HMMFIT: Using Hidden Markov Models to Reconstruct Missing Signals in Multi-Parameter Physiologic Data

Yanen Li*, Yu Sun, Chengxiang Zhai and Lui Sha

Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, United States

Real-time monitoring of physiologic signals is an important clinical tool for intensive care of patients. However, transient corruption or loss of one or more signals could prevent accurate interpretation of the signals and mislead the downstream analysis. Robust reconstruction of physiologic signals utilizing multi-parameter information remains a research challenge. The successful reconstruction of physiologic signals relies on the accurate detection of changes in patient state as well as the precise estimation of intervals of signal corruption. To address this problem, we assume that the time series of multi-parameter signals follow Common Hidden States (Normal, Abnormal and Critical), which can be modeled by the continuous-density Hidden Markov Models (HMMs). The multi-modal signals are modeled as conditional probability density functions given the Common Hidden States of a patient. Yet effective, traditional HMMs are inherently limited in that the given parametric form for the probability density functions is arbitrary. Moreover, there is no successful discriminative training procedure and regularization technique for standard HMMs parameter estimation. Here we propose HMMFIT, a Non-parametric Hidden Markov Model to jointly model the correlation be-tween different modality of multi-parameter signals as well as the temporal correlation of the signal to be reconstructed. A fast learning algorithm using stochastic gradient decent is designed to estimate the model. Results on the Physionet/CiC Challenge 2010 data set A and B (100 series of records for each set) show that the HMMFIT model significantly outperforms the baseline models. We achieve average score of 0.58 and 0.66 at event 1 and event 2 in Set A by 10-fold cross validation; and we achieve score of 0.55 and 0.63 in average at event 1 and event 2 in 100 series of records of Set B, respectively.