



ABSTRAK

SINTESIS NYANYIAN BERBAHASA INDONESIA BERBASIS KONTROL F0 MENGGUNAKAN MBROLA

Oleh

GALIH PRABASIDI

12/334715/PA/14947

Penelitian ini bertujuan untuk melakukan pembuatan sistem suara nyanyian yang berbasis kontrol F0 dengan menggunakan metode penyambung diphone MBROLA. Sintesis nyanyian berbahasa Indonesia dibuat dengan cara menggunakan backend TTS (Text to Speech) yaitu MBROLA. Masukan backend TTS yang digunakan untuk menyediakan informasi-informasi fonem seperti not dan lirik, yaitu berkas MIDI. Saat manusia bernyanyi, grafik F0 yang dihasilkan memiliki beberapa macam fluktuasi, diantaranya yaitu overshoot, vibrato, dan preparation. Agar hasil sintesis suara lebih alami maka informasi F0 perlu ditambah fluktuasi-fluktuasi di atas. Proses penambahan fluktuasi ini yaitu kontrol F0 (frekuensi dasar). Respon impuls untuk masing-masing fluktuasi F0 menggunakan rumus yang dikembangkan oleh Saitou dkk. Selain fluktuasi F0, terdapat juga waktu reaksi perubahan F0, yaitu waktu yang dibutuhkan untuk perubahan dari satu frekuensi ke frekuensi lainnya. Untuk waktu reaksi ini menggunakan persamaan yang dikembangkan oleh Xu dan Sun. Hasil penelitian menunjukkan bahwa sintesis nyanyian berbahasa Indonesia berbasis kontrol F0 menggunakan MBROLA telah berhasil dilakukan. Hasil sintesis nyanyian berbahasa Indonesia dengan menggunakan kontrol F0 adalah lebih baik jika diperdengarkan pada responden yang terlatih musik jika dibandingkan dengan tanpa menggunakan kontrol F0. Responden yang tidak terlatih musik tidak dapat membedakan antara sintesis nyanyian dengan menggunakan kontrol dan tanpa menggunakan kontrol F0.

Kata kunci : sintesis nyanyian berbahasa Indonesia, kontrol F0, MBROLA



ABSTRACT

INDONESIA VOICE SINGING SYNTHESIZER BASED ON F0 CONTROL USING MBROLA

By

GALIH PRABASIDI

12/334715/PA/14947

The purpose of this research was to make an Indonesian voice singing synthesizer based on F0 control using the MBROLA diphone connector method. The synthesis of Indonesian songs was made by using the TTS (Text to Speech) backend, which is MBROLA. Enter backen TTS which was used to provide phoneme information such as notes and lyrics, namely MIDI files. When humans sing, the resulting F0 graph had several kinds of fluctuations, including overshoot, vibrato, and preparation. In order for sound synthesis results to be more natural, then F0 information needs to be added to the above fluctuations. This process of adding fluctuations was control F0 (basic frequency). The impulse response for each of the F0 fluctuations used a formula developed by Saitou et al. In addition to F0 fluctuations, there was also a reaction time of F0, which was the time needed to change from one frequency to another. For this reaction time used the equation developed by Xu and Sun. The results of Indonesian singing synthesizer with F0 control is better if heard by respondents who are trained in music compared to without F0 control. Respondents who are not trained in music cannot differentiate between synthesized song with F0 control and synthesized song without F0 control.

Keywords : Indonesian voice singing synthesizer, F0 control, MBROLA