Nonparametric Mean Shift Functional Detection in the Functional Space for Task and Resting-state fMRI

Jian Cheng[12], Feng Shi[3], Kun Wang[3], Ming Song, Jiefeng Jiang, Lijuan Xu, Tianzi Jiang[1]

[1]LIAMA, NLPR, Institute of Automation, Chinese Academy of Sciences, China

[2]Odyssée, INRIA Sophia Antipolis, France

[3]Department of Radiology and BRIC, University of North Carolina at Chapel Hill

LIAMA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE — INRIA — centre de recherche SOPHIA ANTIPOLIS - MÉDITERRANÉE

Abstract

In functional Magnetic Resonance Imaging (fMRI) data analysis, normalization of time series is an important and sometimes necessary preprocessing step in many widely used methods. The space of normalized time series with n time points is the unit sphere S^{n-2}, named the functional space. Riemannian framework on the sphere, including the geodesic, the exponential map, and the logarithmic map, has been well studied in Riemannian geometry. In this paper, by introducing the Riemannian framework in the functional space, we propose a novel nonparametric robust method, namely Mean Shift Functional Detection (MSFD), to explore the functional space. The first merit of the MSFD is that it does not need many assumptions on data which are assumed in many existing method, e.g. linear addition (GLM, PCA, ICA), uncorrelation (PCA), independence (ICA), the number and the shape of clusters (FCM). Second, MSFD takes into account the spatial information and can be seen as a multivariate extension of the functional connectivity analysis method. It is robust and works well for activation detection in task study even with a biased activation reference. It is also able to find the functional networks in resting-state study without a user-selected “seed” region. Third, it can enhance the boundary between different functional networks. Experiments were conducted on synthetic and real data to compare the performance of the proposed method with GLM and ICA. The experimental results validated the accuracy and robustness of MSFD, not only for activation detection in task study but also for functional network exploration in resting-state study.

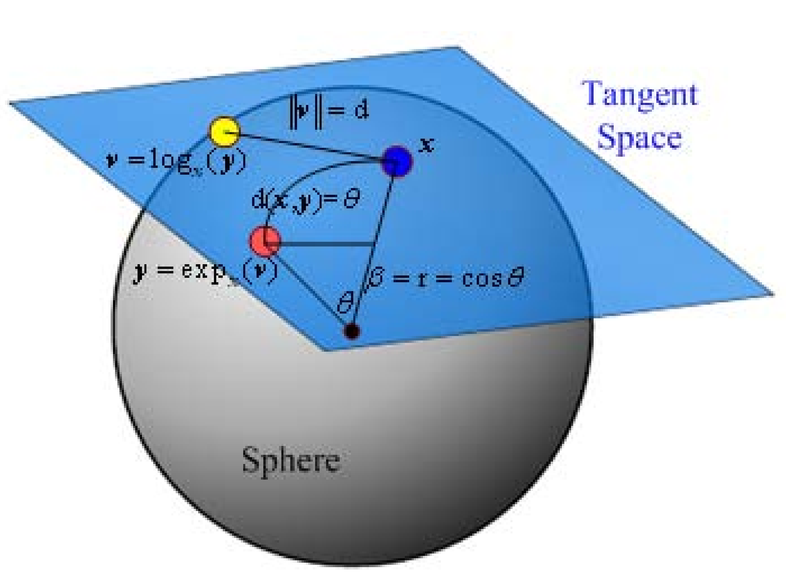
Riemannian Framework in the Functional Space

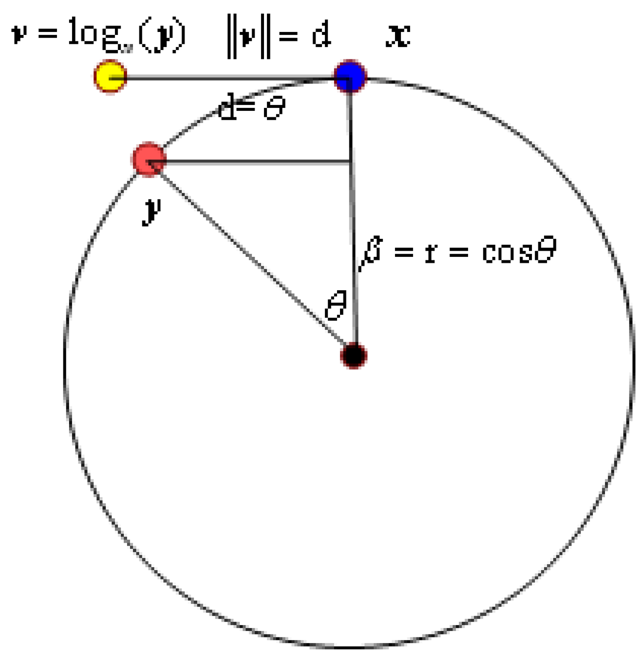
Time course normalization: subtracted by its mean (centering) and then divided by its L2 norm (scaling)

- It is an important pre-processing step in many algorithms.
- Many algorithms are actually “normalization invariant”.
- Functional Space: the space of the normalized time courses is S^{n-2}.
- Riemannian framework could provide mathematical tools for calculations on the functional space S^{n-2}.

The intrinsic Riemannian framework on the Functional Space

- Geodesic: $d(\mathbf{x},\mathbf{y})=\theta=\arccos(r),\quad r=\cos(\theta)=\sum_i x_i y_i$
- Exponential map: $\mathbf{y}=Exp_{\mathbf{x}}(\mathbf{v})=\mathbf{x}\cos\theta+\frac{\mathbf{v}}{\|\mathbf{v}\|}\sin\theta$, where $\theta=\|\mathbf{v}\|$
- Logarithmic map: $\mathbf{v}=Log_{\mathbf{x}}(\mathbf{y})=\frac{\mathbf{y}-\mathbf{x}\cos\theta}{\|\mathbf{y}-\mathbf{x}\cos\theta\|}\theta$, where $\theta=\arccos(\mathbf{x}^T\mathbf{y})$



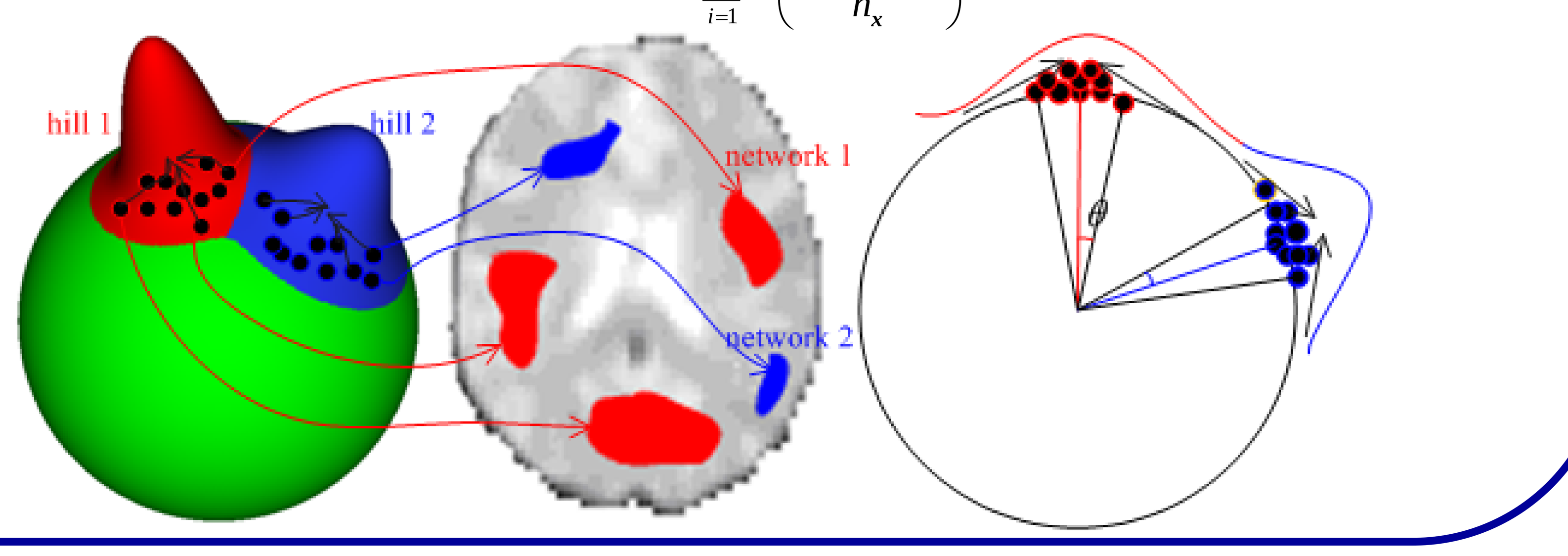


Mean Shift Functional Detection (MSFD)

- Consider the normalized time courses as the observations from the probability distribution function of the time course on S^{n-2}.
- Hills appear in the area with many points, and valleys appear in the area with few points in the distribution.
- The spatial location distribution of the time courses within the same hill could be considered as a functional network.
- The idea of MSFD is to find the optimal representative time courses (peaks of hills) and partition these points based on these peaks of different hills

Adaptive kernel density estimation: $\hat{f}_k(x)=\frac{c_k}{n}\sum_{i=1}^{n}k\left(\frac{d^2(\mathbf{x},\mathbf{x}_i)}{h_x^2}\right)$

Mean shift vector: $m(\mathbf{x})=C\frac{\nabla\hat{f}_k(\mathbf{x})}{\hat{f}_g(\mathbf{x})}=\frac{\sum_{i=1}^{n}\log_{\mathbf{x}}(\mathbf{x}_i)g\left(\frac{d^2(\mathbf{x},\mathbf{x}_i)}{h_x^2}\right)}{\sum_{i=1}^{n}g\left(\frac{d^2(\mathbf{x},\mathbf{x}_i)}{h_x^2}\right)}$ $\qquad g(\cdot)=-k'(\cdot)$



MSFD for Task and Resting fMRI Study

- Mean Shift Iteration (MSI)

```
Algorithm: Mean Shift iteration (MSI)
Input:  time courses, x_i, i = 1,...,N
        normalize x_i, i = 1,...,N
        for i <- 1,...,N
            x <- x_i  dist(i) = 0
            Repeat :
                Determine h_x (KNN)
                m(x) = [Σ_{i=1}^n log_x(x_i) g(d^2(x,x_i)/h_x)] / [Σ_{i=1}^n g(d^2(x,x_i)/h_x)],
                x <- exp_x(m(x)),   dist(i) <- dist(i) + ||m(x)||
            Until ||m(x)|| < ε
Output:  save every new x_i and dist(i), i = 1,...,N
```

- Post-processing:
  - Over-segmentation using single-link clustering with very small threshold
  - Find the most representative time courses $\{\mathbf{r}_k\}$.
  - For each representative time course $\mathbf{r}_k$, we can get a T map $\{T_{ki}\}$.

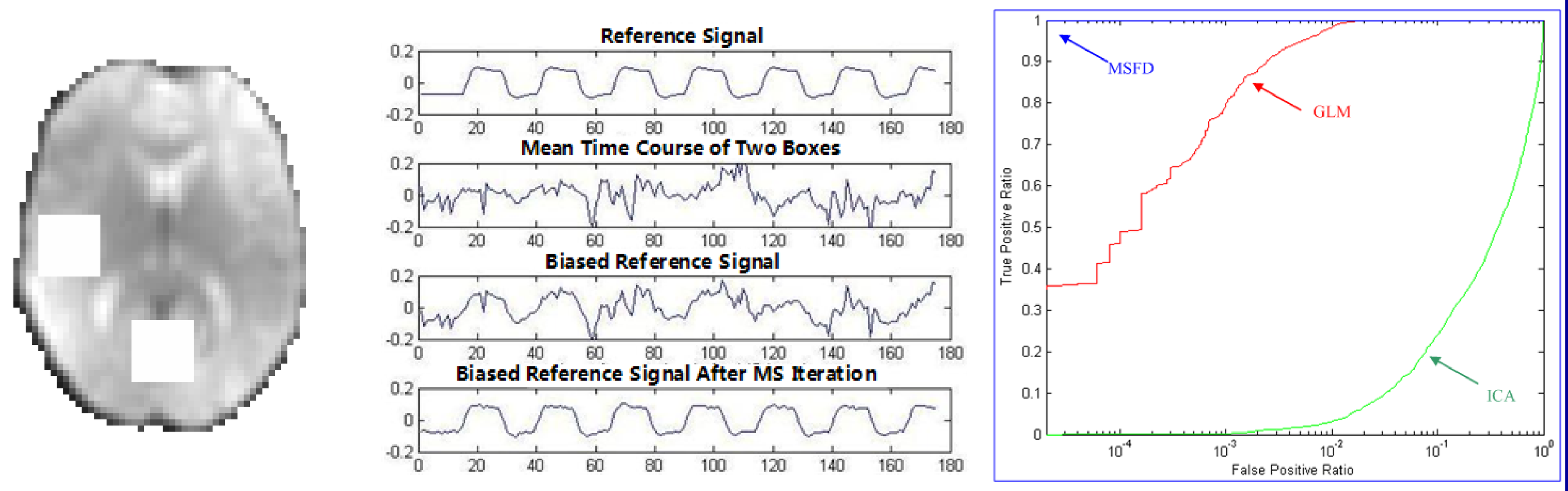
$$d_{ki}=d(x_i,r_k)+dist \qquad T_{ki}=\frac{\sqrt{n-2}\cos(d_{ki})}{\sqrt{1-\cos^2 d_{ki}}}$$

- For task study, we just need to find the hill related with the activated signal.
- the given hemodynamic response (HR) is actually biased.
- We consider the given biased reference as a real time course in brain, and update it through MSI. The corrected expected HR is the real activation reference for this subject.

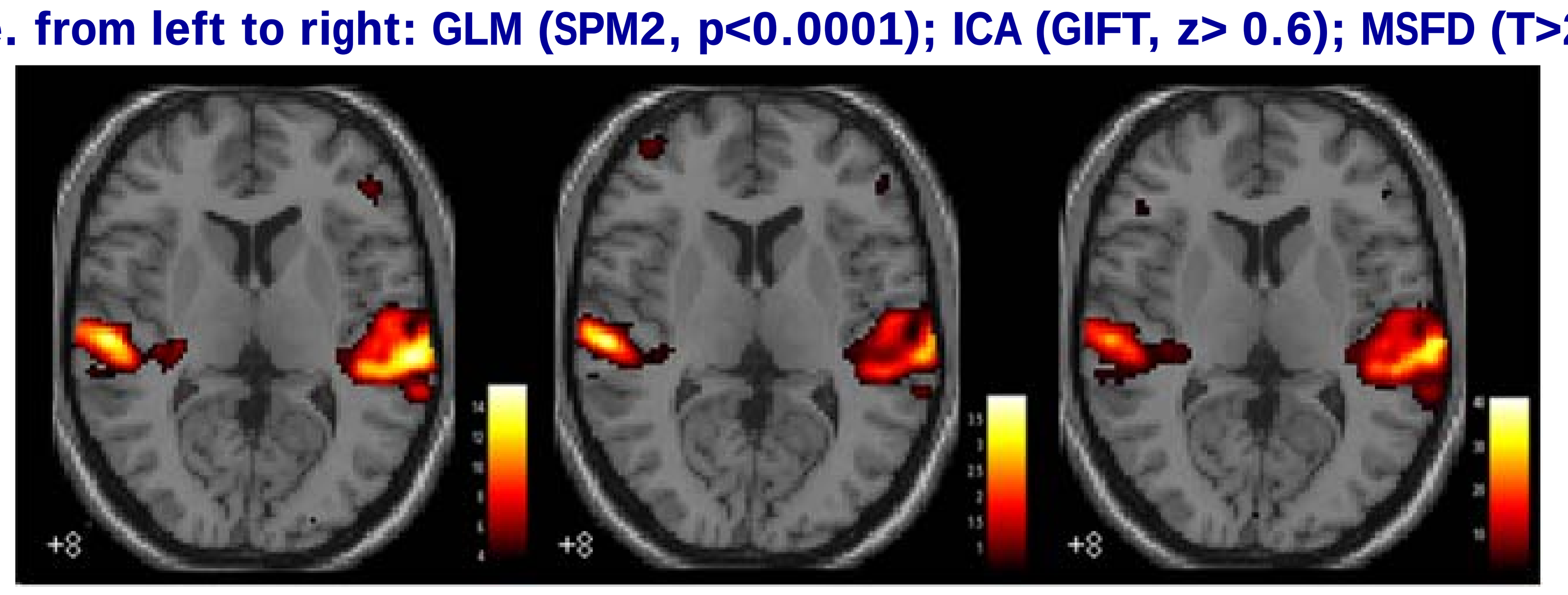
Experiments (compare MSFD, GLM, ICA)

I. Synthetic data: replace the resting signal in the boxes with the nosy reference with CNR=0.4. With the accurate estimation of reference, MSFD, GLM and ICA all achieve good performance (perfect classification).

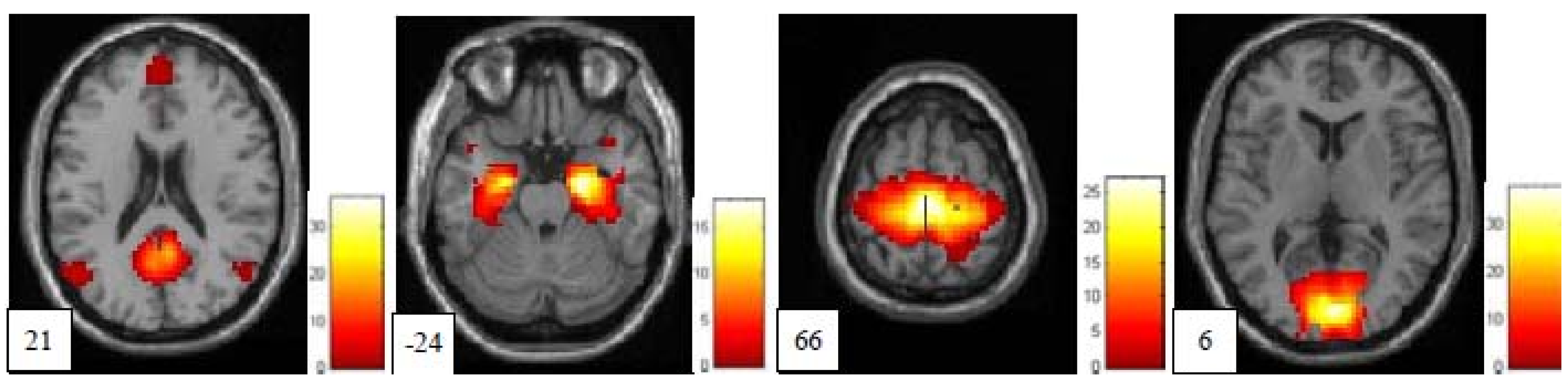
II. Synthetic data: With the biased estimation of reference, MSFD is robust and gets the best result. The ROC curves are shown.



III. Real data (task study): the auditory bi-syllabic data from the SPM public dataset (http://www.fil.ion.ucl.ac.uk/spm/data/auditory.html) . Results of one slice. from left to right: GLM (SPM2, p<0.0001); ICA (GIFT, z> 0.6); MSFD (T>2).



III. Real data (resting study): the null data in the synthetic data. We show 4 T maps for the default-mode network, hippocampus, sensorimotor and visual area (from left to right).



Contact: jiancheng@nlpr.ia.ac.cn, Jian.Cheng@sophia.inria.fr