# Unsupervised Clustering of fMRI Time Series with the Granger Causality Metric

**S. B. Katwal[1,2], J. C. Gore[2,3], and B. P. Rogers[2,3]**

[1]Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, United States, [2]VUIIS, Nashville, TN, United States, [3]Biomedical Engineering, Vanderbilt University, Nashville, TN, United States

## INTRODUCTION

Unsupervised clustering methods such as Self-Organizing Map (SOM) or Hierarchical Clustering (HC) are data-driven techniques, which have been used successfully in fMRI data analyses [Ngan, Peltier]. In conventional SOM or HC methods, the Euclidean distance is used as the similarity metric to compare signals [Kohonen]. However, Euclidean distance does not accurately delineate the interference of noise points in fMRI signals. In [Liao], a correlation-based spatio-temporal measure has been introduced that outperforms the Euclidean distance metric in fMRI data analysis. Both Euclidean distance and correlation metrics can distinguish data obtained from similar task paradigms (block-design or event-related) having large and discernible timing variability, of the order of a few seconds [Liao]. However, if the timing difference is small, of the order of a few tens of milliseconds, neither metric is very useful. We propose a new approach based upon Granger Causality to cluster fMRI data having small timing variability. High field fMRI provides high signal-to-noise ratio (SNR) measurements. A low TR during image acquisition at high field helps to detect small differences in latency of the blood oxygenation level-dependent (BOLD) response. In some cases, the temporal difference between signals may imply causal relationship which measures the directional influence (= effective connectivity) [Goebel]. In other cases, the timing difference indicates the presence of delayed signals from macrovasculatures. In any case, the temporal difference can be interpreted in terms of Granger Causality, which, in principle, gives the directional influence. We use this directional influence measure as a similarity metric in an Agglomerative Hierarchical Clustering method to cluster fMRI data into different latency groups. This metric also may segregate brain regions that have different patterns of effective connectivity.

## METHODS

FMRI data simulating event-related acquisitions from human visual cortex [Katwal] were generated as follows: Two second wide boxcar-shaped stimuli were created at a regular interval of 16 seconds (rest period) for a total duration of five minutes (figure 1(a)). The stimuli were convolved with a hemodynamic response function based on a gamma variate function with a repeat time of 250 ms. Appropriate levels of random noise were added to the results to simulate real fMRI time series data. As shown in figure 1(b), the two second wide stimuli were shifted in time to introduce known latencies (temporal differences) in the BOLD responses (figure 1(c)). We created 40 time series of four latency groups with 10 time series in each group. A bi-variate Autoregressive (AR) model was fit to each pair of time series data and the Granger Causality between them was calculated, following [Goebel]. The overall directed influence of first time series (X) on the other (Y) was calculated as the difference between the first-to-second and second-to-first directed causal measures (FXtoY-FYtoX). This measure of directed influence will be different for signals with different latencies relative to the reference signal (0 latency). At zero latency, the overall influence is zero. At other latencies, the overall influence from leading signal to the lagging is always positive and unique for the given latency. This overall influence, which is also the measure of temporal difference between signals, was used as a similarity metric in Hierarchical Clustering to cluster signals exhibiting timing variability.

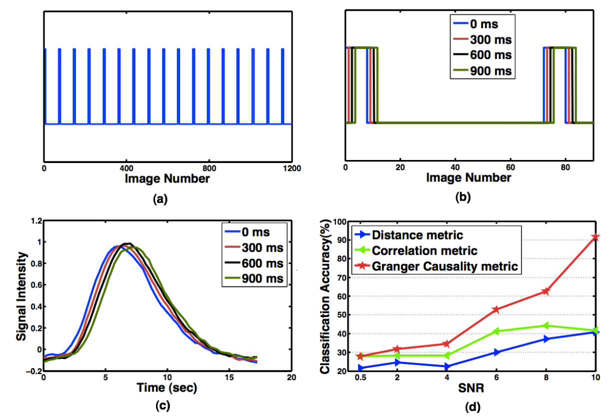


*Figure 1: (a) Stimulus signal. (b) Blow-up of stimuli with four latencies: 0 ms, 300 ms, 600 ms and 900 ms. (c) BOLD responses from four stimuli. (d) Performance (average of six simulations) of three Hierarchical Clustering metrics on the simulated fMRI data-set at various SNRs (Latencies in the signals : 0, 100, 200, 300 ms).*

Visual stimuli were generated by a two-second flashing of checkerboard images at an 8Hz contrast reversal rate followed by a 16 second fixation cross, both repeating for five minutes. The stimuli were presented to the left and right hemifields of the human subject with a delay of 112 ms between the hemifield onsets. Gradient echo planar images of 1mm x 1 mm x 2 mm voxel size at a 250 ms TR were acquired on a Philips Achieva 7T. Regions-of-interest (ROIs) were defined on a single slice in the calcarian fissure of the right and left primary visual cortex.

## RESULTS

Figure 1(d) shows the comparison of the performance of HC with three different metrics: Euclidean distance, correlation and Granger Causality on the simulated data-set having 0, 100, 200 and 300 ms latencies at various SNRs (0.5 – 10). All three metrics performed better with the increase in SNR. As expected, the correlation metric performed better than the Euclidean distance metric as correlation delineates noise points in fMRI signal better than the Euclidean distance. The Granger Causality metric achieved a better classification than the other two metrics at all SNRs. At SNR=10, the Granger Causality metric achieved more than 90% accuracy which was more than two times with the other two metrics. Figure (2) shows the comparison between the Euclidean distance and the Granger Causality metrics on clustering the right and left primary visual cortex fMRI data acquired at 7T. The Euclidean distance achieved a classification accuracy of 46.67% where the Granger Causality metric achieved 93.33% accuracy.

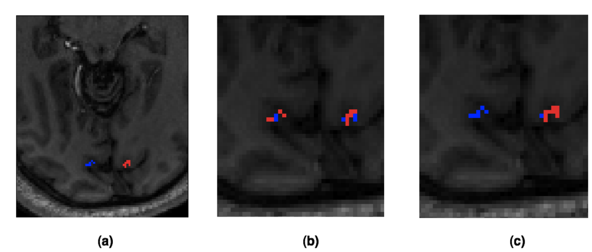


*Figure 2: Classification of right (red) and left (blue) hemisphere ROI voxels with a latency of 112 ms between the corresponding BOLD responses. (a) Theoretical classification. (b) Results from the Euclidean distance Metric based HC (Classification Accuracy: 46.67%). (c) Results from the Granger Causality Metric (Classification Accuracy: 93.33% ). In (b) and (c), blue blobs on right hemispheres and red on left indicate misclassified voxels from the clustering method.*

## CONCLUSIONS

The Granger Causality metric works better than other common metrics to cluster fMRI data with timing differences. With image acquisition at a low TR, it is capable of discerning small timing differences, down to 100 milliseconds. It is the most suitable of all existing metrics at high SNRs to delineate effectively connected regions in brain. The metric can be used in data reduction and clustering methods such as SOM and HC for fMRI data analysis.

## REFERENCES

**Ngan** SC et al. *MRM.* 41:939-46 (1999). **Peltier** SJ et al. *HBM* 20:220-26 (2003). **Liao** L et al. *IEEE Med. Img.* 27:1472-83 (2008). **Goebel** R et al. *MRI* 21:1251-61 (2003). **Katwal** SB et al. *17th Proc. ISMRM* (2009). **Kohonen** T. *IEEE Proc.* 1464-80 (1990)