NeuroImage 147 (2017) 360–380

Contents lists available at ScienceDirect

NeuroImage

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

Towards a Holistic Cortical Thickness Descriptor: Heat Kernel-Based Grey Matter Morphology Signatures☆

Gang Wang[a,*], Yalin Wang[b,*], for the Alzheimer's Disease Neuroimaging Initiative

[a] School of Information and Electrical Engineering, Ludong University, Yantai, Shandong 264025, China
[b] Arizona State University, School of Computing, Informatics, Decision Systems Engineering, 699 S. Mill Avenue, Tempe, AZ 85281, United States

ARTICLE INFO

Keywords:
Magnetic resonance imaging (MRI)
Alzheimer's disease
Computer-Aided Diagnosis
Shape analysis
Heat Kernel

ABSTRACT

In this paper, we propose a heat kernel based regional shape descriptor that may be capable of better exploiting volumetric morphological information than other available methods, thereby improving statistical power on brain magnetic resonance imaging (MRI) analysis. The mechanism of our analysis is driven by the graph spectrum and the heat kernel theory, to capture the volumetric geometry information in the constructed tetrahedral meshes. In order to capture profound brain grey matter shape changes, we first use the volumetric Laplace-Beltrami operator to determine the point pair correspondence between white-grey matter and CSF-grey matter boundary surfaces by computing the streamlines in a tetrahedral mesh. Secondly, we propose multi-scale grey matter morphology signatures to describe the transition probability by random walk between the point pairs, which reflects the inherent geometric characteristics. Thirdly, a point distribution model is applied to reduce the dimensionality of the grey matter morphology signatures and generate the internal structure features. With the sparse linear discriminant analysis, we select a concise morphology feature set with improved classification accuracies. In our experiments, the proposed work outperformed the cortical thickness features computed by FreeSurfer software in the classification of Alzheimer's disease and its prodromal stage, i.e., mild cognitive impairment, on publicly available data from the Alzheimer's Disease Neuroimaging Initiative. The multi-scale and physics based volumetric structure feature may bring stronger statistical power than some traditional methods for MRI-based grey matter morphology analysis.

1. Introduction

A multitude of morphometric studies using magnetic resonance imaging (MRI) have examined the brain structural marker changes associated with Alzheimer's disease (AD), including cortical atrophy, hippocampal atrophy or ventricular enlargement, which may serve as indicative signs of early diagnosis of AD. Despite evidence that medial temporal atrophy is associated with AD progression, the evaluation of medial temporal atrophy is still not sufficiently accurate on its own to serve as a definitive diagnostic specification for the clinical diagnosis of AD at the mild cognitive impairment (MCI) stage (Frisoni et al., 2010). Missing at this time is some structural features which are able to capture subtle grey matter morphometry differences between different clinical groups and thus have a high discriminant power. For instance, the thickness of cortex is an important feature which has been applied to detect localized anatomical differences (MacDonald et al., 2000; Fischl and Dale, 2000; Jones et al., 2000; Miller et al., 2000; Kabani et al., 2001; Sowell et al., 2004; Chung et al., 2005; Kochunov et al., 2012). At present, there are two different computational paradigms on cerebral cortical thickness estimation, with methods generally classified as either surface or voxel based (Clarkson et al., 2011; Jones et al., 2000; Fischl and Dale, 2000). Besides, there is another line of research that has been focused on grey matter density (GMD) (e.g. Wright et al., 1995; Ashburner and Friston, 2000; Thompson et al., 2003; Sowell et al., 1999). Basically, GMD can be defined as the proportion of voxels classified as grey matter falling within a sphere centered at some points on registered cortical surfaces (Thompson et al., 2004). Prior research (e.g. Narr et al., 2004) has shown that GMD is highly correlated with cortical thickness measures. However, the main disadvantages of current cortical thickness estimation methods are either computation complexity on constantly correcting the weights of various evolutionary parameters or inaccuracy on the discrete grid. In addition, all measured

☆ Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

* Corresponding authors.
E-mail addresses: gangwang1970@ldu.edu.cn (G. Wang), ylwang@asu.edu (Y. Wang).

http://dx.doi.org/10.1016/j.neuroimage.2016.12.014
Received 20 September 2016; Accepted 7 December 2016
Available online 26 December 2016
1053-8119/

distances are unitary distances between boundary surface points and they indicate only global trends and cannot measure topological variations (e.g., the local topological characteristics along the streamlines are not studied).

In this paper, we propose a diffusion geometry method to obtain multi-scale intrinsic grey matter morphology signatures (GMMS). This class of methods offers the advantages of inelastic deformation invariance and robustness to topological noise. Mathematically, diffusion kernels (Coifman et al., 2005) express the transition probability by random walk of time t, $t \geq 0$. It allows for defining a scale space of kernels with the scale parameter t. Such heat kernel-based spectral analysis induces a robust and multi-scale metric to compare different shapes and has strong theoretical guarantees. It has achieved great success in machine learning, computer vision and medical imaging research (Coifman et al., 2005; Reuter et al., 2006; Nain et al., 2007; Yu et al., 2007; Sun et al., 2009; Lombaert et al., 2013). Some prior work (Rustamov, 2011) studied volumetric heat kernel but their work mainly relied on regular grid mesh, thus suffering from the limited grid resolution which cannot precisely characterize the curved cortical surfaces from MR images. In contrast, we model the grey matter structure by tetrahedral meshes. Based on the volumetric Laplace-Beltrami operator (Shi et al., 2015; Wang et al., 2015; Wachinger et al., 2015; Wang and Wang, 2015), we introduce multi-scale heat kernel shape descriptors to depict the heat-transfer probability by random walk between some pre-determined boundary point pairs (Wang et al., 2015) and obtain sub-voxel numerical accuracy. The defined shape features rely upon a solid theoretical background and are robust to noise. Our work may provide accurate quantitative measures of grey matter morphology changes which are important for a variety of neuroimaging studies.

In our work, a new set of morphological descriptors are introduced to describe the grey matter morphology changes. They depend on heat transmission time and are also influenced by the topological properties on the heat transmission paths. Following that, a point distribution model (PDM) is applied to reduce the feature dimensionality. To further reduce feature dimension and improve classification accuracy, we adopt a sparse learning (Hastie et al., 2015) approach, sparse linear discriminant analysis (Sparse LDA) (Clemmensen et al., 2011), which is built upon a solid theoretical foundation and has demonstrated its strong practical values in imaging research. Combining heat kernel shape features with sparse LDA, we hypothesize that our framework may provide robust, informative and biologically meaningful grey matter morphology measures, and therefore, we may make an important advancement towards a holistic cortical thickness descriptor. We tested our hypothesis on the Alzheimer' s Disease Neuroimaging Initiative (ADNI) dataset. We studied the classification of AD and its prodromal stage, i.e. mild cognitive impairment (MCI), comparing our new method to cortical thickness feature estimated by FreeSurfer software (Fischl et al., 1999a).

2. Methods

2.1. Definitions

2.1.1. Heat operator and heat kernel

The heat kernel diffusion on differentiable manifold M with Riemannian metric is governed by the following heat equation:

$$\triangle_M f(x, t) = \frac{\partial f(x, t)}{\partial t} \tag{1}$$

where $f(x, t)$ is the heat distribution of the volume at the given time. Assuming an initial heat distribution $F: M \to \mathcal{R}$, let $H_t(F)$ denotes the heat distribution at time t, and $\lim_{t\to 0} H_t(F) = F$. $H(t)$ is called the *heat operator*. Both $\triangle_M$ and H_t share the same eigenvectors, and if λ_i is an eigenvalue of $\triangle_M$, then $e^{-\lambda_i t}$ is an eigenvalue of H_t corresponding to the same eigenvector.

For any compact Riemannian manifold, there exists a function $l_t(x, y)$: $\mathcal{R}^+ \times M \times M \to \mathcal{R}$, satisfying the formula

$$H_t F(x) = \int_M l_t(x, y) F(y) dy \tag{2}$$

where dy is the volume form at $y \in M$. The minimum function $l_t(x, y)$ that satisfies Eq. (2) is called the *heat kernel* (Coifman et al., 2005). It can be explained as the amount of heat that is transferred from x to y in time t given a unit heat source at x. In other words, $l_t(x, \cdot) = H_t(\delta_x)$ where δ_x is the Direc delta function at x: $\delta_x(z) = 0$ for any $z \neq x$ and $\int_M \delta_x(z) = 1$.

According to the theory of the spectral analysis, for compact M, the heat kernel has the following eigen-decomposition heat diffusion distance:

$$l_t(x, y) = \sum_{i=0}^{\infty} e^{-\lambda_i t} \phi_i(x) \phi_i(y) \tag{3}$$

where λ_i and ϕ_i are the i^{th} eigenvalues and eigenvectors of the Laplace-Beltrami operator, respectively. The heat kernel $l_t(x, y)$ can be interpreted as the transition density function of the Brownian motion on the manifold (Sun et al., 2009).

2.1.2. Laplace-beltrami operator

Let f be a real-valued function, with $f \in C^2$, defined on a Riemannian manifold M (differentiable manifold with Riemannian metric). The Laplace-Beltrami Operator Δ is:

$$\Delta f = div(grad f) \tag{4}$$

with grad f the gradient of f and div the divergence on the manifold. The Laplace-Beltrami operator is a linear differential operator. It can be calculated in local coordinates. In the non-Euclidean case for a Riemannian metric, the Laplacian eigenvalue problem is given by

$$\Delta f = -\lambda f \tag{5}$$

Since the Laplace-Beltrami operator is self-adjoint and semi-positive definite (Rosenberg, 1997), it admits an orthonormal eigensystem $\mathbb{B} := (\lambda_i \varphi_i)$, which is a basis of the space of square integrable function, with $\Delta \varphi_i = \lambda_i \varphi_i$, $\lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_i \leq \ldots \leq \infty$. Detailed treatments on harmonic map, heat kernel and the Laplace-Beltrami operators can be found in (Schoen and Yau, 1997; Rosenberg, 1997). There are also intensive studies on discrete harmonic map and Laplace-Beltrami operator (e.g. Pinkall and Polthier, 1993; Wang et al., 2004a; Coifman et al., 2005; Levy, 2006; Reuter et al., 2006).

2.1.3. Discrete volumetric laplace-beltrami operator

The solution to Eq. (5) on volume can be approximated by a piecewise linear function $f: \mathcal{T} \to \mathcal{R}$ over a Tetrahedralization $\mathcal{T}$ with vertices P: p_u, $u = 1, \ldots, n$, n is the vertex number in a tetrahedron. In this work, we propose to refine our prior discrete Laplace-Beltrami operator definition (Wang et al., 2015) by using a normalization factor, which takes into account the volume of all tetrahedra at each vertex. The lumped discrete Laplace-Beltrami operators can be represented as:

$$\Delta f(p_u) = \frac{1}{d_u} \sum_{v \in N(u)} k_{u,v}(f(p_u) - f(p_v)) \tag{6}$$

where $N(u)$ represents the index set of the 1-ring of the vertices p_u, i.e., the indices of all neighbors connected to p_u by an edge. The normalization factor, which takes into account the total tetrahedral volume $a(u)$ of all tetrahedra at vertex u, is defined as $d_u = a(u)/4$. We define the vector $f = [f(p_1, \ldots, f(p_n)]^T$ of the function values at the vertices, the weighted adjacency matrix $K = (k_{u,v})$ (our definition of $k_{u,v}$ will be discussed in Section 2.2.3) and the diagonal matrix $W = diag(w_1, \ldots, w_n)$ which contains the diagonal elements $w_u = \Sigma_{v \in N(u)} k_{u,v}$. Then the stiffness matrix can be defined as $A = W - K$,