SINGLE-TRIAL DENOISING OF EEG WITH A WAVELET DOMAIN HIDDEN MARKOV TREE

Adam Gerson, Paul Sajda
Laboratory for Intelligent Imaging and Neurocomputing
Department of Biomedical Engineering
Columbia University

We have investigated the application of a wavelet domain hidden Markov tree (wHMT) for denoising electroencephalgraphy (EEG), the aim being to improve single trial detection of mental motor imagery. A wHMT is a directed graphical model that captures statistical dependencies between wavelet coefficients across scale. In the wHMT, the probability density function of wavelet coefficients is approximated with a Gaussian mixture model. Observations of the distributions of wavelet coefficients of EEG evoked potentials reveal that the coefficients have near zero mean and long tails--i.e. they are supergaussian. Such a distribution may be approximated using a two state, zero mean, Gaussian mixture model in which the large number of small coefficients are modeled with a low variance Gaussian, and the small number of large coefficients are modeled with a high variance Gaussian.

The wavelet transform of evoked potentials, and indeed most types of natural signals, inherently exhibits the persistence of large or small coefficients across scale and the clustering of coefficients within scale. To efficiently describe the statistics of wavelet coefficients, each coefficient is associated with a hidden state variable that describes whether the coefficient is in either a high or low variance state as described by the Gaussian mixture model. In order to model the conditional relationships described by clustering and persistence properties of wavelet coefficients, hidden state variables are linked within and across scales. A model that contains links between state variables within scales captures the clustering properties of the wavelet transform and is referred to as a hidden Markov chain. Likewise, a model that contains links *across* scales captures the persistence properties of the wavelet transform and is referred to as a hidden Markov tree. Since modeling the local dependencies within scale becomes computationally expensive we consider the simplified case of a wHMT in which all coefficients within scale are assumed to have the same probability density function; an approximation referred to as tying within scale. A modified Baum-Welch upward-downward expectation maximization algorithm is used to determine the parameters of the hidden Markov tree.

Once the parameters of the wHMT have been determined, denoising is straightforward. Assuming that additive noise increases the variance of all wavelet coefficients, denoising is accomplished by removing the variance due to noise from the variance of each noisy coefficient, conditioned on the hidden state. An estimate of noise variance is derived from the variance of the finest scale wavelet coefficients. The denoised wavelet coefficients are estimated using a hidden state weighted Wiener filter. The inverse wavelet transform is then used to reconstruct the signal. The value of the wHMT over traditional wavelet denoising methods is that it captures the statistical dependencies between wavelet coefficients.

To evaluate the utility of the wHMT for denoising EEG signals and improving single-trial detection of mental activity we denoised single-trial EEG signals collected for a synchronized mental motor imagery task.[1] Data from 9 subjects was used in the evaluation. Separate wHMT's were trained for each subject and for each sensor using data from a time window centered over a period during which the subject was instructed to imagine pressing a button with their left or right index finger. Note that only one trial was used to train each wHMT. Subsequently, 90 left and 90 right trials were denoised for each subject. Logistic regression was used to classify trials as left or right imagined movement. Leave-one-out performance showed a statistically significant increase in the ROC area (A_z) and percent correct (PC) before and after denoising (before A_z=0.64; after A_z=0.66 $p<0.005$, before PC=62%; after PC=64% $p<0.005$). We conclude that denoising using a wHMT is able to improve single trial classification performance of our linear classifier for this set of mental motor imagery data. We are currently investigating the wHMT as a model for directly classifying single-trial EEG.

---

[1] Data were kindly provided by Allen Osman from the Psychology Department of the University of Pennsylvania. A description of the data can be found at http://newton.bme.columbia.edu/competition.htm