# Ranking and Averaging Independent Component Analysis by Reproducibility (RAICAR)

**Z. Yang[1,2], S. LaConte[1], X. Weng[2], and X. Hu[1]**

[1]Department of Biomedical Engineering, GA Tech/Emory Univ., Atlanta, Georgia, United States, [2]Institute of Psychology, Chinese Academy of Sciences, Beijing, China, People's Republic of

## Introduction

Independent component analysis (ICA) is a data-driven approach that is widely used in functional magnetic resonance imaging (fMRI). The widely used gradient-based ICA algorithm has three limitations: (1) inability to determine the number of components; (2) inability to provide an order of the components and (3) instability due to the dependence on the initial value. Here we introduce a method, ranking and averaging ICA by reproducibility (RAICAR), to alleviate these limitations. RAICAR makes use of the stochastic nature of the gradient-based ICA algorithm to provide a reproducibility ranking of ICA components, estimate the number of important components, and generate more stable result.

## Method

*Ranking ICA*

The FastICA [1] is applied on fMRI data for K times (K realization), each starts at different initial values and C target components (C equals to the dimension of the data, i.e. the number of time points in spatial ICA), resulting in K•C component maps. The K•C component maps are then correlated with each other to form a cross-realization correlation matrix (CRCM), which contains K by K blocks, each comprising of C by C correlation coefficients. Within the CRCM, an iterative alignment procedure is applied using the following steps. (1) Find for the global maximum. Denote its location as element (m,n) in block (a,b). (2) Search for the block-wise maxima in the $m^{th}$ row in blocks (a, $i$), $i$ = 1, 2…, K, and the $n^{th}$ column in blocks ($i$, b), $i$ = 1, 2..., K. (3) The maxima in each pair of blocks, say (m, p) in block (a, $i$) and (n, q) in block ($i$, b), should usually have the same location, i.e., p = q. In this case, the component corresponding to p or q is selected for the alignment. If p ≠ q, the one corresponding to the larger cross-realization correlation is selected for the alignment. By going through all values of $i$, ($i$ = 1, 2..., K), K realization-components are aligned together to represent one final component. This procedure is repeated until C components are aligned. From the histogram of the CRCM elements, a valley can be found and used as a threshold in calculating the reproducibility indices of the aligned components. For each aligned component, all pair-wise correlation coefficients of its realizations are thresholded and summed to generate the reproducibility index. The aligned components are then ranked according to the reproducibility, and their component maps and corresponding mixing time course are obtained by averaging the representing components. The number of important components is estimated according to the drop-off in the reproducibility rank. Although the drop-off is generally sharp enough to allow users inspect "by eye", here we employ (K-1)K/4 as an empirical cut-off for the sake of consistency.

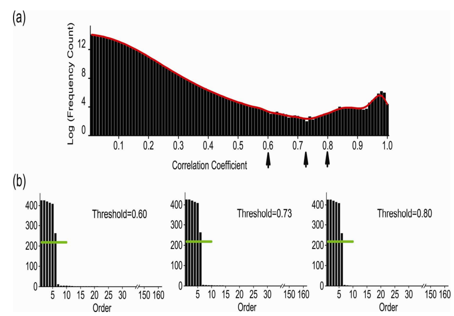


**Fig 1. Reproducibility ranking of simulation**

*Simulation and Experimental Data*

A simulated data consisting of six equal-area spatial sources, one global varying background and white noise was generated. The CNR (CNR ≡ $\Delta s/\sigma_{noise}$) of the sources was controlled by adjusting the variance of their mixing time courses, leading to a range of 0.92~3.87, which is in the range of fMRI data [2]. Besides testing RAICAR in simulation, two other common used methods, MELODIC [3] and ICASSO [4], are also compared by varying the dimension of the simulated data. Two experimental fMRI datasets were used to evaluate RAICAR. One used a delayed motor task [5], the other a constant force grip task [6]. To test the RAICAR's sensitivity to the threshold used, different thresholds were used. In addition, the stability of RAICA was tested by repeating RAICAR several times.

## Results and Discussion

Figure 1(a) shows the histogram of the CRCM for the simulated data. In the reproducibility rank shown in Fig 1(b), it is easy to see that RAICAR correctly estimated the number of sources, and it is not sensitive to the selection of the threshold. Figure 2 shows the ROC analysis of the individual ICA and RAICAR. The curves are generated using the source with the lowest SNR. The black curves provide evidence that individual ICA results are not stable. The narrow pink region shows the spread of ten repetitions of RAICAR, indicating that RAICAR result is stable and more accurate than most of the individual ICAs. In the delayed motor data, the estimated number of components and the rank of the task related components do not change significantly using different thresholds or across different repetitions. Figure 3 shows several interesting components and their positions in the reproducibility rank, extracted from the constant force grip task. The task related components tend to be ranked at the top. The reproducibility rank reveals the "strength" of each component, thus it provides a new index to compare different activations. Figure 4 shows the comparison of three methods in terms of estimating the number of components of simulated data. RAICAR is stable and accurate in different situations, while MELODIC and ICASSO estimations depend on the number of time points in the data (data dimension).

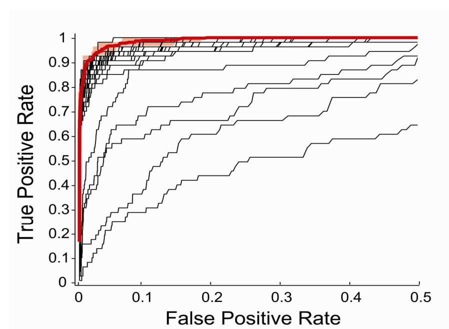


**Fig 2. Detectability of RAICAR and individual ICA**

## Conclusions

In this work, we present a new method, RAICAR, to rank the ICA components by reproducibility and estimate the number of important components. RAICAR improves the decomposition and interpretation of fMRI data with ICA. As shown in both simulation and experimental datasets, it is not sensitive to the choice of the parameters.

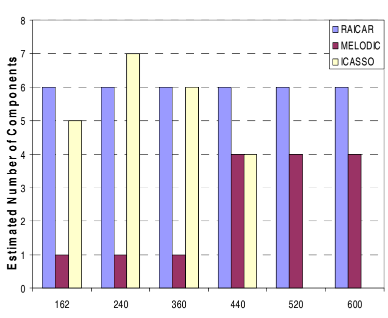


**Fig4. Comparison of estimations of the number of components across RAICAR, MELODIC and ICASSO**

**Acknowledgement**

This work was supported in part by the National Institutes of Health (RO1EB002009) (USA), the National Science Foundation of China (30425008, 30670674) and the Chinese Ministry of Science and Technology (2003CB515400).

**Reference**

**1.** Hyvarinen A., 1999. IEEE Trans Neural Networks 10(3):626-634. **2.** Huettel SA et al., 2004. Functional magnetic resonance imaging. **3.** Beckmann CF et al., 2004. IEEE Trans Med Imaging 23(2):137-52. **4.** Himberg J et al., 2004. Neuroimage 22(3):1214-22. **5.** Catalan MJ et al., 1998. Brain 121(2):253-264. **6.** Peltier SJ et al., 2005. Brain Res 1057(1-2):10-6.

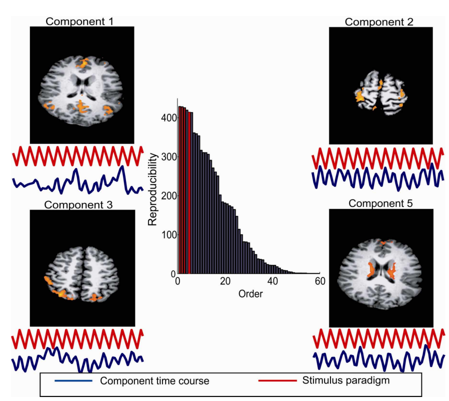


**Fig 3. Components in the constant force grip task, extracted by RAICAR**