20 Competing Hidden Markov Models on the Self-Organizing Map

Panu Somervuo

Models associated with the nodes of the Self-Organizing Map (SOM) can learn to become selective to the segments of temporal input sequences. Using the probability as a similarity measure between the input and the models leads to the concept of hidden Markov models (HMMs) as the nodes.
HMMs are stochastic signal models which have commonly been used in speech recognition. Their benefit is to tie separate observations in time together and utilize the time-dependency and order of acoustic phenomena in recognition while at the same time represent the speech patterns in a compact form as a state network. Besides speech recognition, HMMs have also been used in various other tasks, like natural handwriting recognition, text analysis, coding theory, ecology, and molecular biology.
Usually the training of the HMMs is supervised which requires that the segment units to be modeled are pre-defined. However, it might be advantageous to let the system choose the segment units itself. This was experimented in the present work. Input data may consist of unsegmented feature vector sequences with arbitrary lengths.
The unsupervised training of the segment models proceeds by utilizing the competitive-learning principles of the SOM. This is illustrated in Fig. 37.

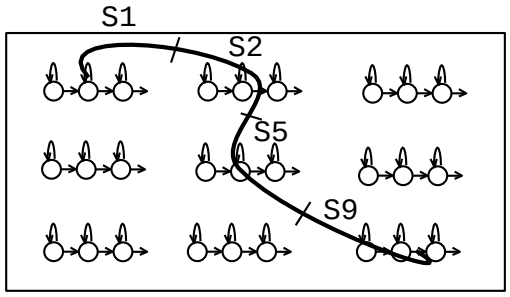

Figure 37: Competitive-learning of segment models on the SOM. Each map node is associated with an HMM having one or more states (three states in this case). The thick line represents the Viterbi segmentation of one input sequence. This corresponds to the best matching unit (BMU) search. The models of the BMUs and neighboring units are then updated by the corresponding segments.

Unsupervisedly derived segment models were experimented in the word recognition task. The recognition rate was 99.1% for a speaker-dependent system with the vocabulary of 350 Finnish words. This was equal to the best results of the supervisedly trained linguistic speech unit models.
The main result of this work was to demonstrate that the SOM gives a framework to train emergent state models by using unsupervised learning. A two-dimensional SOM array offers also a convenient way to visualize the state space of the recognition system.