NEW METHODOLOGY FOR TRAINING HIDDEN MARKOV MODELS FOR BRAIN COMPUTER INTERFACES

M. Gibbs, S. Roberts, I. Rezek
Pattern Analysis and Machine Learning Group
Department of Engineering Science
University of Oxford, Oxford, U.K.

Abstract

We introduce a new method of training hidden Markov models (HMM) for use in a brain computer interface (BCI). Additional information is used in the train processes by generalising the HMM to a coupled hidden Markov Model (CHMM). The additional parts of the model are removed after training to leave the desired HMM. This method significantly improves the classification results.

In this poster we outline a new method for training a brain computer interface (BCI) for the classification of movement planning. The interface makes use of a hidden Markov model (HMM) for classification of the Electroencephalogram (EEG) Data. A HMM is a statistical model which utilises temporal information both in learning its parameters and in classification. In a HMM the probability of an observation, O_t, in this case the EEG data, is dependent on some hidden state, S_t, and the hidden states are related to each other by a first order Markov process, i.e. $P(S_t|S_t\text{-}1)$.

The training algorithm outlined here uses the more general Coupled Hidden Markov Model (CHMM) [RGRed]. These are two, or more, HMMs which are coupled via their hidden states, as can be seen in figure 1. This allows us to model the interaction of the two data sources in a probablistic manner.

HMM training has to be performed by an unsupervised method, which traditionally is performed by the Expectation Maximisation (EM) algorithm [PR98]. The drawback with the EM algorithm is the impossibility of ensuring that the algorithm picks up the relevant *state transitions* in the EEG data. A method that tackles this problem is Maximum a Posteriori (MAP) [RGRed], an extension to EM, which makes use of priors over the model parameters. These priors bias the learning algorithm towards a favoured area in the parameter space. In addition to priors we can introduce further information from data sources, which are related through their hidden states to the EEG (not necessarily a 1 to 1 mapping), to help with the learning process. In this case we use data recorded simultaneously from muscle movement.

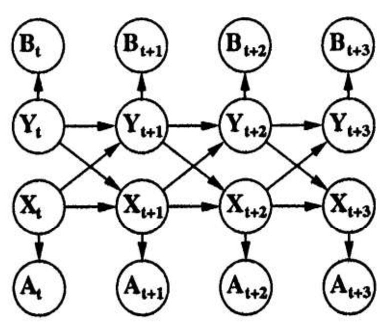

Figure 1: A Directed Acyclic Graph (DAG) of a Coupled Hidden Markov Model with two chains and a unitary lag. Where A_t and B_t are the observation models over the feature spaces and X_t and Y_t are the associated hidden states

In terms of the graphical model one of these chains is the EEG data and the other is the electromyogram (EMG) data from the muscles of interest. These additional sources of information are removed from the model, by marginalisation, after training to leave a model which is based purely on EEG data for classification. This trained model performs significantly better than one trained in the traditional manner.

References

[PR98] W.D Penny and S.J. Roberts. Hidden Markov Models with extended observation densities. Technical Report TR_98_15, Imperial College London, 1998.

[RGRed] Iead Rezek, Michael Gibbs, and Stephen J. Roberts. Maximum a posterior estimation of coupled hidden markov models. *Journal of VLSI Signal Processing_Systems for Signal, Image, and Video Technology*, to be published.