3D Surface Matching and Registration through Shape Images

Zhaoqiang Lai and Jing Hua

Department of Computer Science,
Wayne State University,
Detroit, MI 48202, USA

Abstract. In this paper, we present a novel and efficient surface matching framework through shape image representation. This representation allows us to simplify a 3D surface matching problem to a 2D shape image matching problem. Furthermore, we present a shape image diffusion-based method to find the most robust features to construct the matching and registration of surfaces. This is particularly important for inter-subject surfaces from medical scans of different subjects since these surfaces exhibit the inherited physiological variances among subjects. We conducted extensive experiments on real 3D human neocortical surfaces, which demonstrate the excellent performance of our approach in terms of accuracy and robustness.

Key words: Surface Matching, Conformal Map, Shape Image

1 Introduction

3D surface matching and registration remains a very challenging research problem in many medical image analysis fields since surfaces may have very flexible free-form shape characteristics which are difficult to capture and use for matching and registration purpose. The key issue in surface matching has been the shape representation scheme. In recent years, a large amount of research effort has been devoted on finding appropriate shape representations for 3D surface matching. Different approaches, such as curvature-based representations [1], regional point representations [2], shape distributions [3], spherical thin-plate splines [4], etc., have been proposed for this purpose. However, many of these representations are not robust and cannot perform well for intersubject surface registration under such circumstance as noise, clutter, and physiological variances.

This paper presents a novel and efficient surface matching framework through shape image representation. Basically, a 2D representation is constructed to describe a 3D surface. In practice, a family of geometric maps can be adopted to create the 2D representation. This simplifies a surface matching problem to a 2D image matching problem. When constructing shape images, Riemann geometric maps, which can encode the shape information of the surface into the 2D image, provide a viable solution to the mapping of a 3D surface onto a 2D domain. Theoretically, in Riemannian geometry, any surface admits a Riemannian metric

of constant Gaussian curvature. In fact, one can find a metric with constant Gaussian curvature in any given conformal class. Certain 3D surfaces can be mapped to a 2D domain through a global optimization and the resulting map is a diffeomorphism [5–8]. As a result, the 3D surface-matching problem can be rigorously simplified to a 2D shape image-matching problem through the Riemannian geometric maps. These maps are stable, insensitive to resolution changes and robust to noise. The analysis of 2D image which integrates geometric and appearance information is a better understood problem [9–11]. Based on the Riemannian geometric maps, we further provide a diffusion-based algorithm for detecting the most robust shape features from the shape images. Therefore, highly accurate and efficient 3D shape matching algorithms can be achieved by matching these features. The rest of this paper will discuss the conformal maps, describe how to construct the shape image given a 3D surface, and then the matching framework based on shape image diffusion. The conducted experiments demonstrates the excellent performance of the proposed approach.

2 Shape Image Construction with Conformal Mapping

A good shape image should be able to fully represent the global geometric characteristics of a given surface and also serves a domain for modeling and indexing of other heterogenous attributes. Thus, a 3D surface can be converted to a multidimensional vector image for effective processing. Technically, we resort to Riemannian geometric mapping to accomplish the task.

In the theorem of differential geometry, a diffeomorphism $f : M \rightarrow N$ is conformal if and only if, for any surface patch σ_m on M, the first fundamental forms of σ_m and $\sigma_n = f \circ \sigma_m$ are proportional. Mathematically, this means that

$$f \circ ds_m^2 = \lambda ds_n^2, \tag{1}$$

where λ is called the conformal factor, ds_m^2 and ds_n^2 are the first fundamental form on M and N. By minimizing the harmonic energy over the surface, we can calculate the conformal map [12, 13].

To match 3D shapes accurately and efficiently, a new 2D representation, shape image, is developed in our framework using conformal mapping. We have previously shown that, given a surface patch M, its conformal image I_c can be created using conformal mapping. There is one-to-one correspondence between the vertices in M and the vertices in I_c. With the involvement of the shape attributes at each vertex of M, attribute values can be interpolated and computed for each pixel of the conformal shape image. As discussed in [15, 16], conformal representation surface $S(u, v)$ can be represented by conformal factor function, $\lambda(u, v)$, and mean curvature function, $H(u, v)$. We assign these two attributes to the I_c to form a vector image I, where the pixel attributes are represented by a vector $[H, \lambda]^\top$. Figure 1 shows the human neocortical surface (Figure 1(a)) and its corresponding mean curvature channel (Figure 1(b)) and conformal factor channel (Figure 1(c)). The composite shape image is shown in Figure 1(d).

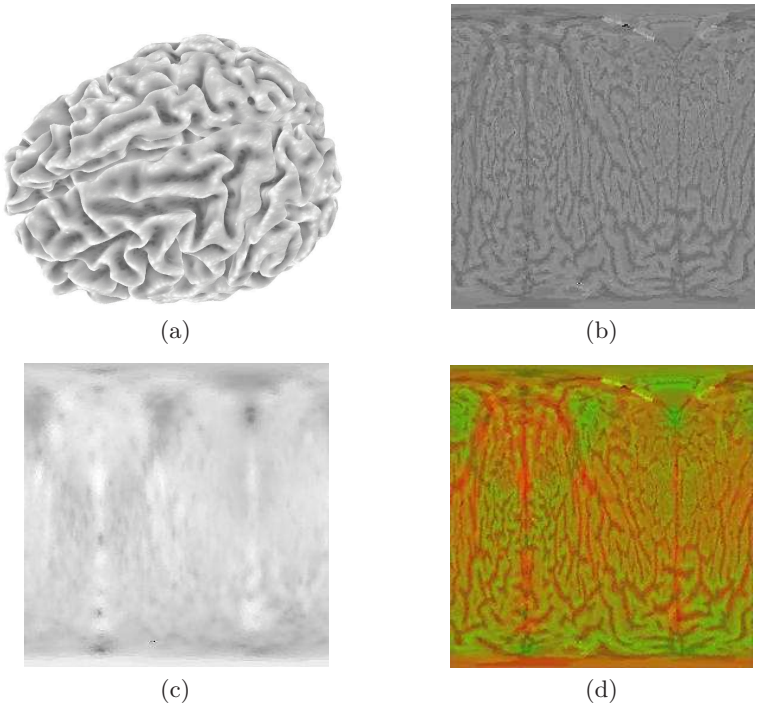

Fig. 1. Shape Image. (a) shows the neocortical surface extracted from MR scans; (b) and (c) are the mean curvature channel and conformal factor channel of the shape images of the neocortical surface; (d) is the composite shape image including both channels.

3 Feature Extraction through Shape Image

Since a surface can be represented as a unique shape image composed of conformal factors and curvatures, many algorithms suitable for image computing may be used in the analysis and feature extraction of this type of images. For the purpose of matching and registration, the main task is to find the stable keypoints or regions and their local image features for alignment. This section describes a novel diffusion-based algorithm to extract distinctive features from the shape images. Through the shape image diffusion, we can identify the robust keypoints and their scales from the computed diffusion extrema, which are ideal for matching purpose.

3.1 Shape image diffusion

Our shape image is a multichannel image, similar to a color image. The simplest way to do the diffusion filtering of the shape images is to deal with channels separately and independently from each other. For a vector image $I =$

$(I_1, I_2, ..., I_m)^\top$, the anisotropic diffusion is performed by

$$\frac{\partial I_i}{\partial t} = div\Big(g\nabla I_i\Big) \qquad (i = 1, ..., m), \tag{2}$$

where div indicates the divergence operator, and ∇ is the gradient operator. However, this method leads to an undesirable effect that edges may be formed at different locations for each channel. In our framework, we employs a common diffusivity g which combines information from all channels. For the case in which g is a constant for a specific channel I_i, it reduces to the isotropic heat diffusion equation,

$$\frac{\partial I_i}{\partial t} = c\triangle I_i, \tag{3}$$

where $\triangle$ is the Laplacian operators. Its solution is Gaussian smoothing. Gaussian smoothing has a typical disadvantage: Gaussian smoothing does not only reduce noise, but also blurs important features such as edges and, thus, makes them harder to identify.

One model to improve this is inhomogeneous linear diffusion filtering,

$$\frac{\partial I_i}{\partial t} = div\left(g\Big(|\nabla f_{I_i}|^2\Big)\nabla I_i\right), \tag{4}$$

where f_{I_i} is the original image, I_i is the actual image which we are trying to calculate at a specific scale, and g is the diffusivity function. We extend this to our shape image. Particularly, our approach sums up the diffusivity of each channel to a common diffusivity. This may be regarded as collecting the contrast information of all channels. Thus, for a shape image $I = (I_1, I_2, ..., I_m)^\top$, the vector diffusion is performed by

$$\frac{\partial I_i}{\partial t} = div\left(g\Big(\sum_{j=1}^{m}|\nabla f_{I_j}|^2\Big)\nabla I_i\right) \qquad (i = 1, ..., m). \tag{5}$$

The numerical solution for Eq. 5 can be referred to Perona-Malik's model [14]. By solving the equations in Eq. 5, what we obtain is,

$$\begin{pmatrix} \mathbf{I}_1 \\ \mathbf{I}_2 \\ \vdots \\ \mathbf{I}_m \end{pmatrix} = \begin{pmatrix} \mathbf{I}^{t_0} & \mathbf{I}^{t_1} & \dots & \mathbf{I}^{t_n} \end{pmatrix} = \begin{pmatrix} {I_1}^{t_0} & {I_1}^{t_1} & \dots & {I_1}^{t_n} \\ {I_2}^{t_0} & {I_2}^{t_1} & \dots & {I_2}^{t_n} \\ \vdots & \vdots & \ddots & \\ {I_m}^{t_0} & {I_m}^{t_1} & \dots & {I_m}^{t_n} \end{pmatrix}, \tag{6}$$

which is a sequence of vector image with t as the scale in matrix format, i.e., each row of the matrix is the sequence images of a specific channel with t as the scale, and each column of the matrix is the vector image at a specific scale t.

3.2 Extrema detection

Eq.5 satisfies the maximum principle, which states that all the maxima of the solution of the equation in space and time belong to the initial condition (the original image) [14].

Instead of calculating the Difference of Gaussian (DoG) as in the SIFT algorithm [9], we compute the Difference of Diffusion (DoD) using the following vector-based equation,

$$\mathbf{DoD}^{t_i} = \mathbf{I}^{t_{i+1}} - \mathbf{I}^{t_i} \qquad (i = 1, ..., n-1). \tag{7}$$

Once DoD vector images have been obtained, keypoints are identified as local minima/maxima of the DoD images across scales. This is done by comparing each pixel in the DoD images to its eight neighbors at the same scale and nine corresponding neighboring pixels in each of the neighboring scales. If the pixel value is the maximum or minimum among all compared pixels, it is selected as a candidate keypoint. This algorithm is carried out through all the channels of the vector image: $\mathbf{DoD}_i$, (i=1,...,m). The maximum and minimum which are found in every channel will be considered as the interest points, which are very robust points suitable for matching purpose.

3.3 Descriptor Construction

After localizing the interest points, feature descriptors are built to characterize these points. These descriptors should contain the necessary distinct information for their corresponding interest points. Like SIFT, the local gradient-orientation histograms for the same-scale neighboring pixels of an interest point are used as the key entries of the descriptor. The descriptor is calculated channel by channel. The feature descriptor is computed as a set of orientation histograms on (4×4) pixel neighborhoods. The orientation histograms are relative to the keypoint orientation. Histograms contain 8 bins each, and each descriptor contains an array of 4×4 histograms around the keypoint. This leads it to be a feature vector with $4 \times 4 \times 8 = 128$ elements.

In our case, m channel vector image, for a specific keypoint, it has m descriptors which we combine as a vector, $\mathbf{des} = [\mathbf{des}_1, \mathbf{des}_2, \ldots, \mathbf{des}_m]^\top$, where m is the dimension of the vector images. Hence, the descriptor **des** of the vector image is a $m \times 128$ dimension-based vector. This descriptor will be used for matching, and all the descriptors from all the interest points form a feature descriptor database.

4 Shape Matching and Registration

In our framework, shape matching is to match the interest points in different objects since the interest points are considered as most reliable feature points presented in surfaces. In Section 3, we have explained how the feature descriptor for the each interest points is calculated and how a feature descriptor database is formed.

Descriptor matching is performed for the constructed local descriptors by comparing the distance of the closest neighbor (DIS_{CN}) to that of the second-closest neighbor (DIS_{SCN}). The distance of two descriptors, **des**1 and **des**2

which are m dimension vectors, is calculated by,

$$DIS = \sum_{i=1}^{m} \parallel \mathbf{des1}_i - \mathbf{des2}_i \parallel . \tag{8}$$

Once the DIS_{CN} and the DIS_{SCN} are found, the DIS_{CN} and the DIS_{SCN} are compared to decide whether they are matched or not. The judge function for the comparison is

$$threshold \times DIS_{CN} <= DIS_{SCN}. \tag{9}$$

If this inequation holds, the points are matched; Otherwise, they are not matched.

Since the 3D data is usually coarsely aligned through affine transformations, we can use the uniform subdivision grid to speed up the matching. The Euclidean distance bound (ED) of two potentially matched interest points is calculated and can be used in efficiently finding the the closest and the second-closest neighbors grids. After finding all the matched points, registration can be easily achieved using thin-plate splines deformation with the matched points as point constraints.

5 Experiments

To evaluate the proposed approach, we have applied our algorithm on matching human neocortex surfaces, which were extracted from high-resolution SPGR scans. For the neocortex surface, a genus zero surface, conformal mapping is performed to transfer it to a sphere. The reparameterization which we used to map the sphere to a 2D domain is

$$\sigma(\theta, \varphi) = (\cos\theta\cos\varphi, \cos\theta\sin\varphi, \sin\varphi),$$

where θ and φ are the row and column in the 2D domain image. Based on this planar parameter domain, we construct the shape image by assign the H and λ values to each corresponding image pixel. Hence, the shape image is a two dimensional vector image $[I_1, I_2]^\top$, where $I_1 = H$ and $I_1 = \lambda$.

After the shape image is generated, we use the vector diffusion to create the **DoD** matrix, of which each row is a sequence of images in different scale in the each channel.

$$\begin{pmatrix} \mathbf{DoD}_1 \\ \mathbf{DoD}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{DoD}^{t_0} & \mathbf{DoD}^{t_1} & \dots & \mathbf{DoD}^{t_{n-1}} \end{pmatrix} = \begin{pmatrix} {DoD_1}^{t_0} & {DoD_1}^{t_1} & \dots & {DoD_1}^{t_{n-1}} \\ {DoD_2}^{t_0} & {DoD_2}^{t_1} & \dots & {DoD_2}^{t_{n-1}} \end{pmatrix} \tag{10}$$

By finding the maximum and minimum in each row of the matrix as the interest points, the descriptor is computed for each point. Each descriptor is a 2*128=256 dimension vector and all these descriptors form the descriptor database. The matching algorithm is performed to find the matched points which satisfies the inequation (9). Fig. 2 shows the matching result of two different subjects. To

clearly see the matched points between the two shape images, only 10% of the points are shown in the figure. We have conducted the evaluation on intersubject matching of 20 brain surfaces. The results are evaluated quantitatively in terms of major landmark overlap (the central sulcus, the sylvian fissure, and the posterior sulcus). Table.1 shows the average mismatch distance of the above three major landmark regions for 20 surfaces. The average mismatch distance error of the total 20 different subjects is only 4.08 mm.

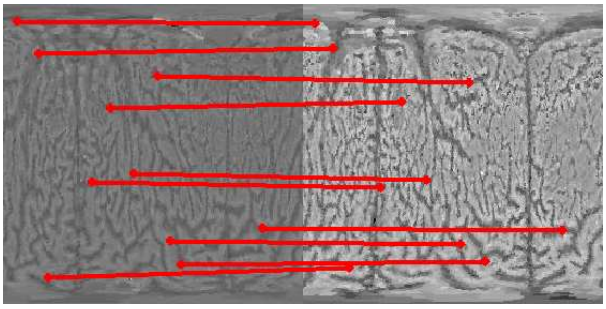

Fig. 2. Matching of two different subjects' brain surfaces. 10% of matched points are shown above using the linked lines.

Subject Number	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10
Average Mismatch distance(mm)	4.32	4.43	4.05	3.97	4.03	4.07	4.38	4.09	4.25	4.11
Subject Number	s11	s12	s13	s14	s15	s16	s17	s18	s19	s20
Average Mismatch distance(mm)	3.92	3.67	3.98	4.12	4.26	3.86	3.78	4.15	4.22	3.95

Table 1. Mismatch distance in terms of major landmark overlap.

6 Conclusion

In this paper, we have presented a novel and efficient surface matching framework through diffusion over shape images. Our framework converts 3D surface matching problem to a 2D shape image matching problem based on stable keypoints and local features. The most robust features facilitate a reliable registration as observed from our experiments. This is particularly important for inter-subject surfaces from medical scans of different subjects since these surfaces exhibit the inherited physiological variances among subjects. The 2D representation also allows easier statistical analysis of other modality features directly on the matched 2D domain. Future work will be focused on multimodality data integration over

the 2D domain to support more accurate localization of brain disorder regions from population study.

Acknowledgments: This work is supported in part by the research grants awarded to Dr. Jing Hua, including the National Science Foundation grants IIS-0713315 and CNS-0751045, the National Institute of Health grants 1R01NS058802-01A2 and 2R01NS041922-05A1, the Michigan Technology Tri-Corridor grants MTTC05-135/GR686 and MTTC05-154/GR705, and the Michigan 21st Century Jobs Funds 06-1-P1-0193.

References

1. B. Vemuri, A. Mitiche, and J. Aggarwal, Curvature-based representation of objects from range data, Image and Vision Computing, no. 4, pp. 107–114 (1986)
2. S. Ruiz-Correa, L. Shapiro, and M. Meila, A new paradigm for recognizing 3-D object shapes from range data, in ICCV, pp. 1126–1133 (2003)
3. R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, Shape distributions, ACM Transactions on Graphics, vol. 22, pp. 807–832 (2002)
4. G. Zou, J. Hua, and O. Muzik, Non-rigid Surface Registration Using Spherical Thin-plate Splines, In MICCAI, pp. 367-374 (2007)
5. M. S. Floater and K. Hormann, Surface parameterization: a tutorial and survey, Advances in Multiresolution for Geometric Modelling, pp. 157–186, Springer (2004)
6. R. Schoen and S. T. Yau, Lectures on Harmonic Maps. International Press, Harvard University, Cambridge MA (1997)
7. E. Sharon and D. Mumford, 2D-shape analysis using conformal mapping, in Proc. CVPR, pp. II: 350–357 (2004)
8. B. Levy, S. Petitjean, N. Ray, and J. Maillot, Least squares conformal maps for automatic texture atlas generation, in SIGGRAPH, pp. 362–371 (2002)
9. D. G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110 (2004)
10. V. Athitsos, J. Alon, S. Sclaroff, and G. Kollios, Boostmap: a method for efficient approximate similarity rankings, in Proc. CVPR, pp. pages II: 268–275 (2004)
11. G. Mori, S. Belongie, and J.Malik, Efficient shape matching using shape contexts, IEEE Transactions on Pattern Analysis and Machine Intelligence, no. 27(11), pp. 1832–1837 (2005)
12. D. Zhang and M. Hebert, Harmonic maps and their applications in surface matching, in Proc. CVPR, pp. 524–530 (1999)
13. X. Gu, Y. Wang, T. Chan, P. Thompson, and S. T. Yau, Genus zero surface conformal mapping and its application to brain surface mapping, IEEE Trans. Med. Imaging, vol. 23, no. 8, pp. 949–958 (2004)
14. P. Perona and J. Malik, Scale space and edge detection using anisotropic diffusion, IEEE Transactions on pattern analysis and machine intelligence, vol. 12, no. 7, pp. 629–639 (1990)
15. X. Gu, Y. Wang and S. Yau, Geometric Compression using Riemann surface structure, Communications in Information and Systems, vol. 3, no. 3, pp. 171–182 (2004)
16. L. Lui, Y. Wang, T. Chan and P. Thompson, A Landmark-Based Brain Conformal Parametrization with Automatic Landmark Tracking Technique, in MICCAI, pp. 308–315 (2006).