

1. Pendahuluan

1.1 Latar belakang

Perkembangan teknologi dari hari ke hari semakin pesat. Sebagai contoh, di negara dengan koneksi internet cepat dan menyeluruh seperti Korea Selatan, penggunaan sistem navigasi sangat membantu dalam kehidupan sehari-hari. Tentunya sistem navigasi yang baik harus memudahkan penggunaannya, selain menunjukkan arah yang tepat, cara penyampaian informasi yang interaktif juga sangat membantu seperti penggunaan teknologi *Text-To-Speech* (TTS) dimana mesin navigasi akan mengucapkan petunjuk arah sehingga pengguna tidak perlu lagi membaca informasi petunjuk arah, pengguna cukup mendengarkan instruksi yang disampaikan oleh mesin. Fitur ini sangat cocok untuk pengguna yang sedang menggunakan kendaraan. Tidak hanya di bidang navigasi, penggunaan TTS juga bisa dilihat di bidang kesehatan, TTS sangat membantu para penderita tunanetra guna memperoleh informasi khususnya dari teks.

TTS adalah system berbasis komputer yang mampu mengucapkan teks secara otomatis [10]. Dalam TTS dikenal istilah prosodi yaitu aspek pengucapan kalimat yang mampu menggambarkan kondisi emosional pembicara, penekanan kata, pemberian jeda kalimat kedalam frasa-frasa, pengaturan ritme dan pengontrolan intonasi dari ucapan [10]. Prosodi merupakan faktor yang sangat mempengaruhi kualitas sebuah TTS. Namun, penentuan prosodi tidaklah mudah karena sangat bergantung pada konteks kalimat, sehingga dibutuhkan sebuah sistem yang secara otomatis mampu memprediksi prosodi yang sesuai berdasarkan konteks kalimatnya. Salah satu aspek prosodi yang memegang peranan penting adalah jeda frasa dalam kalimat. Penentuan jeda frasa ini sangat mempengaruhi intonasi pengucapan sebuah kalimat yang tentunya akan mempengaruhi kualitas pengucapan kalimat.

Berbagai riset telah dilakukan untuk menghasilkan prosodi dengan menggunakan berbagai macam pendekatan atau metode, diantaranya adalah pemodelan prosodi berbasis aturan (*rule based*) [6], dan pemodelan prosodi dan sintesis berbasis *Hidden Markov Model* (HMM) [5, 8]. Pemodelan berbasis aturan tidak cocok untuk digunakan secara global karena sangat tergantung pada bahasa yang digunakan sehingga tidak bisa digunakan untuk TTS multibahasa (*multilingual* TTS). Selain itu, pemodelan berbasis aturan memiliki fleksibilitas yang rendah sehingga kurang cocok untuk digunakan sebagai metode untuk membangkitkan prosodi karena pengucapan (prosodi) yang bisa berbeda antara satu orang dengan orang yang lain. Pemodelan berbasis HMM memiliki fleksibilitas yang relatif lebih baik dibanding pemodelan berbasis aturan karena kemampuan HMM yang sangat baik untuk memodelkan data melalui proses *training*. Selain itu, HMM juga bisa digunakan sebagai *tools* prediksi data yang belum pernah diberikan sebelumnya (tidak terdapat di data *training*).

1.2 Perumusan masalah

Permasalahan yang akan dibahas pada tugas akhir ini adalah:

1. Pembangkitan prosodi berbasis HMM untuk TTS bahasa Indonesia
2. Analisis parameter HMM yang dapat mempengaruhi penentuan jeda frasa dalam sebuah kalimat
3. Analisis kualitas TTS dengan parameter kejelasan, durasi dan nada, antara kalimat yang menggunakan prosodi dengan kalimat yang tidak menggunakan prosodi, dibandingkan dengan kalimat hasil rekaman suara manusia.

Adapun batasan masalah dalam tugas akhir ini antara lain:

1. Tugas akhir ini khusus membahas proses pembangkitan prosodi dari teks
2. Teks *input* terdiri dari karakter alfabet a-z, A-Z, dan tanda baca ‘-’, ‘,’ , ‘!’ , ‘.’ , ‘?’
3. Kalimat *input* harus sesuai dengan struktur kalimat bahasa Indonesia yang baik dan benar
4. Jumlah kalimat input hanya satu kalimat
5. Aspek prosodi yang dibangkitkan adalah jeda frasa, durasi, dan nada.

1.3 Tujuan

Tujuan yang akan dicapai dari tugas akhir ini adalah:

1. Merancang dan mengimplementasikan sistem pembangkit prosodi berbasis HMM untuk TTS bahasa Indonesia
2. Menganalisis pengaruh parameter HMM terhadap penentuan jeda frasa dalam sebuah kalimat
3. Menganalisis kualitas TTS antara kalimat menggunakan prosodi dengan kalimat yang tidak menggunakan prosodi dibandingkan dengan kalimat hasil rekaman suara manusia dari segi kejelasan, durasi dan nada.

1.4 Metodologi penyelesaian masalah

Metodologi yang dilakukan untuk menyelesaikan permasalahan yang dibahas pada tugas akhir ini yaitu:

1. Studi literatur
Pencarian referensi yang menunjang penelitian tugas akhir ini seperti dasar teori TTS, HMM, cara penggunaan pensintesa ucapan MBROLA, dan hal-hal yang berhubungan dengan topik tugas akhir ini.
2. Analisis dan perancangan kebutuhan sistem
Pada tahap ini, dilakukan analisis dan perancangan kebutuhan sistem pembangkitan prosodi dari teks input.
3. Implementasi sistem
Melakukan implementasi menggunakan bahasa pemrograman Java untuk membangun sistem sesuai rancangan pada tahapan sebelumnya.

4. Pengujian sistem
Pengujian sistem dilakukan dengan metode *Mean Opinion Score* (MOS), antara hasil ucapan sistem menggunakan prosodi dengan hasil ucapan sistem tanpa menggunakan prosodi.
5. Analisis hasil pengujian
Pada tahap ini, dilakukan analisis penggunaan HMM untuk menentukan jeda frasa dalam sebuah kalimat yang akan berpengaruh pada penentuan durasi dan intonasi.
6. Penyusunan laporan Tugas Akhir
Dilakukan penyusunan laporan akhir dan dokumentasi yang diperlukan.