

MULTIMODAL SUBSPACE INDEPENDENT VECTOR ANALYSIS BETTER CAPTURES HIDDEN RELATIONSHIPS IN MULTIMODAL NEUROIMAGING DATA

Xinhui Li^{*†} Tulay Adali[‡] Rogers F. Silva^{*} Vince D. Calhoun^{*†}

^{*} Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State University, Georgia Institute of Technology, and Emory University, Atlanta, GA, USA

[†] Department of Electrical and Computer Engineering,
Georgia Institute of Technology, Atlanta, GA, USA

[‡] Department of Computer Science and Electrical Engineering,
University of Maryland Baltimore County, Baltimore, MD, USA

ABSTRACT

We introduce *multimodal subspace independent vector analysis (MSIVA)*, a methodology to capture both joint and unique latent sources across data modalities by defining shared and modality-specific subspaces. We compared MSIVA to a unimodal analysis (UA) baseline and tested both methods with four distinct subspace structures on synthetic and multimodal neuroimaging datasets. We demonstrated that both approaches can identify and distinguish the correct subspace structures from incorrect ones on synthetic datasets. We then showed that MSIVA can better capture the subspace structures across two neuroimaging modalities. Results from subsequent per-subspace canonical correlation analysis and brain-phenotype modeling showed that the sources from the optimal subspace structure are significantly associated with phenotype measures including age and sex.

Index Terms— multimodal fusion, MSIVA, MISA, IVA

1. INTRODUCTION

Multiple neuroimaging techniques such as magnetic resonance imaging (MRI) have been developed to understand the structural and functional relationships of the brain. However, each neuroimaging modality has its own strengths and weaknesses, and only captures certain aspects of the brain. For example, structural MRI (sMRI) can reveal high-resolution anatomical structure of the brain but cannot capture temporal dynamics, while functional MRI (fMRI) can measure blood-oxygenation-level-dependent (BOLD) signals across time at the cost of lower spatial resolution. To jointly analyze multiple data modalities and capture multifaceted information of the brain, a multidataset independent subspace analysis (MISA) [1] framework has been developed encompassing many blind source separation methods such as independent

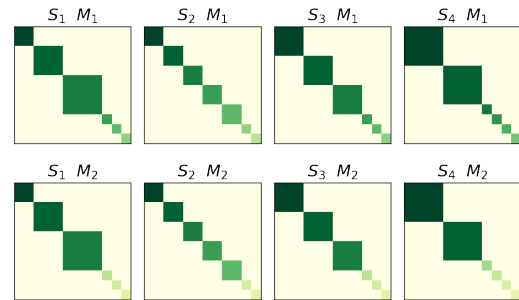


Fig. 1. Proposed four subspace structures (S) in two modalities (M). Same-color blocks are linked across modalities. Block size reflects number of linked sources within modality.

component analysis (ICA) [2], independent vector analysis (IVA) [3, 4] and independent subspace analysis (ISA) [5]. Recently, a multimodal IVA (MMIVA) fusion method has been proposed using MISA to identify linked biomarkers in multimodal neuroimaging datasets, revealing joint biomarkers of age and sex in two large studies [6]. MMIVA assumes that sources are independent within each modality, but there may exist linkage among sources in neuroimaging data, potentially grouped by their anatomical or functional properties.

Aiming to detect the linkage across sources and modalities, we propose a novel methodology, *multimodal subspace independent vector analysis (MSIVA)*, to identify cross-modal linkage of source *groups* by defining joint and unique subspaces. MSIVA is built on MMIVA by defining a block diagonal matrix as the subspace instead of the identity matrix used in MMIVA. MSIVA is designed to simultaneously capture two types of latent sources, one is shared across all modalities and the other is unique to a specific modality.

We first demonstrate that both MSIVA and a unimodal analysis (UA) baseline can successfully reveal the correct

Corresponding author: Xinhui Li, xinhui.li@gatech.edu.
RFS and VDC contributed equally.

subspace structures in multiple synthetic datasets. We then run both UA and MSIVA methods on a large multimodal neuroimaging dataset [7]. Our results indicate that MSIVA can capture shared and modality-specific sources in the neuroimaging data. Using canonical correlation analysis (CCA) [8], we conduct a follow-up assessment of each identified subspace separately and find projections within the optimal subspace structure yielding shared sources that are significantly associated with age and sex. We finally performed prediction tasks to validate the association between phenotype measures and shared sources.

2. METHODS

2.1. Subspace Structure

Subspaces with two to four dimensions are commonly used to cluster functional networks [9, 10]. We propose four subspace structures (S) in two modalities (M), with 12 sources grouped by different subspace dimensions in each modality (Fig. 1):

- S_1 : One 2-dimensional (2D) shared subspace, one 3D shared subspace, one 4D shared subspace, and three 1D modality-specific subspaces per modality.
- S_2 : Five 2D shared subspaces and two 1D modality-specific subspaces per modality.
- S_3 : Three 3D shared subspaces and three 1D modality-specific subspaces per modality.
- S_4 : Two 4D shared subspaces and four 1D modality-specific subspaces per modality.

For each subspace structure, we generated a synthetic dataset $\mathbf{X}^{[m]} \in \mathbb{R}^{V \times N}$, where m is the modality index ($m \in \{1, 2\}$), V is the feature dimensionality ($V = 20000$) and N is the number of samples ($N = 3000$). $\mathbf{X}^{[m]}$ was a linear mixture of 12 sources spanning the defined subspaces. Each subspace was independently sampled from a multivariate Laplace distribution (the distribution marginals correspond to different sources). Sources in the same subspace, but assigned to different modalities, are dependent (or linked) with a correlation coefficient ranging from 0.65 to 0.85. Sources in the 1D subspaces are independent from all others.

We also utilized a large multimodal neuroimaging dataset from UK Biobank [7] including two image modalities: T1-weighted sMRI (M_1) and fMRI (M_2). 2907 subjects (mean age $\pm$ std: 62.09 ± 7.32 years; 1452 males, 1455 females) were used after excluding subjects missing phenotype measures. After subject selection, we preprocessed these two imaging modalities to obtain the gray matter (GM) and amplitude of low frequency fluctuations (ALFF) feature maps, respectively. Data preprocessing details can be found at [6].

2.2. Multimodal subspace independent vector analysis

The UA approach subsequently applied principal component analysis (PCA) and ICA on each modality separately to obtain 12 sources $\mathbf{S}^{[m]}$ per modality. Greedy combinatorial op-

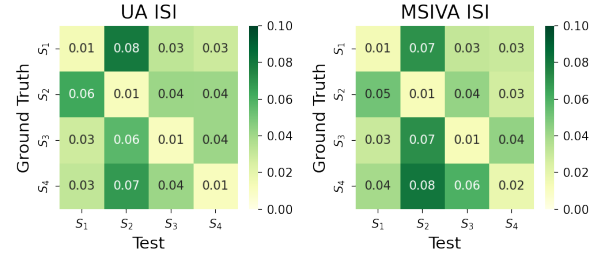


Fig. 2. Synthetic data experiment result: ISI (lower is better). The diagonal has the smallest ISI values in both UA and MSIVA cases, demonstrating that the correct subspace structures yield the best results for both UA and MSIVA.

timization (CO) and MISA were then run sequentially for 10 and 20 iterations on synthetic and neuroimaging data, respectively. The MSIVA approach first used multimodal group principal component analysis (MGPCA) to identify common principal components across the two modalities and then applied ICA on the MGPCA reduced data from each modality separately to get 12 sources $\mathbf{S}^{[m]}$ per modality. Then we ran greedy CO and MISA for 10 and 20 iterations on synthetic and neuroimaging data, respectively. Finally, we identified the optimal subspace structure based on the lowest loss value. The loss function $\mathcal{L}$ [1] is defined as the Kullback-Leibler (KL) divergence between the joint Kotz distribution [11] of all sources $p(\mathbf{S})$ and the product of the joint Kotz distribution of sources at each of K subspaces $q(\mathbf{S}) = \prod_{k=1}^K p(\mathbf{S}_k)$. Subspaces are to be statistically independent and may include multimodal sources. We want to minimize the loss $\mathcal{L}$ by solving the following optimization problem:

$$\begin{aligned}
\min \mathcal{L} &= \min \mathbb{E} \left[\ln \frac{p(\mathbf{S})}{q(\mathbf{S})} \right] \\
&= \min \mathbb{E} [\ln p(\mathbf{S})] - \sum_{k=1}^K \mathbb{E} [\ln p(\mathbf{S}_k)] \\
&= \min_{\mathbf{W}, \mathbf{P}_k, K} \mathbb{E} [\ln p(\mathbf{W}\mathbf{X})] - \sum_{k=1}^K \mathbb{E} [\ln p(\mathbf{P}_k \mathbf{W}\mathbf{X})],
\end{aligned}$$

where $\mathbf{W}$ is the demixing matrix such that $\mathbf{S} = \mathbf{W}\mathbf{X}$ and $\mathbf{P}_k$ is the k -th subspace assignment matrix. Code is available at <https://github.com/trendscenter/MSIVA/releases/tag/v0.0.0>.

2.3. Experiments

For each of four subspace structures, we first generated the synthetic dataset from the ground truth subspace structure. For each dataset, we performed UA and MSIVA experiments with all four subspace structures and measured the normalized multidataset Moreau-Amari intersymbol interference

Table 1. Synthetic data: Final MISA loss values (lower is better). Rows are the ground truth structures used to generate the synthetic data and columns are the test structures.

UA	S_1	S_2	S_3	S_4
S_1	42.69	42.88	42.76	42.99
S_2	42.65	42.30	42.85	42.87
S_3	42.72	42.86	42.64	43.10
S_4	43.09	43.24	43.17	42.98
MSIVA	S_1	S_2	S_3	S_4
S_1	42.68	42.86	42.75	43.04
S_2	42.66	42.23	42.63	42.76
S_3	42.69	42.86	42.62	43.04
S_4	42.69	42.40	41.12	39.94

Table 2. Neuroimaging data: Final MISA loss values.

	S_1	S_2	S_3	S_4
UA	47.74	47.81	47.77	47.78
MSIVA	46.79	46.77	46.80	46.89

(ISI) [1, 12, 13] to evaluate the difference between the recovered sources and the ground truth sources, as well as the loss values. The synthetic data experiments aim to verify whether UA and MSIVA can identify and distinguish the correct subspace structure used for data generation from incorrect ones.

We then performed UA and MSIVA experiments on the multimodal neuroimaging dataset using these four candidate subspace structures, and identified the optimal structure yielding the lowest final MISA loss. Separate follow-up CCA of each subspace recovered projections with maximum cross-modal correlation, for ease of interpretation. MANCOVA was used to check for association of the post-CCA sources with age and sex. To further evaluate that association, we performed an age prediction task and a sex classification task. Specifically, we trained a ridge regression model to predict the age and a linear support vector machine to classify the sex. 2907 subjects were divided into a training set of 2000 subjects and a hold-out test set of 907 subjects. We performed 10-fold cross-validation to choose the best hyperparameters (regularization parameter range: $[0.1, 1]$) on the training set, then trained the model using all 2000 training subjects and evaluated it on the hold-out test set.

3. RESULTS

As shown in Table 1 and Fig. 2, the ISI and loss values along the diagonal are the lowest (row-wise) in the synthetic data experiments, demonstrating that both UA and MSIVA can correctly identify and distinguish the ground truth subspace structures from the incorrect ones. Also, note the MSIVA

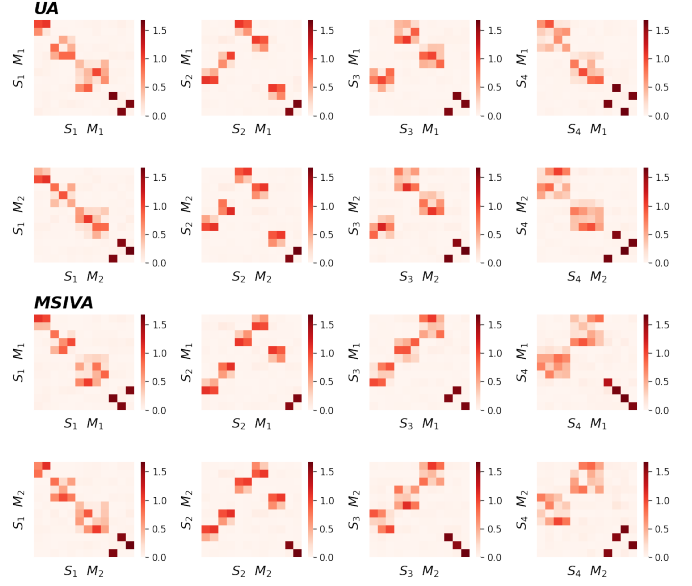


Fig. 3. Synthetic data results: Interference matrices corresponding to diagonal ISI values in Fig. 2.

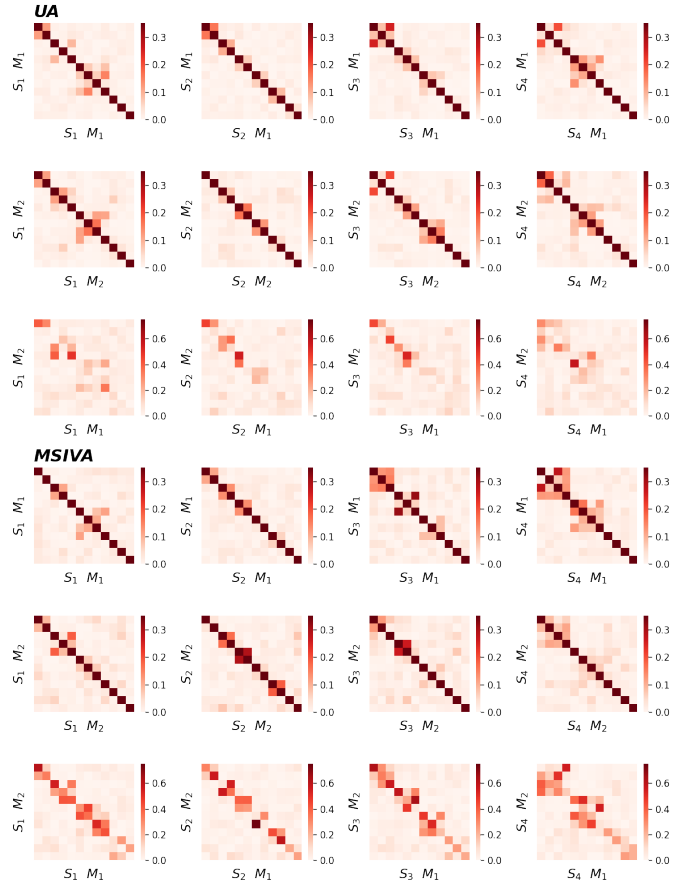


Fig. 4. Neuroimaging data result: Correlation of the recovered sources within each modality and across two modalities.

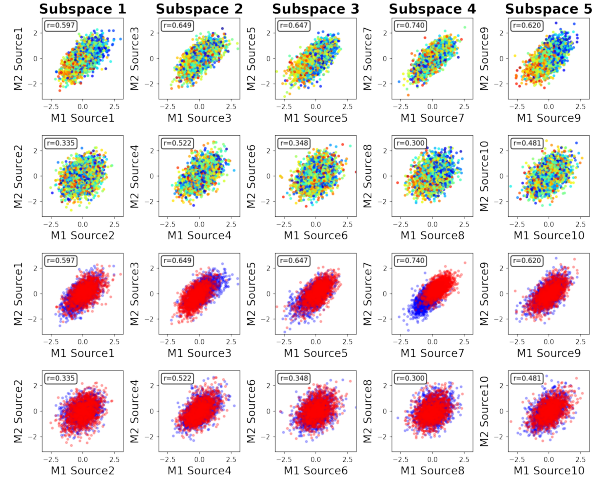


Fig. 5. MSIVA S_2 shared sources, color coded by phenotype. Rows I and II show age effect (warm color: older group; cold color: younger group). Rows III and IV show sex effect (blue: male; red: female). M_1 : sMRI GM, M_2 : fMRI ALFF.

loss value is smaller than the UA loss value for each of the four correct subspace structures, suggesting that MSIVA can fit the data a bit better than UA. The recovered subspace structures align with the proposed ground-truths (Fig. 3). Block permutation of the subspaces are acceptable, as long as alignment between modalities is retained, which is the case.

In the neuroimaging dataset, we observe within-modal self-correlation patterns align with the predefined subspace structures (Fig. 4, rows I-II and IV-V). We note that MSIVA recovers stronger cross-modal correlation than UA for all predefined subspace structures (Fig. 4, rows III and VI). MSIVA S_2 yields the lowest final MISA loss of 46.77 across all cases tested (Table 2), suggesting the S_2 subspace structure best fits the latent structure of this neuroimaging dataset.

We then identified the associations between phenotype measures and the sources captured by MSIVA S_2 . MANCOVA results suggest that the recovered subspaces are significantly associated with age and sex. Visual inspection of individual variability from the cross-modal CCA projections in each shared subspace (Fig. 5) suggests that subspaces 1, 3, 4 and 5 are associated with aging (especially cross-modal source 9 in subspace 5), while subspaces 2 and 4 show sex effect (especially cross-modal source 7 in subspace 4). The age regression and sex classification performance also confirmed this finding (Table 3). Specifically, the age prediction mean absolute error (MAE) in subspace 5 is the lowest (5.400 years), and sex classification accuracy is the highest in subspace 4 (0.812). The spatial maps from sources 7 and 9 from both modalities are presented in Fig. 6. The sex effect can be found in the parietal lobe, the occipital lobe (sMRI) and the cerebellar region (sMRI). The age effect is shown in sensorimotor and occipital areas in both sMRI and fMRI.

Table 3. Age regression MAE and sex classification accuracy (acc.) using shared subspaces estimated with MSIVA S_2 .

Subspace	1	2	3	4	5
Age MAE	5.757	6.270	5.936	5.888	5.400
Sex acc.	0.594	0.610	0.576	0.812	0.530

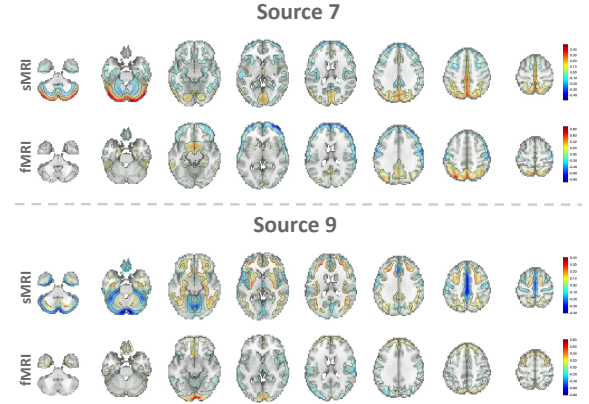


Fig. 6. Spatial maps of MSIVA S_2 CCA projections related to sex (source 7) and age (source 9).

4. DISCUSSION

We proposed a novel approach MSIVA to identify two types of sources, one is linked across modalities and one is specific to a modality. We first showed that both UA and MSIVA can correctly identify the subspace structure from incorrect ones and verified that the correct subspace structure always results in the lowest loss value from synthetic data experiments. We then applied both approaches on a large multimodal neuroimaging dataset and illustrated that MSIVA can identify linked subspace structures across two imaging modalities. Among all cases, MSIVA S_2 yields the lowest loss and, thus, is considered as the best fit to the latent structures in this dataset. The CCA projections within each shared subspace are significantly associated with age and sex, as verified through the prediction tasks. The age- and sex-related spatial maps align with previous findings [6]. Future work will include comparing MSIVA with MMIVA and applying data-driven subspace structures such as NeuroMark template [14].

5. CONCLUSION

Our proposed approach MSIVA can capture shared and modality-specific sources on both synthetic and neuroimaging datasets and yield a lower loss value compared with the unimodal baseline. The sources in the shared subspaces are significantly associated with age and sex according to statistical analysis and brain-phenotype modeling.

6. ACKNOWLEDGMENTS

The work was supported by the NSF CREST D-MAP Program (NSF2112455) and the Georgia Tech/Emory NIH/NIBIB Training Program in Computational Neural-engineering (T32EB025816).

7. COMPLIANCE WITH ETHICAL STANDARDS

This research study was conducted retrospectively using human subject data made available in open access by the UK Biobank Resource under Application Number 34175. Ethical approval was not required as confirmed by the license attached with the open access data.

8. REFERENCES

- [1] Rogers F Silva, Sergey M Plis, Tülay Adalı, Marios S Pattichis, and Vince D Calhoun, “Multidataset independent subspace analysis with application to multimodal fusion,” *IEEE Transactions on Image Processing*, vol. 30, pp. 588–602, 2020.
- [2] Pierre Comon, “Independent component analysis, a new concept?,” *Signal processing*, vol. 36, no. 3, pp. 287–314, 1994.
- [3] Taesu Kim, Torbjørn Eltoft, and Te-Won Lee, “Independent vector analysis: An extension of ica to multivariate components,” in *International conference on independent component analysis and signal separation*. Springer, 2006, pp. 165–172.
- [4] T. Adalı, M. Anderson, and G.-S. Fu, “Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging,” *IEEE Signal Proc. Mag.*, vol. 31, no. 3, pp. 18–33, May 2014.
- [5] J-F Cardoso, “Multidimensional independent component analysis,” in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP’98 (Cat. No. 98CH36181)*. IEEE, 1998, vol. 4, pp. 1941–1944.
- [6] Rogers F Silva, Eswar Damaraju, Xinhui Li, Peter Kochonov, Aysenil Belger, Judith M Ford, Sarah McEwen, Daniel H Mathalon, Bryon A Mueller, Steven G Potkin, et al., “Multimodal iva fusion for detection of linked neuroimaging biomarkers,” *bioRxiv*, 2021.
- [7] Karla L Miller, Fidel Alfaro-Almagro, Neal K Bangerter, David L Thomas, Essa Yacoub, Junqian Xu, Andreas J Bartsch, Saad Jbabdi, Stamatios N Sotiropoulos, Jesper LR Andersson, et al., “Multimodal population brain imaging in the uk biobank prospective epidemiological study,” *Nature neuroscience*, vol. 19, no. 11, pp. 1523–1536, 2016.
- [8] H. Hotelling, “Relations between two sets of variates,” *Biometrika*, vol. 28, no. 3/4, pp. 321–377, Dec. 1936.
- [9] Sai Ma, Nicolle M Correa, Xi-Lin Li, Tom Eichele, Vince D Calhoun, and Tülay Adalı, “Automatic identification of functional clusters in fmri data using spatial dependence,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 12, pp. 3406–3417, 2011.
- [10] Sai Ma, Xi-Lin Li, Nicolle M Correa, Tülay Adalı, and Vince D Calhoun, “Independent subspace analysis with prior information for fmri data,” in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2010, pp. 1922–1925.
- [11] Samuel Kotz, “Multivariate distributions at a cross road,” in *A Modern Course on Statistical Distributions in Scientific Work*, pp. 247–270. Springer, 1975.
- [12] O. Macchi and E. Moreau, “Self-adaptive source separation by direct or recursive networks,” in *Proc ICDSIP 1995*, Cyprus, 1995, pp. 122–129.
- [13] S.-I. Amari, A. Cichocki, and H. H. Yang, “A New Learning Algorithm for Blind Signal Separation,” *Proc NIPS 1996*, vol. 8, pp. 757–763, 1996.
- [14] Yuhui Du, Zening Fu, Jing Sui, Shuang Gao, Ying Xing, Dongdong Lin, Mustafa Salman, Anees Abrol, Md Abdur Rahaman, Jiayu Chen, et al., “Neuromark: An automated and adaptive ica based pipeline to identify reproducible fmri markers of brain disorders,” *NeuroImage: Clinical*, vol. 28, pp. 102375, 2020.