Mutual Information Rate as an Objective Criterion for Comparison of fMRI Experimental Designs

B. D. Ward[1], Y. Mazaheri[2]

[1]Department of Biophysics, Medical College of Wisconsin, Milwaukee, WI, United States, [2]Center for Functional Magnetic Resonance Imaging, UCSD, La Jolla, CA, United States

INTRODUCTION: Information Theory provides a framework to analyze time-series with few assumptions about the structure of the signal and the nature of the uncertainty (noise). By considering the fMRI experiment as a communication system, various results from Information Theory [1] can be applied. Specifically, we have applied mutual information rate (MIR) to threshold fMRI data sets. The corresponding p-values were calculated without making noise distributional assumptions.

THEORY: (a) Entropy and Mutual Information: The input stimulus function takes on values from the source alphabet *X*. The entropy of the input stimulus function is the average information content of a symbol from *X*: $H(X) = -\sum_{X} P(x_i)\log_2 P(x_i)$ (1)

The reconstructed output takes values from the output alphabet *Y*. The conditional entropy of *X*, given the observed output from *Y*, is: $H(X \mid Y) = -\sum_{X,Y} P(x_i, y_j)\log_2 P(x_i \mid y_j)$ (2)

The mutual information, which indicates the information about *X* obtained from an observation of *Y*, is defined using (1) and (2): $MI(X;Y) = H(X) - H(X \mid Y)$ (3)

(b) Theoretical Bound on Mutual Information Rate: The theoretical bound on the mutual information rate for communication over a noisy channel is defined:

$$C = \max_{P(x)} MI(X;Y) \quad (4)$$

According to the Hartley-Shannon Law [1], the channel capacity, C (bits/sec), is a function of system bandwidth, B (Hz), and S/N = signal-to-noise power ratio:

$$C = B\log_2(1 + S/N) \quad (5)$$

METHODS: fMRI images covering the visual cortex were obtained with the following imaging parameters: TR = 1000 ms, N = 240. The input stimulus function was a random binary sequence (p(ON) = p(OFF) = 0.5). The ON and OFF blocks, within each experiment, were of equal length L, with L = 1, 2, 3, 4, 5, 8, and 20 sec. For L = 20 sec., a non-random periodic design was used.

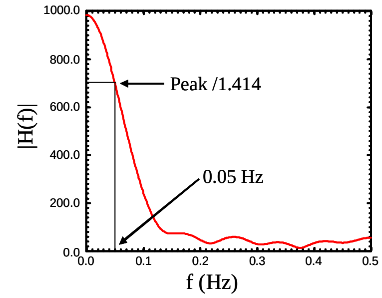

Figure 1: Magnitude of System Transfer Function H(f) for a Single Voxel

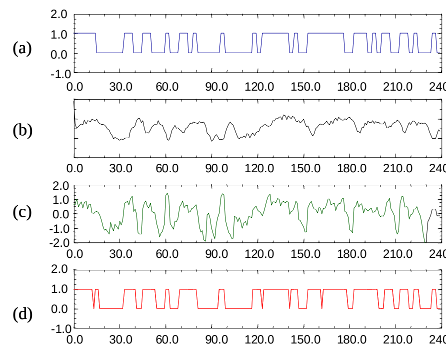

Figure 2: Reconstruction of the Input Stimulus Function

RESULTS AND DISCUSSION: Deconvolution [2] was used to estimate the system impulse response function h(t) for each voxel. The system transfer function H(f) is the Fourier transform of h(t), as illustrated in Fig.1 for a particular voxel. From Fig. 1, we see that the voxel hemodynamic response function acts as a low-pass filter. For this particular voxel, the system bandwidth is approximately B = 0.05 Hz, which corresponds to a channel capacity of approx. C = 0.17 bits/sec by the Hartley-Shannon Law. A Kalman filter was used to reconstruct the input stimulus function [3] for each voxel (see Fig. 2), which can be used to estimate the MIR. Shown in Fig. 2 is the input stimulus function (Fig. 2a), the fMRI signal z(t) (Fig. 2b), the Kalman filter state estimate x(t) (Fig. 2c), and the reconstructed input stimulus function obtained by thresholding the Kalman filter output (Fig. 2d). As shown in Fig. 3, the channel capacity (blue horizontal line) represents the theoretical limit on the rate at which information can be transmitted through the channel (i.e., single voxel of fMRI experiment). The source information rate (green curve) is the rate at which information is being input to the channel from the source (i.e., the input stimulus function). MIR (red curve) is calculated based on the actual error probabilities in reconstructing the input stimulus function. For this particular experimental paradigm, the maximum MIR occurs at a block length of about 3 seconds. In Fig.4, MIR was used to threshold the functional image (MIR_{thr} = 0.072 bits/sec, corresponding to P_{thr} = 0.000001).

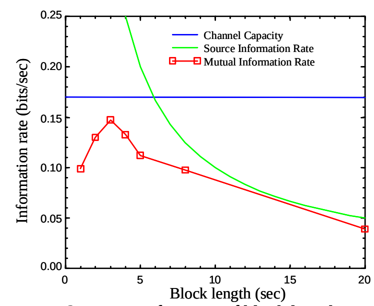

Figure 3:MIR as a function of block length

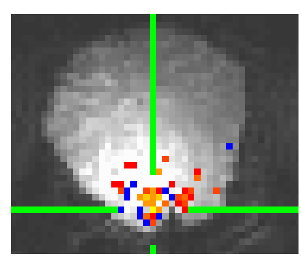

Figure 4: MIR thresholding of activation map.

CONCLUSIONS: MIR is limited by channel capacity for small block lengths, and by the rate at which information is being transmitted by the source for large block lengths. Hence, maximum MIR is achieved at some intermediate block length. MIR can be used to threshold fMRI data sets. The corresponding p-values can be calculated without making noise distributional assumptions.

REFERENCES

[1] Taub H. and Schilling D., Principles of Communication Systems, 1971.
[2] Ward, B.D., Deconvolution Analysis of fMRI Time Series Data, AFNI Documentation, 2002.
[3] Ward, B.D., *et al.*, *Proc. ISMRM*, p. 896, 2003.