

ABSTRAK

Viseme merupakan bentuk visual dari sebuah *phoneme*. Untuk menghasilkan sebuah *talking head model* yang mengucapkan sebuah kalimat, *viseme* merupakan unit terkecil yang dibutuhkan saat proses produksinya dilakukan. *Viseme* sangat bergantung pada bahasa yang dipakai dan tidak bisa diaplikasikan secara umum terhadap semua bahasa. Selain itu, *viseme* juga bergantung pada ekspresi yang ditampilkan model saat berbicara. Hal ini dikarenakan adanya kondisi *conflicting muscle* pada kontraksi otot wajah yang membentuk *viseme* tersebut. Hal ini berdampak pada banyaknya kombinasi *viseme* yang dibutuhkan saat membangun *talking head model*.

Thesis ini membahas tentang klasifikasi dari kombinasi *viseme* tersebut. *Viseme* yang digunakan pada penelitian ini berbasis pada Bahasa Indonesia dan hanya mengambil 1 bentuk pola pengucapan yaitu, konsonan-vokal. Klasifikasi yang dilakukan menggunakan 19 titik penting sebagai representasi dari otot wajah dan 1 titik acuan sebagai standar pada proses normalisasi. *Facial animation model* akan dibangun berdasarkan grup dari hasil klasifikasi. Pada proses ini *free form deformation* (FFD) digunakan untuk mendeformasi model, dan *Bezier curve* diaplikasikan untuk menghasilkan gerakan antar frame acuan.

Penelitian ini mengklasifikasikan 315 kombinasi *viseme* ke dalam 26 grup. Hasil pengujian kesamaan gerakan antara *facial animation model* dan *real human video* mencapai nilai kesamaan sebesar 92,3%. Dengan kata lain, grup *viseme* pada thesis ini telah terbukti efektif untuk menghasilkan persepsi realistis pada *speech animation model*.

Keyword: *realistic speech model*, klasifikasi *viseme*, *co-articulation*, *expression*.