# Detection of neural activities in FMRI using Jensen-Shannon Divergence

**Jayanta Basak**                                                    *basak@netapp.com*
*NetApp India Private Limited*                              *basakjayanta@yahoo.com*
*Advanced Technology Group*
*Bangalore, India.*

## Abstract

In this paper, we present a statistical technique based on Jensen-Shanon divergence for detecting the regions of activity in fMRI images. The method is model free and we exploit the metric property of the square root of Jensen-Shannon divergence to accumulate the variations between successive time frames of fMRI images. Theoretically and experimentally we show the effectiveness of our algorithm.

## 1. INTRODUCTION

Functional Magnetic Resonance Imaging (fMRI) is increasingly gaining in popularity as a non-invasive technique for assessing various clinical situations and in better understanding of the functioning of human brain [1, 2, 3, 4, 5]. The basis of fMRI is the different magnetic properties of oxygenated and deoxygenated blood. Due to a stimulus, increased flow of oxygenated blood into regions of brain activity causes the changes in the MR signal. This results in the corresponding changes in MRI map which are captured as four dimensional (x, y, z, t) fMRI images. Automated, robust and fast detection of the activated brain regions from the entire sequence of fMRI images is a challenging task [6]. First, images based on blood oxygenation level dependent (BOLD) contrast [7, 8] have a very low signal-to-noise (SNR) ratio. Second, adequately high temporal resolution (smaller time between successive frames) restricts the spatial resolution in the image registration process. As a result, each (x, y) plane in the image sequence is only about $64 \times 64$ or $128 \times 128$ with regions of activity occupying a few (dozen or so) pixels. Therefore, it is difficult to use traditional image processing operators and spatial constructs (such as traditional image segmentation, checking connectivity, shape detection, etc.) in the localization of activity in these images. Consequently, various statistical and signal processing methods [9, 10] are used to make statistical inferences about the regions of activity in fMRI images.

One of the most widely used approach for detecting active regions in fMRI images is performed by the computation and subsequent thresholding of a statistical parameter map subjected to the t-test based on the assumption of Gaussian temporal noise. The unpaired Student's t-statistic with pooled normal error is commonly used [11, 12] to estimate the true variance using the sample variance. Many other methods [6] of producing statistical parameter maps have also been proposed (for example, using correlation analysis [13, 14] or the non-parametric Kolmogorov-Smirnoff test [8]). In this class of methods a threshold has to be chosen (empirically or theoretical) and the results obtained are dependent on the threshold that is used. Various other methods based on using the wavelet transform [15, 16], principal component analysis [17], independent component analysis [18, 19], subspace modeling [20] and clustering [21] have also been developed. In parallel, methods for improving sensitivity (for example, by including spatial extent of the region of activation) [22] have also been developed.

In this article, we introduce a statistical method for detecting the regions of activity in fMRI images based on the Jensen-Shannon divergence [23, 24, 25]. This particular method differs from the conventional t-test or ANOVA techniques in the sense that it does not depend on the general linear model. Due to the robustness and insensitivity to noise, Jensen-Shannon divergence is gaining popularity in the statistician community and has been successfully applied in image

segmentation [26] earlier. However, the possibility of using the Jensen-Shannon divergence in detecting activity regions in fMRI images has not been explored so far. Here we provide a method for detecting the activities in fMRI images using the Jensen-Shannon divergence.

The rest of the article is organized as follows. In Section 2, we describe our algorithm which includes a description of the Jensen-Shannon divergence, the way we apply this measure to detect the regions of activities in fMRI images, and an empirical analysis to show the validity of our algorithm. In Section 3, we demonstrate the effectiveness of our algorithm on some synthetic and real-life images. Finally, we conclude in Section 4.

## 2. ALGORITHM

### 2.1 Description of JS Divergence
Jensen-Shannon divergence [23, 24, 25] measures the difference between two discrete distributions. Let two different discrete probability distributions p and q are given as $p = [p_1, p_2, \ldots, p_n]$ and $q = [q_1, q_2, \ldots, q_n]$ where $p_i$ denotes the probability of a random variable X taking the *i-th* value. For example, if we have two different coins then their probability distributions of 'Head' and 'Tail' can be represented as $[p_1, p_2]$ and $[q_1, q_2]$.

The divergence between the two discrete distributions p and q is given as

$$JS(p, q) = -\alpha_p H(p) - \alpha_q H(q) + H(\alpha_p p + \alpha_q q) \tag{1}$$

where $\alpha_p, \alpha_q \in [0,1]$ are two positive constants indicating the respective weights for the distributions subject to $\alpha_p + \alpha_q = 1$. H(.) denotes the Shannon entropy, i.e.,

$$H(p) = -\sum_i p_i \log p_i \tag{2}$$

For $\alpha_p = \alpha_q = 0.5$, $JS(p, q)$ is symmetric unlike the Kullback-Leibler divergence. Although Jensen-Shannon divergence does not guarantee the triangular inequality of a metric, the square root of the divergence follows the metric property (as shown in [27, 28]).

### 2.2 Application of JS divergence to fMRI signal detection
The four dimensional fMRI images (x, y, z, t) can be considered as the spatio-temporal signals, where in each time frame, the activation occurs over a few pixels, and it propagates over a sequence of time frames depending on the hemodynamic response function.

In the case of Jensen-Shannon divergence (JS), since $\sqrt{JS}$ is a metric, we have

$$\sqrt{JS(w_l(t_i), w_l(t_j))} + \sqrt{JS(w_l(t_j), w_l(t_k))} \geq \sqrt{JS(w_l(t_i), w_l(t_k))} \tag{3}$$

for any $t_i < t_j < t_k$. $w_l(t)$ represents the pixel statistics over a chosen window at a certain location l at a time frame t. For example, we can choose a 7 x 7 x 5 window at a specific location (x, y, z) at different time frames. Equation (3) reveals that

$$\sqrt{JS(w_l(t_1), w_l(t_n))} \leq \sqrt{JS(w_l(t_1), w_l(t_2))} + \sqrt{JS(w_l(t_2), w_l(t_3))} + \ldots + \sqrt{JS(w_l(t_{n-1}), w_l(t_n))} \tag{4}$$

for any n. Thus we can add the square root of the divergence ($\sqrt{JS}$) between every consecutive pair of time frames and preserve the activation if there exists any.

The overall algorithm is described in Figure 1. First we define an accumulator array A(x, y, z) and initialize A = 0 for every (x, y, z). Then for every $t \in \{1, 2, \ldots, n-1\}$ (assuming that there are n

time frames available) and for every location (x, y, z), we compute the Jensen-Shannon divergence $JS\big(w_{x,y,z}(t), w_{x,y,z}(t+1)\big)$ where $w_{x,y,z}$ denotes the three dimensional window centered at (x, y, z). We then accumulate the variations between successive time frames in terms of the square root of the Jensen-Shannon divergence. Finally we threshold the accumulator array with certain user defined threshold and obtain the regions of activity. Note that, it is also possible to recover the time frames where exactly the stimulus has started by adding one more dimension to the accumulator A.

---

Input : L slices of M x N fMRI images at each time frame. There are T such time frames.
Output :L slices of M x N output image.
**begin**
   Initialize an accumulator array A(x, y, z) = 0
   where $x \in \{1,2,\ldots,M\}$, $y \in \{1,2,\ldots,N\}$, $z \in \{1,2,\ldots,L\}$
   Define a window size (2m+ 1, 2n + 1, 2l + 1) where $m,n,l \geq 1$.
   **for** every $t \in \{1,2,\ldots,T-1\}$
      **for** every $(x,y,z) \in \{(m,n,l),\ldots,(M-m,N-n,L-l)\}$
      **begin**
         get the window $w_{x,y,z}(t)$ centered at (x, y, z) from time frame t;
         compute p $\leftarrow$ normalized histogram of $w_{x,y,z}(t)$ ;
         get the window $w_{x,y,z}(t+1)$ centered at (x, y, z) from time frame t + 1;
         compute q $\leftarrow$ normalized histogram of $w_{x,y,z}(t+1)$ ;
         Update $A(x,y,z) \leftarrow A(x,y,z) + \sqrt{JS(p,q)}$
      **end**
   **end**
   Threshold A with a user defined threshold; output thresholded A.
**end**

---

**FIGURE 1**: The algorithm based on Jensen-Shannon divergence for detecting activation regions in fMRI images.

## 2.3 Analysis

In this section, we empirically analyze the effectiveness of the proposed method of applying Jensen-Shannon divergence. We approximate the distribution over a window volume by a histogram. It is necessary because by definition, JS-divergence considers only the discrete distribution. Let the distribution in the original window be represented as

$$p(x) = \{p_1, p_2, \ldots, p_n\} \tag{5}$$

subject to $\sum_i p_i = 1$. $p_i$ represents the probability of the pixels taking the i-th intensity level. After stimulation, let a fraction of pixels be moved from the i-th intensity level to the j-th intensity level. Thus the modified discrete distribution after stimulation is

$$q(x) = \{p_1, p_2, \ldots, p_i - \Delta, p_{i+1}, \ldots, p_j + \Delta, p_{j+1}, \ldots, p_n\} \tag{6}$$

where $\Delta$ represents the change in the density of the i-th intensity level. It may be possible that due to activation at any time point, voxels at different intensity levels change their intensity values. However, here we assume that the activation is local in nature and affect the voxels with similar intensity values such that the activated voxels belong in the same bin of the histogram or at most neighboring two or three bins. The Jensen-Shannon divergence is given as

$$JS = \frac{1}{2}\left(p_i \log p_i + p_j \log p_j\right) + \frac{1}{2}\left((p_i - \Delta)\log(p_i - \Delta)\right) + \frac{1}{2}\left((p_j + \Delta)\log(p_j + \Delta)\right)$$
$$- \left((p_i - \Delta/2)\log(p_i - \Delta/2)\right) - \left((p_j + \Delta/2)\log(p_j + \Delta/2)\right) \tag{7}$$

In Equation (7), all other bins apart from i and j do not contribute to the measure. Considering that $\Delta = \alpha p_i = \beta p_j$,

$$JS = \frac{p_i}{2}\log\left(\frac{1-\alpha}{(1-\alpha/2)^2}\right) + \frac{\alpha p_i}{2}\log\left(\frac{1-\alpha/2}{1-\alpha}\right) + \frac{p_j}{2}\log\left(\frac{1+\beta}{(1+\beta/2)^2}\right) + \frac{\beta p_j}{2}\log\left(\frac{1+\beta}{1+\beta/2}\right) \tag{8}$$

where $\alpha$ represents the fractional decrease in the number of pixels having i-th intensity and $\beta$ is the fractional gain in the number of pixels having the j-th intensity value. Considering that $\alpha \leq 1$ for all i, we neglect the higher order terms in $\alpha$ such that

$$JS = \frac{\alpha^2 p_i}{4} + \frac{p_j}{2}\log\left(\frac{1+\beta}{(1+\beta/2)^2}\right) + \frac{\beta p_j}{2}\log\left(\frac{1+\beta}{1+\beta/2}\right) \tag{9}$$

The Jensen-Shannon divergence (Equation (9)) behaves in two different ways in two cases for (i) $\beta < 1$, and (ii) $\beta > 1$. Let us analyze these two cases separately.

**Case I**: Since $\beta < 1$, we neglect the higher order terms in $\beta$ (similar to that of $\alpha$) such that

$$JS = \frac{\alpha^2 p_i}{4} + \frac{\beta^2 p_j}{4} \tag{10}$$

i.e., $JS = (\alpha + \beta)\dfrac{\Delta}{4}$ \hfill (11)

$$\text{i.e., } JS = (\alpha + \beta)\frac{\Delta}{4} \tag{11}$$

Therefore, $\sqrt{JS} = \sqrt{\dfrac{1}{p_i} + \dfrac{1}{p_j}}\dfrac{\Delta}{2}$ \hfill (12)

$$\text{Therefore, } \sqrt{JS} = \sqrt{\frac{1}{p_i} + \frac{1}{p_j}}\,\frac{\Delta}{2} \tag{12}$$

In other words, given $p_i$ and $p_j$, $\sqrt{JS}$ varies linearly with $\Delta$ independent of the condition that $i < j$ or $i > j$.

**Case II**: Since $\beta \geq 1$, we can approximate Equation (9) as

$$JS = \frac{\alpha^2 p_i}{4} + (\beta - \log\beta + 2\log 2)\frac{p_j}{2} \tag{13}$$

If $\beta \geq 1$, we have

$$JS = \left(1 + \frac{\alpha}{2}\right)\frac{\Delta}{2} \tag{14}$$

Thus when $p_j$ is very small such that $\beta \gg 1$, Equation (14) reveals the fact that JS measure is independent of $\beta$ and depends on $\alpha$ and $\Delta$. The dependency of JS is approximately linear with $\beta$. However, the measure is independent of the condition whether i < j or i > j.

Thus in both the cases (Equations (12) and (14)), $\sqrt{JS}$ behaves symmetrically to the rising and falling part of the hemodynamic response curve. The behavior of the divergence is also independent of any assumption on the distribution.

## 3. RESULTS

### 3.1 Synthetic Images

In order to establish the effectiveness of Jensen-Shannon divergence, first we considered synthetically generated random data. A random noise of amplitude in the range [50–110] has been generated over a sequence of 80x80 images of sequence length 25 (thus the synthetic images are three dimensional (x, y, t) instead of four-dimensional images in the fMRI). We then added synthetic activation to the random noisy images. Synthetic activation is generated by convolving a synthetic stimulus (which is a step function) with a hemodynamic response function (Figure 2) given as [29]

$$h(t) = (t/t_1)^{d_1} \exp\left(-(d_1/t_1)(t-t_1)\right) - c(t/t_2)^{d_2} \exp\left(-(d_2/t_2)(t-t_2)\right) \qquad (15)$$

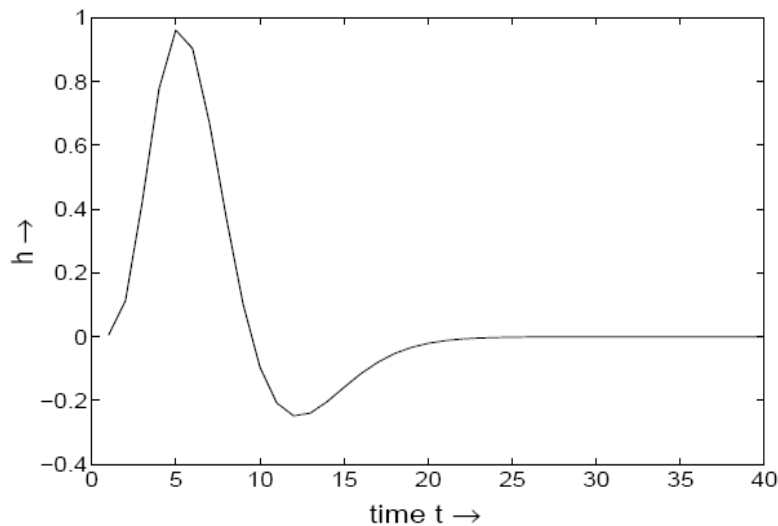with five parameters $t_1, t_2, d_1, d_2,$ and $c$.



**FIGURE 2**: A typical hemodynamic response function h(t) for the auditory cortex.

We added two synthetic activation at the (x, y) locations (30, 30) and (50, 50). The starting and stopping times of the stimuli are (3, 10) and (10, 20) respectively. We tested the effectiveness of our algorithm with two different types of hemodynamic response functions, one for the auditory cortex and the other for the motor cortex. For the auditory cortex, the parameter values are approximated as [29] $t_1 = 5.4$, $d_1 = 6$, $t_2 = 10.8$, $d_2 = 12$, and $c = 0.35$. For motor cortex, the parameter values are $t_1 = 5.5$, $d_1 = 5$, $t_2 = 10.8$, $d_2 = 12$, and $c = 0.4$. Figure 3 illustrates the regions of activity detected for two different hemodynamic responses and for different amplitude of synthetic stimuli ranging from 30 to 60. Note that, we consider the stimulus amplitude to be much less than the noise amplitude in order to make a low signal-to-noise ratio.
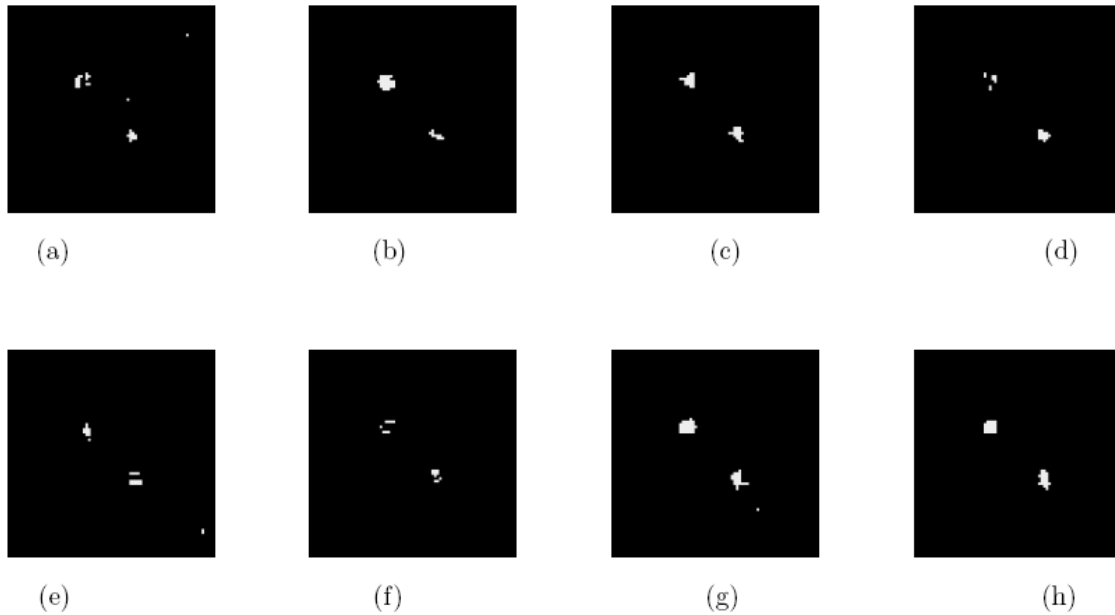
**FIGURE 3**: Regions of activation detected by our algorithm in synthetic noisy images of size 80 × 80 with noise amplitude in the range [50–110]. Activations corresponding to hemodynamic response to a stimulus (step function) are added at the locations (30, 30) and (50, 50). (a),(b),(c), and (d) correspond to the regions detected for stimulus amplitude 30,40,50, and 60 respectively with auditory cortex hemodynamic response. (e),(f),(g), and (h) correspond to the regions detected for stimulus amplitude 30,40,50, and 60 respectively with motor cortex hemodynamic response.

## 3.2   fMRI Images

We tested the effectiveness of Jensen-Shannon divergence in detecting the activation regions in fMRI images. We considered the fMRI data from the fMRI data center [30] particularly the dataset used by Hirsch, Rodriguez, and Kim [31]. Each data set in this experiment consists of sequences of 128 × 128 images with sequence length 21 over 36 time frames. In the experiment by Hirsch et al. [31], subjects performed three cognitive tasks namely, object naming, integer computation and same-different discrimination. We considered fMRI images for the first task i.e., object naming. As mentioned by Hirsch et al. [31], the brain areas involved in the object-naming task (object-naming subsystem) are left inferior frontal gyrus (Brodmann's areas 44 and 45), left superior temporal gyrus (Brodmann area 22) and left medial frontal gyrus (Brodmann Area 6). Figure 4 illustrates the results obtained by our algorithm using the Jensen-Shannon divergence (with a window size 7 × 7 × 5). The t-test results are provided by the fMRI data center [30].

Note that, in the proposed method, we compute the statistics over a window in the fMRI images. If we observe a difference in the distribution of the gray values in a window over successive time frames in fMRI images as measured by the JS divergence, we consider that there is certain activity in that window location (center of the window). Therefore, due to the effect of blocking, certain activities are detected outside the brain region (liquor and the CSF around the brain). This can be eliminated by restricting the activity to be detected within the brain region. The brain region can be obtained by segmenting the brain images.

Moreover, the proposed method has one inherent drawback. The method is not based on the generalized linear model (GLIM) and accumulates the statistical differences over successive time frames. Therefore, if the fMRI images are not properly registered and there exist statistical differences (over successive time frames) due to various reasons such as patient motion then certain false active regions may be detected. We did not address this issue in this article.
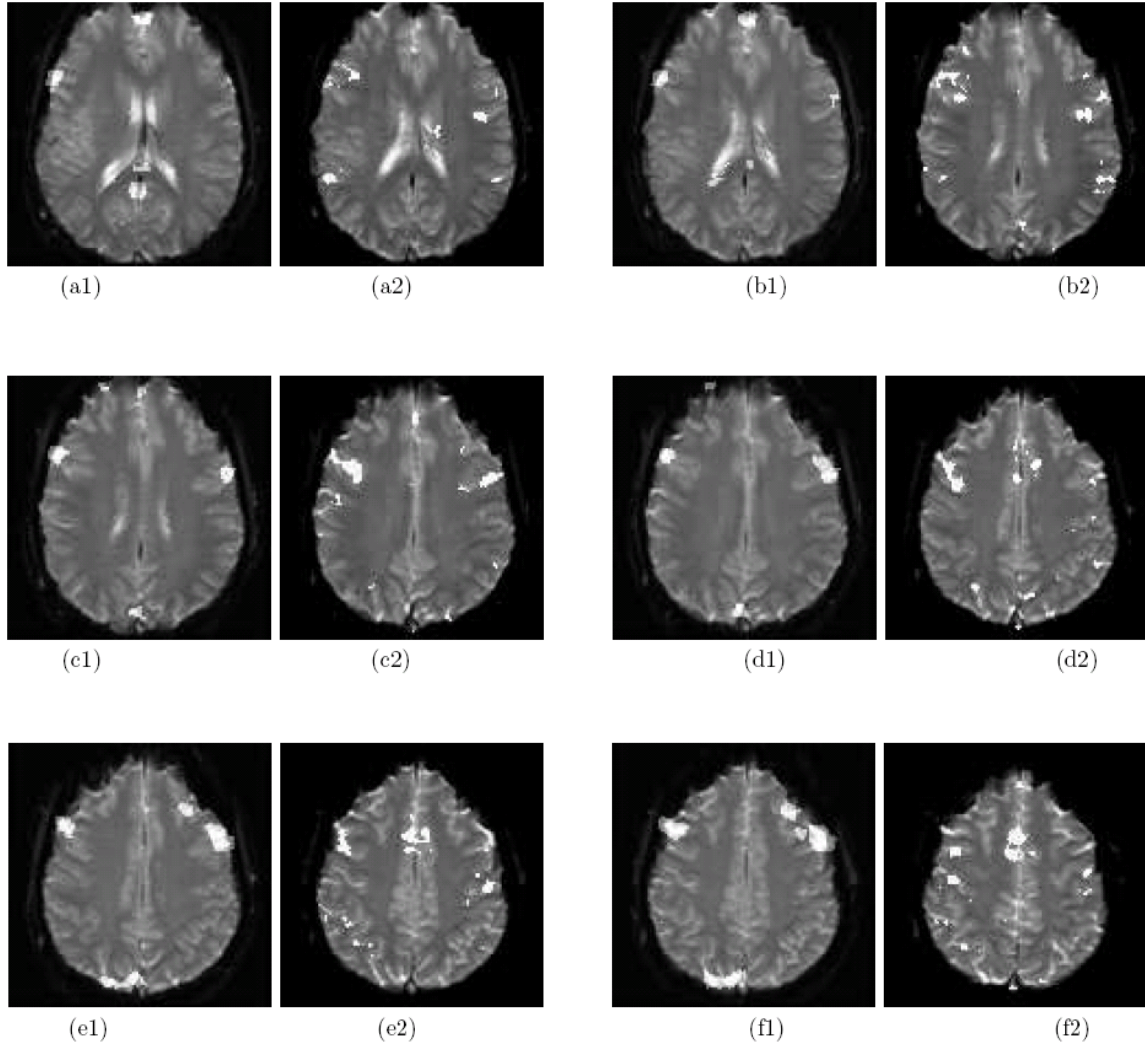
**FIGURE 4**: The regions of activity detected in the fMRI images captured when subject performs object naming task [31]. The left panel of each pair ((a1), (b1), (c1), (d1), (e1), and (f1)) shows the regions of activity detected by our algorithm, and the right panel ((a2), (b2), (c2), (d2), (e2), and (f2)) shows that by t-test.

## 4. CONCLUSIONS

We presented a statistical technique based on Jensen-Shannon divergence for detecting the regions of activity in fMRI images. We exploited the metric property of the square root of Jensen-Shannon divergence to accumulate the variations between successive time frames of fMRI images. Use of Jensen-Shannon divergence makes our algorithm independent of the assumption of any statistical distribution. Jensen-Shannon divergence has been used in the context of image segmentation [26] before, but the use of the same in spatio-temporal data analysis has not been explored, and fMRI is one such example. In the proposed method, we consider a window around each voxel in a M x N x L (say) image and compute statistics over T such time frames. Since the computation of $\sqrt{JS}$ metric is linear in time with the number of pixels (considering a fixed number of bins in the histogram) in a window, the overall computation requires $O(MNLTw^2h)$ time where the size of the window is w×w×h. A smarter computation can be performed by considering a shifting window. In that case, we require $O(w^2)$ time for computation in each window instead of $O(w^2h)$ time. The overall computation, in that case, will take $O(MNLTw^2)$ time. As mentioned

before, we do not address the issues of false activity detection due to improper registration in this article. This can be pursued as one of the future work. The output of our algorithm can possibly further be improved by processing the regions of activity with some other techniques such as clustering [12].

## 5.  REFERENCES

[1] B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI," Magn. Reson. Med., vol. 34, pp. 537–541, 1995.

[2] W. Richter, P. M. Andersen, A. P. Georgopoulos, and S. G. Kim, "Sequential activity in human motor areas during a delayed cued finger movement task studied by time-resolved fMRI," Neuro Report, vol. 24, pp. 1–15, 1997.

[3] J. B. Brewer, J. E. Desmond, G. H. Glover, and J. D. E. Gabrieli, "Making memories: Brain activity predicts how well visual experience will be remembered," Science, vol. 281, pp. 1185–1187, 1998.

[4] A. D. Wagner, D. L. Schacter, M. Rotte, W. Koutstaal, A. Marial, A. M. Dale, B. R. Rosen, and R. L. Bucker, "Building memories: Remembering and forgetting of verbal experiences and predicted by brain activity," Science, vol. 281, pp. 1188–1190, 1998.

[5] S. Y. Bookheimer, M. H. Strojwas, M. S. Cohen, A. M. Saunders, M. A. Pericak-Vance, J. C. Mazziotta, and G. W. Small, "Patterns of brain activation in people at risk for alzheimer's disease," New England Journal of Medicine, vol. 343, pp. 450–456, 2000.

[6] S. Gold, B. Christian, S. Arndt, G. Zeien, T. Cizadlo, D. L. Johnson, M. Flaum, and N. C. Andreasen, "Functional MRI statistical software packages : A comparative analysis," Human Brain Mapping, vol. 6, pp. 73–84, 1998.

[7] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," Proc. National Academy of Science, USA, vol. 87, pp. 9868–9872, 1990.

[8] K. K. Kwong, "Functional resonance imaging with echoplanar imaging," Magn. Reson. Q., vol. 11, pp. 1–20, 1995.

[9] E. Bullmore, S. C. Brammer, M. Williams, S. Rabe-Hesketh, N. Janot, A. David, J. Mellers, R. Howard, and P. Sham, "Statistical methods of estimation and inference for functional MR image analysis," Magn. Resonance Med., vol. 35, pp. 261–277, 1996.

[10] S. Clare, Functional Magnetic Resonance Imaging: Methods and Applications. PhD thesis, University of Nottingham, 1997.

[11] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," Journal of the Royal Statistical Society, Series B, vol. 57, pp. 289–300, 1995.

[12] E. Salli, H. H. Aronen, S. Savolainen, A. Korvenoja, and A. Visa, "Contextual clustering for analysis of functional fMRI data," IEEE Transactions on Medical Imaging, vol. 20, pp. 403–414, 2001.

[13] P. A. Bandettini, A. Jesmanowicz, E. C. Wong, and J. S. Hyde, "Processing strategies for time-course data sets in functional MRI of the human brain," Magn. Reson. Med., vol. 30, pp. 161–173, 1993.

[14] K. Kuppusamy, W. Lin, and E. M. Haacke, "Statistical assesment of crosscorrelation and variance methods and the importance of electrocardiogram gating in functional magnetic resonance imaging," Magn. Resonance Imaging, vol. 15, pp. 169–181, 1997.

[15] U. E. Ruttimann, M. Unser, R. R. Rawlings, D. Rio, N. F. Ramsey, V. S. Mattay, D. W. Hommer, J. A. Frank, and D. R. Weinberger, "Statistical analysis of functional MRI data in the wavelet domain," IEEE Transactions on Medical Imaging, vol. 17, pp. 142–154, 1998.

[16] M. J. Brammer, "Multidimensional wavelet analysis of functional magnetic resonance images," Human Brain Mapping, vol. 6, pp. 378–382, 1998.

[17] W. Backfrieder, R. Baumgartner, M. Smal, E. Moser, and H. Bergmann, "Quantification of intensity variations in functional MR images using rotated principal components," Phys. Med. Biol., vol. 41, pp. 1425–1438, 1996.

[18] M. M. J., S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Analysis of fMRI data by blind separation into independent spatial components," Human Brain Mapping, vol. 6, pp. 160–188, 1998.

[19] J. V. Stone, J. Porrill, C. Buchel, and K. Friston, "Spatial, temporal, and spatiotemporal independent component analysis of fMRI data," in 18th Leeds Statistical Research Workshop on Spatio-temporal modeling and its applications, University of Leeds, 1999.

[20] B. A. Ardekani, J. Kershaw, K. Kashikura, and I. Kanno, "Activation detection in functional MRI using subspacemodeling and maximum likelihood estimation," IEEE Trans. Medical Imaging, vol. 18, pp. 101–114, 1996.

[21] C. Goutte, P. Toft, E. Rostrup, F. A. Nielsen, and L. K. Hansen, "On clustering fMRI time series," NeuroImage, vol. 9, pp. 298–310, 1999.

[22] K. J. Friston, K. J. Worsley, R. S. J. Frackowiak, J. C. Mazziotta, and A. C. Evans, "Assessing the significance of focal activations using their spatial extent," Human Brain Mapping, vol. 1, pp. 214–220, 1994.

[23] C. R. Rao, "Diversity : Its measurement, decomposition, appointment and analysis," Sankhya: The Indian Journal of Statistics, vol. 11(A), pp. 1–22, 1982.

[24] A. K. C. Wong and M. You, "Entropy and distance of random graphs with application to structural pattern recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 7, pp. 599–609, 1985.

[25] J. Lin, "Divergence measures based on the Shannon entropy," IEEE Trans. Information Theory, vol. 37, pp. 145–151, 1991.

[26] C. Atae-Allah, J. F. G´omez-Lopera, J. Mart´ınez-Aroza, Rom´an-Rold´an, and P. Luque-Escamilla, "Image segmentation by Jensen-Shannon divergence : Application to measurement of interfacial tension," in Proc. Int. Conference on Pattern Recognition (ICPR00), Barcelona, Spain, vol. 3, 2000.

[27] D. M. Endres and J. E. Schindelin, "A new metric for probability distributions," IEEE Trans.Information Theory, vol. 49, pp. 1858–60, 2003.

[28] F. ¨Osterreicher and I. Vajda, "A new class of metric divergences on probability spaces and its statistical applications," Ann. Inst. Statist. Math., vol. 55, pp. 639–653, 2003.

[29] G. H. Glover, "Deconvolution of impulse response in event-related bold fmri," NeuroImage, vol. 9, pp. 416–429, 1999.

[30] fMRIDC, "fmri data center," http://www.fmridc.org/f/fmridc.

[31] J. Hirsch, D. Rodriguez Moreno, and K. H. S. Kim, "Interconnected large-scale systems for three fundamental cognitive tasks revealed by fMRI," Journal of Cognitive Neuroscience, vol. 13, pp. 389–405, 2001.