-emphasis

Secara umum bentuk sinyal ucapan yang didigitalisasi memiliki rentang dinamik yang tinggi

dan mendapatkan *noise* tambahan [9]. Dengan adanya *noise* ini maka kemungkinan terjadinya *error* kuantisasi akan semakin besar. Untuk mengurangnya dilakukan *pre-processing*, salah satunya adalah *pre-emphasis*. Pada proses ini, sinyal yang telah dibagi ke dalam *frame-frame* ditapis dengan menggunakan filter orde satu. Jadi *pre-emphasis* ini mempunyai efek meratakan spektral. Fungsi transfer *pre-emphasis* adalah sebagai berikut [9].

$$s(n) = s(n) - 0,9s(n-1) \quad (3.1)$$

3.3. Windowing

Untuk meminimalkan ketidakkontinuan sinyal pada awal dan akhir setiap *frame*, maka dilakukan *windowing*. Proses ini membuat sinyal menuju nol pada awal dan akhir setiap *frame*. Jenis *window* yang digunakan adalah *hamming window* [9], yaitu :

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

$$0 \leq n \leq N-1 \quad (3.2)$$

Alasan penggunaan *hamming window* ini karena bentuk spektral yang dihasilkannya lebih *smooth* dan rapat [3].

3.4. Post Processing

Untuk mengembalikan sinyal ke sifat spektral sinyal suara asli, perlu dilewatkan pada *post processing* melalui *de-emphasis*. Persamaan untuk pengembalian sinyal tersebut adalah sebagai berikut [9].

$$s(n) = s(n) + 0,9s(n-1) \quad (3.3)$$

3.5. Algoritma RPE-LTP

A. Pada Encoder

1. *Sampling*.
2. Inisialisasi : set nilai tetapan yang digunakan pada sistem seluler.
3. Segmentasi suara.
4. *Pre-emphasis*.
5. *Windowing*.
6. Analisis LPC, karena orde filter LPC yang digunakan adalah 16, maka pada setiap *frame*, sampel-sampel direpresentasikan oleh 16 nilai koefisien prediksi.
7. Konversi dari *reflection coefficient* atau *partial correlation* menjadi LAR.
8. Interpolasi LAR dan konversi ke bentuk RC atau PARCOR.
9. Filter invers LPC dan menghasilkan sinyal residu LPC.
10. Analisis *pitch*.
11. Dilakukan eksitasi oleh RPE dengan terlebih dahulu membagi *frame* menjadi beberapa *subframe*, dalam setiap *subframe* dibagi lagi menjadi beberapa *grid*.
12. Pemilihan *grid* dan pulsa RPE (termasuk 1 pulsa maksimum).

13. Transmisi secara serial dengan terlebih dahulu melewati *multiplexer*.

B. Pada Decoder

14. Pada *decoder* RPE, dilakukan penggabungan subsegmen, reposisi grid dan pulsa RPE, termasuk pulsa maksimum.
15. Filter sintesis LTP dan penggabungan subsegmen.
16. Interpolasi LAR dan konversi ke bentuk RC atau PARCOR.
17. Lakukan invers filter LPC dengan sinyal input merupakan sinyal hasil *decoder* RPE dan koefisien prediksi RC atau PARCOR.
18. *De-emphasis*.

Orde filter LPC yang digunakan pada perancangan sistem pengkode suara *Wideband RPE-LTP* adalah orde 16. Berdasarkan standar pengkodean suara pita lebar (*wideband speech coding*), jumlah orde ini merupakan jumlah orde filter LPC minimum dengan frekuensi sampling 16 kHz.

4. HASIL DAN PEMBAHASAN

Simulasi pengkodean sinyal suara diawali dengan memberikan input berupa hasil rekaman suara melalui fasilitas perekam suara yang terdapat pada sistem operasi windows, yaitu *sound recorder*. *Sound recorder* akan menghasilkan sinyal suara dalam bentuk file berekstensi “.wav”. Format perekaman dapat dipilih untuk menghasilkan sinyal suara dengan frekuensi sampling yang berbeda. Untuk sinyal suara wideband, format perekaman yang digunakan adalah dengan frekuensi sampling 16 kHz (8 bit - mono). Sedangkan untuk sinyal suara narrowband, format perekaman yang digunakan adalah dengan frekuensi sampling 8 kHz (8 bit - mono). File hasil rekaman ini dijadikan input bagi program simulator yang menggunakan bahasa pemrograman MATLAB versi 7.0.1. Sebagai outputnya, program simulator akan menghasilkan sinyal suara sintesis yang juga dalam bentuk file berekstensi “.wav” dan gambar tampilan berupa grafik sinyal suara input, grafik sinyal suara sintesis hasil pengkodean dan grafik MSE yang dihasil dalam proses pengkodean sinyal suara.

4.1. Encoder Pengkode Wideband RPE-LTP

Sinyal suara yang akan diolah terlebih dahulu dibagi ke dalam bentuk durasi yang lebih singkat, biasanya disebut *frame*. Setiap *frame* berukuran 20 ms yang masing-masingnya berisi 320 *sample*. Dengan demikian dalam satu detik terdapat 50 *frame*. Selanjutnya dilakukan proses *pre-emphasis* terhadap sinyal yang telah dibagi ke dalam *frame-frame*, yaitu menapisnya dengan filter orde satu. Proses *pre-emphasis* akan memberikan efek perataan

spektral. Untuk meminimalkan ketidakkontinuan sinyal pada awal dan akhir setiap frame, digunakan window hamming.

Proses analisis LPC dilakukan pada setiap *frame* dengan menggunakan metode *autocorrelation* yang menghasilkan koefisien LPC. Filter LPC yang digunakan adalah orde 16, yang merupakan standar minimum orde filter LPC untuk *wideband speech coding* dengan frekuensi sampling 16 kHz. Dengan demikian akan dihasilkan 16 koefisien LPC. Kemudian 16 koefisien LPC ini dikonversi menjadi 16 koefisien LAR dan ditransmisikan. Selanjutnya koefisien LAR kembali dikonversi menjadi koefisien LPC guna menghasilkan sinyal residu filter LPC.

Sinyal residu dilewatkan melalui blok LTP untuk dilakukan analisis *pitch*, kemudian dieksitasi oleh RPE menjadi *frame* eksitasi (subframe). Panjang setiap *frame* eksitasi adalah 80 cuplikan (5 ms), sehingga setiap *frame* terdiri dari 4 *frame* eksitasi. Pada blok RPE, sinyal dianalisis pada setiap *frame* eksitasi. Setiap *frame* eksitasi tersebut dibagi menjadi 3 *grid*, sehingga terdapat 3 calon sinyal eksitasi, yaitu sebagai berikut :

- $Grid\ 1 = \{ 1, 4, 7, 10, 13, \dots, 73, 76 \}$
- $Grid\ 2 = \{ 2, 5, 8, 11, 14, \dots, 74, 77 \}$
- $Grid\ 3 = \{ 3, 6, 9, 12, 15, \dots, 75, 78 \}$

Angka dalam tanda kurung pada setiap *grid* menyatakan posisi pulsa maksimum (amplitudo impuls) yang diasumsikan periodik.

4.2. Decoder Pengkode Wideband RPE-LTP

Pada bagian *decoder* dilakukan proses rekonstruksi sinyal suara yang diterima *decoder*. Proses ini berfungsi untuk menghasilkan sinyal suara sintesis yang identik dengan sinyal suara input sistem. Proses sintesis melibatkan 3 blok utama, yaitu *decoder* RPE, LTP sintesis dan filter sintesis LPC. Pada *decoder* RPE dilakukan reposisi *grid*. Kemudian pulsa-pulsa RPE, *Gain* dan *Delay* disintesis pada blok LTP sintesis dan dilakukan penggabungan subsegmen. Pada blok filter sintesis LPC dilakukan interpolasi terhadap koefisien LAR dan mengkonversinya menjadi koefisien LPC. Selanjutnya filter sintesis LPC akan membentuk sinyal suara sintesis berdasarkan keluaran LTP sintesis dan koefisien LPC. Untuk mengembalikan sinyal ke sifat spektral sinyal suara aslinya, sinyal dilewatkan pada *post processing* melalui *de-emphasis*. Sehingga akhirnya dapat diperoleh sinyal suara sintesis yang identik dengan sinyal suara input sistem.

4.3. Kinerja Sistem

Untuk melihat kinerja pengkode suara hasil penelitian, dilakukan perbandingan hasil pengkodean sinyal suara pita lebar (*wideband speech*) menggunakan pengkode *Wideband RPE-LTP*

LTP terhadap hasil pengkodean sinyal suara untuk ucapan yang sama dengan input pada pengkode *Wideband PRE-LTP* yang dikodekan dengan 2 jenis pengkode suara lainnya, yaitu pengkode *Narrowband RPE-LTP* dan *Low Delay Narrowband RPE-LTP*. Pada pengkode *Narrowband RPE-LTP*, sinyal suara yang diinputkan berupa sinyal suara pita sempit (*narrowband speech*) yang disampling dengan frekuensi sampling 8 kHz (8 bit - mono). Sedangkan pada pengkode *Low Delay Narrowband RPE-LTP*, sinyal suara yang diinputkan berupa sinyal suara pita lebar (*wideband speech*) yang disampling dengan frekuensi sampling 16 kHz (8 bit - mono). Kemudian kualitas sinyal suara sintesis yang dihasilkan oleh masing-masing jenis pengkode suara dibandingkan satu dengan yang lainnya. Tabel 4.1 berikut ini merangkum parameter pada masing-masing jenis pengkode suara.

Tabel-4.1 Parameter Masing-masing Jenis Pengkode Suara.

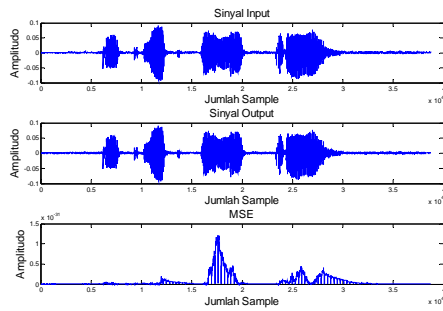
NO	Parameter	Nilai		
		Wideband RPE-LTP	Low Delay Narrowband RPE-LTP	Narrowband RPE-LTP
1	Frekuensi sampling	16 kHz	16 kHz	8 kHz
2	Panjang <i>frame</i>	320 cuplikan (20 ms)	160 cuplikan (10 ms)	160 cuplikan (20 ms)
3	Orde filter LPC	16	8	8
4	Panjang <i>frame</i> Eksitasi (subframe)	80 cuplikan (5 ms)	40 cuplikan (2,5 ms)	40 cuplikan (5 ms)

Untuk pengkode suara *Wideband RPE-LTP*, input yang diberikan adalah sinyal suara dengan nama file *wideband-1.wav*. Sinyal ini merupakan sinyal suara pita lebar (*wideband speech*) hasil rekaman dari ucapan “teknik elektro” dengan format perekaman frekuensi sampling 16 kHz (8 bit - mono). Tabel 4.2 berikut merupakan data tentang sinyal suara input yang digunakan pada pengkode *Wideband RPE-LTP*.

Tabel-4.2 Data Sinyal Suara Input Pengkode *Wideband RPE-LTP*.

Nama File Input	Jenis Suara	Frekuensi Sampling	Jenis Pengkode Suara
wideband-1.wav	Wideband	16 kHz	Wideband RPE-LTP
Ucapan : “teknik elektro”			

Gambar-4.1 berikut merupakan gambar tampilan yang dihasilkan pengkode suara *Wideband RPE-LTP*.



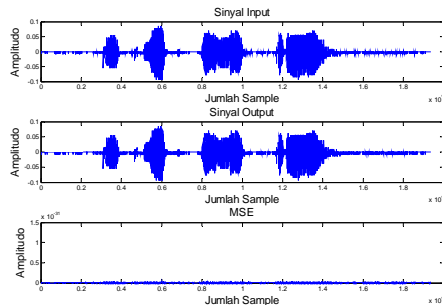
Gambar-4.1 Tampilan hasil pengkode suara *Wideband* RPE-LTP.

Untuk pengkode suara *Narrowband* RPE-LTP, input yang diberikan adalah sinyal suara dengan nama file *narrowband-1.wav*. Sinyal ini merupakan sinyal suara pita sempit (*narrowband speech*) hasil rekaman dari ucapan “teknik elektro” dengan format perekaman frekuensi sampling 8 kHz (8 bit - mono). Tabel-4.3 berikut merupakan data tentang sinyal suara input yang digunakan pada pengkode *Narrowband* RPE-LTP.

Tabel-4.3 Data Sinyal Suara Input Pengkode *Narrowband* RPE-LTP.

Nama File Input	Jenis Suara	Frekuensi Sampling	Jenis Pengkode Suara
<i>narrowband-1.wav</i>	<i>Narrowband</i>	8 kHz	<i>Narrowband</i>
Ucapan : “teknik elektro”			

Gambar-4.2 berikut ini merupakan gambar tampilan yang dihasilkan pengkode suara *Narrowband* RPE-LTP.



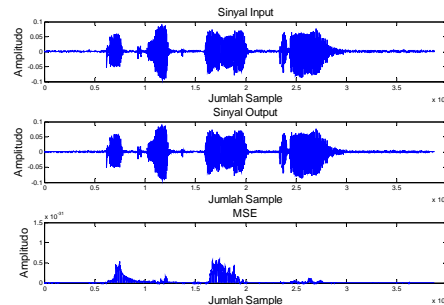
Gambar- 4.2 Tampilan hasil pengkode suara *Narrowband* RPE-LTP.

Untuk pengkode suara *Low Delay Narrowband* RPE-LTP, input yang diberikan adalah sinyal suara dengan nama file *lowdelay-1.wav*. Sinyal ini merupakan sinyal suara pita lebar (*wideband speech*) hasil rekaman dari ucapan “teknik elektro” dengan format perekaman frekuensi sampling 16 kHz (8 bit - mono). Tabel 4.4 berikut merupakan data tentang sinyal suara input yang digunakan pada pengkode *Low Delay Narrowband* RPE-LTP.

Tabel-4.4 Data Sinyal Suara Input Pengkode *Low Delay Narrowband* RPE-LTP.

Nama File Input	Jenis Suara	Frekuensi Sampling	Jenis Pengkode Suara
<i>lowdelay-1.wav</i>	<i>Wideband</i>	16 kHz	<i>Low Delay Narrowband</i>
Ucapan : “teknik elektro”			

Gambar- 4.3 berikut ini merupakan gambar tampilan yang dihasilkan pengkode suara *Low Delay Narrowband* RPE-LTP.



Gambar- 4.3 Tampilan hasil pengkode suara *Low Delay Narrowband* RPE-LTP.

Hasil dari beberapa percobaan lainnya yang dilakukan untuk mendapatkan kinerja sistem dapat dilihat pada Lampiran II. Untuk mendukung gambaran hasil yang didapatkan melalui tampilan grafik, maka dilakukan penilaian kualitas suara sintesis hasil pengkodean berdasarkan penilaian subjektif MOS (*Mean Opinion Score*). Nilai MOS didapatkan melalui hasil survey, yaitu penilaian yang diberikan oleh 10 orang pendengar terhadap masing-masing suara sintesis hasil pengkodean. Penilaian MOS dari hasil survey selengkapnya dapat dilihat dalam tabel pada Lampiran III. Tabel 4.5, Tabel 4.6 dan Tabel 4.7 berikut ini merupakan rangkuman nilai MOS dan MSE yang diperoleh untuk masing-masing jenis pengkode suara.

Tabel-4.5 Nilai MOS dan MSE Sinyal Suara Sintesis Hasil Pengkode *Wideband* RPE-LTP.

NO	Nama File	MOS	MSE
1	<i>wideband-1.wav</i>	5	3.1087×10^{-40}
2	<i>wideband-2.wav</i>	3,9	1.9600×10^{-35}
3	<i>wideband-3.wav</i>	3,5	2.2655×10^{-27}
4	<i>wideband-4.wav</i>	3,8	6.0769×10^{-30}
5	<i>wideband-5.wav</i>	3,7	1.1448×10^{-28}
Jumlah		19,9	$2,3861 \times 10^{-27}$
Rata-rata		3,98	4.7721×10^{-28}

Tabel-4.6 Nilai MOS dan MSE Sinyal Suara Sintesis

NO	Nama File	MOS	MSE
1	narrowband-1.wav	3,5	3.7628×10^{-55}
2	narrowband-2.wav	3,2	3.9135×10^{-35}
3	narrowband-3.wav	3,1	4.4251×10^{-29}
4	narrowband-4.wav	3,2	3.6579×10^{-29}
5	narrowband-5.wav	3,1	2.9202×10^{-31}
Jumlah		16,1	8.1122×10^{-29}
Rata-rata		3,22	1.6224×10^{-29}

Hasil Pengkode *Narrowband* RPE-LTP.

Tabel- 4.7 Nilai MOS dan MSE Sinyal Suara Sintesis

Hasil Pengkode *Low Delay Narrowband* RPE-LTP.

NO	Nama File	MOS	MSE
1	Lowdelay-1.wav	4,8	2.3801×10^{-40}
2	Lowdelay-2.wav	3,6	3.7981×10^{-31}
3	Lowdelay-3.wav	3,3	1.4124×10^{-31}
4	Lowdelay-4.wav	3,6	4.9450×10^{-30}
5	lowdelay-5.wav	3,5	4.7243×10^{-30}
Jumlah		18,8	1.0190×10^{-29}
Rata-rata		3,76	2.0381×10^{-30}

Dari segi tingkat kesalahan (nilai MSE), tabel rangkuman diatas menunjukkan bahwa masing-masing jenis pengkode suara menghasilkan nilai MSE yang sangat kecil sekali, yaitu mendekati nol. Nilai MSE yang sangat kecil ini tidak memberikan pengaruh yang signifikan terhadap kualitas sinyal suara sintesis hasil pengkodean. Artinya, sinyal suara sintesis hasil pengkodean sangat identik dengan sinyal suara aslinya. Hal ini memberikan gambaran bahwa kinerja masing-masing jenis pengkode suara adalah sangat baik sekali.

Namun dari segi kualitas sinyal suara sintesis hasil pengkodean yang diperoleh, terdapat perbedaan yang cukup signifikan antara pengkode *Wideband* RPE-LTP dengan 2 pengkode lainnya. Hasil survey menunjukkan bahwa pengkode *Wideband* RPE-LTP menghasilkan kualitas sinyal suara sintesis hasil pengkodean yang lebih baik dari pada 2 pengkode lainnya, yaitu dengan rata-rata nilai MOS 3,98. Penyebabnya adalah karena kualitas sinyal suara pita lebar (*wideband speech*) lebih baik dari pada kualitas sinyal suara pita sempit (*narrowband speech*) sehingga suara sintesis hasil pengkodean yang didapatkan juga lebih baik. Sedangkan nilai rata-rata MOS untuk suara pita lebar (*wideband speech*) yang dikodekan dengan pengkode *Low Delay Narrowband* RPE-LTP, berada sedikit dibawah nilai rata-rata MOS yang dihasilkan pengkode *Wideband* RPE-LTP, yaitu 3,76. Perbedaan ini terjadi akibat sinyal suara pita lebar (*wideband speech*) dikodekan dengan pengkode untuk sinyal suara pita sempit (*narrowband speech*), yang hanya menggunakan filter LPC orde 8. Sehingga tidak memenuhi standar jumlah orde filter LPC minimum untuk pengkodean suara pita lebar (*wideband speech coding*). Akibatnya, kualitas suara sintesis hasil pengkodean menjadi berkurang.

dimana sinyal suara sintesis hasil pengkodean yang diperoleh dari pengkode *Wideband* RPE-LTP terdengar lebih jernih, lebih alami, dan lebih nyata. Hal ini dibuktikan dengan lebih tingginya nilai rata-rata MOS sinyal suara sintesis yang dihasilkan pengkode *Wideband* RPE-LTP dibandingkan nilai MOS sinyal suara sintesis yang dihasilkan oleh 2 pengkode lainnya. Sehingga secara keseluruhan menunjukkan bahwa kualitas sinyal suara pita lebar (*wideband speech*) lebih baik dari pada kualitas sinyal suara pita sempit (*narrowband speech*) dan pengkodean suara pita lebar (*wideband speech coding*) menggunakan pengkode suara *Wideband* RPE-LTP dengan filter LPC orde 16 dapat mempertahankan kejernihan dan kealamian sinyal suara pita lebar (*wideband speech*) yang dikodekan serta mampu menghasilkan sinyal suara sintesis hasil pengkodean yang terdengar lebih nyata.

5. PENUTUP

5.1. Kesimpulan

Beberapa kesimpulan yang dapat diambil dari pelaksanaan Tugas Akhir ini adalah sebagai berikut :

1. Penggunaan sinyal suara pita lebar (*wideband speech*) dengan *bandwidth* 50 – 7000 Hz dapat meningkatkan kualitas dan kejernihan suara sintesis hasil pengkodean serta terdengar lebih nyata.
2. Pengkode suara *Wideband* RPE-LTP dengan menggunakan filter LPC orde 16 tidak hanya mampu mengkodekan sinyal suara pita lebar (*wideband speech*) dengan baik, tetapi juga dapat mempertahankan kejernihan dan kealamian sinyal suara pita lebar (*wideband speech*) yang dikodekan serta menghasilkan suara sintesis hasil pengkodean yang identik dengan suara aslinya.
3. Rata-rata nilai MOS menunjukkan bahwa kualitas suara sintesis hasil pengkodean yang diperoleh dari pengkode suara *Wideband* RPE-LTP, yaitu 3,98, lebih baik dibandingkan kualitas suara sintesis yang dihasilkan oleh pengkode suara *Narrowband* RPE-LTP dan *Low Delay Narrowband* RPE-LTP.

5.2. Saran

Beberapa saran yang dapat diberikan untuk pengembangan atau penelitian sejenis dengan pelaksanaan Tugas Akhir ini adalah sebagai berikut :

1. Dianjurkan menggunakan orde filter LPC yang lebih tinggi sehingga kualitas sinyal suara sintesis yang dihasilkan lebih baik lagi.
2. Guna mendapatkan hasil yang lebih optimal, sebaiknya dilakukan proses kuantisasi dalam melakukan pengkodean sinyal suara pita lebar agar dalam implementasinya dapat mengurangi nilai bit rate sehingga hasil pengkodean yang didapatkan menjadi lebih efisien.

DAFTAR PUSTAKA

- [1] VoiceAge, Wideband Speech Coding Standards and Applications.
- [2] Elfritri I., Pengkode Suara Pita Lebar, Jurnal Teknik, No.27, Vol.3, Thn.XIV, Universitas Andalas, Padang, 2007.
- [3] Kondoz, A.M., Digital Speech Coding for Low Bit Rate Communication System, Wiley Publisher, New York, 1995.
- [4] Mohammad M.A. Khan, Coding of Excitation Signal In a Waveform Interpolation Speech Coder, Master Thesis, McGill University, Montreal, 2001.
- [5] Nasrun, Nayusrizal, Perancangan Differential Vector Quantizer dan Penggunaannya pada RPE-LTP Speech Coder 11,6 kbps, Skripsi, Universitas Andalas, Padang, 2003.
- [6] Islam, Tammana., Interpolation of Linear Prediction Coefficients for Speech Coding, Master Thesis, McGill University, Montreal, 2000.
- [7] Yoanes, Algoritma Mixed Excitation Linear Prediction (MELP) Untuk Pemampatan Sinyal Suara dengan Kualitas Tinggi dan Laju Bit Rendah, Institut Teknologi Bandung, 1999.
- [8] Orio, Nicola, Speech Coding, Journal, Department of Information Engineering, University of Padova, Modena, 2006.
- [9] Chandra, Dody, Simulasi Pengkode Suara RPE-LPC 9,8 kbps, Skripsi, Universitas Andalas, Padang, 2006.
- [10] Spanias, Andrea, Multimedia Signal Processing (Linear Prediction – *lecture 13*), Arizona State University, Arizona, 2001
- [11] Tjondronegoro, Suhartono, Speech Coding for Communication, Tutorial, Edisi Semester I, Bandung, 2000.
- [12] Gibson, J.D., Speech Coding Methods, Standards, and Applications, IEEE Circuits and Systems Magazine, Vol. 5, No. 4, pp. 3049, 2005.
- [13] Adoul, J.-P., Lefebvre, R., Wideband Speech Coding, in : Kleijn, W.B., Paliwal, K.K. (Eds.), Speech Coding and Synthesis, Elsevier. Amsterdam, pp. 289-309, 1995.
- [14] Hellwig, K., Wideband Speech Telephony : from 1984 to 2007, ETSI Workshop on Speech and Noise in Wideband Speech Coding, Sophia Antipolis, France, 2007.
- [15] Quinquis, C., Quality comparison of wideband coders including tandeming and transcoding, ETSI Workshop on Speech and Noise in Wideband Speech Coding, Sophia Antipolis, France, 2007.
- [16] Besette, B. et al, The adaptive multirate wideband speech coder (AMR-WB), IEEE Transaction on Speech and Audio Processing, vol. 10, pp. 620-636, 2002.
- [17] Byun, K.J. et al, An Embedded ACELP Speech Coding Based on The AMR-WB Codec, ETRI Journal, volume 27, Number 2, pp. 231- 234, 2005.
- [18] Biundo G. dkk, Design techniques for spectral quantization in wideband speech coding, Proc. of 3rd COST 276 workshop on Information and Knowledge Management for Integrated Media Communication, Budapest, pp. 114-119, 2002.
- [19] Guibe G. How H.T., Hanzo L., Speech spectral quantisers for wideband speech coding, European Trans. Telecommunication 12 (6), pp. 535-545, 2001.
- [20] W. B. Kleijn dan K. K. Paliwal, eds., *Speech Coding and Synthesis*. Elsevier, 1995.