

ABSTRAK

Perkembangan *speech synthesizer* saat ini meningkat cukup pesat. Berawal dari hasil suara yang didapatkan tidak terdengar alami sama sekali, kemudian menuju ke arah prosodi yang semakin bagus. Salah satu contoh penerapan *speech synthesizer* adalah pada *Text-to-Speech*. *Speech synthesizer* berada pada blok terakhir dalam sistem *Text-to-Speech*. *Speech synthesizer* merupakan sebuah sistem yang mampu menghasilkan suara tiruan manusia dengan sintesa ucapan. Metode *speech synthesizer* yang ada saat ini adalah *formant synthesis*, *articulatory synthesis*, *concatenative synthesis*.

Metode yang digunakan dalam tugas akhir ini adalah metode *diphone concatenation*. Mula-mula sintesa ucapan dibentuk dengan melakukan perekaman suara dan hasilnya disimpan dalam *database*. Kemudian suara rekaman tersebut dipecah menjadi *diphone* yang memiliki transisi antar dua bunyi yang berdekatan (*adjacent phones*) sehingga akan lebih stabil saat digabungkan dengan *diphone* yang lain. Untuk menggabungkan unit ucapan *diphone*, digunakan algoritma *Time Domain Pitch Synchronous Overlap-Add* (TD PSOLA). Dengan menggunakan algoritma TD PSOLA, perangkaian antar *diphone* yang mengandung transisi antar dua bunyi yang berdekatan (*adjacent phones*), menjadi halus tanpa bunyi yang bersifat eksplosif.

Dari metode *diphone concatenation* dan penerapan algoritma TD PSOLA maka sintesis ucapan yang dihasilkan ternyata dapat dimengerti dengan jelas, lancar dalam pengucapan dan datar tanpa intonasi. Hal ini disebabkan adanya sinkronisasi *pitch* terlebih dahulu sebelum dirangkaikan. Terbukti dari hasil uji MOS yang bisa dianggap cukup karena mencapai indeks 3,4. Meskipun untuk beberapa kasus, saat pengucapan kata masih terdengar bunyi “klik” atau “pop”. Semakin beranekaragam unit *diphone* yang terdapat pada *database* akan memudahkan dalam pembentukan sintesis ucapan, namun semakin besar memori yang dibutuhkan untuk menyimpan unit-unit *diphone* tersebut.

Kata kunci : *diphone concatenation*, *pitch*, algoritma TD PSOLA, *speech synthesizer*