

Geodesic Distance-weighted Shape Vector Image Diffusion

Jing Hua, *Member, IEEE*, Zhaoqiang Lai, Ming Dong, Xianfeng Gu, and Hong Qin

Abstract—This paper presents a novel and efficient surface matching and visualization framework through the geodesic distance-weighted shape vector image diffusion. Based on conformal geometry, our approach can uniquely map a 3D surface to a canonical rectangular domain and encode the shape characteristics (e.g., mean curvatures and conformal factors) of the surface in the 2D domain to construct a *geodesic distance-weighted shape vector image*, where the distances between sampling pixels are not uniform but the actual geodesic distances on the manifold. Through the novel geodesic distance-weighted shape vector image diffusion presented in this paper, we can create a multiscale diffusion space, in which the cross-scale extrema can be detected as the robust geometric features for the matching and registration of surfaces. Therefore, statistical analysis and visualization of surface properties across subjects become readily available. The experiments on scanned surface models show that our method is very robust for feature extraction and surface matching even under noise and resolution change. We have also applied the framework on the real 3D human neocortical surfaces, and demonstrated the excellent performance of our approach in statistical analysis and integrated visualization of the multimodality volumetric data over the shape vector image.

Index Terms—Surface Matching, Shape Vector Image, Multiscale Diffusion, Visualization.

1 INTRODUCTION

To date, 3D surface matching and data visualization is still a very challenging research problem in many visual data processing and analysis fields. This is partly because surfaces may have highly flexible free-form shape characteristics which are difficult to be captured and used for matching and registration purposes, especially under noisy conditions. One of the fundamental issues in surface matching is the shape representation scheme. In recent years, different shape representations, such as curvature-based representations [25], regional point representations [20], shape distributions [18], etc., have been proposed for 3D surface matching. These representations and matching algorithms are directly dependent on surface meshes, and therefore, are not robust and do not perform well for inter-subject surface registration under such circumstances as noises, resolution changes, and variances.

Recently, converting the problem of surface analysis to canonical domains has gained increasing attention [6, 28]. A family of geometric maps has recently been adopted to create a shape image that does not suffer from aforementioned problems such as noise, etc. This simplifies the shape-analysis problem to a 2D image-analysis problem. When constructing shape images, geometric mapping provides a viable solution to map a 3D surface to a 2D domain and encode the shape information of the surface in the 2D image. According to the conformal geometry theory, each 3D shape can be mapped to a 2D domain through a global optimization and the resulting map is a diffeomorphism [23, 8]. Consequently, the 3D shape analysis problem can be simplified to a 2D image analysis problem of the conformal geometric maps. These maps are stable, insensitive to resolution changes and robust to occlusion and noises. Therefore, accurate and efficient 3D shape analysis may be achieved through 2D image analysis. Along this direction, Wang et al. [26] presented a global correlation-based method for matching least-squares conformal maps, which works for the recognition purpose, but is not effective for exact matching and registration of different surfaces. In general, there is a lack of analysis tools for processing of this type of special shape images, especially based on local features.

In this paper, we present a novel and efficient surface matching and visualization framework based on the geodesic distance-weighted shape vector image diffusion. Firstly, our framework conformally maps a to-be-analyzed surface to a canonical 2D domain. The surface curvatures and conformal factors are then interpolated and encoded into the rectangular 2D domain, which we call *shape vector image* in this paper. As the surface curvatures and conformal factors can uniquely define the surface, the vector image composed by curvature and conformal factors can serve as the shape signature. In the shape vector image, the distances between sampling pixels are the actual geodesic distances on the manifold. Since the mapping is independent of the mesh resolution, the resulting shape vector image is robust to different samplings of the surface. In order to extract the most robust and salient features to abstract the shape vector image, we propose to create a multiscale vector-valued diffusion space through our novel geodesic distance-weighted shape vector image diffusion. As a result, analysis of the shape vector image in its diffusion space is similar to the direct diffusion analysis of the 3D model. A valuable point here is that our computation is executed in a regular 2D domain, which is much simpler than in the 3D domain.

In the diffusion space, we can then extract distinctive features used for matching and analysis. A rich set of scale-aware features can be extracted from the diffusion space representation. Similar to the feature extraction technique in [10], our approach detects the extrema across the scales as keypoints. We then calculate the orientation histograms around the keypoints as feature descriptors, which provide distinctive bases for representing the 3D geometry of the original shape. These scale-aware geometric features can directly be used for robust matching and registration against the noises and distortions. Therefore, statistical analysis and visualization of surface properties across subjects become readily available. This is important for many real-world applications. For example, it is very useful for processing inter-subject brain surfaces from medical scans of different subjects since these surfaces exhibit the inherited physiological variances among subjects. We have conducted extensive experiments on scanned real-world surface models and real 3D human neocortical surfaces, through which we demonstrate the excellent performance of our approach in surface matching and registration, statistical analysis, and integrated visualization of the multimodality volumetric data over the shape vector image.

Our contributions in this paper can be summarized as follows:

-
- J. Hua, Z. Lai and M. Dong are with Wayne State University. Email: {jinghua|kevinlai|mdong}@wayne.edu
 - X. Gu and H. Qin are with Stony Brook University. Email: {gu|qin}@cs.sunysb.edu.
 - Correspondence to Prof. J. Hua.

Manuscript received 31 March 2008; accepted 1 August 2008; posted online 19 October 2008; mailed on 13 October 2008.

For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

- We present a novel geodesic distance-weighted shape vector image diffusion method. It allows the construction of the multiscale diffusion space and scale-space processing of the shape vector image composed of the intrinsic geometric characteristics, e.g., mean curvatures and conformal factors.

- By detecting the cross-scale extrema in the diffusion space, we can create a set of distinctive scale-aware geometric features. These robust features are well suitable for surface matching and registration.
- We apply the framework to the matching of scanned face data and multimodality brain surface data, which demonstrates the excellent performance of the proposed geodesic distance-weighted shape vector image diffusion method.
- With thin-plate spline deformation of matched features among different subjects, we develop a framework for comparative analysis and visualization of cross-subject multimodality neuroimaging data. It is a powerful tool for analyzing surfaces with multidimensional textures.

1.1 Prior work

Geometric maps have been widely applied in the fields of computer graphics and computer vision. Zhang and Hebert used harmonic maps to construct scalar shape images to match the surfaces [29]. In [27], harmonic maps were used to track dynamic 3D surface. However, calculating harmonic maps needs to identify the surface boundary and create the boundary mapping from 3D surfaces to the 2D domain, which becomes unreliable when there are noises and occlusions in the 3D original data. Since the interior feature points are often more robust in the 3D original data, conformal maps, which do not need boundary information, can be a natural choice to overcome the difficulty. Using several feature constraints instead of the boundary condition, conformal maps have many appealing properties [26]. For example, if the parameterization is conformal, the surface can be uniquely determined by the mean curvatures with area stretching factors defined on the parametric domain. In [7], genus zero surface conformal mapping was discussed and it was adopted for the brain surface mapping [31]. The conformal parameterization can be uniquely determined by two corresponding points. Conformal parameterization depends on the geometry itself, not the triangulation of the surfaces. Hence, conformal mapping is a viable solution for 3D shape image construction. This motivated us to encode a shape vector image using conformal mapping for surface representation.

After mapping the 3D surface to the 2D image plane, extracting features from the shape image becomes one of the most important tasks. Generally speaking, the analysis of 2D image is a better understood problem [10, 2, 13]. However, the analysis of the shape image that integrates geometric and appearance information has its own special challenges mainly due to the non-uniform sampling and different pixel properties (i.e., geometric characteristics instead of grey-scale intensities). Hence, conventional image analysis techniques may not work well. For example, line segments [5], groupings of edges [9, 15], and regions [3] have all been proposed as features for image matching. Nevertheless, these feature extraction techniques are not reliable for shape images due to the aforementioned special characteristics of shape images and they only work well under certain circumstances. Recently, there has been a great deal of research work on developing more reliable features for conventional images. One of them is to take advantage of corner detectors. Zhang et al. [30] used the Harris corner detector to identify feature locations. Harris corner detector was also used by Schmid and Mohr [22] to identify interest points for the object recognition problem. Other approaches have been proposed for appearance-based matching, including eigenspace analysis [14], color histograms [24], and receptive field histograms [21]. However, it is difficult to extend them to match variant inter-subject images because of their more global features. Ohba and Ikeuchi [17] successfully applied the eigenspace approach to cluttered images by using many small local eigen-windows, but this is very time consuming and not practical. Most recently, SIFT, scale-invariant features, has been proposed [10] for 2D regular images and proven to be the most robust among many other local invariant feature descriptors with respect to different geometrical changes [12, 1]. SIFT is based on the scale space theory. Scale-space processing usually convolves an image with

Gaussian filters, generating a sequence of images, and then the difference of successive Gaussian-blurred images are calculated to create the Difference of Gaussians (DoG) for further analysis. Since scale-space theory was mainly developed for gray-scale images, the usage of the scale-space processing on the special shape images or shape vector images are under-explored and its performance remains unknown.

2 CONFORMAL SHAPE VECTOR IMAGE CONSTRUCTION

A good shape image should be able to fully represent the geometric characteristics of a given surface, and thus serves as a domain for indexing other heterogenous attributes. Thus, a 3D surface can be converted to a multidimensional vector image for effective processing. We employ conformal 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 $f \circ ds_m^2 = \lambda ds_n^2$, where λ is called the conformal factor, ds_m^2 and ds_n^2 are the first fundamental form on M and N . If M and N are surfaces, a diffeomorphism $f: M \rightarrow N$ is said to be conformal if, whenever f takes two intersecting curves γ_m and $\tilde{\gamma}_m$ on M to curves γ_n and $\tilde{\gamma}_n$ on N , the angle of intersection of γ_m and $\tilde{\gamma}_m$ is equal to the angle of intersection of γ_n and $\tilde{\gamma}_n$. In short, f is conformal if it preserves angles.

In order to match 3D shapes accurately and efficiently, we develop a 2D representation, shape vector image, using conformal mapping. 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 . Based on the shape attributes computed at each vertex of M , attribute values can be interpolated and computed for each pixel of the conformal shape vector image. In practice, we compute the conformal parameterization by a nonlinear optimization method carried out in the tangential space of a sphere as proposed in [7].

Suppose K denotes the simplicial complex and there is a piecewise linear embedding $\vec{l}: |K| \rightarrow \mathbf{R}^3$. Then a triangular mesh can be represented as (K, \vec{l}) . For the purpose of implementation, surfaces are usually approximated by triangular meshes. We use p, q to denote the vertices and $\{p, q\}$ to denote the edge spanned between p and q . The surface and its parametric domain are modeled as piecewise linear functions \vec{f} and \vec{g} in accordance with (K, \vec{l}) , respectively. The mean curvature at vertex q is estimated as in [11] by

$$H(q) = \frac{1}{4A} \left\| \sum_{p \in N_1(q)} (\cot \alpha_{p,q} + \cot \beta_{p,q}) (\vec{f}(q) - \vec{f}(p)) \right\|^2, \quad (1)$$

where $\alpha_{p,q}$ and $\beta_{p,q}$ are the two opposite angles of edge p, q in the two triangles sharing this edge, and $N_1(q)$ is the set of 1-ring neighbor vertices of vertex q . A is the area of the associated surface patch (so-called finite volume in Mechanics), which is given by

$$A(q) = \frac{1}{8} \sum_{p \in N_1(q)} (\cot \alpha_{p,q} + \cot \beta_{p,q}) \|\vec{f}(q) - \vec{f}(p)\|^2, \quad (2)$$

under the condition that the triangles in the 1-ring neighbors are non-obtuse. In case of obtuse triangles, refer to [11] for solutions. Following this path, we define the discrete conformal factor operator as

$$\lambda(q) = \frac{A_g(q)}{A_f(q)}, \quad (3)$$

where the $A_f(q)$ and $A_g(q)$ are the averaging areas for each homotopic vertex q on surface \vec{f} and \vec{g} , respectively.

As conformal surface representation $S(u, v)$ is parameterized by conformal parameters (u, v) on a domain D , the conformal factor function, $\lambda(u, v)$, and mean curvature function, $H(u, v)$, defined on D satisfy the Gauss and Codazzi equation. If $\lambda(u, v)$ and $H(u, v)$ are given together with the boundary conditions, $S(u, v)$ can be uniquely reconstructed. Since the mean curvature and the conformal factor are two

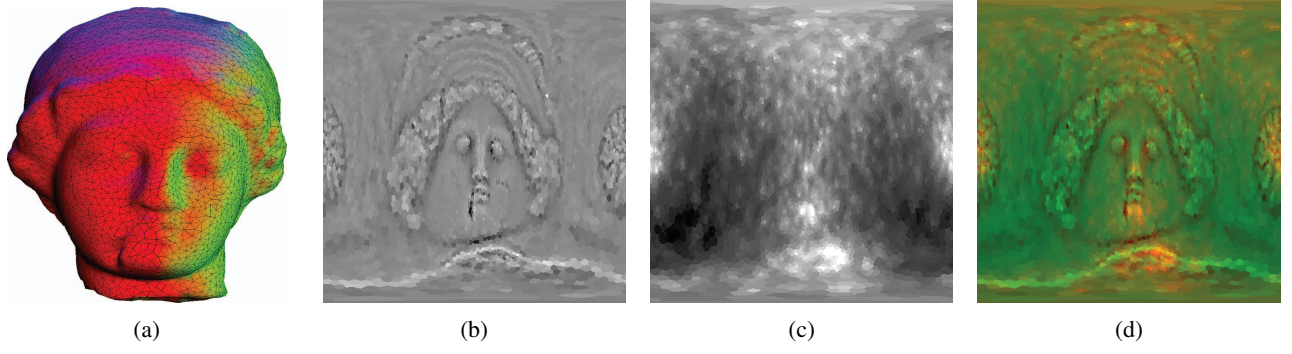


Fig. 1. Shape Vector Image. (a) shows the Igea (5002 vertices) surface and mesh; (b) and (c) show the mean curvature channel and conformal factor channel of the shape vector image representation of the Igea model; (d) is the composite shape vector image including both channels.

important attributes, we assign these two attributes to $\vec{I}_c(u, v)$ to construct a vector image I , where the pixel attributes are represented by a vector $[H, \lambda]^\top$. Other features such as normal and texture can be considered if necessary. We use barycentric interpolation for sampling $\vec{I}_c(u, v)$ to I . Fig. 1 shows the Igea surface model with 5002 vertices (Fig. 1(a)) and its corresponding mean curvature channel (Fig. 1(b)) and conformal factor channel (Fig. 1(c)). The composite shape vector image is shown in Fig. 1(d).

3 GEODESIC DISTANCE-WEIGHTED DIFFUSION

Since a surface can be represented as a unique shape vector image composed of conformal factors and curvatures, many algorithms suitable for image computing may be used for feature extraction from this type of images. For the purpose of matching and visualization of cross-subject data, the main task is to find the stable keypoints or regions and their local image features for alignment. Since the shape vector image representation that we propose consists of the mean curvature and area distortion, it provides important signature of the local geometry, which is transformation invariant and suitable for shape matching. This section describes a novel diffusion-based algorithm to extract distinctive features from the shape vector images. Through the geodesic distance-weighted shape vector image diffusion, we can identify the robust keypoints and their scales from the computed diffusion extrema, which are suitable for the matching purpose.

3.1 Shape vector image diffusion and diffusion space

Our shape vector image is a multichannel image. The simplest way to perform the diffusion filtering of the shape vector image is to deal with each channel separately and independently. However, this method leads to an undesirable effect that edges may be formed at different locations for each channel since the curvature and conformal factor channels must take effect simultaneously in order to accurately determine the local geometry. In our framework, we employ a diffusivity g which combines information from all channels. For a vector image $I = (I_1, I_2, \dots, I_m)^\top$, the diffusion is performed by

$$\frac{\partial I_k}{\partial t} = \text{div}(g \nabla I_k) \quad (k = 1, \dots, m), \quad (4)$$

where div indicates the divergence operator, ∇ is the gradient operator, and $g(x) = \frac{1}{\sqrt{1+(\frac{x}{l})^2}}$ (l is a constant). For the case in which g is a constant for a specific channel I_k , it reduces to the isotropic heat diffusion equation, $\frac{\partial I_k}{\partial t} = c \Delta I_k$, where Δ is the Laplacian operators and the solution is Gaussian smoothing. However, Gaussian smoothing has a typical disadvantage, especially for the shape vector image. Gaussian smoothing does not only reduce noise, but also blurs important geometric features such as sharp edges, hence making them harder to be identified.

To solve the problem, we propose to perform geodesic distance-weighted inhomogeneous linear diffusion of the shape vector image,

$$\frac{\partial I_k}{\partial t} = \text{div} \left(g \left(\|\nabla f_k\| \right) \nabla I_k \right), \quad (5)$$

where I_k is the actual diffused image, f_k is the original image and g is the diffusivity function. For a specific channel $P = I_k$, the numerical solution for Equation 5 can be computed, similar to the 4-nearest-neighbors discretization [19], as follows,

$$P_{i,j}^{t+1} = P_{i,j}^t + \rho [c_N \cdot \nabla_N P + c_S \cdot \nabla_S P + c_E \cdot \nabla_E P + c_W \cdot \nabla_W P], \quad (6)$$

where $0 \leq \rho \leq 1/4$, N, S, E, W are the subscripts for the North, South, East, and West, t is the scale, i and j are the indices of the image pixel. Since the shape vector image encodes geodesic distance information, the symbol ∇ is defined as follows with the consideration of geodesic distances:

$$\begin{aligned} \nabla_N P_{i,j} &= \frac{P_{i-1,j} - P_{i,j}}{\text{GeoD}[N]}, & \nabla_S P_{i,j} &= \frac{P_{i+1,j} - P_{i,j}}{\text{GeoD}[S]}, \\ \nabla_E P_{i,j} &= \frac{P_{i,j+1} - P_{i,j}}{\text{GeoD}[E]}, & \nabla_W P_{i,j} &= \frac{P_{i,j-1} - P_{i,j}}{\text{GeoD}[W]}, \end{aligned} \quad (7)$$

where

$\text{GeoD}[N] = \text{GeoD}([i-1, j], [i, j])$, $\text{GeoD}[S] = \text{GeoD}([i+1, j], [i, j])$,
 $\text{GeoD}[E] = \text{GeoD}([i, j+1], [i, j])$, $\text{GeoD}[W] = \text{GeoD}([i, j-1], [i, j])$,
and $\text{GeoD}([pixelA], [pixelB])$ is the normalized geodesic distance between the $pixelA$ and $pixelB$ on the manifold, which is normalized by dividing the averaged GeoD over the image. And the c is defined as:

$$\begin{aligned} c_{N,i,j} &= g(|\nabla_N f_{i,j}|), & c_{S,i,j} &= g(|\nabla_S f_{i,j}|), \\ c_{E,i,j} &= g(|\nabla_E f_{i,j}|), & c_{W,i,j} &= g(|\nabla_W f_{i,j}|), \end{aligned}$$

where $\nabla_N f_{i,j}$, $\nabla_S f_{i,j}$, $\nabla_E f_{i,j}$ and $\nabla_W f_{i,j}$ are computed by Equation 7. Therefore, the final numerical solution is

$$P_{i,j}^{t+1} = P_{i,j}^t + \rho \left[\frac{c_N \cdot \nabla_N P + c_S \cdot \nabla_S P}{\text{GeoD}_{NS}} + \frac{c_E \cdot \nabla_E P + c_W \cdot \nabla_W P}{\text{GeoD}_{EW}} \right], \quad (8)$$

where

$$\begin{aligned} \text{GeoD}_{NS} &= \frac{\text{GeoD}[N] + \text{GeoD}[S]}{2}, \\ \text{GeoD}_{EW} &= \frac{\text{GeoD}[E] + \text{GeoD}[W]}{2}. \end{aligned}$$

Using this numerical solution, we can construct a discrete geometric diffusion space which encodes the surface geometric information.

Particularly for our shape vector image scheme, 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 vector image $I = (I_1, I_2, \dots, I_m)^\top$, the vector diffusion is performed by

$$\frac{\partial I_k}{\partial t} = \text{div} \left(g \left(\sum_{n=1}^m \|\nabla f_{I_n}\| \right) \nabla I_k \right) \quad (k = 1, \dots, m). \quad (9)$$

By solving Equations 9, we obtain the diffusion space,

$$\left(\mathbf{I}^{t_0} \quad \mathbf{I}^{t_1} \quad \dots \quad \mathbf{I}^{t_n} \right) = \begin{pmatrix} I_1^{t_0} & I_1^{t_1} & \dots & I_1^{t_n} \\ I_2^{t_0} & I_2^{t_1} & \dots & I_2^{t_n} \\ \vdots & \vdots & \ddots & \vdots \\ I_m^{t_0} & I_m^{t_1} & \dots & I_m^{t_n} \end{pmatrix}, \quad (10)$$

which is a sequence of shape vector images with t as the scale in a 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 . Fig. 2 shows the diffused shape vector images of the Igea model with increasing diffusion scale t .

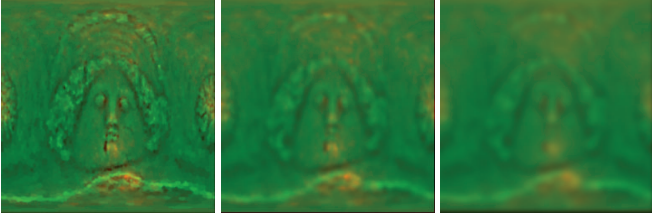


Fig. 2. The diffused shape vector images, consisting both curvature and conformal factor channels, of the Igea model at different diffusion scales, t , computed by the geodesic distance-weighted diffusion.

3.2 Properties of diffusion space

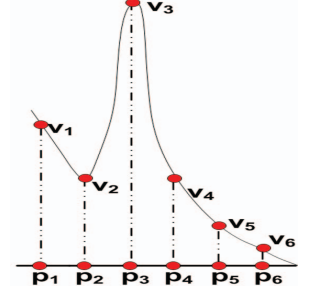
Suppose a mapping $f: M \rightarrow I$ and the diffusion $D: I \rightarrow S$, where M is a 3D surface, I is the shape vector image and S is the diffusion space. In this section, we firstly show that D satisfies the criteria of the multiple scale descriptions. Secondly, we show that $f \circ D$, together, creates the diffusion space appropriate for scale-space processing of the 3D surface geometry.

The diffusion space construction is to create a multiscale ‘‘semantically meaningful’’ descriptions of images. As we know, a scale-space representation must have the property that no spurious detail will be generated passing from finer to coarser scales. This is so-called causality, which means any feature at a coarse level of resolution is required to possess a ‘‘cause’’ at a finer level of resolution although the reverse need not be true. The causality criterion can be established by showing the used diffusion equation satisfies the maximum principle, that is to say, all the maxima of the solution of the equation in space and time belong to the initial condition (i.e., the original image). A proof of the maximum principle for the diffusion equation can be found in [19, 16]. Therefore, for the diffusion D , the satisfaction of the maximum principle leads to the satisfaction of the causality for the diffusion space. Consequently, D satisfies the criteria of the multiple scale descriptions.

For our shape vector image, we use the geodesic distance-weighted method, in which the distance can be retrieved by the mapping, f , for the computation of the diffusion to construct a geometric diffusion space. Therefore, constructing and analyzing the geometric diffusion space is similar to analyzing a direct diffusion space of the 3D surface. As a result, $f \circ D$ is able to construct a multiple scale space and multiscale descriptions for the 3D surface.

The geodesic distance-weighted anisotropic diffusion has the advantages of preserving and identifying true features as well as preventing dislocated false features in the diffusion space when taking the actual geodesic distance as a-priori information. The adjacent figure shows an illustrative one-dimensional example, where

the curve is the actual shape object and the line segment is the 1D shape ‘‘image’’. The sampling vertices (from v_1 to v_6) on the curve are mapped to the pixels (from p_1 to p_6) on the 1D shape ‘‘image’’ shown at the bottom. As we can see, the sampling vertices of the curve, v_2 and v_3 , have relatively high curvature values while other sampling vertices have low curvatures. Without considering the geodesic distance among the pixels, p_4 ’s curvature value will be increased while moving from finer to coarser scales. For example, at the immediately next level, p_3 will boost the diffusion of p_4 since its distance to p_4 is considered the same as the distance between p_4 and p_5 (a unit length). The situation will get worse when diffusing further since p_2 will also dramatically affect the diffusion of p_4 . However, v_2 is far away from the diffusion vertex, v_4 . In fact, p_2 is not suppose to have significant influence on p_4 . On the other side, it is also difficult to preserve the high curvature values for p_2 and p_3 . Therefore, the procedure will dislocate keypoints when moving from finer to coarser scales. So the keypoints detected at a coarse scale do not give the correct location in the original image, which will result in instability and incorrect matching. The situation is much improved when considering geodesic distances in the diffusion.



3.3 Keypoint-based shape descriptors

The maximum principle states that all the maxima of the solution of the equation in space and time belong to the initial condition (i.e., the original image). Therefore, we propose to detect the extrema across the diffusion space as our keypoints since they are most robust points at the specific scales which are able to represent collectively the original image. We sample the diffusion space by computing a sequence of diffused shape vector images at discrete scales, t . For each diffusion scale, we use Equation 9 to calculate its diffused images which can be expressed in a matrix form like Equation 10.

In order to extract the cross-scale extrema, we compute the Difference of Diffusion (DoD) using the following vector-based equation,

$$\mathbf{DoD}^t = \mathbf{I}^{t+1} - \mathbf{I}^t \quad (i = 0, \dots, n-1). \quad (11)$$

Consequently, we can obtain,

$$\begin{pmatrix} \mathbf{DoD}_1 \\ \mathbf{DoD}_2 \\ \vdots \\ \mathbf{DoD}_m \end{pmatrix} = \begin{pmatrix} DoD_1^{t_0} & DoD_1^{t_1} & \dots & DoD_1^{t_{n-1}} \\ DoD_2^{t_0} & DoD_2^{t_1} & \dots & DoD_2^{t_{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ DoD_m^{t_0} & DoD_m^{t_1} & \dots & DoD_m^{t_{n-1}} \end{pmatrix}. \quad (12)$$

Fig. 3(c-e), (g-i), and (k-m) show the computed intermediate curvature channel images of the DoDs across scales.

Once DoD vector images have been obtained, keypoints are identified as local minima/maxima of the DoD images across scales t . For each channel \mathbf{DoD}_i , it is done by comparing each pixel in the $\mathbf{DoD}_i^{t_j}$ images to its eight neighbors at the same scale t_j and nine corresponding neighboring pixels in each of the neighboring scales t_{j-1} and t_{j+1} . If the pixel value is the maximum or minimum among all compared pixels, it is selected as a keypoint. This algorithm is carried out through all the channels of the vector images: \mathbf{DoD}_i , ($i = 1, \dots, m$). The maxima and minima found in all the channels will be considered as the keypoints. Fig. 3(a) and (b) show all the detected keypoints on the Igea model. Fig. 3(c-e), (g-i) and (k-m) show the detected extrema (shown with points) on the corresponding DoDs at different scales. Fig. 3(f), (j) and (n) show the scale sizes, at which the extrema in (e), (i) and (m), respectively, are detected, with the corresponding sizes of circles. One valuable point is that the detected keypoints have the associated scales computed by the algorithm, which are very important to construct scale-aware feature descriptors.

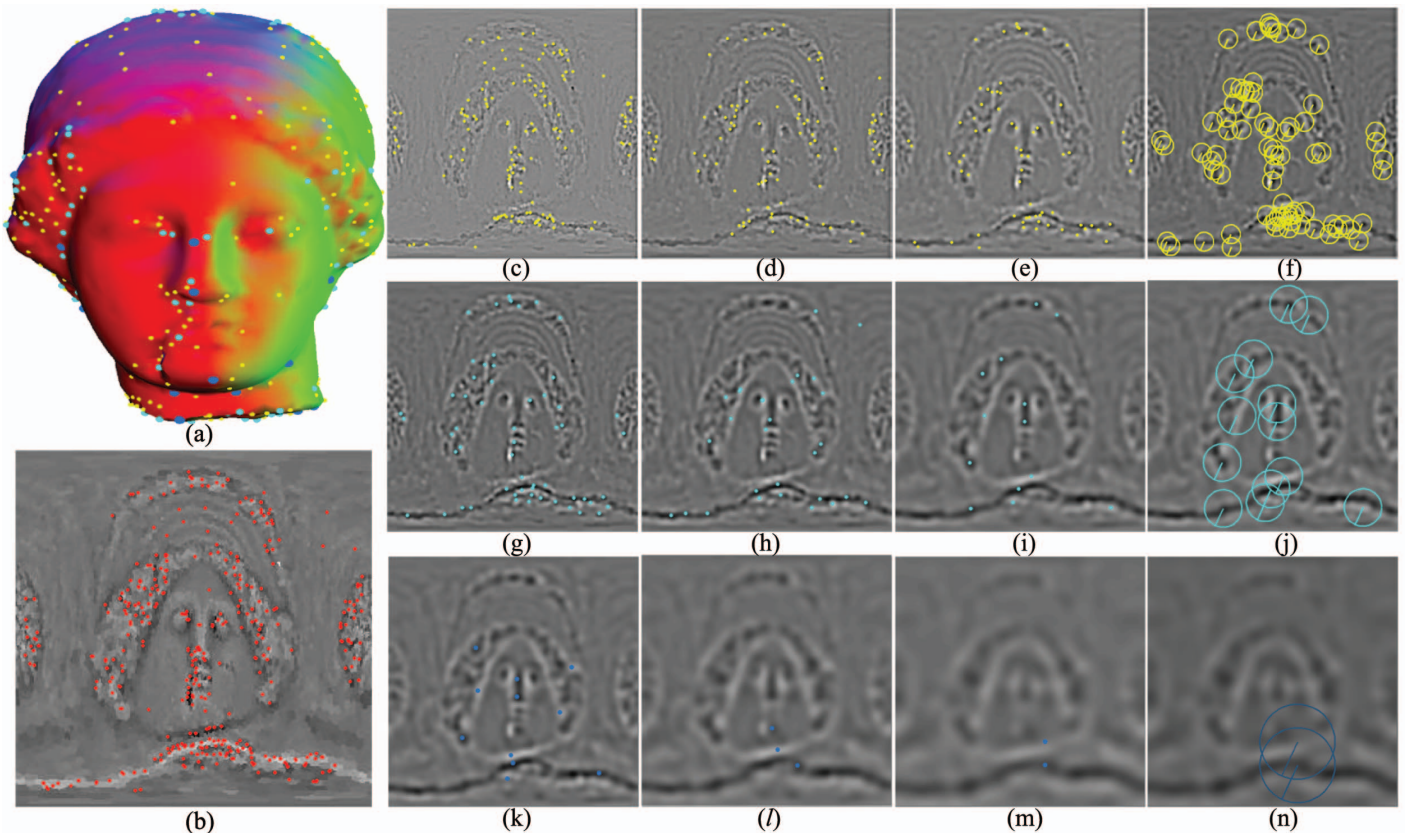


Fig. 3. Keypoint detection in the diffusion space. (a) The Igea model with all the detected keypoints at different scales indicated by the points of different colors and sizes. (b) All the detected keypoints shown on the curvature channel of the shape vector image. (c-e), (g-i) and (k-m) show the intermediate curvature channel images of the DoDs across scales t and the detected extrema (shown by points) on the corresponding DoDs at different scales. (f), (j) and (n) show again the extrema detected at (e), (i) and (m), respectively, with the different sizes of circles indicating the sizes of scales at which these extrema are detected.

After localizing the keypoints, feature descriptors are built to characterize these points at the scales where they are identified. These descriptors should contain the necessary distinct information for their corresponding keypoints. In our framework, the descriptor is calculated channel by channel.

For each channel, the local gradient-orientation histograms of the same-scale neighboring pixels of a keypoint are used as the key entries of the descriptor. In this work, we construct a keypoint descriptor with 4×4 subdescriptors computed from a 16×16 sample pixel array, which is shown on the left side of Fig. 4. That is to say, a feature descriptor is computed as a set of orientation histograms on 4×4 pixel neighborhoods or subregions. The coordinates of the subdescriptors and the gradient orientations are rotated relative to the keypoint orientation (defined by the gradient vector at the keypoint location) so that it can achieve orientation invariance. One of the subdescriptors is shown on the right panel of Fig. 4, which gives eight directions of the orientation histogram with the length of each arrow corresponding to the magnitude of that histogram entry. Since the descriptor is computed with a 4×4 array of histograms with 8 orientation bins in each, this leads to a feature vector with $4 \times 4 \times 8 = 128$ elements.

In the case of an m channel vector image, a keypoint has m descriptors which are combined as a vector, $\mathbf{des} = [\mathbf{des}_1, \mathbf{des}_2, \dots, \mathbf{des}_m]^T$, where m is the dimension of the vector image. Hence, the descriptor \mathbf{des} of a keypoint in the shape vector image is a $m \times 128$ dimension-based vector. This descriptor will be used for matching, and all the descriptors computed for all the keypoints form a feature descriptor database, which abstracts the original surface with a small number of robust keypoints and their local descriptors. The robust keypoints and constructed local shape descriptors together are well suitable for the matching purpose as demonstrated by our experiments in Section 5.

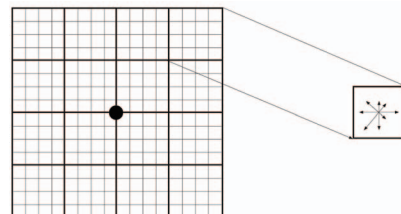


Fig. 4. A keypoint descriptor is generated by computing the gradient magnitude and orientation at every pixel around the keypoint (16×16 sample pixels). These samples are then accumulated into orientation histograms summarizing the contents over 4×4 subregions, indicated with thicker framed boxes. The right panel shows one subregion with the length of each arrow corresponding to the sum of the gradient magnitudes along that direction within the region.

4 SHAPE MATCHING AND REGISTRATION

In our framework, shape matching is to match the keypoints and their associated scale-aware local shape descriptors among different objects. Since the keypoints detected from the diffusion space are very reliable feature points presented in the original surface, matching these keypoints with thin-plate spline deformation will lead to accurate registration of the entire surfaces as well.

Descriptor matching is performed for a keypoint by comparing the distance from its constructed local descriptor to the descriptor of its closest matched point (DIS_{CN}) with the distance from the keypoint descriptor to the descriptor of its second-closest matched point (DIS_{SCN}) found on the to-be-matched object. The distance of two descrip-

tors, \mathbf{des}_1 and \mathbf{des}_2 which are m dimension vectors, is calculated by, $DIS = \sum_{i=1}^m \|\mathbf{des}_1 - \mathbf{des}_2\|$. 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

$$threshold \times DIS_{CN} \leq DIS_{SCN}. \quad (13)$$

If this inequality holds, the points are matched; Otherwise, they are not matched. This inequality ensures that only distinctive keypoints having prominent similarity are matched. Since the 3D data can be coarsely aligned easily through affine transformations during the pre-processing, we can use the uniform subdivision grid to speed up the matching. The Euclidean distance bound (ED) of two potentially matched keypoints is calculated and can be used in efficiently finding the closest and the second-closest matched points within grids.

After finding all the matched points, registration can be achieved using thin-plate splines deformation with the matched points as point constraints [32]. In our shape vector image registration, the keypoints (x_i, y_i) are taken as landmarks and $V = (x_i, y_i)_i$ is a set of the matched keypoints on the other shape vector image. After computing the thin-plate splines with the above point constraints, the deformation function $f(x, y) = [f_x(x, y), f_y(x, y)]$ can be obtained to map each point (x_i, y_i) to its homolog (x'_i, y'_i) . The other unconstrained areas will follow the deformation. At the end, we can register the two shape vector images (i.e., the two 3D surfaces).

5 EXPERIMENTS AND APPLICATIONS

To evaluate the proposed approach, we have conducted extensive experiments. We have applied our algorithm on real scanned face models and human neocortex surfaces extracted from high-resolution MR scans. The surface matching is demonstrated first, followed by the application of the framework in the multimodality image analysis and visualization. Our system is implemented with C++ for the computationally intensive algorithms and VTK/OpenGL for rendering and visualization. The experiments are conducted on a Dell Precision Workstation T7400, which has a Xeon CPU with Quad Cores and 4GB RAM.

5.1 Repeatability under noise

We have tested the repeatability of keypoints detection with noises. The Igea surface model is added up to 10% Gaussian noise directly on the mesh. The perturbed surfaces with different noise levels are converted to the shape vector images and then the keypoints are detected with our geodesic distance-weighted diffusion method. We compare the repeatability of the detected keypoints with the ones detected without additive noise. The repeatability result is shown in Fig. 5. Compared to the repeatability results by regular anisotropic diffusion method and isotropic diffusion method, our method is much more robust under noise. The main reason is that those two methods have instabilities when moving from finer to coarser scales as described in Section 3.2. They are easier to be affected by noise during the diffusion procedure. Therefore, more keypoints originally detected without noises cannot be repeatedly detected across scales under noisy circumstances.

5.2 Surface matching

For scanned face models, we create the shape vector images using conformal mapping. Based on this planar parameter domain, we construct the shape vector image by assigning the H and λ values to each corresponding image pixel. Hence, the shape vector image is a two dimensional vector image $[I_1, I_2]^T$, where $I_1 = H$ and $I_2 = \lambda$. The geodesic distance is computed and encoded as well. 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 scales in each channel.

$$\begin{pmatrix} \mathbf{DoD}_1 \\ \mathbf{DoD}_2 \end{pmatrix} = \begin{pmatrix} DoD_1^{t_0} & DoD_1^{t_1} & \dots & DoD_1^{t_{n-1}} \\ DoD_2^{t_0} & DoD_2^{t_1} & \dots & DoD_2^{t_{n-1}} \end{pmatrix} \quad (14)$$

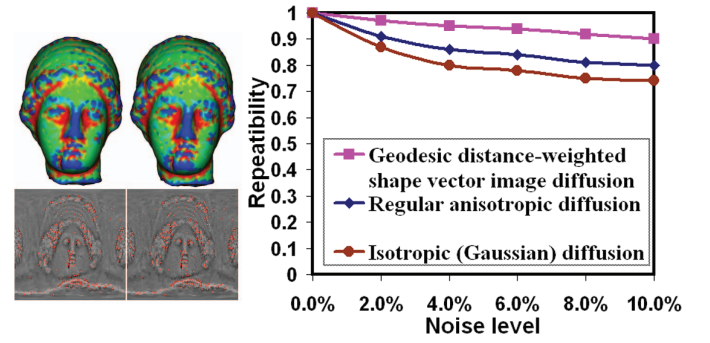


Fig. 5. Repeatability of keypoint features when the Igea model is under different Gaussian noise levels. The left panel shows the Igea models (with the computed curvature colormaps) with 4% and 10% additive Gaussian noise and their corresponding shape vector images. The detected keypoints are shown in the shape vector images. The right panel shows the repeatability of the feature points extracted by our geodesic distance-weighted shape vector image diffusion method. The comparison to the conventional anisotropic and isotropic diffusion methods is demonstrated.

By finding the maxima and minima in each row of the matrix as the keypoints, the descriptor is computed for each point. Each descriptor is a $2 \times 128 = 256$ dimension vector and all these descriptors form the descriptor database. The matching algorithm is performed to find the matched points which satisfies the inequality 13. Fig. 6 shows the matching result of two faces with different expressions from the same human subject. The average matching accuracy of 10 such experiments on 10 different subjects reaches 95% in terms of correct keypoint correspondence.

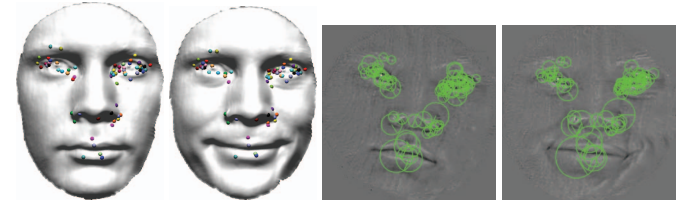


Fig. 6. Matching of face models with different expressions from the same subject. The left panel shows all the matched keypoints between the two surfaces. The right panel shows the scales of the matched keypoints.

For the neocortex surface, a genus zero surface, conformal mapping is performed to transfer it to a sphere. The sphere can be mapped to a 2D domain through a reparameterization as follows,

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

where θ and φ are the rows and columns in the 2D domain image. Then, we follow the same procedure as we use for face models to find the matching keypoints. Fig. 7(a) shows the matching result of two different subjects. In order to allow readers clearly see the matched points between the two shape images, only 10% of the matched points are shown in the figure. After matching, we use the matched points as landmarks to register the shape vector images using the thin-plate spline technique. We have conducted the evaluation on intersubject matching of 20 brain surfaces. The results are evaluated quantitatively in terms of major landmark (e.g., the central sulcus, the sylvian fissure, the posterior sulcus) overlaps. Fig. 7(b) shows one region that we used to test the registration accuracy, where the green color indicates completely correct overlap while the red color indicates mismatched areas among all the subjects. The average mismatch distance error for total 20 different subjects is only 3.98 mm, which outperforms the latest reported results on inter-subject brain surface registration [4].

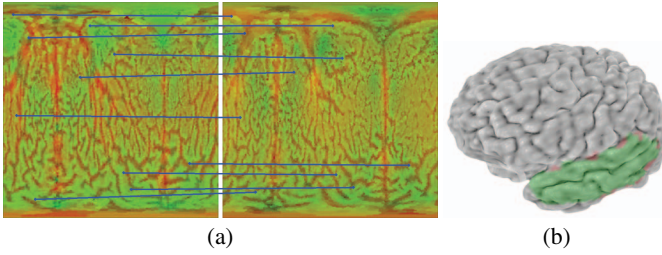


Fig. 7. Matching of two different subjects' brain surfaces. (a) 10% of matched points are shown using the linked lines. (b) The overlap test on one registered brain region.

To further show the efficacy of our approach, we have compared our approach with the closely-related methods, anisotropic diffusion method and SIFT. Since SIFT can only work on scalar image, we only input the curvature channel to the SIFT processing. The regular anisotropic diffusion method is applied on both the curvature and conformal factor channels. We randomly select a pair of brain surfaces among 20 subjects. Then, our geodesic distance-weighted shape vector image diffusion method, the regular SIFT method, and the anisotropic diffusion method are performed for matching and registration. The comparison results are shown in Fig. 8. The main advantage of our method is to introduce geodesic distance into diffusion space. Therefore, it increases the stability of extrema detection as described in Section 3.2 and the robustness of shape descriptors. The experimental results confirm that the keypoints and constructed local shape descriptors together are very robust features well suitable for the matching purpose. The computational time of the geodesic distance-weighted shape vector image diffusion-based feature extraction and matching is recorded for the tested models in Table 1. Note that, the geodesic distance information is pre-computed offline and is not included in the recorded time.

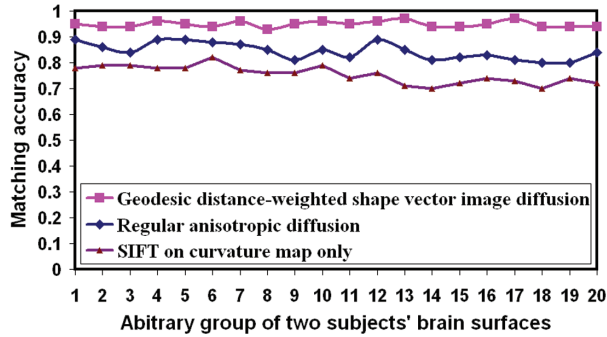


Fig. 8. The experimental results on matching arbitrary two different brain surfaces randomly selected from 20 subjects. The comparison shows that our method constantly outperforms the regular anisotropic diffusion method and the SIFT method.

Model	# vertices	Generate SVI (s)	SVI Size	SVI Diffusion & Matching (s)
Igea	5,002	12	256 × 256	43
Brain	15,102	40	512 × 512	172
Face	20,376	35	256 × 256	45

Table 1. Runtime of the shape vector image (SVI) construction, feature extraction and matching.

5.3 Multimodality Analysis

The developed framework is ideal for cross-subject analysis and visualization of multimodal brain surface properties. In order to facilitate

effective analysis of medical imaging data, especially related to the human neocortical surface, a combination of noninvasive anatomical and functional imaging, such as Magnetic Resonance Imaging (MRI), Diffusion Tensor Imaging (DTI) and Positron Emission Tomography (PET), is frequently used. These modalities provide important, complementary information over the cortex regions. During the preprocessing, a brain surface can be extracted from the MRI volume data. The registration of PET and DTI volumes to the same subject's MRI volume can be done with the mutual information registration algorithm provided in Insight Segmentation and Registration Toolkits (ITK). The registration is easy since the data is for the same subject. In order to integrate PET and MRI data, a normal fusion approach is applied in the native space of the registered MRI and PET volumes of each subject.

In this analysis, we choose one normal brain as the template in our framework. All other normal or abnormal individuals are registered to the template shape vector image (SVI) using the methods described in Section 3 and 4. Fig. 9 shows the idea and flow of our framework. The last two columns in the figure show the maps of the brain surface PET texture and DTI texture, computed from PET and DTI volumes, which are also registered across subjects because their alignments are already registered to their corresponding MRI volumes during the aforementioned preprocessing. Based on the registered SVIs, PET and DTI maps, statistical analysis of PET and DTI across subjects can be achieved.

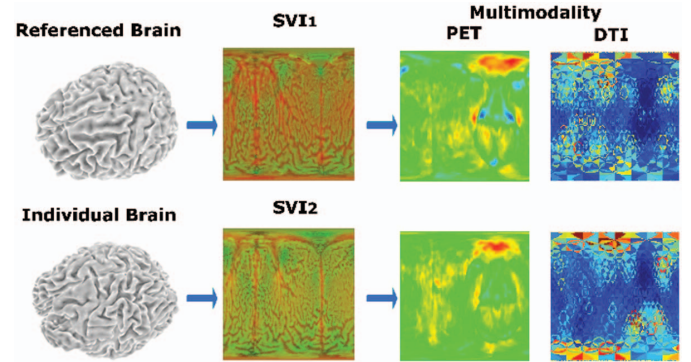


Fig. 9. The multimodality image analysis pipeline. The referenced brain is used as the template SVI (TSVI), and then all other brain SVIs are registered based on this TSVI. Based on the registered shape vector images, multimodality data such as the PET and DTI, can be integrated over the SVI images to perform the multimodality analysis.

By comparing a patient's PET texture with a group of normal subjects based upon matched SVIs, we can identify abnormal PET regions which significantly vary from the normal distribution. Fig. 10(a) shows two detected abnormal regions on the PET shape image. Because we know the mapping and parameterization, we can easily find out the abnormal regions in the actual brain surface. Fig. 10(b) shows the corresponding abnormal regions on the brain PET data. The same scheme can be applied to the population-based DTI analysis. During the preprocessing, the cortico-cortical fiber tracts can be extracted using the brute-force fiber tracking method as shown in Fig. 10(c). Then, these fiber connectivity can be converted to the fiber connectivity strength ratio and plotted in the shape vector image domain to form a DTI texture image. The analysis framework can be used to detect the abnormal regions based on statistical comparison of DTI information between a patient and a group of normal subjects. Fig. 10(d) shows a DTI fiber connectivity strength image of an abnormal subject and the detected abnormal region as highlighted with a red contour.

6 CONCLUSION

In this paper, we have presented a novel and accurate surface matching method based on the geodesic distance-weighted diffusion of shape vector images. Through the detected stable keypoints and their local shape descriptors in the diffusion space, our method converts a 3D surface matching problem to a 2D shape vector image matching problem.

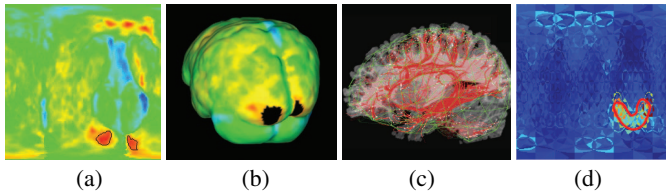


Fig. 10. Population-based PET and DTI image analysis. (a) The regions within the black contours are the detected abnormal regions in the PET texture image; (b) The regions in black are the corresponding abnormalities on the individual's brain cortical surface. (c) shows a 3D rendering of a normal DTI fiber connectivity. (d) shows the abnormal DTI map where the abnormality is contoured in red.

The robust features facilitate the reliable matching and registration as demonstrated by our experiments. The 2D representation allows easier statistical analysis of other modality features directly computed in the matched 2D domain. The applications to medical image analysis and visualization are demonstrated through multimodality data integration in the 2D domain to support more accurate localization of brain disorder regions using population study.

There are some remaining research issues in our current framework. In some situation, when a patch of a surface model containing rich, complex features is squeezed into a very small patch of the domain during the conformal mapping, there may not have enough samples to capture the geometric characteristics with the uniform sampling through the entire domain. This issue may be addressed through controlling the distortion during the mapping or adaptive sampling of the domain according to the distortions and feature richness. Our future work will be focused on this development. Another direction for future work is to apply our framework to very high-genus shapes. In this case, a robust, consistent topological cut algorithm needs to be used to cut and map the high-genus shape to a planar domain. Then, the boundaries of the cuts need to be glued together. Interesting research issues may arise from these boundaries and singular points where diffusion behaviors need to be specially handled. Our future research work will look into these issues.

ACKNOWLEDGEMENTS

This work is supported in part by the research grants: NSF IIS-0713315, NSF CNS-0751045, NIH 1R01NS058802-01A2, NIH 2R01NS041922-05A1, MTTC 05-135/GR686, MTTC 05-154/GR705, and MEDC 06-1-P1-0193.

REFERENCES

- [1] A. Abdel-Hakim and A. Farag. CSIFT: A SIFT descriptor with color invariant characteristics. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1978–1983, 2006.
- [2] V. Athitsos, J. Alon, S. Sclaroff, and G. Kollios. Boostmap: a method for efficient approximate similarity rankings. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 268–275, 2004.
- [3] R. Basri and D. W. Jacobs. Recognition using region correspondences. *International Journal of Computer Vision*, 25(2):141–162, 1996.
- [4] I. Eckstein, A. Joshi, C. Kuo, R. Leahy, and M. Desbrun. Generalized surface flows for deformable registration and cortical matching. In *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 1, pages 692–697, 2007.
- [5] E. Grimson and T. Lozano-Pérez. Localizing overlapping parts by searching the interpretation tree. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):469–482, 1987.
- [6] X. Gu, S. J. Gortler, and H. Hoppe. Geometry images. *ACM Transactions on Graphics*, 21(3):355–361, 2002.
- [7] X. Gu, Y. Wang, T. Chan, P. Thompson, and S. T. Yau. Genus zero surface conformal mapping and its application to brain surface mappings. *IEEE Transactions on Medical Imaging*, 23(8):949–958, 2004.

- [8] B. Levy, S. Petitjean, N. Ray, and J. Maillot. Least squares conformal maps for automatic texture atlas generation. *ACM Transactions on Graphics*, 21(3):362–371, 2002.
- [9] D. G. Lowe. Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, 31(3):355–395, 1987.
- [10] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [11] M. Meyer, M. Desbrun, P. Schroder, and A. Barr. Discrete differential geometry operators for triangulated 2-manifolds. *Proceedings of International Conference on Visualization and Mathematics*, pages 35–57, 2002.
- [12] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 257–263, 2002.
- [13] G. Mori, S. Belongie, and J. Malik. Efficient shape matching using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(11):1832–1837, 2005.
- [14] H. Murase and S. K. Nayar. Visual learning and recognition of 3-D objects from appearance. *International Journal of Computer Vision*, 14(1):5–24, 1995.
- [15] R. C. Nelson and A. Selinger. Large-scale tests of a keyed, appearance-based 3-D object recognition system. *Vision Research*, 38(15):2469–88, 2004.
- [16] L. Nirenberg. A strong maximum principle for parabolic equations. *Communications on Pure and Applied Mathematics*, 6(2):167–177, 1953.
- [17] K. Ohba and K. Ikeuchi. Detectability, uniqueness, and reliability of eigen windows for stable verification of partially occluded objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(9):1043–48, 1997.
- [18] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Shape distributions. *ACM Transactions on Graphics*, 21(4):807–832, 2002.
- [19] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence*, 12(7):628–639, 1990.
- [20] S. Ruiz-Correa, L. Shapiro, and M. Meila. A new paradigm for recognizing 3-D object shapes from range data. In *Proceedings of International Conference on Computer Vision*, pages 1126–1133, 2003.
- [21] B. Schiele and L. James. Object recognition using multidimensional receptive field histogram. In *Proceedings of the 4th European Conference on Computer Vision*, pages 610–619, 1996.
- [22] C. Schmid and R. Mohr. Local grayvalue invariants for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5):530–534, 1997.
- [23] E. Sharon and D. Mumford. 2D-shape analysis using conformal mapping. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 350–357, 2004.
- [24] M. Swain and D. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.
- [25] B. Vemuri, A. Mitche, and J. Aggarwal. Curvature-based representation of objects from range data. *Image and Vision Computing*, 4(2):107–114, 1986.
- [26] S. Wang, Y. Wang, M. Jin, X. Gu, and D. Samaras. Conformal geometry and its applications on 3D shape matching, recognition and stitching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(7):1209–1220, 2007.
- [27] Y. Wang, M. Gupta, S. Zhang, X. Gu, D. Samaras, and P. Huang. High resolution tracking of non-rigid 3D motion of densely sampled data using harmonic maps. In *Proceedings of International Conference on Computer Vision*, pages 388–395, 2005.
- [28] C.-Y. Yao and T.-Y. Lee. Adaptive geometry image. *IEEE Transactions on Visualization and Computer Graphics*, 14(4):948–960, 2008.
- [29] D. Zhang and M. Hebert. Harmonic maps and their applications in surface matching. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 524–530, 1999.
- [30] Z. Zhang, R. Deriche, O. Faugeras, and Q. T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1-2):87–119, 1995.
- [31] G. Zou, J. Hua, X. Gu, and O. Muzik. An approach for intersubject analysis of 3d brain images based on conformal geometry. In *Proceedings of International Conference on Image Processing*, pages 1193–1196, 2006.
- [32] G. Zou, J. Hua, and O. Muzik. Non-rigid surface registration using spherical thin-plate splines. In *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 367–374, 2007.