

HIDDEN MARKOV MODELS BASED INDONESIAN VISEME MODEL FOR NATURAL SPEECH WITH AFFECTION

^aEndang Setyati, ^bJoan Santoso, ^cSurya Sumpeno, ^dMauridhi Hery Purnomo

a,b,c,d Electrical Engineering Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

a,b Informatics Department, Sekolah Tinggi Teknik Surabaya, Surabaya, Indonesia

E-Mail: a endang@stts.edu, b joan@stts.edu, c surya@ee.its.ac.id, d hery@ee.its.ac.id

Abstract

In a communication using text input, viseme (visual phoneme) is derived from a group of phonemes having similar visual appearances. For speech emotion recognition, a HMM is trained for each emotion and an unknown sample is classified according to the model which illustrate the derived feature sequence best. This study was based on Indonesian viseme models, which are derived from the phoneme-viseme mapping for Indonesian Language based on blend shape animation, which are resulted of previous author's research. They come from 49 phonemes and 12 Indonesian visemes have been produced, including silent. Until now there has been no other researcher who is interested in using text input from affective Indonesian sentences. Therefore, in this paper we will get a viseme sequence for natural speech that can accept a parameter expressions based on text input Indonesian sentence. We used a HMM for natural viseme sequence. There are two main processes in this study. The first process is tagging model with trigram HMM for the separation of training data from text input Indonesian sentence with affection. The second process is decoding process using viterbi algorithm to obtain a viseme sequence that is used as a synchronization by speech from mouth shape and lip movements. The accuracy is about 83.73%.

Keywords: Hidden Markov Model, Indonesian viseme, Natural speech with affection