A METHOD FOR GROUP DIFFERENCE ENHANCEMENT BY CONSTRAINING MIXING COEFFICIENTS OF ICA FRAMEWORK

J. Sui[1], and V. Calhoun[1,2]

[1]The Mind Research Network, Albuquerque, New Mexico, United States, [2]Department of ECE, University of New Mexico, Albuquerque, New Mexico, United States

Introduction

Independent component analysis has been successfully utilized to analyze brain imaging data and to extract features which may significantly differentiate healthy controls and schizophrenia patients. However, when working with real fMRI data which is noisy, Infomax, a popular ICA algorithm, does not always produce the optimum results. For example, the components showing significant group difference are identified by a 2 sample *t*-test on the mixing coefficients to test whether the mean is different between healthy controls and patients. We found that the extracted independent component (IC) exhibiting the largest group difference does not always have the smallest *p* value for the 2 sample *t*-test, and its *J* divergence (a measure of the difference between two distributions) between the back-reconstructed sources for each group is not always the largest one among all components, which may lead to confusion when choosing the IC that can best differentiate the groups.

Method

We modified a well established ICA approach, Infomax [1], by adding an adaptive constraint to the mixing coefficients, and call our approach constrained coefficients ICA (CCICA). Prior group member information in terms of *T* statistic from a 2 sample *t*-test is introduced into the normal entropy cost function *H* to emphasize the components that may distinguish the two groups. The expanded cost function *C* is defined as $C = Max\{H + \lambda \cdot \sum T_i^2\}$. When maximizing *C*, the constrained components are flexible and the constraint strength λ is adaptively adjusted to ensure the maximum independence achieved.

Results

CCICA is then applied to hybrid fMRI data in form of joint ICA framework [2]. The hybrid data is created as figure 1(a) shown, the group-varying known sources are added to features of Auditory oddball and Sternberg task with shared mixing coefficients (the average of patients are set to be higher than that of the controls). The features are contrast images obtained by processing the original fMRI data with SPM5 and are stacked side by side to obtain the observed data (subjects by voxels). Since the superimposed sources have a known pattern, it is straightforward to extract them from the observed data so as to evaluate the performance of CCICA and Infomax.

For both ICA algorithms, we use 8 components and 5 runs to separate the observed data. Since the superimposed sources between groups have different distributions, its corresponding component should have the smallest p value and the largest J divergence among all the extracted ICs. Figure 1(b) illustrates the first two extracted ICs (sorted by the p values of two sample *t*-test from low to high) that showed the largest group difference. Obviously component 5, ranked first in CCICA, and component 8, ranked second in Infomax, is our desired IC with known sources showing at the right positions of both task features in Figure 1(a). The correlation of their loading parameters versus ground truth is plotted in Figure 1 (c) (patients are coded in yellow squares, controls in blue circles). As expected, the controls showed a lower mean in mixing coefficients than the patients did, which can also explain why the shape of the activations is almost identical to the sources added to the patients. Further, the desired IC shows the highest correlation among all components, confirming both algorithms select out the coupled source into a separate component successfully. Note that CCICA outperformed Infomax both by increasing the significance ($p < 1.2e-4$ vs. $p < 0.0042$) and also the coefficient accuracy (0.8225 vs. 0.8049). For the desired IC, we also calculated its J divergence of the back-reconstructed source distributions. The results and their sort order among all extracted ICs are listed in table 1. In contrast with Infomax, CCICA showed consistent results by identifying the IC that can best differentiate patients from controls on both criterions.

We can also examine the efficacy of CCICA more directly by applying it to real fMRI data (SPM contrast images from a patient and controls group). The sort orders of the desired component on above measures are listed in table 1. CCICA consistently resulted in the desired IC ranked first for both criterions, and hence identified the desired component differentiating controls and patients most significantly on both the mixing coefficients and the back-reconstructed source distributions.

Discussion/Conclusion

The presented results support the claim that, by incorporating prior statistical group information in our CCICA algorithm, we can improve the capability for identifying the IC showing the largest group difference related to mixing coefficients. Because ICA assumes a linear relationship between the observed data and sources, consequently, the accuracy of estimated sources for each group is also enhanced so as to better reflect the group difference. Standard Infomax provides less consistent results, which may lead to confusion when choosing the desired IC. Therefore CCICA is a promising method for identification of ICs that can best distinguish patients and controls, and further, it may also be more sensitive for group classification and for identifying features that may serve as potential brain imaging biomarkers of disease.

References

[1]M.J. McKeown, et al. HBM 6(3) p.160 (1998).
[2]V.D.Calhoun, et al. HBM 27(7) p.203 (2005).

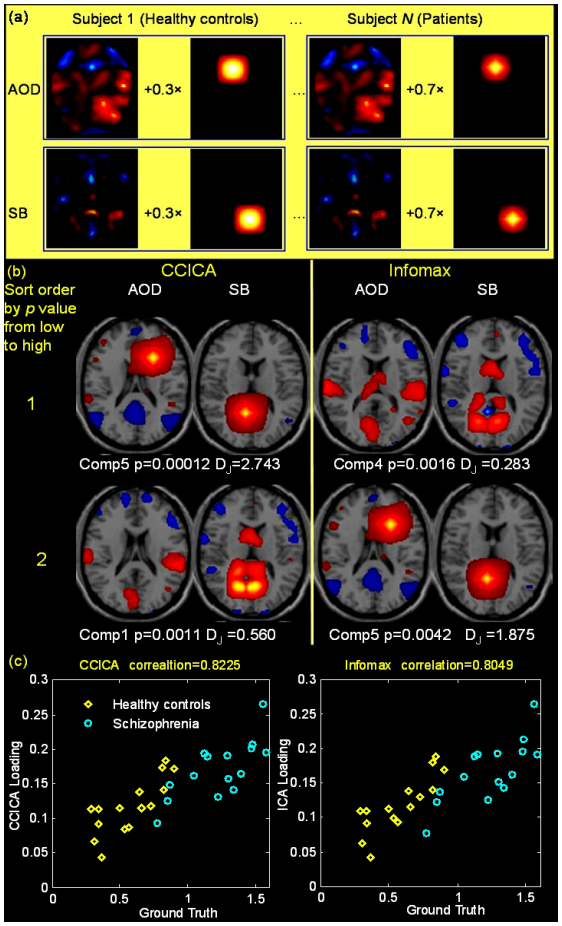

Figure 1 (a) generation of hybrid fMRI data;(b) the first two extracted ICs showing the largest group difference sorted by *p* values of two sample *t*-test; (c) correlation of the desired IC's loading parameters versus ground truth.

Table 1. Sort order on two criterions of the desired IC that can best differentiate the groups

The Desired Independent Component		*p* of 2t-test (sort order)	J divergence (sort order)
Hybrid fMRI data	CCICA	0.00012 (1)	2.743 (1)
	Infomax	0.0042 (2)	1.875 (2)
Real fMRI data	CCICA	0.00076 (1)	0.9962 (1)
	Infomax	0.00332 (1)	0.1437 (8)