Neuroscience Letters 513 (2012) 141–145

Contents lists available at SciVerse ScienceDirect

Neuroscience Letters

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

ISOMAP induced manifold embedding and its application to Alzheimer's disease and mild cognitive impairment

Hyunjin Park*, the ADNI[1]

Department of Biomedical Engineering, Gachon University, Republic of Korea

ARTICLE INFO

Article history:
Received 30 September 2011
Received in revised form 6 February 2012
Accepted 7 February 2012

Keywords:
Manifold embedding
ISOMAP
Alzheimer's disease
Mild cognitive impairment
Shape quantification
Morphometry

ABSTRACT

Neuroimaging data are high dimensional and thus cumbersome to analyze. Manifold learning is a technique to find a low dimensional representation for high dimensional data. With manifold learning, data analysis becomes more tractable in the low dimensional space. We propose a novel shape quantification method based on a manifold learning method, ISOMAP, for brain MRI. Existing work applied another manifold learning method, multidimensional scaling (MDS), to quantify shape information for distinguishing Alzheimer's disease (AD) from normal. We enhance the existing methodology by (1) applying it to distinguish mild cognitive impairment (MCI) from normal, (2) adopting a more advanced manifold learning technique, ISOMAP, and (3) showing the effectiveness of the induced low dimensional embedding space to predict key clinical variables such as mini mental state exam scores and clinical diagnosis using the standard multiple linear regression. Our methodology was tested using 25 normal, 25 AD, and 25 MCI patients.

Introduction

Many neurodegenerative diseases cause unique morphological changes in brain anatomy. Only certain structures of the brain are selectively affected by the diseases, while the rest of the brain remains the same. Alzheimer's disease (AD) is known to cause atrophy in the hippocampus region. Computational anatomy (CA) is a research field which applies a computer algorithm to quantify such changes in shape information [8]. The task of measuring shape is not simple and has been a matter of significant controversy [2,4]. There are two main approaches for measuring shape. The first approach, deformation based morphometry (DBM), assumes that all shape information is encoded in the deformation fields, which relates one brain scan with another scan [3]. The second approach, voxel based morphometry (VBM), assumes that all shape information is encoded in some scalar function of the registered scans [1]. Two scans are segmented and then linearly registered so that both scans are in the same spatial coordinates in the VBM approach. Shape information is assessed by voxel-wise difference in the labels after registration. DBM uses deformation fields obtained from registrations of a population and identifies differences in the relative positions of structures within the region of interest (ROI). One major weakness of DBM is that it requires a very accurate registration algorithm for computation of the displacement field. We adopt the DBM approach and assume all the necessary shape information is given by the displacement fields. This study focuses on AD and mild cognitive impairment (MCI). Many patients with MCI often convert to AD later in the disease progression. MCI has been extensively studied as detection of MCI is highly related to early diagnosis of AD. We apply a computer algorithm to measure shape differences so that AD or MCI patients may be distinguished from normal patients. The shape information is computed from the displacement field and the displacement field is burdened with high dimensionality. The dimension of a displacement field is as high as the number of voxels of the given scan, which may number in the millions. Many displacement fields must be considered when studying a population. Hence, the dimension of the overall data is quite large. One way to ease the burden of high dimensionality is to apply manifold learning techniques. Manifold learning is a technique used for finding a low dimensional representation for high dimensional data [10]. Researchers applied manifold learning methods to neuroimaging data in order to effectively represent shape information in a low dimensional embedding space [7].

Park et al. [12] applied multidimensional scaling (MDS) combined with bending energy of the displacement field to discriminate shape information between AD from normal controls [12]. Their approach reported effective separation of AD from normal controls and showed robustness to errors in displacement fields improving the major weakness with DBM based shape quantification. We

* Tel.: +82 32 820 4432; fax: +82 32 820 4059.
E-mail address: hyunjinp@gachon.ac.kr

[1] Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI).

0304-3940/$ – see front matter
doi:10.1016/j.neulet.2012.02.016

extended Park et al. [12] to include MCI patients in this study. Here we adopted a more advanced manifold learning technique called ISOMAP instead of MDS. Shape information lies with non-Euclidean Riemannian space and the induced distance is known to be geodesic. ISOMAP better represents geodesic distances and hence is better suited for studying shape information. In the previous study, the induced low dimensional embedding space was used as a feature space for a kNN classifier to distinguish between AD and normal. We used the low dimensional embedding space as a feature space to perform classical statistical tests for clinical variables, including score of mini mental state exam (MMSE) and clinical diagnosis. Using multiple linear regression, we performed statistical testing in order to determine how well the low dimensional embedding space was able to predict MMSE score and clinical diagnosis. Support vector machine (SVM) classifier combined with advanced feature selection has been successfully applied to classify normal and AD for PET, resulting up to 90% classification accuracy [11]. The SVM based method is one of the state of the art methods for classifying AD and normal. We compared the performance of the method in this study with that of SVM based method in the results section. In summary, this study builds on the work of Park et al. [12] and we extended the methodology by (1) testing its applicability to MCI patients, (2) adopting a more advanced manifold learning technique, and (3) testing the effectiveness of the low dimensional embedding space as a platform to carry out statistical tests for key clinical variables.

Materials and methods

Our shape quantification methodology is very similar to the previous work [12]. Here we briefly describe the overall framework emphasizing the differences.

Registration framework

Registration is a task of finding the geometric mapping between two images, so that one image can be mapped onto the other. This study used mutual information (MI) as the similarity measure and thin-plate splines (TPS) as the geometric interpolant [9]. There are many definitions of MI and we adopt the metric MI in the equation below.

$$\mathrm{MI}(X, Y) = H(X|Y) + H(Y|X) \tag{1}$$

where X is the intensity distribution of scan X, Y is the intensity distribution of scan Y, $H(X|Y)$ is the conditional entropy of X given Y, and $H(Y|X)$ is the conditional entropy of Y given X. This particular variant of MI satisfies the metric property including symmetry and triangular inequality. Many manifold learning techniques work best with metric distance measures.

Distance measure

Registration between two scans yields a geometric transform optimized for maximization of a certain cost function. The displacement field is a collection of evaluations of the geometric transform at all voxel locations. The entire deformation field, whose order is equal to the number of voxels, is compressed to a single scalar value. The scalar value is the geometric distance, hereafter called distance, which measures the roughness of the geometric transform that associates the coordinate spaces of two scans. If the geometric transform is complex between two images, then the distance measure will be high. If the geometric transform is simple, then the distance measure will be small. We adopted the integral of the squared value of the second-order derivative of the geometric transform as the distance measure. We chose the second-order derivative in order to ensure invariance to affine transforms. The formulation for the distance measure is given below.

$$d^2 = \iint \left(\frac{\partial^2 f_x}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f_x}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f_x}{\partial y^2}\right)^2 dxdy + \iint \left(\frac{\partial^2 f_y}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f_y}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f_y}{\partial y^2}\right)^2 dxdy \tag{2}$$

where f_x displacement in x and f_y displacement in y.

Formulation in (2) is for 2D and can be easily extended for 3D. This distance is called the bending energy. The proposed distance measure is based on the displacement field. Others adopted a distance measure based on grayscale information of the registered scans [14]. If two scans can be registered with a high MI value, it is likely that two scans are similar hence the distance between the two should be small. Thus, the inverse of MI (i.e., 1/MI) may be used as a distance measure. We compared our distance measure based on the displacement field (i.e., bending energy) with the distance measure based on grayscale values (i.e., metric MI) in the results section.

ISOMAP

ISOMAP is a manifold learning technique based on pair-wise distances derived from high dimensional data [13]. Compared to traditional manifold learning methods such as MDS, it approximates the geodesic distances using weighted neighborhood graphs. Shape information occupies non-Euclidean Riemannian space and the induced distance is geodesic, thus ISOMAP is well equipped to deal with shape information. Given a set of distances in the distance matrix D, ISOMAP outputs a set of coordinates in a user-specified dimension. The dimension of ISOMAP output is determined based on the eigenstructure of the distance matrix. The output coordinates are in the standard Euclidean space of the user-chosen dimension. ISOMAP considers only distance measures from nearby objects and approximates large distances from distant objects by composition of small scale distances. It basically trusts only small distance values and approximates the large distance values using composition of small distances. For scans that are relatively similar and thus easy to register, the resulting bending energy is likely to be small and ISOMAP places high confidence on such distances. For scans that are vastly different and thus difficult to register, the resulting bending energy is likely to be large and ISOMAP does not trust such large bending energy values. Instead, ISOMAP approximates the difficult registration between two vastly different scans by composition of small scale and easy to do registration tasks.

Framework for shape quantification

ISOMAP produces relative positional locations from a collection of pair-wise distances, which in turn assigns a low dimensional coordinate for each MRI scan. The key idea is to use a distance measure that quantifies distances between MRI scans. We adopted a distance measure called bending energy, which is based on the displacement field. Output of ISOMAP is often visualized on a scatter plot, where each dot represents a scan. In the scatter plot, the relative positions of all scans are plotted in the Euclidean space of a user-chosen dimension. We hypothesize that scans of the same type will be placed adjacent and scans of different types will be placed separately. Therefore, we expect a scatter plot in which two distinct clusters can be observed. The low dimensional embedding space (i.e., ISOMAP output coordinates) may be used as a feature space for quantification of shape information. In this study, the ISOMAP embedding space was used as the feature space for