Comparative Study of Continuous Hidden Markov Models (CHMM) and Artificial Neural Network (ANN) on Speaker Identification System

Sawit Kasuriya[1], Chai Wutiwiwatchai[2], Varin Achariyakulporn[3] and Chularat Tanprasert[4]
[1,2]Information Research and Development Division,
[4]High Performance Computing Research and Development Division,
National Electronics and Computer Technology Center,
National Science and Technology Development Agency
[3]Asia Bank Co.,Ltd.

ABSTRACT -- This paper reports a comparative study between a continuous hidden Markov model (CHMM) and an artificial neural network (ANN) on a text dependent, closed set speaker identification (SID) system with Thai language recording in office and telephone environment. Thai isolated digit "0-9" and their concatenation are used as speaking text. Mel frequency cepstral coefficients (MFCC) are selected as the studied features. Two well-known recognition engines, CHMM and ANN, are conducted and compared. The ANN system (multilayer perceptron network with backpropagation learning algorithm) is applied with a special design of input feeding methods in avoiding the distortion from the normalization process. The general Gaussian density distribution HMM is developed for CHMM system. After optimizing some system's parameters by performing some preliminary experiments, CHMM gives the best identification rate at 90.4%, which is slightly better than 90.1% of ANN on digit "5" in office environment. For telephone environment, ANN gives the best identification rate at 88.84% on digit "0", which is higher than 81.1% of CHMM on digit "3". When using 3-concatenated digit, the identification rate of ANN and CHMM achieves 97.3% and 95.7% respectively for office environment, and 92.1% and 96.3% respectively for telephone environment.

KEYWORDS -- Speaker identification (SID), Thai language, Continuous hidden Markov model (CHMM), Artificial neural network (ANN), telephone environment

National Electronics and Computer Technology Center (NECTEC)