Methods 73 (2015) 43–53

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Eigenanatomy: Sparse dimensionality reduction for multi-modal medical image analysis

Benjamin M. Kandel [a,b,*], Danny J.J. Wang [c], James C. Gee [a,d], Brian B. Avants [a,d]

[a] Penn Image Computing and Science Laboratory, University of Pennsylvania, Philadelphia, PA, United States
[b] Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, United States
[c] Department of Neurology, University of California, Los Angeles, Los Angeles, CA, United States
[d] Department of Radiology, Hospital of the University of Pennsylvania, Philadelphia, PA, United States

ARTICLE INFO

Article history:
Received 9 April 2014
Received in revised form 2 October 2014
Accepted 15 October 2014
Available online 22 October 2014

Keywords:
Multi-modal
Magnetic resonance imaging
Sparse
Pediatric

ABSTRACT

Rigorous statistical analysis of multimodal imaging datasets is challenging. Mass-univariate methods for extracting correlations between image voxels and outcome measurements are not ideal for multimodal datasets, as they do not account for interactions between the different modalities. The extremely high dimensionality of medical images necessitates dimensionality reduction, such as principal component analysis (PCA) or independent component analysis (ICA). These dimensionality reduction techniques, however, consist of contributions from every region in the brain and are therefore difficult to interpret. Recent advances in sparse dimensionality reduction have enabled construction of a set of image regions that explain the variance of the images while still maintaining anatomical interpretability. The projections of the original data on the sparse eigenvectors, however, are highly collinear and therefore difficult to incorporate into multi-modal image analysis pipelines. We propose here a method for clustering sparse eigenvectors and selecting a subset of the eigenvectors to make interpretable predictions from a multi-modal dataset. Evaluation on a publicly available dataset shows that the proposed method outperforms PCA and ICA-based regressions while still maintaining anatomical meaning. To facilitate reproducibility, the complete dataset used and all source code is publicly available.

1. Introduction

Modern imaging datasets are increasingly multimodal. Virtually all modern large-scale imaging studies, even those that concentrate on a given modality, such as resting state fMRI [1], include a variety of imaging measures [2,3]. Although some groups have reported improvements in classification accuracy in Alzheimer's Disease when using multimodal data [4], others have claimed that multimodal classification does not tend to outperform a single sensitive test [5]. This trend towards multimodal data presents challenges in data processing, visualization, and statistical inference. In particular, the extremely high dimensionality of medical imaging data presents challenges to classical linear model-based statistical analyses, which assume that there are more subjects than measured variables ($n > p$). Several approaches exist to deal with the high-dimensional nature of medical imaging datasets.

1.1. Mass-univariate approaches

One of the most widely used methods to perform statistical analyses on medical images is to use voxel-based morphometry (VBM) [6]. VBM performs a statistical test on each voxel in the image, producing a spatial map that describes how closely the values at a given voxel are correlated with an outcome measure. The massive number of multiple comparisons conducted when using VBM necessitate appropriate corrections [7]. In addition, because brain function is spread over regions larger than a single voxel [8], multivariate approaches are more naturally suited to leveraging the spatially distributed information contained in medical imaging data [9].

When examining multimodal data, univariate approaches are further restricted because they do not provide insight into the relationships between the various modalities. One way of using univariate approaches to analyze multimodal data is to perform separate mass-univariate analyses on each modality and examine the degree of spatial overlap between the resulting statistical maps [10–12]. A drawback of this method is that spatial overlap alone does not give insight into the subject-wise interactions or

* Corresponding author at: Penn Image Computing and Science Laboratory, University of Pennsylvania, Philadelphia, PA.
E-mail address: bkandel@seas.upenn.edu (B.M. Kandel).

http://dx.doi.org/10.1016/j.ymeth.2014.10.016
1046-2023/

correlations of the various modalities. To take a somewhat extreme example, if half the experimental population have increased cortical thickness as compared to controls and the other half have increased BOLD activation, a spatial map may show overlapping significant areas, even though no individual subject actually has increased cortical thickness and increased BOLD activation. To provide greater insight into the biological mechanisms underlying observed changes, several studies have begun investigating multivariate approaches to multimodal data [11,13,14], looking at, for example, the correlation between cortical thickness and BOLD activation in a given region.

One challenge of integrating large multimodal datasets is the difficulty in visualizing and interpreting the results, especially when performing multivariate analyses of data. Interpretation of multivariate data is often made easier by sparse methods, which ensure that only a small part of the data set is used for predicting an outcome variable. Sparse methods have enjoyed a resurgence in popularity in recent years, with several groups proposing sparse methods tuned for neuroimaging data [15–24]. Applying sparse techniques to multi-modal data enables specific and biologically interpretable statements to be made about data; for example, "Decreased cortical thickness in the left parietal lobe is correlated with decreased perfusion in the left and right parietal lobes, and this network together predicts a decrease in verbal ability."

1.2. Data-driven dimensionality reduction

Many clinical studies using multimodal imaging data average image values over predefined regions of interest (ROI's) to reduce the dimensionality of the data so that it will be more amenable to standard statistical analyses. Although this approach may be ideal if the ROI's are already known and have anatomically meaningful boundaries, this is not ideal for exploratory analyses which have minimal prior knowledge. Traditionally, linear regression from a high-dimensional dataset is performed after a dimensionality reduction step, such as principal component analysis (PCA) [25]. However, PCA-derived eigenvectors have global support and therefore do not provide anatomical specificity. Sparse techniques can provide more local specificity. In particular, a recently introduced sparse dimensionality reduction technique, "eigenanatomy," has proven to provide greater power to detect group differences than either voxel-based morphometry (VBM) [26] or pre-defined ROI's [27] while maintaining anatomical interpretability. Here, we extend the eigenanatomy approach to a multi-modal setting. Although the sparse eigenvectors are orthogonal in the image space, orthogonality is not enforced on the low-dimensional coefficients generated by projecting the imaging data onto the sparse eigenvectors. Therefore, care must be taken to prevent excessive collinearity among the predictor variables. We demonstrate that even with collinearity in the predictor variables, our method of extending eigenanatomy to multi-modal datasets produces a more accurate prediction of age in a pediatric population than principal component regression, independent component regression, or regression on average values within regions defined by the AAL atlas.

The eigenanatomy objective function is not new to this work. Here, we focus on the practical challenges, including validation, interpretation, and visualization of predictive models, involved in multimodal data analysis, and demonstrate the advantages of the eigenanatomy framework for multi-modal neuroimaging studies as compared to either classical dimensionality reduction techniques or predefined regions of interest (ROI's). The release of all data and code used to generate the paper will facilitate the use of this technique as a template for future studies, as well as encourage reproduction of similar evaluations with different datasets.

2. Methods

2.1. Reproducibility

To facilitate the use of this study as a template for other multimodal population studies, we have attempted to make it as reproducible as possible. All the data is available from an open-access data repository. The paper itself is written using the R package `knitr` [28], which facilitates on-the-fly production of figures from data, enhancing reproducibility and documenting all data processing steps. The full code for producing the paper, including raw data and code for producing figures, is available from https://bitbucket.org/bkandel/multimodaleanat.

2.2. Dimensionality reduction techniques

Dimensionality reduction is a technique to reduce the complexity of input data into a relatively small number of summary measures. Linear dimensionality techniques can be written as a matrix factorization problem. We assume the input data is given in an $n \times p$ input matrix $\mathbf{X}$, where n is the number of observations or subjects and p is the number of variables associated with each observation. In the context of medical imaging, n typically ranges from a few tens to a few hundred, and p is on the order of 10^3–10^6, depending on the size of images. Dimensionality reduction seeks to find a factorization of $\mathbf{X}$ into an $n \times k$ *coefficient* matrix $\mathbf{U}$ and a $p \times k$ *loading* or eigenvector matrix $\mathbf{V}$ so that $\mathbf{X} \approx \mathbf{U}\mathbf{V}^{\mathrm{T}}$. The most well-established method for dimensionality reduction is principal component analysis (PCA), which finds an orthogonal matrix (i.e. $\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{1}$) that projects the input matrix to a lower-dimensional subspace. More recently, independent component analysis (ICA, e.g. [29]) has become widely used in the neuroimaging community. ICA seeks a decomposition of $\mathbf{X}$ in which the components are independent, which is a stronger condition than orthogonality.

One drawback of standard PCA and ICA is that the eigenvectors often cover the entire input matrix, meaning that each entry in the coefficient matrix is a weighted average of all the voxels in the image. This makes interpretation of the output difficult for two related reasons. First, the lack of spatial specificity of the eigenvectors makes it difficult to use the coefficients to investigate anatomically-informed biological hypotheses. For example, it is impossible to use the coefficients from a PCA decomposition to look at the relation between left precuneal atrophy and age. In addition, because the eigenvectors have both positive and negative components, interpreting the weights in a linear regression model that relates the coefficients to an outcome measure is not intuitive. This is for two reasons. First, the PCA eigenvectors can contain negative weights, even if the input data is strictly positive, as is the case in cortical thickness, perfusion, and fractional anisotropy (FA) images. If the weight of a given PCA coefficient in a linear model is positive but the corresponding entry in the eigenvector is negative, it follows that an *increase* in the input matrix corresponds to a *decrease* in the outcome variable. Second, this problem is compounded by the overlapping nature of the eigenvectors: Because a given brain region can contribute positively or negatively to each eigenvector, it is very difficult to go back from the coefficient weights to the biological meaning of the weight. Furthermore, interpreting the eigenvectors themselves without accounting for the weights is ill-advised [30], as the eigenvectors can be confounded by biases in the data in non-obvious ways [31]. Therefore, although PCA may be used for predicting age in unseen data, it is not as useful for testing biological hypotheses.

Sparse dimensionality reduction techniques [32,33] deal with the problems of global support of PCA eigenvectors by enforcing