
Abstrak

Penulisan ini bertujuan untuk mengetahui pengaruh sample rate terhadap hasil voice cloning pada sistem text-to-speech (TTS). Sistem TTS pada penulisan ini berbasis Tacotron untuk cloning suara dan Soundex untuk penukaran silabel. Pengucapan silabel Bahasa Inggris dengan silabel Bahasa Indonesia memiliki perbedaan sehingga diperlukan penukaran silabel supaya sistem yang dibangun berbasis Bahasa Inggris ini mampu mengucapkan silabel Bahasa Indonesia. Penukaran silabel ini berdasarkan fonetik yang sama antara kedua bahasa. Sampel audio yang akan dikloning dapat bersumber dari rekaman suara, video kamera, atau dari sumber lain. Setiap sampel audio memiliki sample rate yang berbeda tergantung format dan sumber audio seperti rekaman dengan format Audio CD menggunakan sample rate 44100 Hz, cam-recorder menggunakan sample rate 32000 Hz pada umumnya, dan sumber lain dengan sample rate yang berbeda. Dari hasil pengujian ditemukan semakin tinggi sample rate yang digunakan maka hasil klon suara akan sangat mirip dengan sampel, namun pendengar semakin sulit untuk mengenali kata-kata yang didengar pada audio hasil dari TTS.

Kata kunci : TTS, sample rate, fonetik, kloning suara, soundex