Prediction and correction of physiological noise in fMRI using machine learning

T. Ash[1], J. Suckling[2], M. Walter[3], C. Ooi[2], C. Tempelmann[4], A. Carpenter[1], and G. Williams[1]

[1]Wolfson Brain Imaging Centre, University of Cambridge, Cambridge, United Kingdom, [2]Brain Mapping Unit, University of Cambridge, Cambridge, United Kingdom, [3]Department of Psychiatry, University of Magdeburg, Magdeburg, Germany, [4]Department of Neurology, Otto v. Guericke University, Magdeburg, Germany

Introduction:

Multiple tools have been created for the removal of periodic noise originating with the physiological cycles in BOLD fMRI. Many of these rely on data acquired by physiological monitoring apparatus used concurrently with the fMRI acquisition. Unfortunately this physiological monitoring data is sometimes partially or fully corrupted or lost, due to subject motion, equipment failure, lack of recording capabilities or similar problems, making physiological noise removal problematic. To address this problem, we created a physiological data prediction technique that uses the fMRI image data to predict the lost physiological monitoring data, allowing standard physiological detrending tools to be used.

Our physiological data prediction technique uses a multi-class support vector machine algorithm on each fMRI image volume to classify each scan time point as belonging to a certain range of phases, with interpolation used to achieve finer resolution. The training of this classifier can either be on subject's own fMRI and physiological data or on other subjects' data, both cases are discussed below.

Methods:

Data Acquisition

27 healthy volunteers were scanned whilst resting with their eyes closed. At each acquisition echo-planar imaging (EPI) was used to collect 488 volumes with the following parameters: Repetition time, TR = 1.25 seconds, field of view 220mm x 220mm x 120mm, 26 slices, giving 5mm isotropic resolution. Total scan time was 10 minutes and 10 seconds. Concurrent with fMRI scanning, a respiratory bellows and pulse oximeter measured respiratory and cardiac cycles respectively. The physiological monitoring device sampled each cycle at a frequency of 49.82Hz with a time stamp on the output allowing temporal registration to the fMRI time series. T1-weighted, MPRAGE structural scans (with 1mm isotropic resolution) were also acquired for registration to standard coordinate space.

Prior to analysis all data was motion corrected and the first 18 scans removed to allow for T_1-magnetization stabilization.

Physiological Cycle Prediction

We first computed phase of each physiological cycle at the time of each volume acquisition using the phase definitions of RETROICOR [1], as calculated by AFNI's RETROICOR function [2] using the recorded physiological monitoring data. For each physiological cycle, fMRI training volumes were then divided into 6 groups, each group having members spanning 1/6 of phase space, as determined by the concurrent physiological recording. Fifteen separate Support Vector Machine classifiers were then trained to differentiate between each possible group pair using solely the whole-brain image data. On the test image data, we applied all fifteen classifiers and used a pairwise voting scheme [3] to compute the probability that a given image volume belonged to each phase class. Spline interpolation between class probabilities was then used to find the most probable exact phase value.

Training data was either the first half of one subject's recorded fMRI and physiological dataset with testing on the second half of the fMRI data, to test the situation of partially lost data (due to subject motion, for example), or training on 26 of the subjects' fMRI and physiological data, followed by registering to and testing on the left out subject's fMRI data to represent the case of fully absent physiological data (caused, for example, by equipment failure or lack of recording facilities).

Physiological Detrending

After prediction of the physiological monitoring data, we applied RETROICOR, a widely cited physiological noise correction tool, using either the predicted or the recorded physiological data.

Results and discussion:

Using the datasets with partially absent physiological data, predicted phase values for each fMRI time point matched recorded values with an average error of 0.43±0.16 for the cardiac cycle and 0.79±0.08 for the respiratory cycle, with R values of 0.96±0.03 and 0.92±0.02 respectively. In data sets with fully absent physiological data, results were not as accurate, with average errors of 0.73±0.07 and 1.20±0.04 and Rs of 0.93±0.01 and 0.85±0.01 (Figure 1).

When RETROICOR was applied to the fMRI data, the Fourier transform showed similar changes to the frequency spectrum using either recorded or predicted physiological values (Figure 2), showing that the predicted values are accurate enough for useful detrending of data. Detrending of the respiratory cycle in the case of fully absent data was noticeably poorer than cardiac cycle or respiratory cycle in the partially absent case.

Further work is being conducted to ascertain whether this technique is as effective in data sets with different scan parameters, and also whether models trained on scans taken with one set of scan parameters can be applied to data taken with differing parameters.

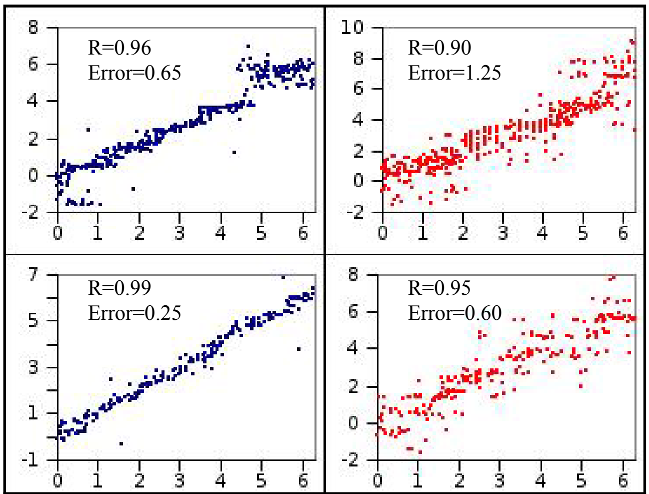

Figure 1: *Plots of predicted phase against recorded phase for the cardiac cycles (blue) and respiratory cycles (red) in fully absent (top) and partially absent data tests (bottom). In each case recorded phase is on the horizontal axis and predicted phase on the vertical axis. Data is shown for a randomly selected representative subject.*

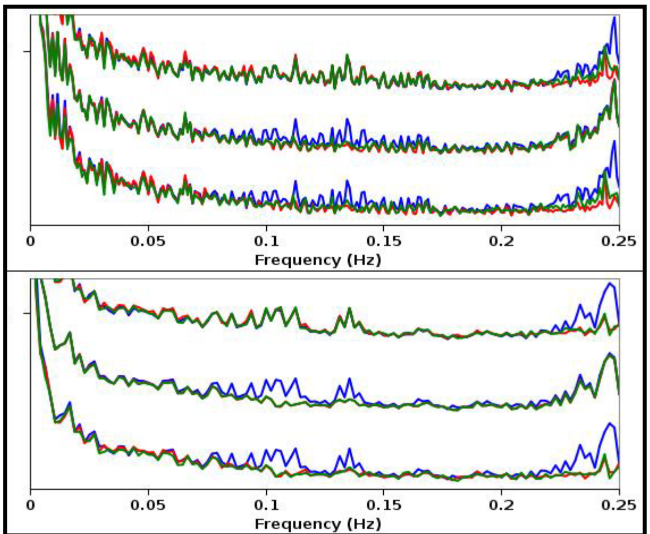

Figure 2: *Average Fourier transforms of voxel time courses before detrending (blue), after RETROICOR with recorded values (red) or RETROICOR with predicted values (green). Graph a) results from fully absent data, graph b) shows partially absent data. In each case respiratory detrending only at the top, cardiac only in the middle and both at the bottom. Data is from same subject as in figure 1.*

Conclusions:

We have demonstrated a machine learning technique that can be used to recreate absent physiological monitoring data from fMRI images. This has been shown to give similar results to genuine recorded physiological data when used with RETROICOR, a widely cited physiological noise removal tool. We therefore believe this could become a useful tool for the preprocessing of fMRI data with partially or fully absent physiological recording data.

References:

1. **Glover, G.H., et.al.,** Magn Reson Med, 2000. 44(1): p. 162-7.
2. **Cox, R.W.,** Computers and Biomedical Research, 1996. 29(3): p. 162-173.
3. **Duan, K.B. and S.S. Keerthi,** Multiple Classifier Systems, 2005. 3541: p. 278-285.