APLIKASI SMS BERBASIS TEXT TO SPEECH PADA POCKET PC (SMS APPLICATION BASED ON TEXT TO SPEECH FOR POCKET PC)

Arisman[1], -[2]

[1]Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Saat ini teknologi SMS membantu manusia dalam melakukan komunikasi, namun sebatas membaca pesan teks saja tidak cukup, oleh karena perlu pengembangan aplikasi SMS berbasis Text To Speech (TTS) yang mampu digunakan tanpa harus membaca pesan.

Pada Tugas Akhir ini, telah dibuat aplikasi simulasi teks SMS ke suara menggunakan metoda sintesa suara OLA. Aplikasi TTS memiliki dua blok yaitu NLP dan DSP. Input blok NLP adalah teks yang akan diolah dengan melakukan proses parsing kata dan analisa per kata untuk dicocokkan dengan kamus suara. Output blok NLP adalah kata yang sudah dinormalisasi dalam bentuk unit ucapan diphone. Blok DSP bertugas mengakses sinyal suara tiap diphone, menggabungkan dan mengolah sinyal tersebut sehingga menghasilkan suara yang sesuai dengan kaidah gramatikal Bahasa Indonesia. Metoda OLA dilakukan dengan melakukan windowing, menentukan batas overlap, daerah overlap dan non-overlap. Metoda OLA mampu memberikan komputasi yang cepat.

Tujuan akhir dari Tugas Akhir ini adalah membangun sebuah aplikasi simulasi untuk SMS pada Pocket PC berbasis TTS yang memiliki nilai MOS di atas 3 (cukup) dari skala 5.

Kata Kunci : TTS, NLP, DSP, OLA, diphone, overlap

Abstract

Sincerely, SMS Technology helps human being to do communication, but as reading its text isn't enough. Because of that, SMS application development based on Text To Speech which is used without read it is needed.

In This final Task, have been developed SMS to Speech simulation application with OLA Synthesizer. Text To Speech has two blocks, they are NLP and DSP. NLP input is text which will be caprice with word parsing process and words analyze for looking up in voice dictionary. NLP output is normalized words in diphone unit. DSP function is accesses each diphone voice signal, combines and caprices it in order to produce Indonesian language grammatical rule voice. OLA method is conducted by windowing, determine overlap limit, overlap and non-overlap region. OLA Method gives fast computation.

Final aim from this final task is build a simulation application for SMS on pocket PC based on TTS that has 3 (Enough) MOS Values above from 5 scale

Keywords : TTS, NLP, DSP, OLA, diphone, overlap

BAB I
PENDAHULUAN

1.1 Latar Belakang

Seiring dengan perkembangan teknologi telekomunikasi saat ini, tentunya telah sangat membantu kita sebagai umat manusia yang senantiasa membutuhkan kemudahan dalam melakukan kegiatan sehari-hari. Berbagai inovasi dan pengembangan teknologi sudah dilakukan saat ini baik oleh para pakar atau ahli di bidangnya atau melalui hasil skripsi atau tugas akhir yang dilakukan oleh mahasiswa. Pengiriman pesan melalui Short Message Service merupakan cara pengiriman pesan yang paling populer saat ini. Berbagai macam inovasi dan modifikasi dilakukan pada fasilitas SMS ini agar pengiriman pesan menjadi lebih mudah dan cepat.

Aplikasi SMS saat ini hanya sebatas membaca pesan teks saja, sehingga masih ada pengguna tertentu yang belum mampu memanfaatkan teknologi SMS. Dengan berkembangnya teknologi Text To Speech saat ini, maka teknologi ini dapat digabungkan dengan aplikasi SMS, sehingga suatu saat orang tidak perlu lagi membaca SMS cukup dengan mendengarkan saja. Dalam pembangunan aplikasi yang berbasis Text To Speech perlu diperhatikan faktor-faktor kualitas suara. Hal ini bertujuan apakah suara yang diucapkan dapat dimengerti oleh manusia atau tidak. Untuk itu berbagai metoda dilakukan, misalnya dengan memilih berbagai metoda NLP, DSP dan pemilihan unit ucapan. Faktor ini sangat menentukan kualitas suara yang akan dihasilkan [PEL93].

1.2 Maksud dan Tujuan

Maksud dan Tujuan dari Tugas Akhir ini adalah :

1. Membuat aplikasi simulasi yang dapat membaca inputan teks berbahasa Indonesia dan mengeluarkankan suara berbahasa Indonesia.

2. Menghasilkan aplikasi simulasi dengan metoda sintesa OLA dengan unit ucapan diphone yang mampu mengeluarkan output suara dengan tingkat akurasi tinggi dan kualitas suara yang mampu dimengerti oleh manusia.

1.3 Rumusan Masalah

Masalah yang dibahas dan dicari solusinya dalam tugas akhir ini adalah :

1. Bagaimana mengimplementasikan simulasi teks SMS to Voice dengan metoda OLA?
2. Bagaimana pengaruh penggunaan unit ucapan diphone dalam aplikasi simulasi ini?
3. Menganalisa dengan mengukur hasil implementasi dalam bentuk suara yang dikeluarkan (dalam skala 1- 5) dengan melihat hasil uji MOS-nya.

1.4 Batasan Masalah

Masalah yang akan dibatasi pada Tugas Akhir ini adalah :

1. Bahasa Indonesia yang digunakan harus dengan Ejaan Yang Disempurnakan (EYD) dan tidak menangani prosody (lagu kalimat).
2. Jenis suara yang dihasilkan hanya 1 (satu) jenis suara, yaitu suara pria dengan unit ucapan diphone..
3. Sistem tidak menangani proses komunikasi data SMS dan sistem hanya aplikasi simulasi saja.
4. Sistem tidak menangani homograf.

1.5 Metoda Pemecahan Masalah

Adapun metode pengembangan perangkat lunak ini dilakukan dengan menggunakan metode berorientasi objek, dengan tahap penelitian sebagai berikut:

1. Studi Literatur

   Mempelajari konsep-konsep Natural Language Processing serta Digital Signalling Process secara global yang diperoleh dari buku-buku, jurnal baik dari dalam maupun luar negeri.

2. Pendalaman materi

   Mendalami materi yang akan digunakan, yaitu : konsep NLP (aturan bahasa baku Indonesia) , format file wav dan DSP (Konsep Windowing, OLA).

3. Analisis

   Menganalisa kebutuhan-kebutuhan yang diperlukan untuk pengembangan perangkat lunak, seperti kebutuhan fungsional, kebutuhan performansi, dan kebutuhan antar muka.

4. Perancangan

   Melakukan desain pembuatan perangkat lunak berdasarkan spesifikasi Perangkat Lunak dan Perangkat Keras, sebagai tahap persiapan sebelum melakukan coding.

5. Implementasi

   Mengimplementasikan hasil rancangan dengan menggunakan tool pemrograman Microsoft Visual Studio .Net.

6. Pengujian

   Menganalisa dengan mengukur kualitas dan akurasi suara yang dihasilkan dalam bentuk table dengan melihat nilai MOS.

1.6 Sistematika Penulisan

BAB I Pendahuluan

Membahas latar belakang masalah, perumusan masalah, batasan masalah, tujuan tugas akhir dan metodologi penulisan yang digunakan dalam tugas akhir ini.

BAB II Landasan Teori

Menjabarkan teori-teori tentang Bahasa Indonesia dengan EYD, Sistem Text To Speech (TTS) yang dijadikan landasan untuk pengembangan perancangan perangkat lunak Text To Speech berbahasa Indonesia.

BAB III Analisa dan Perancangan Sistem

Menjelaskan tahap perancangan dengan pemodelan berbasis objek (UML) dan implementasi aplikasi SMS berbasis Text To Speech dalam ujaran bahasa Indonesia yang dibuat dalam tugas akhir ini.

BAB IV Implementasi dan Analisa Hasil Implementasi

Menjelaskan hasil implementasi aplikasi SMS berbasis TTS yang dibuat dengan menganalisa hasil uji dengar responden menggunakan metoda MOS.

BAB V Penutup

Berisi tentang kesimpulan hasil implementasi dan saran untuk pengembangan selanjutnya.

Telkom University

BAB V
PENUTUP

5.1 Kesimpulan

Dari penjelasan bab sebelumnya, maka penulis dapat menyimpulkan :

1. Dari hasil pengujian MOS pada *Text-to-Speech*, menunjukkan kualitas sinyal sintesa dan proses perekaman sangat menentukan *intelligibility* (kejelasan kata yang didengar).
2. Berdasarkan nilai MOS yang diperoleh, aplikasi yang dibuat sudah dikatakan cukup. Terlihat dari hasil MOS 3.68, sesuai dengan syarat minimal nilai MOS untuk syarat komunikasi yaitu minimal 3 nilai MOS.
3. Unit ucapan diphone dengan gabungan dua buah fonem ternyata tidak perlu membutuhkan kapasitas besar karena dari keseluruhan diphone ternyata hanya membutuhkan kapasitas penyimpanan ± 7 MB, bila dibandingkan dengan kapasitas hardisk sekarang min 20 GB. Namun jika dibandingkan dengan unit ucapan fonem, kapasitas yang diperlukan memang harus lebih besar.
4. Karena sintesis yang digunakan sederhana, dengan menggunakan OLA (*OverLap-Add*) dan unit ucapan diphone masih terdengar suara yang mengganda bunyinya atau terjadi sebaliknya dimana ada suku kata yang tidak jelas ucapannya karena durasi diphone yang terlalu pendek.

5.2 Saran

Berikut saran penulis bagi pembaca yang ingin mengembangkan sistem TTS ini :

1. Agar dilakukan perekaman suara dengan peralatan yang memadai (bebas noise) untuk mendapatkan kualitas diphone yang berkualitas.
2. Pada perekaman suara, file diphone terkadang pemotongannya kurang pas, maka diperlukan keakuratan pemotongan file diphone.

DAFTAR PUSTAKA

[ALI91] Alieva,N.F. 1991. *Bahasa Indonesia : Deskripsi dan Teori.* Yogyakarta : Kanisius.

[ALW00] Alwi, Hasan.2000. *Tata Bahasa Baku : Bahasa Indonesia*. Jakarta : Balai Pustaka.

[ARM00] Arman , Arry Akhmad. 2000 .*IndoTTS* , ITB Bandung

[DUT98] Dutoit , Thierry .1998. *A Short Introduction to Text-to-Speech Synthesis* , Belgia , TCTS Laboratory

[LIU05] Liu, Guipeng. 2005. *Mean Opinion Score*. http://umsis.miami.edu/~gliu/. Miami

[MOU88] Moulines, and F. Charpentier.1988. TTS algorithms based on FFT synthesis. ICASSP 88, pp. 667-670.

[MOU90] Moulines, and F. Charpentier.1990. *Pitch Synchronous Waveform Processing techniques for Text-To-Speech Synthesis using diphones*. Speech Communication Vol. 9

[HEL87] Helbert, Schildt. 1987. *Artificial Intelligence Using C*. McGraw-Hill.

[CHA86] Charpentier, M.J. and M.G. Stella .1986. *Diphone Synthesis using an Overlap-Add Technique for Speech Waveforms Concatenation*. ICASSP 86, pp. 2015-2018.

[KON01] Kondoz,A.M . 2001. *Digital Speech 'Coding for Low Bit Rate Communication System '* . England : John Wiley & Sons

[PEL93] Pelton, Gordon. 1993. *Voice Processing*. New York. McGraw-Hill.

[PRE01] Pressman, Roger S. 2001. Software Engineering : A Practitioner's Approach. New York: McGraw Hill.

[PSD04] PSD, Lab. 2004. *Modul Pelatihan TTS dengan Matlab*. Bandung : STT Telkom Bandung.

[SRI01] Mulyati, Sri. 2001. *Implementasi Linguistic Analysis Pada Text To Speech Bahasa Indonesia Dengan Unit Ucapan Diphone*. Bandung : STT Telkom.

[SUN00] Sun, Xuejing. 2000 .*Voice Quality Conversion in TD-PSOLA Speech.* Evanston: Northwestern University

[TET91] Tetschner Walt. 1991. *Voice Processing*. Norwood : Artech House.