Stony Brook University

Shape Registration and Analysis Framework for Computer Vision and Graphics

A Dissertation Presented

by

Sen Wang

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Computer Science

Stony Brook University

December 2008

Stony Brook University
The Graduate School

Sen Wang

We, the dissertation committee for the above candidate for
the degree of Doctor of Philosophy, hereby recommend
acceptance of this dissertation.

Hong Qin, Dissertation Advisor
Professor, Computer Science Department

Joseph Mitchell, Chairperson of Defense
Professor, Computer Science Department

Xianfeng Gu
Assistant Professor, Computer Science Department

Lijun Yin
Associate Professor, Computer Science Department, Binghamton University(SUNY)

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Shape Registration and Analysis Framework for Computer Vision and Graphics

by

Sen Wang

Doctor of Philosophy

in

Computer Science

Stony Brook University

2008

As 3D digital photographic and scanning devices produce higher resolution images, acquired geometric data sets grow more complex in terms of the modeled objects' size, geometry, texture, and topology. To use and analyze such data, developing new algorithms and techniques for shape registration and analysis has become a common and long-term mission in the computer vision and graphics field. In this dissertation, we propose a novel framework for shape matching, registration, and scientific analysis especially for 3D facial data and biomedical data. In particular, we address the challenges of 3D shape registration and analysis with noise, occlusion, resolution variation and non-rigid deformation.

Firstly, we analyze a family of quasi-conformal maps including harmonic maps, conformal maps, and least squares conformal maps with regards to 3D shape matching. As a result, we propose a novel and computationally efficient shape matching framework by using least squares conformal maps. The robustness of least squares conformal maps is evaluated and analyzed comprehensively in 3D shape matching with occlusion, noise, and resolution variation. In addition to the above conformal geometry approaches, we also propose a framework of shape registration and analysis using Ricci flow. Previous methods based on conformal geometries, such as harmonic maps and least squares conformal maps, which can only handle 3D shapes with simple topology are subsumed by our Ricci flow based

method which can handle surfaces with complex topology. Furthermore, we introduce a method that constrains Ricci flow computation using feature points and feature curves. We also demonstrate the applicability of this intrinsic shape representation through standard shape analysis problems, such as 3D shape matching and registration.

As 3D scanning technologies continue to improve, dynamic densely-sampled 3D data is becoming more and more prevalent for analysis and synthesis. To study and analyze such huge data, an efficient non-rigid registration algorithm is necessary to establish one-to-one inter-frame correspondences automatically. Toward this goal, we present a new framework for automatic non-rigid registration of 3D dynamic facial data. Based on this registration framework, we also develop a new system of facial expression synthesis and transfer.

We have implemented our framework in a wide range of applications which represent the identified challenges in shape registration and analysis. This includes dynamic noise, occluded data, resolution variation, non-rigid deformation, etc. Furthermore, We describe these applications in detail and outline a few new applications which include surface matching, alignment and stitching, dynamic non-rigid deformable shape registration, facial expression synthesis and transfer.

To My Parents, My Wife and My Children with My Love!

Contents

List of Figures

List of Tables

Acknowledgements

I want to express my deep gratitude to my advisor, Professor Hong Qin, for his years of support and encouragement. He gave me a great help and taught me how to set higher goals and accomplish them.

I would also like to thank Professors Joseph Mitchell, Xianfeng Gu for their valuable advices, collaborations, as well as for serving on various committees.

I would like to thank Professor Lijun Yin for taking the time to serve as the external member of my dissertation committee.

I would also like to thank all my collaborators and colleagues in the center of visual computing and computer science department for their help.

Last but not least, I want to thank my family for their endless love and support. This dissertation is dedicated to them.

Publications

Journal Publications

1. Yang Wang, Mohit Gupta, Song Zhang, **Sen Wang**, Xianfeng Gu, Dimitris Samaras, Peisen Huang. "High Resolution Tracking of Non-Rigid Motion of Densely Sampled 3D Data Using Harmonic Maps". *International Journal of Computer Vision (IJCV)*, Volume 76, Issue 3, Pages: 283-300, March 2008.
2. **Sen Wang**, Yang Wang, Miao Jin, Xianfeng Gu, Dimitris Samaras. "Conformal Geometry and Its Applications on 3D Shape Matching, Recognition and Stitching". *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Volume 29, Issue 7, Pages: 1209-1220, July 2007.

Conference Publications

3. **Sen Wang**, Xianfeng Gu, Hong Qin. "Automatic Non-rigid Registration of 3D Dynamic Data for Facial Expression Synthesis and Transfer". In *IEEE Computer Vision Pattern Recognition 2008 (CVPR08)*, Anchorage, Alaska, USA, June 2008.
4. Xianfeng Gu, **Sen Wang**, Junho Kim, Yun Zeng, Yang Wang, Hong Qin, Dimitris Samaras. "Ricci Flow for 3D Shape Analysis". In *International Conference of Computer Vision 2007 (ICCV07)*, Rio de Janeiro, Brazil. pp.1-8, October 2007.
5. **Sen Wang**, Yang Wang, Miao Jin, Xianfeng Gu, Dimitris Samaras. "3D Surface Matching and Recognition Using Conformal Geometry". In *IEEE Computer Vision Pattern Recognition (CVPR06)*, New York, USA, pp. II:2453-2460, June 2006.
6. Yang Wang, Mohit Gupta, Song Zhang, **Sen Wang**, Xianfeng Gu, Dimitris Samaras, Peisen Huang. "High Resolution Tracking of Non-Rigid 3D Motion

of Densely Sampled Data Using Harmonic Maps". In *International Conference of Computer Vision 2005 (ICCV05)*, Beijing, China, pp. 388-395, October 2005.

7. **Sen Wang**, Lei Zhang, Dimitris Samaras. "Face Reconstruction across Different Poses and Arbitrary Illumination Conditions". In *Audio and Video based Biometric Person Authentication (AVBPA05)*, New York, USA, pp. 91-101, July 2005.
8. Lei Zhang, **Sen Wang**, Dimitris Samaras. "Face Synthesis and Recognition under Arbitrary Unknown Lighting using a Spherical Harmonic Basis Morphable Model". In *IEEE Computer Vision Pattern Recognition (CVPR05)*, San Diego, USA, pp. II:209-216, June 2005.
9. Lei Zhang, Yang Wang, **Sen Wang**, Dimitris Samaras. "Image-Driven Retargeting and Relighting of Facial Expressions". In *Computer Graphics International (CGI05)*, New York, USA, pp. 11-18, June 2005.

Chapter 1

Introduction

Shape registration is a fundamental issue in computer vision and graphics with many applications, such as partial scan alignment, 3D object recognition and classification, shape modeling and analysis, etc. Nowadays, as digital photographic and scanning technologies improve, large databases of 3D scans require automated methods for matching and registration. However, registering 3D shapes in noisy and cluttered scenes is a challenging task. Moreover, since most 3D shape scanners can only capture 2.5D data of the target surfaces, aligning and stitching partial 3D surfaces is a fundamental problem in many research areas, such as computer vision and graphics, mechanical engineering, and molecular biology. Examining the rich literature of shape registration and analysis, the topics of non-rigid shape registration, dynamic deformable surface tracking, and shape modeling and analysis appear to be less explored. Developing new algorithms and techniques for shape registration and analysis has become a common and long-term mission in computer vision and graphics field.

Our research tries to build useful and high-fidelity shape matching and registration methods and apply them to many applications. Our research focuses on the following aspects: (1) How to design shape matching and registration methods with noise, occlusions and resolution variations; (2) How to register 3D non-rigid deformable data automatically and efficiently with minimum human labor and manual work; (3) For one kind of specific shape - human face with many degree of freedom, how to build a system to automatically register non-rigid 3D dynamic facial data with expressions.

In this dissertation we present a novel framework for shape modeling, registration and scientific analysis especially for facial data and medical data. Figure 1 shows the complete conceptual hierarchy of the proposed research. Firstly, we analyze a family of quasi-conformal maps including harmonic maps, conformal maps, and least squares conformal maps with regards to 3D shape matching. As a result, we propose a novel and computationally efficient shape matching framework by using least squares conformal maps. The robustness of least squares conformal maps is evaluated and analyzed comprehensively in 3D shape matching with occlusion, noise, and resolution variation. In order to further demonstrate the performance of our proposed method, we also conduct a series of experiments on two computer vision applications, i.e., 3D face recognition and 3D non-rigid surface alignment and stitching. After that, we show that previous methods based on conformal geometries, such as harmonic maps and least squares conformal maps, which can only handle 3D shapes with simple topology are subsumed by our Ricci flow based method which can handle surfaces with arbitrary topology. Because the Ricci flow method is intrinsic and depends on the surface metric only, it is invariant to rigid motion, scaling, and isometric and conformal deformations. The solution to Ricci flow is unique and its computation is robust to noise. Large non-rigid deformations can be registered with feature constraints, hence we introduce a method that constrains Ricci flow computation using feature points and feature curves. We also demonstrate the applicability of this intrinsic shape representation through standard shape analysis problems, such as 3D shape matching and registration.

Automatic non-rigid registration of 3D time-varying data is fundamental in many vision and graphics applications such as facial expression analysis, synthesis, and transfer. Despite many research advances in recent years, it still remains to be technically challenging, especially for 3D dynamic, densely-sampled facial data with a large number of degrees of freedom (necessarily used to represent rich and subtle facial expressions). We present a new method for automatic non-rigid registration of 3D dynamic facial data using least squares conformal maps, and based on this registration method, we also develop a new framework of facial expression synthesis and transfer. Nowadays more and more 3D dynamic, densely-sampled data become prevalent with the advancement of novel 3D scanning techniques. To analyze and utilize such huge 3D data, an efficient non-rigid registration algorithm

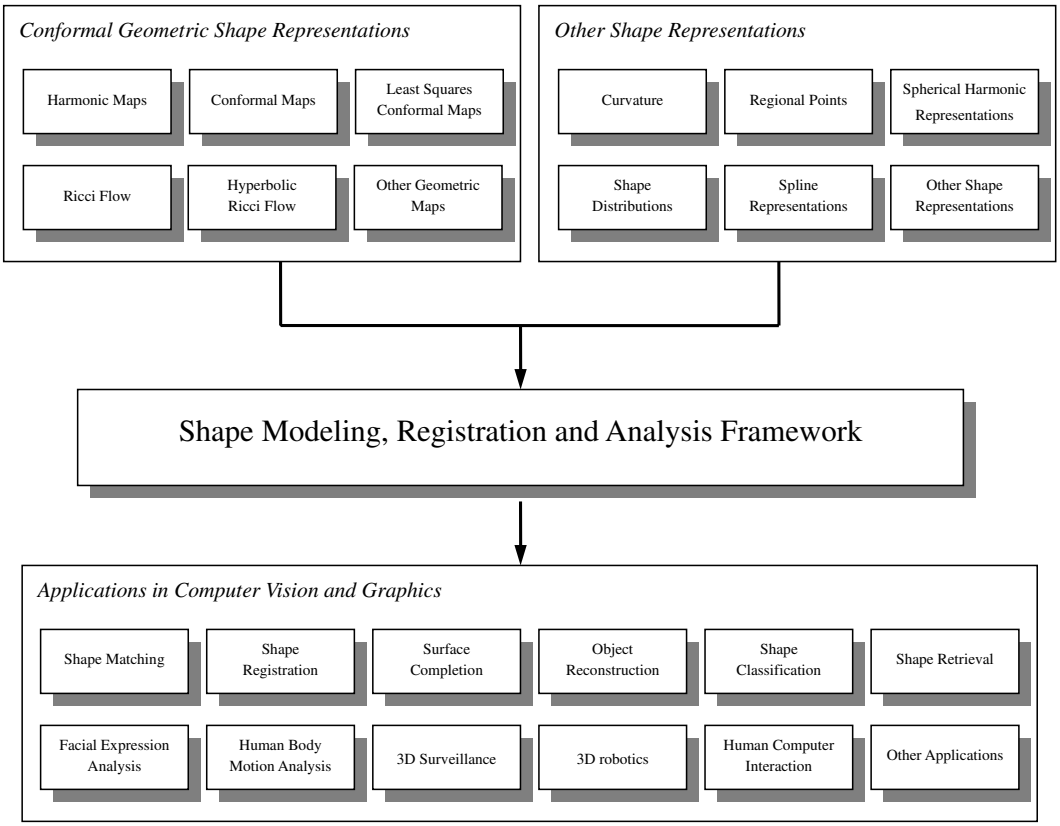


Figure 1: Conceptual hierarchy of the proposed research.

is needed to establish one-to-one inter-frame correspondences. Towards this goal, a non-rigid registration algorithm of 3D dynamic facial data is developed by using least squares conformal maps with additional feature correspondences detected by employing active appearance models (AAMs). The proposed method with additional, interior feature constraints guarantees that the non-rigid data will be accurately registered. The least squares conformal maps between two 3D surfaces are globally optimized with least angle distortion and the resulting 2D maps are stable and one-to-one. Furthermore, by using this non-rigid registration method, we develop a new system of facial expression synthesis and transfer. We perform a series of experiments to evaluate our non-rigid registration method and demonstrate its efficacy and efficiency in the applications of facial expression synthesis and transfer.

1.1 Problem Statement

A fundamental question in 3D shape matching and registration is finding good shape representations: What mathematical description should be chosen to represent the surface of a 3D object on a digital computer? The diversity of the application fields of 3D shape matching is reflected in a wide variety of shape representations that have been proposed in the past. Different approaches include curvature-based representations [77], regional point representations [15, 43, 64, 72], spherical harmonic representations [26, 27, 44], shape distributions [58], spline representations [9] and harmonic shape images [89]. However, many shape representations that use local shape signatures are not stable and cannot perform well in the presence of noise, occlusion and resolution variation. There also has been a lot of research on 3D surface alignment and stitching in recent decades, such as identification and indexing of surface features [24, 70], computing principal axes of scans [20], exhaustive search for corresponding points [13], or iterative closest point(ICP) methods [8, 48, 62, 65]. Compared to matching, there are other additional issues in surface stitching, such as registration and integration [76]. 3D surface alignment and stitching is still a hard problem especially when the transformation between the surfaces to be aligned is non-rigid, e.g., when taking successive scans of humans that might not be standing still.

Automatic non-rigid registration of 3D dynamic data is another hard topic and still remains a challenging task, especially for dynamic facial data with many degrees of freedom. There has been much research on non-rigid registration of 3D facial data in recent decades. Existing approaches to solving this problem typically involve three key techniques: one is to select feature correspondences manually or use markers attached on human faces. The second one is to establish inter-frame correspondences hierarchically using multi-resolution facial data. The third kind of techniques computes correspondences using a 3D deformable model. However, most of existing 3D non-rigid registration methods rely on recovering low dimensional parameters of face model or register 3D faces with local optimization that may not establish accurate one-to-one inter-frame correspondences successfully.

1.2 Contributions

To overcome the above difficulties, in this dissertation we propose an integrated 3D shape registration and analysis framework for surface matching, registration and reconstruction, facial expression analysis, synthesis and transfer. In particular, the contributions of this dissertation are as follows:

1. We present a novel shape registration and analysis framework for computer vision and graphics based on conformal geometry theory which can map 3D surfaces to a 2D planar domain, and thus simply all 3D problems to 2D image problems.
2. We analyze a family of quasi-conformal maps including harmonic maps, conformal maps and least squares conformal maps with regards to 3D shape matching. As a result, we propose a novel and computationally efficient shape matching and registration framework by using least squares conformal maps. The robustness of least square conformal maps is evaluated and analyzed comprehensively in 3D shape matching with occlusion, noise and resolution variation. We also conduct a series of experiments on two computer vision applications, i.e., 3D face recognition and 3D non-rigid surface alignment and stitching.
3. We demonstrate that previous methods based on conformal geometries, such as harmonic maps and least squares conformal maps, which can only handle 3D shapes with simple topology are subsumed by our Ricci flow based method which can handle surfaces with complex topology. Large non-rigid deformations can be registered with feature constraints, hence we propose a method that constrains Ricci flow computation using feature points and feature curves. Finally, we demonstrate the applicability of this intrinsic shape representation through standard shape analysis problems, such as 3D shape matching and registration.
4. We propose a new framework of automatic non-rigid registration for 3D dynamic facial data. The non-rigid registration framework is developed by using least squares conformal maps with additional feature correspondences detected by employing active appearance models (AAMs). Based on this registration method, we also develop a new system of facial expression synthesis

and transfer.

5. Our shape analysis framework is the first time to solve 3D shape registration problems using conformal mapping techniques and can serve as the foundation for a wide range of applications in computer vision and graphics.

1.3 Dissertation Organization

The remainder of the dissertation is organized as follows. In Chapter 2, we briefly review theoretical background of parameterization technology and prior work done in shape registration and facial expression analysis. In Chapter 3, after analyzing a family of quasi-conformal maps including harmonic maps, conformal maps and least squares conformal maps with regards to 3D shape matching, we propose a novel and computationally efficient shape matching framework by using least squares conformal maps. In Chapter 4, we propose a surface registration and analysis system using Ricci flow. In Chapter 5, we present a new method for automatic non-rigid registration of 3D dynamic facial data and based on this registration method, we also develop a new framework of facial expression synthesis and transfer. Finally we conclude this dissertation and outline some future research work in Chapter 6.

Chapter 2

Background Review

Our shape registration and analysis framework is based on previous work in conformal geometry and Ricci flow parameterization methods. We also apply our framework to facial analysis, synthesis and expression transfer. In this chapter, we present a brief survey of the prior work done in these related research fields.

2.1 Theoretical Background

In this section, we briefly introduce theoretical background of three quasi-conformal maps (harmonic maps, conformal maps and least squares conformal maps) and Ricci flow.

2.1.1 Quasi-Conformal Maps

An important merit of quasi-conformal maps, including harmonic maps, conformal maps and least squares conformal maps, is to reduce the 3D shape-matching problem to a 2D image-matching problem, which has been extensively studied. Quasi-conformal mappings, which are almost conformal, do not distort angles arbitrarily and this distortion is uniformly bounded throughout their domain of definition [3]. We are dealing with 3D surfaces, but since they are manifolds, they have an inherent 2D structure, which can be exploited to make the problem more tractable using conformal geometry theory [32, 68]. Most work using conformal geometry theory is done in surface parameterization, which can be viewed as an

embedding from a 3D surface **S** with disk topology to a planar domain **D**. Following the introduction of the notions of harmonic maps, conformal maps and least squares conformal maps, these three parametric maps will be compared in a comprehensive manner.

2.1.1.1 Harmonic Maps

As described in [89], a harmonic map $H : \mathbf{S} \rightarrow \mathbf{D}$ is a critical point for the harmonic energy functional,

$$E(H) = \int_S |\nabla H|^2 d\mu_S, \tag{1}$$

and can be calculated by minimizing $E(H)$. The norm of the differential $|\nabla H|$ is given by the metric on **S** and **D**, and μ_S is the area element on 3D surface **S** [21,23,57,67]. Since the source surface mesh **S** is in the form of a *discrete* triangular mesh, we approximate the harmonic energy as [21, 32, 89],

$$E(H) = \sum_{[v_0,v_1]} k_{[v_0,v_1]} |H(v_0) - H(v_1)|^2, \tag{2}$$

where $[v_0, v_1]$ is an edge connecting two neighboring vertices v_0 and v_1, and $k_{[v_0,v_1]}$ is defined as

$$\frac{1}{2}\left(\frac{(v_0 - v_2)\cdot(v_1 - v_2)}{|(v_0 - v_2)\times(v_1 - v_2)|} + \frac{(v_0 - v_3)\cdot(v_1 - v_3)}{|(v_0 - v_3)\times(v_1 - v_3)|}\right), \tag{3}$$

where $\{v_0, v_1, v_2\}$ and $\{v_0, v_1, v_3\}$ are two adjacent triangular faces.

By minimizing the harmonic energy, a harmonic map can be computed using the Euler-Lagrange differential equation for the energy functional, i.e.,

$$\Delta E = 0, \tag{4}$$

where Δ is the Laplace-Beltrami operator [21, 23, 57, 67]. This will lead to solving a sparse linear least squares system for the mapping H of each vertex v_i [21,32,83,89]. If the boundary condition

$$H|_{\partial S} : \partial\mathbf{S} \rightarrow \partial\mathbf{D}, \tag{5}$$

is given, the solution exists and is unique.

Although harmonic maps are easy to compute, they require satisfaction of the above boundary condition, which becomes unreliable when there are occlusions in the 3D original data. To overcome this problem, the missing boundaries can be approximated [89], which might be enough for rough surface matching. However, since interior feature points are often more robust to occlusion, it is desirable to replace the boundary condition with feature constraints. This can be achieved by conformal maps, another mathematical tool in conformal geometry theory, which only require several feature constraints as an input and obviate the need to specify the boundary condition.

2.1.1.2 Conformal Maps

It can be proven that there exists a mapping from any surface with a disk topology to a 2D planar domain [36], which is one-to-one, onto, and angle preserving. This mapping is called *conformal mapping* and keeps the line element unchanged, except for a local scaling factor [25].

Conformal maps have many appealing properties, one of which is their connection to complex function theory [25, 49]. Consider the 2D case of mapping a planar region **S** to the plane. Such a mapping can be viewed as a function of a complex variable, $d = \mathcal{U}(s)$. Locally, a conformal map is simply any function $\mathcal{U}$ which is analytic in the neighborhood of a point s and such that $\mathcal{U}'(s) \neq 0$. A conformal mapping $\mathcal{U}$ thus satisfies the Cauchy-Riemann equations, which are

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{6}$$

where $d = u + iv$ and $s = x + iy$.

Differentiating one of these equations with respect to x and the other with respect to y, we obtain the two Laplace equations

$$\Delta u = 0, \Delta v = 0. \tag{7}$$

where $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. Any mapping which satisfies these two Laplace equations is called a harmonic mapping. Thus a conformal mapping is also harmonic. However, unlike the harmonic maps described in the previous section, which need the

boundary mapping $H|_{\partial S}$ fixed in advance, conformal maps can be calculated without demanding the mesh boundary to be mapped onto a fixed shape. For a discrete mesh, the main approaches to achieve conformal parameterizations are: harmonic energy minimization [19, 30, 32, 81, 82], Cauchy-Riemann equation approximation [49], Laplacian operator linearization [36], circle packing [41], circle patterns [45], most isometric parameterizations(MIPS) [38] and angle-based flattening method [69]. Here, we compute conformal maps using the harmonic energy minimization method [32].

Riemann's theorem states that for any surface **S** homeomorphic to a disc, it is possible to find a parameterization of the surface satisfying Equation 6 [49], which can be uniquely determined by two points on surface **S**. However, to better handle the errors caused by noise in the data and the inaccuracy of finding feature points, we introduce additional feature constraints, indicating that the corresponding features on two 3D surfaces should be mapped onto the same locations in the 2D domain. However, with these additional constraints, it is not always possible to satisfy the conformality condition. Hence, we seek to minimize the violation of Riemann's condition in the least squares sense.

2.1.1.3 Least Squares Conformal Maps

The Least Squares Conformal Map(LSCM) parameterization algorithm generates a discrete approximation of a conformal map by adding more constraints. Here we give a brief description (see [49] for details using different constraints).

Given a discrete 3D surface mesh **S** and a smooth target mapping $\mathcal{U} : S \rightarrow (u,v)$, then, as described in section 2.1.1.2, $\mathcal{U}$ is conformal on **S** if and only if the Cauchy-Riemann equation,

$$\frac{\partial \mathcal{U}}{\partial x} + i\frac{\partial \mathcal{U}}{\partial y} = 0 \tag{8}$$

holds true on the whole of **S**. However, in general this conformal condition cannot be strictly satisfied on the whole triangulated surface **S**, so the conformal map is constructed in the least squares sense:

$$MinC(S) = \sum_{d \in S} \int_d |\frac{\partial \mathcal{U}}{\partial x} + i\frac{\partial \mathcal{U}}{\partial y}|^2 dA, \tag{9}$$

where d is a triangle on the mesh $\mathbf{S}$. If we suppose the mapping u is linear on d then

$$C(S) = \sum_{d \in S} |\frac{\partial u}{\partial x} + i\frac{\partial u}{\partial y}|^2 A(d), \tag{10}$$

where $A(d)$ is the area of the triangle d. Furthermore let $\alpha_j = u_j + iv_j$ and $\beta_j = x_j + iy_j$, so $\alpha_j = u(\beta_j)$ for $j = 1, 2, ..., n$. Then, we rearrange the vector α such that $\alpha = (\alpha_f, \alpha_p)$ where α_f consists of $n - p$ free coordinates and α_p consists of p constraint point coordinates. Therefore, Equation 10 can be rewritten as

$$C(S) = \|M_f\alpha_f + M_p\alpha_p\|^2, \tag{11}$$

where $M = (M_f, M_p)$, a sparse $m \times n$ complex matrix(m is the number of triangles and n is the number of vertices). The least squares minimization problem in Equation 11 can be efficiently solved using the Conjugate Gradient Method. Thus we can map a 3D surface to a 2D domain with multiple correspondences as constraints by using the LSCM technique.

Since LSCMs have almost all the properties of conformal maps and also provide more correspondences as additional constraints, we expect them to be very useful in 3D shape matching and recognition.

2.1.1.4 Comparison of Quasi-Conformal Maps

Based on conformal geometry theory, harmonic maps, conformal maps and least squares conformal maps(LSCMs) between two topological disks preserve continuity of the underlying surfaces, with minimal stretching energy and angle distortion. All the above quasi-conformal maps are invariant for the same source surface with different poses, thus making it possible to account for global rigid transformations. A very important property, which governs our matching algorithm, is that all the maps can establish a common 2D parametric domain for the two surfaces. Therefore we can simplify the 3D shape-matching problem to a 2D image-matching problem. However, they vary in performance for 3D surface matching as can be seen in table 1.

Compared to the exact solutions for harmonic maps and conformal maps, LSCMs are generated by minimizing the violation of Riemann's condition in the

Table 1: Performance comparison of conformal geometric maps.

	Harmonic Maps	Conformal Maps	Least Squares Conformal Maps
Resolution changes	Not sensitive	Not sensitive	Not sensitive
Boundary constraint	Needed	Not needed	Not needed
Boundary occlusion	Difficult to handle	No significant impact	No significant impact
Interior feature points used in mapping	Do not use	Use 2 Points (from Riemann's theorem)	Use more feature constraints
Error of interior feature points detection	Not sensitive	Sensitive	Not sensitive
Computational Complexity	Linear	Nonlinear (with linear approximation available)	Linear

least squares sense. This optimization-based parameterization method has the following properties:

1. LSCMs have the same properties as conformal maps, e.g., existence and uniqueness which have already been proven in [49].
2. LSCMs can map a 3D shape to a 2D domain in a continuous manner with minimized local angle distortion.
3. LSCMs can handle missing boundaries and occlusion and also allow multiple constraints.
4. LSCMs are independent of mesh resolution.
5. The least squares minimization problem in calculating LSCMs has the advantage of being linear.

For actual 3D surfaces, it is very likely to have noise and missing data. From the above comparison, we can see that LSCMs are the best candidate among all three parametric maps to perform 3D shape matching efficiently. LSCMs do not require the boundary condition explicitly which means they can handle missing boundaries and occlusions. Also, they take multiple feature constraints as input, which allows them to better handle noise introduced by the feature point detection. We will confirm this experimentally in the Chapter 3 by analyzing the robustness of the three parametric maps for 3D shape matching with occlusion, noise and resolution variation.

2.1.2 Ricci Flow

This section briefly introduces the theoretic background of surface Ricci flow (For the details in [37]).

Let S be a smooth surface embedded in $\mathbb{R}^3$, then S has an induced Euclidean metric $\mathbf{g}$. Suppose $u : S \to \mathbb{R}$ is a function on the surface, we can define another metric $\bar{\mathbf{g}} = e^u \mathbf{g}$, which is *conformal* to the original metric $\mathbf{g}$ with an area distortion factor e^{2u}. We call u the *conformal factor*.

Furthermore, when the metric of S is changed from $\mathbf{g}$ to $\bar{\mathbf{g}}$ along the change of u, every intrinsic property (e.g., Gaussian and geodesic curvatures) of S is changed. The Gaussian curvature k of interior points changes by $\bar{k} = e^{-2u}(k - \Delta u)$, where Δ is the Laplace-Beltrami operator [11] induced by the original metric $\mathbf{g}$. The geodesic curvature k_g on the boundary points changes as $\bar{k_g} = e^{-u}(k_g - \frac{\partial u}{\partial \mathbf{n}})$, where $\mathbf{n}$ is the normal to the boundary of the surface ∂S.

Although the curvature value at a point is determined from the Riemannian metric, the sum of the total curvatures solely depends on the topology of the surface, as described in the Gauss-Bonnet formulae [11], such that

$$\int_S k dA + \int_{\partial S} k_g ds = \int_S \bar{k} d\bar{A} + \int_{\partial S} \bar{k_g} d\bar{s} = 2\pi\chi(S), \tag{12}$$

where $\chi(S)$ is the Euler characteristic number of the surface S.

Ricci flow is a powerful tool to compute the desired metric $\bar{\mathbf{g}}$ which satisfies the given target curvature $\bar{k}$, from the induced metric $\mathbf{g}$ in S. Suppose S is a closed surface with a Riemannian metric $\mathbf{g}$, the Ricci flow is defined as

$$\frac{d\mathbf{g}}{dt} = -2k\mathbf{g}, \tag{13}$$

where k is the Gaussian curvature determined by the current metric.

Surface Ricci flow deforms a Riemannian metric $\mathbf{g}$ to another metric $e^{2u(t)}\mathbf{g}$, which is conformal to the original one. When the desired target curvature $\bar{k}$ is given, then the corresponding conformal metric can be achieved by the following general Ricci flow

$$\frac{du(t)}{dt} = 2(\bar{k} - k(t)). \tag{14}$$

Eventually, the limit metric $\mathbf{g}(\infty)$ becomes $\bar{\mathbf{g}}$, which is conformal to $\mathbf{g}$ and satisfies the target curvature $\bar{k}$.

Surface Ricci flow offers a novel means to manipulate shapes by curvatures and Riemannian metrics. Specifically, it can *conformally* transform all shapes to one of the three canonical spaces: the sphere, the plane, and the hyperbolic space. For example, by using Ricci flow we can calculate the metric $\bar{\mathbf{g}}$, which satisfies non-zero curvatures $\bar{k}_g$ only on the boundary of S. Then, we can simply flatten S into a 2D planar domain with the metric $\bar{\mathbf{g}}$. Therefore, all 3D problems can be converted to 2D problems in theses canonical spaces. In Chapter 4, we demonstrate how Ricci flow provides a powerful unified tool for computer vision applications, such as surface matching and shape registration.

2.2 3D Facial Expression Analysis Review

Automatic non-rigid registration of 3D time-varying densely-sampled data is a fundamental and critical issue in 3D vision and graphics which has widespread applications. As 3D scanning technologies continue to improve, 3D dynamic densely-sampled data is becoming more and more prevalent for analysis and synthesis. To study and analyze such huge data, an efficient non-rigid registration algorithm is necessary to establish one-to-one inter-frame correspondences automatically. However, automatic 3D non-rigid registration still remains a challenging task, especially for dynamic densely-sampled facial expression data with many degrees of freedom. There has been much research on registration of 3D facial data in recent decades. Existing approaches to solving this problem typically involve three key techniques: one is to select feature correspondences manually or use markers attached on human faces [34, 56, 71]. The second one is to establish inter-frame correspondences hierarchically using multi-resolution facial data [55, 85]. The third kind of techniques computes correspondences using a low-resolution 3D deformable model [12, 60, 73]. However, most of these existing 3D non-rigid registration methods rely on recovering low dimensional parameters of face model or register 3D faces with local optimization that may not establish accurate one-to-one inter-frame correspondences successfully.

Realistic facial animation and expression analysis remains a fundamental challenge in vision and graphics. Earlier approaches explicitly model the facial anatomy, deriving facial animations from the physical behaviors of the bone, joint,

and muscle structures [47, 86]. Others focus only on the surface of the face, using smooth surface deformation mechanisms to create facial expressions [34, 56, 85]. These approaches make use of existing data for animating a new model. Previous works also use techniques for tracking head motions and facial expression in video [18, 60] and copy deformations from one subject onto the geometry of other faces [12]. Expression cloning [56, 63, 71] improves upon this deformation transfer process with both 3D source and target face data. Recently, facial animation and expression analysis using 3D motion capture becomes available with the advancement of new 3D scan techniques [85, 90]. However, these 3D motion data is not registered in the space-time domain. For this purpose, a number of registration method have been proposed for 3D dynamic facial data. Zhang et al. [90] propose a new tracking method based on optic flow estimation which can be sensitive to noise. Wang et al. [85] use a hierarchical method to track 3D motion facial data with expression transfer at the cost of making the estimation of model parameters more difficult. Moreover, their method requires a lot of manual work by dividing the face model into several deformable regions.

Chapter 3

Shape Registration and Analysis Using Quasi-Conformal Maps

In this chapter, we analyze a family of quasi-conformal maps including harmonic maps, conformal maps and least squares conformal maps with regards to 3D shape matching. As a result, we propose a novel and computationally efficient shape matching framework by using least squares conformal maps. The robustness of least square conformal maps is evaluated and analyzed comprehensively in 3D shape matching with occlusion, noise and resolution variation. We also conduct a series of experiments on two computer vision applications, i.e., 3D face recognition and 3D non-rigid surface alignment and stitching.

This work has been published in the proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition 2006 [79]. An extended version with more applications has been published in IEEE Transection on Pattern Analysis and Machine Intelligence 2007 [80].

3.1 Introduction

3D shape matching is a fundamental issue in computer vision and graphics field with many applications, such as shape registration, partial scan alignment, 3D object recognition and classification [10, 40, 64, 87]. As digital photographic and scanning technologies improve, large databases of 3D scans require automated

methods for matching. However, matching 3D shapes in noisy and cluttered scenes is a challenging task. Moreover, since most 3D shape scanners can only capture 2.5D data of the target surfaces, aligning and stitching partial 3D surfaces is a fundamental problem in many research areas, such as computer vision, mechanical engineering, and molecular biology.

Generally, the crux of 3D shape matching is finding good shape representations, allowing us to match two given free-form surfaces by comparing their shape representations. Different approaches include curvature-based representations [77], regional point representations [15, 43, 64, 72], spherical harmonic representations [26, 27, 44], shape distributions [58], spline representations [9] and harmonic shape images [89]. However, many shape representations that use local shape signatures are not stable and cannot perform well in the presence of noise. In this chapter, we propose to use a family of quasi-conformal maps, including harmonic maps, conformal maps and least squares conformal maps, that does not suffer from such problems. According to conformal geometry theory, each 3D shape with disk topology can be mapped to a 2D domain through a global optimization and the resulting map is a diffeomorphism, i.e., *one-to-one* and *onto*. Consequently the 3D shape-matching problem can be simplified to a 2D image-matching problem of the quasi-conformal maps. These maps are stable, insensitive to resolution changes and robust to occlusion and noise. The 2D maps integrate geometric and appearance information and 2D matching is a better understood problem [5, 51]. Therefore, highly accurate and efficient 3D shape matching algorithms can be achieved using quasi-conformal maps.

The robustness and easy use of the technique we proposed allow us to cope with more challenging problems such as surface alignment and stitching, when only two parts of surfaces could be matched. There has been a lot of research on 3D surface alignment and stitching in recent decades, such as identification and indexing of surface features [24, 70], computing principal axes of scans [20], exhaustive search for corresponding points [13], or iterative closest point(ICP) methods [8, 48, 62, 65]. Compared to matching, there are other additional issues in surface stitching, such as registration and integration [76]. 3D surface alignment and stitching is still a challenging task especially when the transformation between the surfaces to be aligned is non-rigid, e.g., when taking successive scans of humans

that might not be standing still. Based on conformal geometry theory, an important property of Least Squares Conformal Maps(LSCMs) is that they can map a 3D surface to a 2D domain in a continuous manner with minimized local angle distortion. This implies that *LSCMs are not sensitive to surface deformations*, which leads to a natural solution to 3D non-rigid surface alignment and stitching.

Quasi-Conformal maps including harmonic maps, conformal maps and least squares conformal maps have been used in several applications of computer vision and graphics. In [89], Zhang et al. proposed harmonic maps for surface matching. In [83], Wang et al. use harmonic maps to track dynamic 3D surfaces. In [30, 32, 81, 82], conformal maps are used for face and brain surface matching. Moreover, Sharon et al. [68] use conformal maps to analyze similarities of 2D shapes. Least squares conformal maps are introduced by Levy et al. [49] for texture atlas generation and used by Wang et al. [79] to do 3D surface matching. In order to calculate harmonic maps, the surface boundary needs to be identified and a boundary mapping from 3D surfaces to the 2D domain needs to be created, which can be a difficult problem especially when part of the surface is occluded. However, the two other quasi-conformal maps we discuss in this chapter, conformal maps and least squares conformal maps, do not need boundary information and so lend themselves as a natural choice to solve this problem. Moreover, in addition to the advantages of harmonic maps, such as sound mathematical basis and preservation of continuity of the underlying surfaces, conformal maps are also angle preserving, which leads to less distortion and robustness to noise. The differences between conformal maps and harmonics maps are shown in Figure 2.

3.2 Shape Matching and Registration Using Least Squares Conformal Maps

To match 3D shapes accurately and efficiently, a new 2D representation, least squares conformal shape images, is developed in our framework using LSCMs. Therefore, we simplify the original 3D shape-matching problem to a 2D image-matching problem. In particular, our shape matching framework includes two steps: First, interior feature correspondences are detected by using spin-images [43]; After

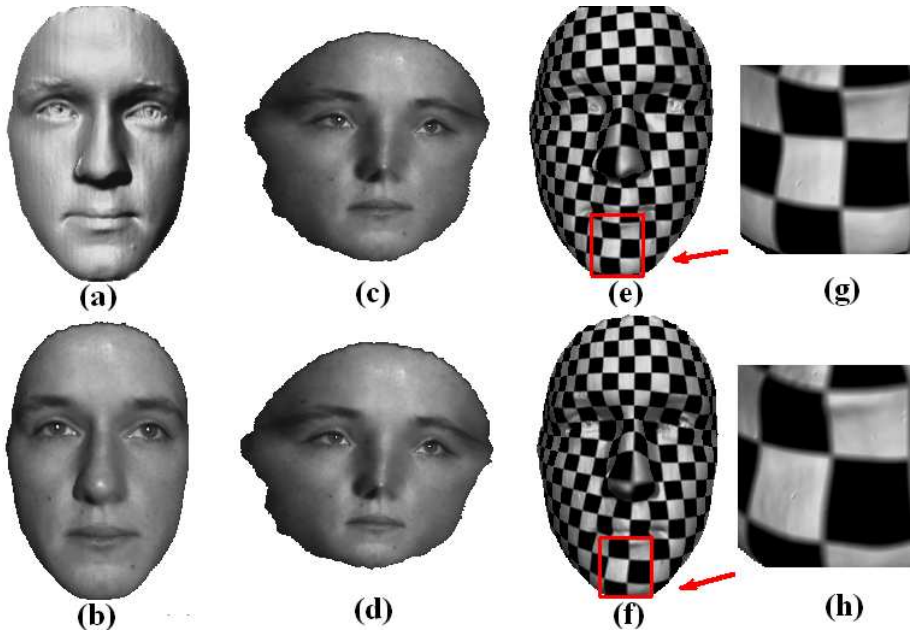


Figure 2: Distortion comparison between a conformal map and a harmonic map. (a) Original surface without texture. (b) Original surface with texture. (c) The 2D conformal map of the surface with texture. (d) The harmonic map of the surface with texture. (e) Checkerbox textured surface by conformal mapping. (f) Checkerbox textured surface by harmonic mapping. Because of angle-preservation, (c) and (e) have less distortions than (d) and (f), which can be clearly seen in the close-up views (g) and (h) of the chin areas in the red boxes respectively.

that, we generate and match least squares conformal shape images.

3.2.1 Correspondence Detection Using Spin-Images

In order to use least squares conformal mappings, we need to establish interior feature constraints between the 3D shapes. For this purpose, we first select candidate points with curvature larger than a threshold T_c, and then compare their spin-images to detect feature correspondences. The spin-image is a well-known technique that has been proven useful for 3D point matching [43]. It encodes the surface shape surrounding an oriented point p by projecting nearby surface points into a 2D histogram, which has cylindrical coordinates of radius r and height h centered at p, with its axis aligned with the surface normal of p. The number of bins and support size in the spin-image histograms are parameters fixed at generation. It

has been shown that the matching results using spin-images are insensitive to the choice of the above parameters [40]. In our experiments, the highest confidence feature correspondences are used. The typical number of selected feature points is 5-6 for 3D face surfaces and 10-12 for brain surfaces.

3.2.2 Least Squares Conformal Shape Images (LSCSIs)

In this section, we will introduce a method to describe 3D surfaces using least squares conformal shape images (LSCSIs). In section 2.1.1.3, we have shown that there exists a least squares conformal mapping that can map each 3D surface with disk topology to the canonical 2D domain. The LSCSIs are generated by associating a shape attribute with each vertex. Mean curvature is a useful geometric attribute that depends only on the surface's intrinsic geometry. In our method, the mean curvature is computed in the same way as in [32]. Moreover, least squares conformal maps can also help generate additional shape representations by associating other attributes, e.g. texture, which leads to a natural solution of combining multiple important cues for 3D surface matching and recognition, such as shape and texture. In our current framework, these cues are weighted equally for surface matching. More elaborate schemes to combine different cues can be done in the future work.

As an example, Figure 3(d) shows the LSCSI of the surface Figure 3(b), with darker color representing larger mean curvature. Figure 3(a) is the original surface with texture information and Figure 3(c) is its LSCM. Figure 3(e) is the LSCM of a lower resolution(25%) version of the original surface. The similarity between Figure 3(c) and Figure 3(e) shows that LSCMs are independent to resolution variation.

3.2.3 Matching Surfaces by Matching LSCSIs

Given two general surfaces S_1 and S_2 with disk tropology, we first detect high curvature correspondences using spin-images. Then, by incorporating interior correspondences as constraints, LSCSIs are generated for both surfaces as described in the above section. After that, the normalized correlation coefficient M_{S_1,S_2} and the similarity criterion $S(S_1,S_2)$ introduced in [42] are computed on the two resulting

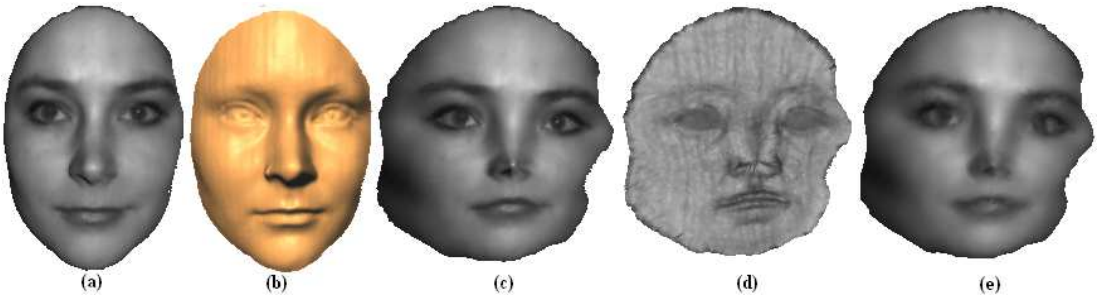


Figure 3: Least Squares Conformal Shape Image: (a) Original surface with texture. (b) Original surface without texture. (c) Least squares conformal maps with texture. (d) Least squares conformal shape image. (e)Least squares conformal maps of the same surface, sub-sampled by a factor of 4, still very similar to (c).

LSCSIs by

$$M_{S_1,S_2} = \frac{N\sum p_i^{S_1} p_i^{S_2} - \sum p_i^{S_1} \sum p_i^{S_2}}{\sqrt{(N\sum (p_i^{S_1})^2 - (\sum p_i^{S_1})^2)(N\sum (p_i^{S_2})^2 - (\sum p_i^{S_2})^2)}}, \tag{15}$$

$$S(S_1,S_2) = (ln\frac{1+M_{S_1,S_2}}{1-M_{S_1,S_2}})^2 - \frac{1}{2N}. \tag{16}$$

where N is the number of overlapping points in the LSCSIs of 3D surface S_1 and S_2, and $p_i^{S_k}$ is the value (e.g., the mean curvature or the texture) of point i in the LSCSI of surface $S_k (k = 1,2)$. In the case of matching surfaces with different resolutions, N is the number of overlapping points in the LSCSIs of the surface with the lower resolution.

According to section 2.1.1.4, an important property of Least Squares Conformal Maps (LSCMs) is that they can map a 3D shape to a 2D domain in a continuous manner with minimized local angle distortion. This implies that *LSCSIs are not sensitive to surface deformations*, e.g., if there is not too much stretching between two faces with different expressions, they will induce similar LSCSIs. As an example, Figure 4 shows a comparison between the LSCSIs of faces with different expressions and of different faces. More specifically, the first, the second, and the third columns of Figure 4 correspond to face scans of one subject with different expressions while the forth column corresponds to another subject. For each column in Figure 4, the bottom row represents the LSCSIs of the surfaces (shown in

the middle row), with darker color representing larger mean curvature. The original surfaces with texture information are also shown in the top row of Figure 4. Based on Equation 15, the normalized correlation coefficient ($M_{i,j}$) between Figure 4(i) and Figure 4(j) and the normalized correlation coefficient ($M_{i,k}$) between Figure 4(i) and Figure 4(k) are 0.92 and 0.86, respectively, while the normalized correlation coefficient ($M_{i,l}$) between Figure 4(i) and Figure 4(l) is only 0.65. As is evident, the normalized correlation coefficients of LSCSIs between the face scans of the same person with different expressions are much larger than the coefficients between face scans of different persons, thus making it possible to match surfaces with small deformations using LSCSIs. This relative expression-invariance is also an important property for shape representations used in face recognition.

However, for 3D surfaces with holes, which violate the disk topology assumption, we can not calculate the LSCMs directly. To overcome this problem, we can simply fill in the holes through interpolation [50] and then use our method to generate the LSCSIs of the new surfaces. The filled-in regions are masked out when we compute the normalized correlation coefficient using Equation 15. As discussed in section 2.1.1.4, LSCMs depend on the geometry in a continuous manner, which leads to robustness to local perturbation. Figure 5 demonstrates the robustness of our method to holes on surfaces. The normalized correlation coefficient of the LSCSIs shown in Figure 5(b,f) is 0.99, which means a very good match between the two surfaces of Figure 5(a,e) after hole filling. If we desire to preserve the non-disk topology of the object during matching, then the object should be partitioned into simpler parts with disk topology [49] which could then be matched. Optimal partitioning will be studied in future work.

3.3 Experimental Results and Performance Analysis

In this section we analyze the robustness of our proposed 3D shape matching method using least squares conformal maps on real data with occlusion, noise and resolution variation. Furthermore, we demonstrate the performance of our method through two applications: 3D face recognition and 3D non-rigid surface alignment and stitching.

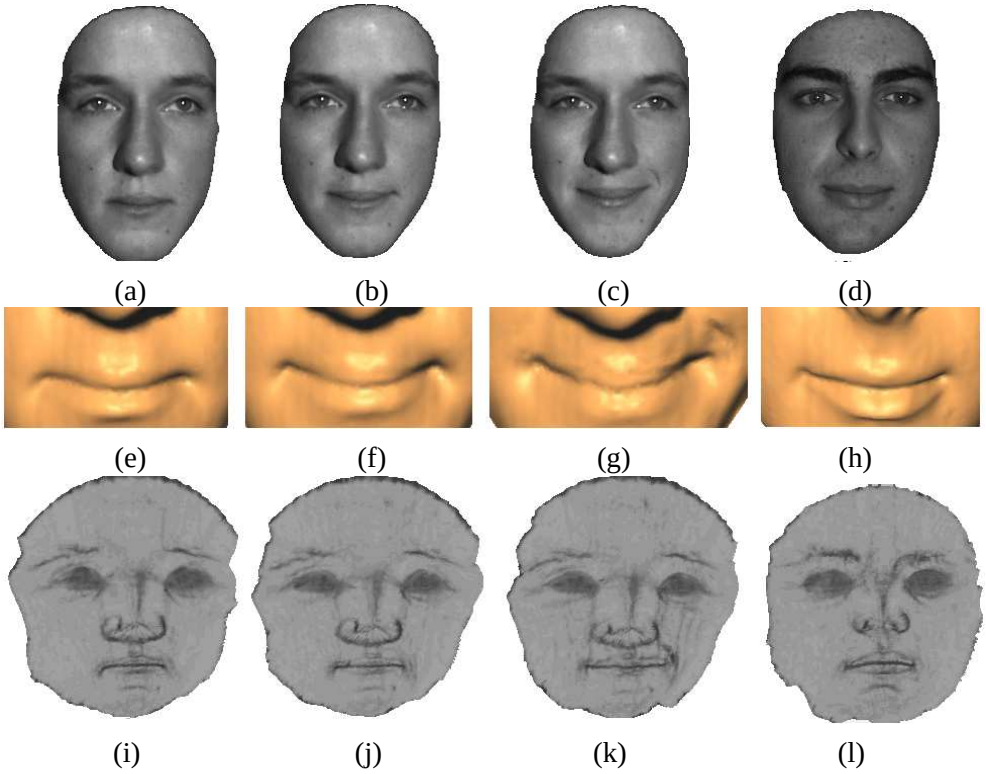


Figure 4: Surface matching with deformation: The original 3D surfaces with texture are in the top row. The detail of the deformed mouth areas are shown in the second row and the LSCSIs of the original surfaces are in the last row. In each row, the first, the second, and the third surfaces are from the same person with different expressions and the forth one is another person. The normalized correlation coefficient ($M_{i,j}$) between (i) and (j) and the normalized correlation coefficient ($M_{i,k}$) between (i) and (k) are 0.92 and 0.86, respectively, while the normalized correlation coefficient ($M_{i,l}$) between (i) and (l) is only 0.65.

3.3.1 Robustness Analysis

In this section we use two surface types: brains (4 instances) and faces (6 instances) to analyze the performance of our proposed 3D shape matching method. We present three experiments in which 3D surface matching is performed under occlusion, noise and resolution variation using least squares conformal maps, followed by a full comparison between several related work of quasi-conformal maps including harmonic maps, conformal maps and least squares conformal maps.

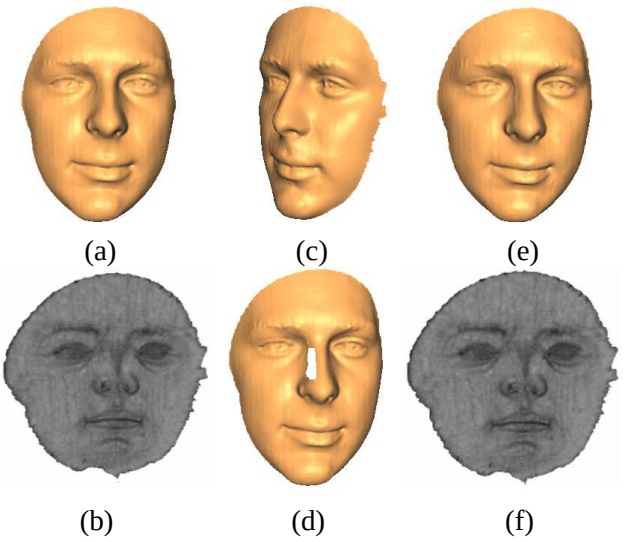


Figure 5: An example of surface matching with holes : (a) A frontal 3D scan. (b) The LSCSI of (a). (c) A side 3D scan of the same subject as in (a), which has a hole illustrated in (d). (e) The same surface of (c,d) after hole filling. (f) The LSCSI of (e).

3.3.1.1 Experiment on Data Occlusion

In this experiment, we test the robustness of Least Squares Conformal Maps (LSCMs) under occlusion for both face and brain surfaces. Such occlusions might be caused by rotation of the object in front of the scanner. Figures 6 and 8 show examples of 3D face and brain surfaces respectively, under different occlusions with their least squares conformal shape images (LSCSIs). For each original surface, partially occluded surfaces were generated with occlusion rates between 5% and 45%. Average matching results of these face and brain surfaces using LSCMs are shown in Figure 7 and 9, respectively. In experiments, we superimpose the matched surfaces with significant occlusions (only 60% of area is common to both). Matching error is very hard to detect visually, which suggests that our framework could be useful for partial scan alignment.

3.3.1.2 Experiment on Noisy Data

The second experiment tests the robustness of Least Squares Conformal Maps (LSCMs) in the presence of noise. We add gaussian noise($\mathcal{N}(0,\sigma)$) on each vertex of the face and brain surfaces. σ increases from 0.0 mm to 2.0 mm while the

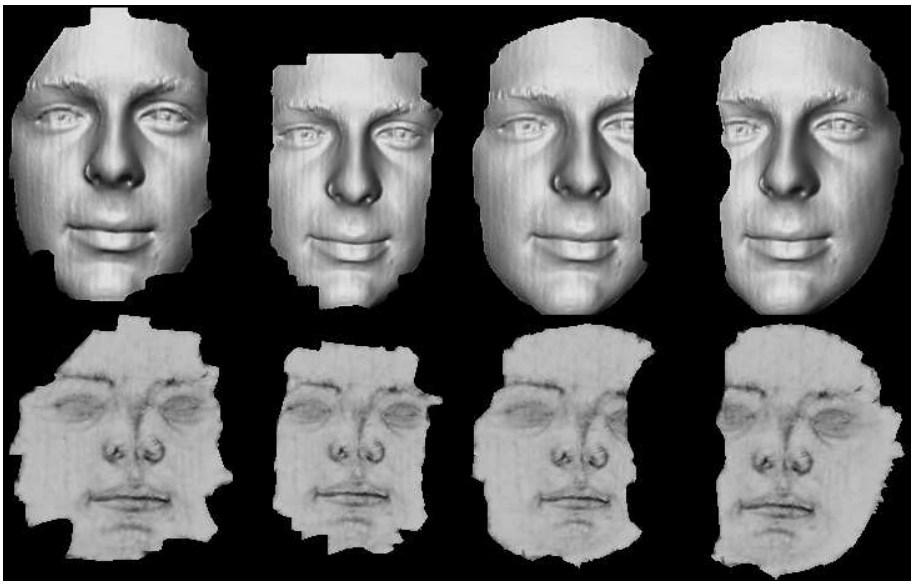


Figure 6: 3D face surfaces and their LSCSIs under occlusion. The original 3D face surfaces with different occlusions are in the top row. Their LSCSIs are in the bottom row.

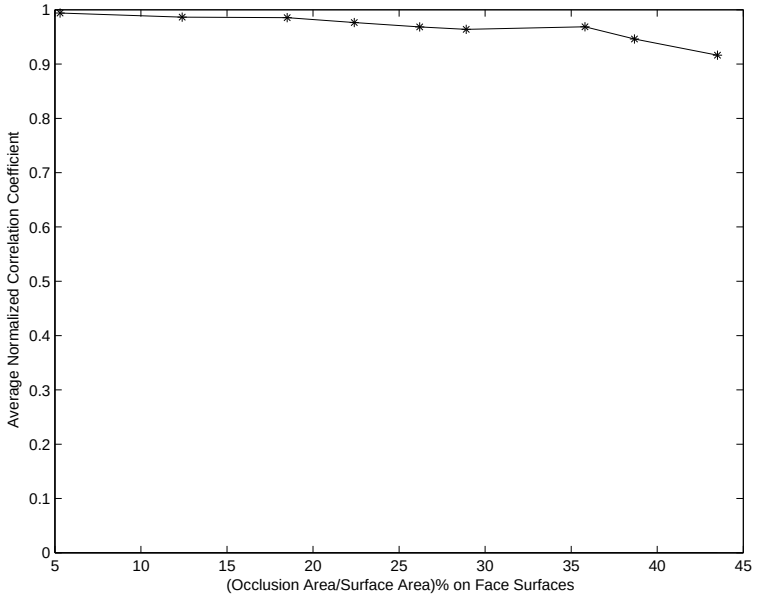


Figure 7: Average matching results of the face surfaces under occlusion using LSCMs.

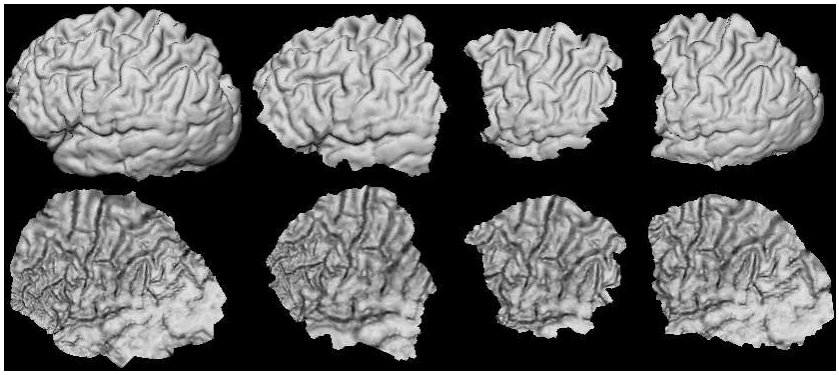


Figure 8: 3D brain surfaces and their LSCSIs under occlusion. The original 3D brain surfaces with different occlusions are in the top row. Their LSCSIs are in the bottom row.

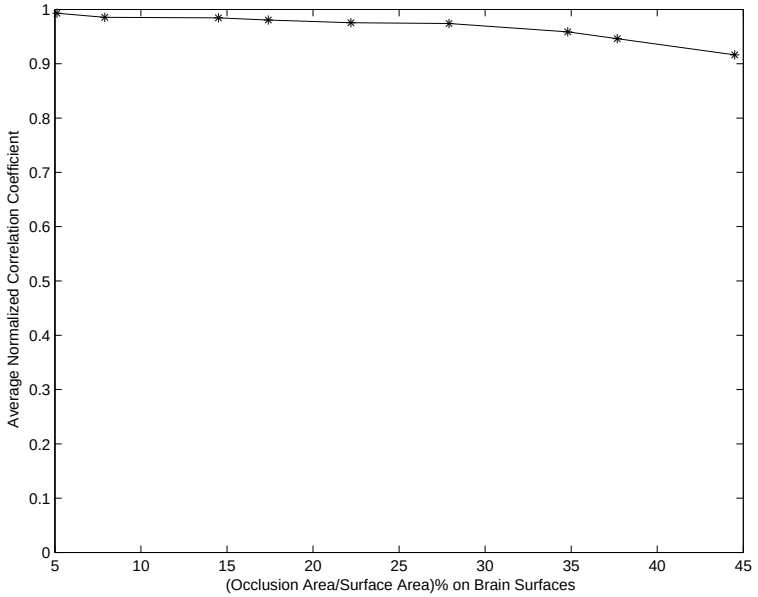


Figure 9: Average matching results of the brain surfaces under occlusion using LSCMs.

window size for computing the curvatures of 3D face and brain surfaces is 10.0 mm. Example surfaces with noise under different σ are shown in Figure 10. We match the various noisy surfaces to the original noise-free surface and the average matching results of the face and brain surfaces are shown in Figure 11 for various σ values. From the results we can see that LSCMs appear robust to gaussian noise.

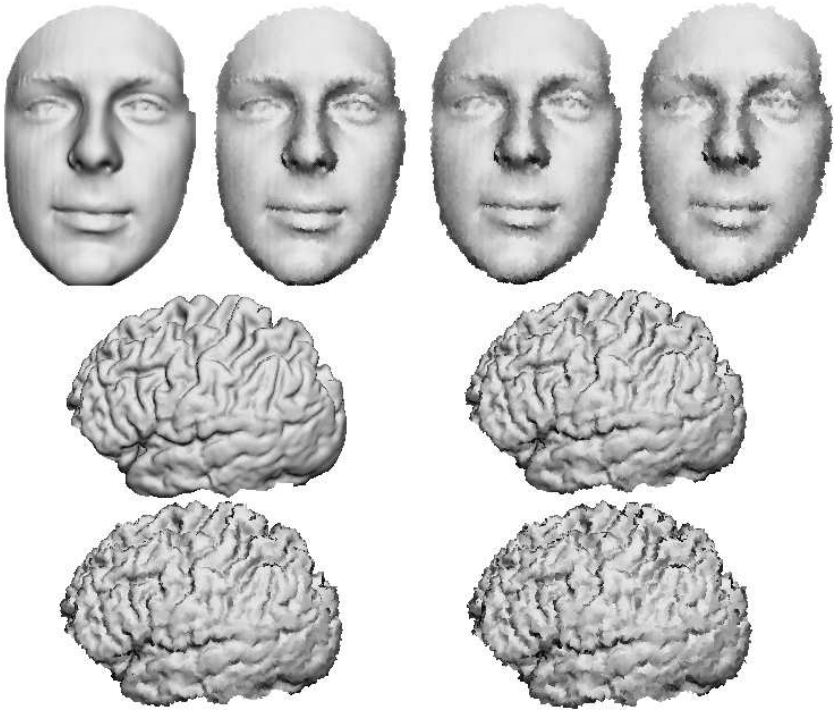


Figure 10: Examples of face and brain surfaces under gaussian noise with different σ set to 0.0, 0.4, 1.0 and 2.0 mm, respectively.

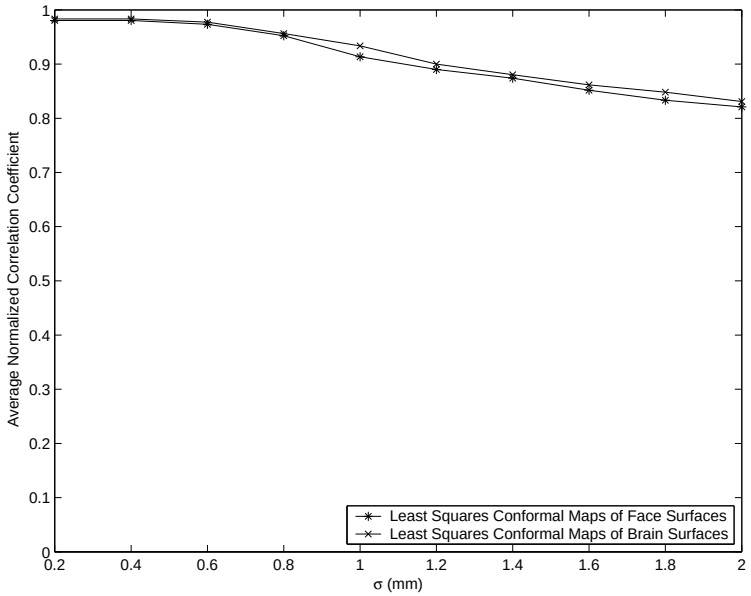


Figure 11: Average matching results of LSCMs under gaussian noise increases. The window size for computing the curvatures of faces surfaces and brain surfaces is 10.0 mm and the σ increases from 0.0 mm to 2.0 mm.

3.3.1.3 Experiment on Resolution Variation

The third experiment tests the robustness of Least Squares Conformal Maps (LSCMs) to resolution changes. Figure 12 shows examples of 3D face and brain surfaces with resolution variation, where all the meshes have the same shape but different resolution. The surfaces with low resolution are matched to the original surfaces and average matching results using the LSCMs are shown in Figure 13. Results show that LSCMs achieve fairly stable matching results and impervious to resolution changes. A small deterioration of the matching results is due to the use of a discrete curvature approximation, since approximation error increases as the resolution drops.

3.3.1.4 Comparison Between Quasi-conformal Maps

For completeness purposes, we also performed comparison experiments between several related work of quasi-conformal maps, including least squares conformal maps, conformal maps [32] and harmonic maps [83, 89], to confirm the

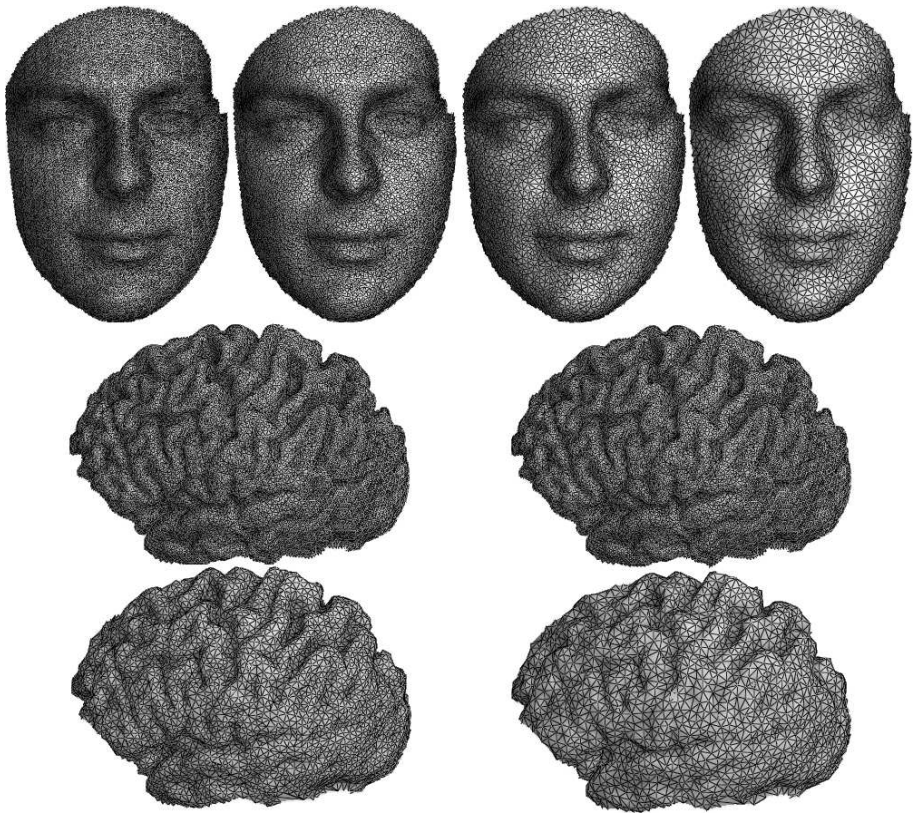


Figure 12: 3D face and brain surfaces with 1, 1/2,1/4 and 1/8 of the original resolution, respectively.

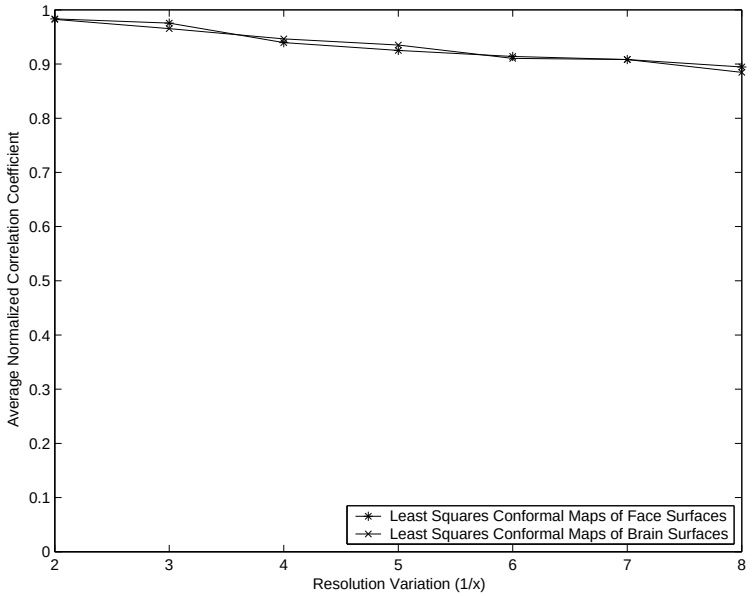


Figure 13: Average matching results of LSCMs under resolution variation.

conclusion in Section 2.1.1.4. Average matching results of the face and brain surfaces using the above three parametric maps under occlusion, noise and resolution variation are shown in Figure 14, 15 and 16, respectively. In Figure 14, since the harmonic maps require satisfaction of the surface boundary condition as discussed in section 2.1.1.1, the performance of harmonic maps is more impacted than the performance of conformal maps and least squares conformal maps. Instead, changes of boundary have very small effects on both conformal maps and least square conformal maps. From the results in Figure 15 we can see that all three maps appear robust to gaussian noise. However, since conformal maps depend on 2 feature points only, which might be detected with errors caused by the noise, they have lower matching rates than the harmonic maps and the least square conformal maps. Finally, Figure 16 shows that the above three parametric maps achieve fairly stable matching results and all of them are impervious to resolution changes.

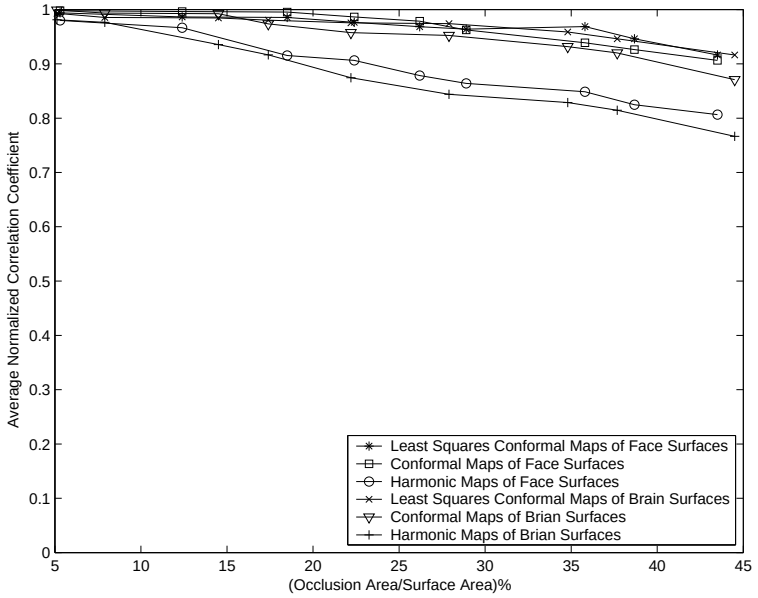


Figure 14: Average matching results of the face and brain surfaces under occlusion using all three parametric maps.

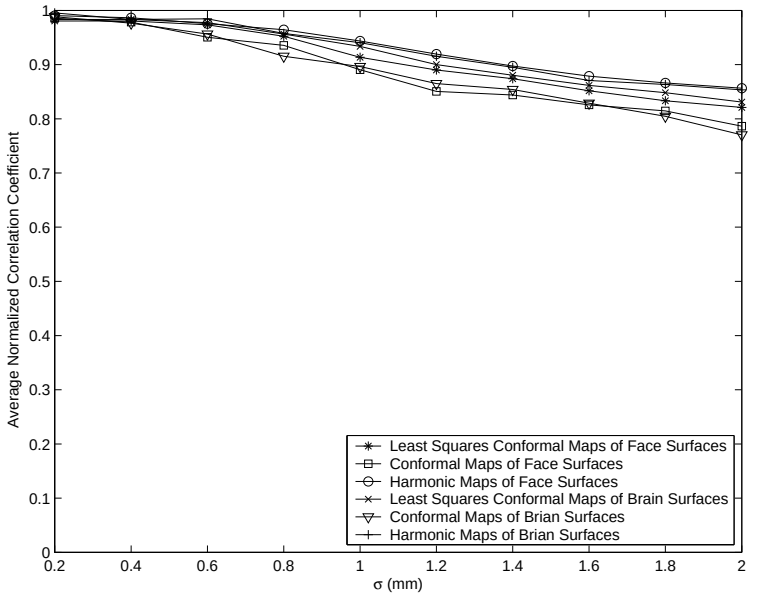


Figure 15: Average matching results of all three parametric maps under gaussian noise increases. The window size for computing the curvatures of faces surfaces and brain surfaces is 10.0 mm and the σ increases from 0.0 mm to 2.0 mm.

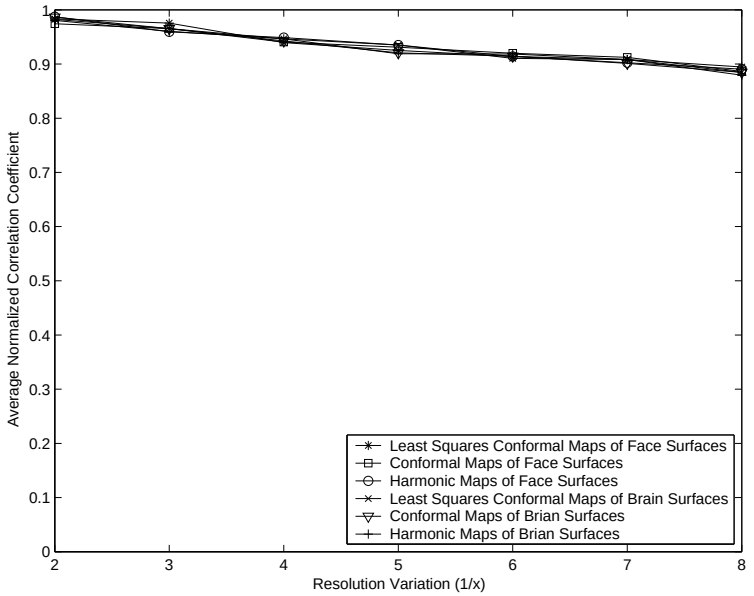


Figure 16: Average matching results of all three parametric maps under resolution variation.

3.3.2 Recognition of 3D Faces

In this section, we apply Least Squares Conformal Maps (LSCMs) to 3D face recognition on a 3D face database which contains 100 3D face scans from 10 subjects. The data are captured by a phase-shifting structured light ranging system in different time [91]. Each face has approximately 80K 3D points with both shape and texture information available (example face data from two subjects in the database are shown in Figure 17). In order to further evaluate our recognition method, we also perform a comparison with other existing methods, including the surface curvature technique [77] and the spherical harmonic shape contexts [26]. For the computation of curvatures from 3D surfaces we had to chose the size of the neighborhood for the surface fit. Clearly, choosing the mask size is a trade-off between reliability and accuracy. When choosing a small mask curvature computation will be strongly affected by noise, due to the small number of points considered for regression. The reliability of the curvature estimation can be improved by increasing the size of the mask. However, a large mask size will produce an incorrect result in the area curvature changes quickly. In our experiments, we used a mask size of 10×10.

The spherical harmonic shape contexts descriptor is computed using the method developed in [26], based on 3D shape contexts. The 3D shape contexts technique is the straightforward extension of 2D shape contexts [54], to three dimensions. The support region for a 3D shape contexts is a sphere centered on the basis point p and its north pole oriented with the surface normal estimate N for p. The support region is divided into bins by equally spaced boundaries in the azimuth and elevation dimensions and logarithmically spaced boundaries along the radial dimension. Based on the histogram from 3D shape contexts, we use the bin values as samples to calculate a spherical harmonic transformation for the shells and discard the original histogram. The descriptor is a vector of the amplitudes of the transformation, which are rotationally invariant in the azimuth direction, thus removing the degree of freedom. We compute the spherical harmonic shape contexts representations in 64×64 grids sampled evenly along the directions of longitude and latitude with bandwidth $b = 16$.

In each experiment, we randomly select a single face from each subject for the gallery and use all the remaining faces as the probe set. The average recognition results from 15 experiments (with different randomly selected galleries) are reported in Table 2. From the recognition results, we can see that the least squares conformal maps perform 10.7% better than the spherical harmonic shape contexts and 14.3% better than the surface curvature technique even if only the shape information is used. Moreover, least squares conformal maps allow to combine both shape and texture information, which improves the accuracy of 3D face recognition.

Table 2: Recognition results of least squares conformal maps, spherical harmonic shape contexts and surface curvature technique.

Recognition Result	Least Squares Conformal Maps	Spherical Harmonic Shape Contexts	Surface Curvature
Using shape information only	97.3%	86.6%	83.0%
Using texture information only	98.0%	N/A	N/A
Using both shape and texture	98.4%	N/A	N/A

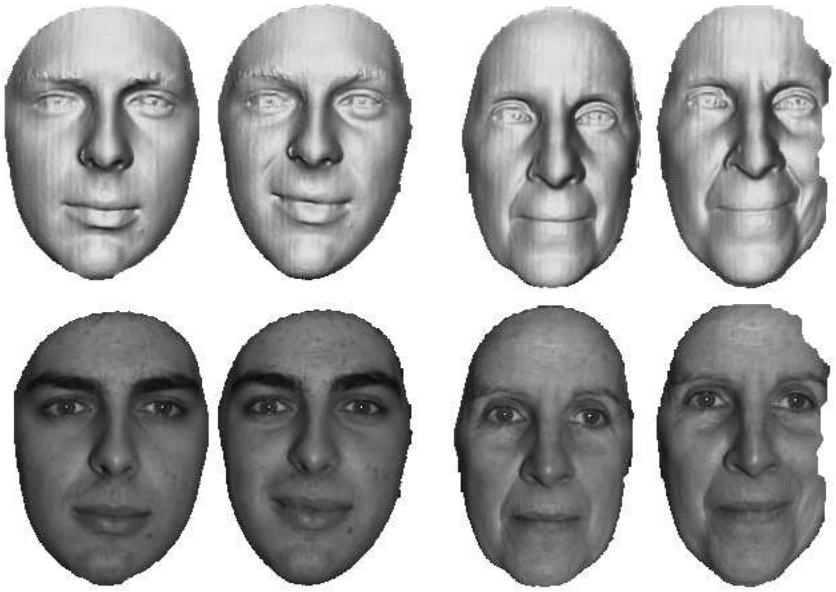


Figure 17: Two subjects in the 3D face database. Shape information is in the first row and texture information is in the second row.

3.3.3 Non-Rigid Surface Alignment and Stitching

In this section, we apply the Least Squares Conformal Maps(LSCMs) to another application: 3D non-rigid surface alignment and stitching. A very important property, which governs our alignment and stitching algorithm, is that the LSCMs can establish a 2D common parametric domain for the 3D surfaces. Therefore we can simplify the 3D surface alignment and stitching problem to a 2D registration and stitching problem. Furthermore, because the LSCMs is a diffeomorphism, i.e., one-to-one and onto, we can detect and remove the duplicated regions in the original 3D surfaces by removing the overlapping areas in the resulting 2D common parametric domain. After that, we can stitch the 3D surface patches by connecting the exclusive regions in the resulting LSCMs. There is a lot of research on 3D surface remeshing [1, 2, 6, 61], but in our case the problem is simplified to a 2D triangulation problem by connecting the neighboring patches in the 2D common parametric domain. As an example, Figure 18 demonstrates the alignment and stitching of two 3D surfaces undergoing non-rigid deformations. 3D faces are captured by a phase-shifting structured light ranging system [91] and each face has approximately 80K 3D points with both shape and texture information available. The subjects were not

asked to keep their head and facial expression still during the 3D face scanning.

Furthermore, Figure 19 shows another example of the accurate face alignment and stitching result of our method on two 3D scans of one face undergoing different transformations and deformations. The leftmost column shows the two input 3D face scans with texture. The same 3D face scans without texture information are shown in the second column. The Least Squares Conformal Shape Images (LSCSIs) of both 3D scans are in the third column. Their aligned LSCSIs and the resulting stitched 3D faces are in the fourth column. Because of the one-to-one mapping between the LSCSI and original face, we can align and stitch 3D faces by registering and stitching 2D LSCSIs.

In order to demonstrate the performance of our method, we also compare our results to the results from the Iterative Closest Point (ICP) method [65] in Figure 20. Figure 20(a) shows a 3D scan of a neutral face, while Figure 20(b) shows a 3D scan of the same face undergoing a large deformation in the mouth area. From Figure 20(c) and (d) which are the front view of (a) and (b), we can see the occlusion area clearly. The face alignment and stitching result of our method is in Figure 20(f) with the close up view of mouth area in Figure 20(h). The result of the ICP method is in Figure 20(e) with the close up view in Figure 20(g). As we can see, in the close up view Figure 20(g), there is a redundant region in the result because the ICP method failed to detect the overlapping areas between deformed surfaces and can only register two surface with rigid transformations. However, as can be seen in Figure 20(g) and (h), our method correctly aligns even at areas of significant local deformations.

3.4 Discussion

In this chapter, we presented a family of quasi-conformal maps, including harmonic maps, conformal maps and least squares conformal maps, and proposed a fully automatic and novel 3D shape matching framework using least squares conformal shape images – a new shape representation which simplified the 3D surface matching problem to a 2D image matching problem. The performance of least squares conformal maps was evaluated vis-a-vis other existing techniques in 3D face recognition and 3D non-rigid surface alignment and stitching. Furthermore,

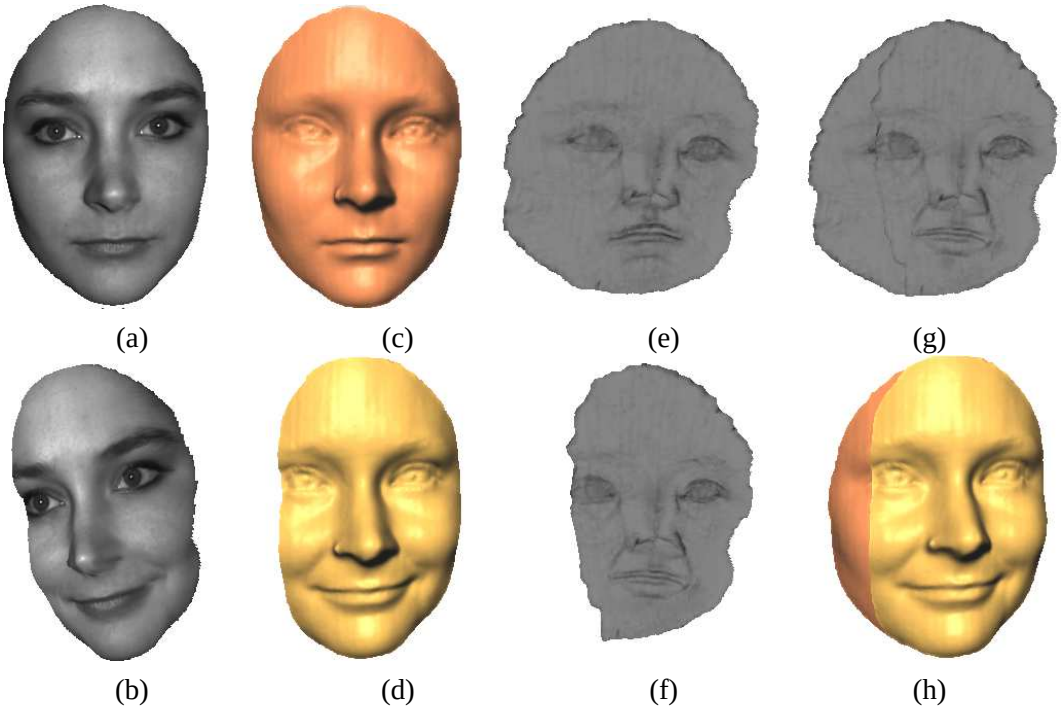


Figure 18: An example of surface alignment and stitching: (a,b) Two original 3D faces with texture in different poses and deformations. (c,d) Original 3D faces without texture. (e,f) The Least Squares conformal Shape Images (LSCSIs) of the faces. (g) The aligned LSCSI of the two faces. (h) The resulting 3D face by stitching a part of (c) into (d). Because of the one-to-one mapping between the LSCSI and original face, we can align and stitch 3D faces by registering and stitching 2D LSCSIs.

our comparison results have shown that all above three parametric maps are robust to occlusion, noise and different resolutions and that the least squares conformal mapping is the best choice for 3D surface matching.

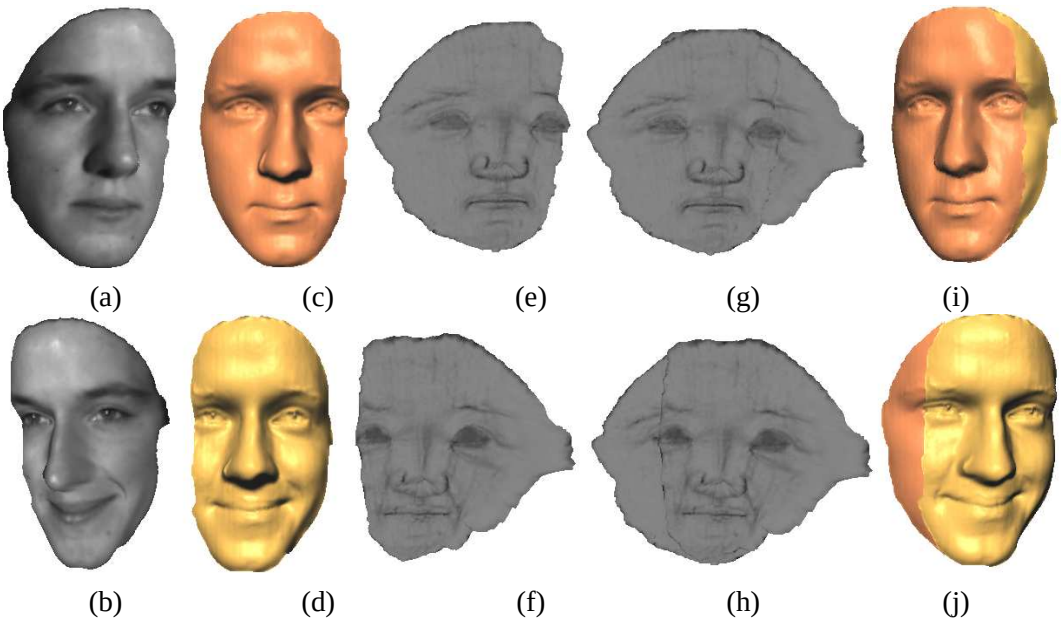


Figure 19: Another example of surface alignment and stitching: (a,b) Two original 3D faces with texture in different poses and deformations. (c,d) Original 3D faces without texture. (e,f) The Least Squares conformal Shape Images (LSCSIs) of the faces. (g) The aligned LSCSI of the two faces by connecting the non-overlapping area in (f) into (e). (h) The aligned LSCSI of the two faces by connecting the non-overlapping area in (e) into (f). (i) The resulting 3D face by stitching a part of (d) into (c). (j) The resulting 3D face by stitching a part of (c) into (d). Because of the one-to-one mapping between the LSCSI and original face, we can detect and remove the duplicated regions in the original 3D surfaces by removing the overlapping areas in the resulting 2D common parametric domain. The user can decide which of the two expressions to keep on the final stitched mesh. In this case (i) has the expression of original (a) and (j) of original (b).

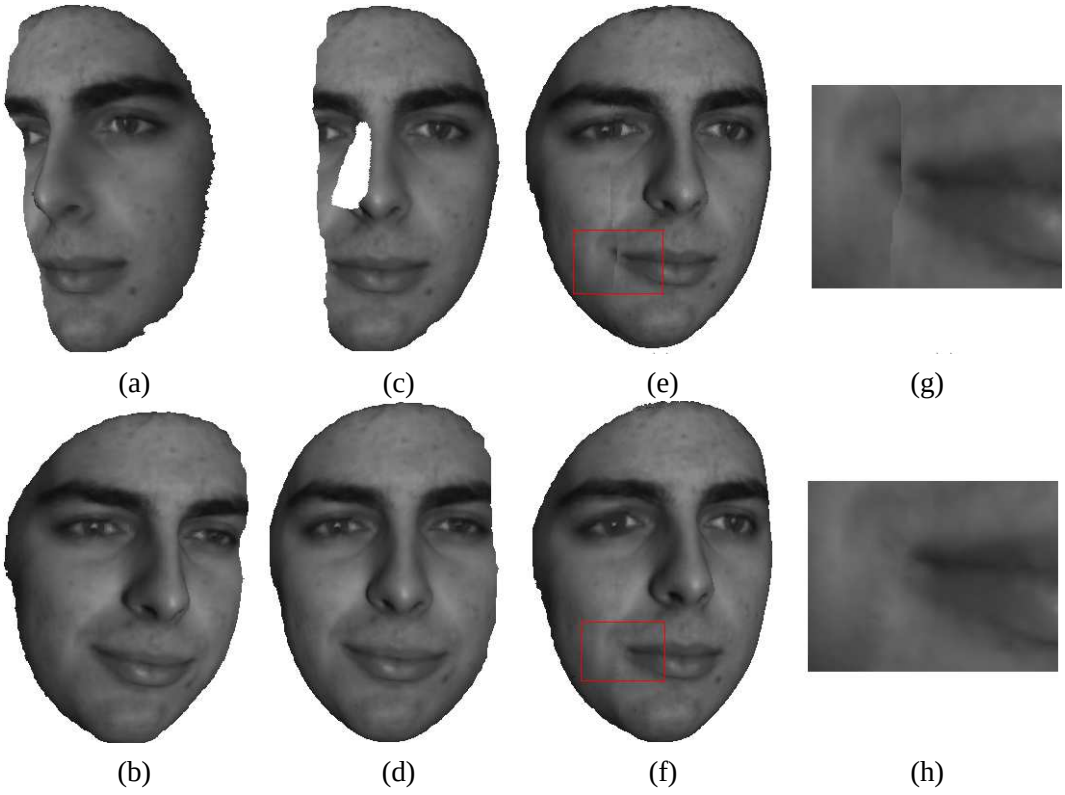


Figure 20: A comparison between the alignment and stitching result of our method and of the ICP method: (a) A 3D scan of a neutral face. (b) A 3D scan of the same face undergoing a large deformation in the mouth area. (c) and (d) are the front view of (a) and (b) with the occlusion area shown clearly. (e) The face alignment and stitching result of the ICP method. (g) The close up view in the mouth area of (e). (f) The face alignment and stitching result of our method. (h) The close up view in the mouth area of (f).

Chapter 4

Shape Registration and Analysis Using Ricci Flow

This work is the first application of surface Ricci flow in computer vision. We demonstrate that previous methods based on conformal geometries, such as harmonic maps and least-square conformal maps, which can only handle 3D shapes with simple topology are subsumed by our Ricci flow based method which can handle surfaces with arbitrary topology. Our Ricci flow based method can convert all 3D problems into 2D domains and offers a general framework for 3D surface analysis. Large non-rigid deformations can be registered with feature constraints, hence we introduce a method that constrains Ricci flow computation using feature points and feature curves. Finally, we demonstrate the applicability of this intrinsic shape representation through standard shape analysis problems, such as 3D shape matching and registration.

This work has been published in the proceedings of the IEEE International Conference on Computer Vision 2007 [31].

4.1 Introduction

Ricci flow is a powerful curvature flow method in Riemannian geometry. In particular, 3-manifold Ricci flow has been successfully applied to prove the Poincaré conjecture recently [59]. In this chapter, we introduce Ricci flow as a

novel 3D shape representation for computer vision and graphics applications. We are motivated by the fact that Ricci flow can handle arbitrary topologies, allowing the mapping of any 3D surface to 2D domain and its ability to handle large 3D shape deformation.

In recent decades, there have been a lot of researches to develop surface representations for 3D surface analysis, which is a fundamental issue for many computer vision and graphics applications, such as 3D shape registration, partial scan alignment, 3D object recognition, and classification [10, 40, 64, 87]. However, matching surfaces undergoing non-rigid deformation is still a challenging problem, especially when data is noisy and with complicated topology. Different approaches include curvature-based representations [77], regional point representations [15, 43, 64, 72], spherical harmonic representations [26, 27], shape distributions [58], harmonic and conformal shape images [79, 81, 89], physics-based deformable models [75], Free-Form Deformation (FFD) [39], and Level-Set based methods [52]. However, many surface representations that use local shape signatures are not stable and cannot perform well in the presence of non-rigid deformation. Conformal geometric maps have been used in several applications of computer vision and graphics. In [89], Zhang et al. propose harmonic maps for surface matching. In [83], Wang et al. use harmonic maps to track dynamic 3D surfaces. However, in order to calculate harmonic maps the surface boundary needs to be identified and a boundary mapping from 3D surfaces to the 2D domain needs to be created which can be a difficult problem. In [32, 81], conformal maps are used for face and brain surface matching. Levy et al. [49] use least squares conformal maps for texture atlas generation, and Sharon et al. [68] analyze similarities of 2D shapes using conformal maps. In [79], Wang et al. analyze a family of quasi-conformal maps for 3D shape matching, such as harmonic maps and least squares conformal maps. However, the conventional conformal geometric methods have the drawbacks such that they can only handle surfaces with simple topologies or compute simple maps. As a result, most existing algorithms are limited to surfaces with simple topology such as genus zero with/without a single boundary. In contrast, our method can handle surfaces with *arbitrary* topologies for shape analysis.

In this chapter, we introduce a new 3D non-rigid surface analysis framework based on Ricci flow conformal mapping. Surface Ricci flow offers a novel means to

manipulate shapes by curvatures and Riemannian metrics. With surface Ricci flow, the curvature evolves like a heat diffusion process on the surface and converges exponentially fast to a constant value. During the whole process, the angle structure on the surface is preserved, and the final surface can be embedded in one of the canonical domain, such as the sphere, the plane, or the hyperbolic space. By computing conformal maps using the Ricci flow method, each 3D surface, even with a complicated topology (e.g. having multiple holes), can be mapped to a 2D domain through a global optimization. The resulting map does not have any singularities and is a diffeomorphism, i.e., *one-to-one* and *onto*. These maps are stable, insensitive to resolution changes, and robust to noise. Hence, the original 3D surface-matching problem simplifies to a 2D image-matching problem of the conformal geometric maps, which is a better understood problem [5, 51, 54].

The previous conformal map methods are subsumed by Ricci flow. Hence, our framework is more general, while we can take advantage of the significant body of work for 3D surface analysis using previous conformal map methods [32, 68, 79, 81, 83, 89]. To integrate feature constraints in Ricci flow computation. Taking advantage of meaningful features is essential for any matching or registration method. In the case of large non-rigid deformations, matched features allow accurate description of the deformations. Thus, in order to make Ricci flow applicable to computer vision problems, we develop: 1) a representation of feature points and feature curves suitable to our framework; 2) a novel feature based metric; 3) an algorithm which, based on features, decomposes the surface into conformal patches; and 4) an algorithm to embed these patches onto the plane (On the target canonical domains, the entire curvature is concentrated on feature points and feature curves are mapped to straight boundary lines). This association of feature points with target domain curvature is novel and has broader implications for geometric modeling and graphics. Finally, we provide initial experiments that demonstrate the potential of our method in a broad range of 3D shape analysis applications such as 3D shape matching and registration in a variety of data sets including face scans and biomedical data.

4.2 Generalization of Conformal Maps

Conventional conformal geometric methods, such as harmonic maps, least squares conformal maps (LSCMs), and methods based on holomorphic forms, can be unified by Ricci flow. In order to clarify this point, we first compare Ricci flow with conventional methods. Then, we briefly introduce the discrete Ricci flow and show the intrinsic connection between Ricci flow and other conformal geometric maps such as harmonic maps.

4.2.1 Comparing with Conventional Conformal Map Methods

In general, conformal mapping algorithms can be classified to the following categories. The first class computes maps from the surface to the plane, such as harmonic map method [83, 89], least squares conformal maps (LSCMs) [49, 79], spherical conformal maps [4, 32]; The second class aims at computing the derivatives of maps, such as the method based on holomorphic forms [33]; The third class computes the conformal metrics to induce conformal maps. The Ricci flow method belongs to the third class, which is more general than the other two classes.

The pull-back metric of a conformal map is a conformal metric on the surface, which induces zero curvature in the entire interior of the target domain. Conversely, the map can be recovered by its pull-back metric directly. Since the curvature is zero everywhere in the map, the pull-back metric can be computed using Ricci flow by specifying the target curvature 0 everywhere. Therefore, any conformal maps (with zero target curvature) which can be computed using either harmonic maps, LSCMs or holomorphic 1-form method can be computed by Ricci flow.

The main difference between Ricci flow and conventional methods is in how much complicated topologies they can handle. In particular, the algorithms in the first class can handle surfaces with genus 0 with/without a single boundary, but can not handle high genus cases. The algorithms in the second class can handle all topologies, but they can not compute the conformal maps between multi-holed annuli, as shown in Figure 25, which are frequently encountered in the scanning process. Furthermore, if the target surface has arbitrary curvature, only Ricci flow can find the map. Therefore, Ricci flow is much more general, and all algorithms in the other two categories can only handle a subset problems handled by Ricci flow.

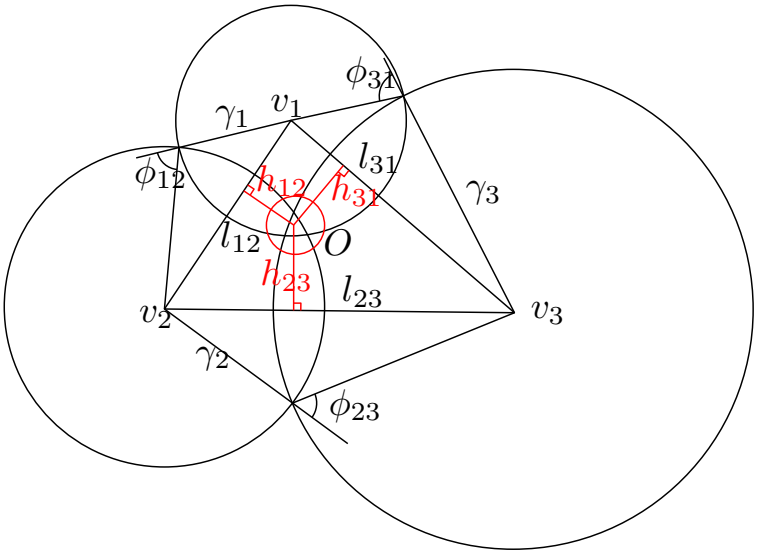


Figure 21: Circle packing metric for a triangle. The dual circle (the red one) in orthogonal to the other 3 circles.

Ricci flow can address situations that can not be handled by other existing algorithms, such as Iterative Closest Point (ICP) [66] and level set based methods [52]. ICP can find a good rigid alignment in $\mathbb{R}^3$, but for non-rigid surface deformations, such as the bending deformation shown in Figure 23, ICP can not find a good registration between two surfaces. Level set based methods are powerful tools for surface analysis. However, it is required that the surface deformation process performed explicitly in $\mathbb{R}^3$ at each step, which will be difficult for surfaces with complex topology since the deformed surface may not be able to be embedded in $\mathbb{R}^3$.

4.2.2 Discrete Ricci Flow

Conventional Ricci flow is defined on C^2 smooth surfaces. In this section, we focus on the discrete approximation of Ricci flows on triangular meshes [14], which is robust for polygonal meshes with sharp corners. Discrete Ricci flow is useful for handling noisy data sets in real applications, as shown in the heart registration example in Section 4.5.

The key observation about the discrete Ricci flow is that the conformal metric

deformation can be treated as a local scaling transformation, which preserves angles and transforms an infinitesimal circle to an infinitesimal circle. Therefore, the general idea of the discrete Ricci flow is to cover the mesh by many circles centered at the vertices. Each circle has a cone angle at the vertex, which can be treated as the discrete curvature. Then, by adjusting the circle radii, we can deform the Riemannian metric of the mesh in a discrete conformal way [14]. The change of the circle radii is the analogy to the change of the conformal factor u. The relation between the discrete curvature and the discrete conformal factor is *exactly same* as that in the smooth case.

Suppose M is a mesh with boundary ∂M, a circle packing metric (M,Γ,Φ) is shown in Figure 21, where Γ and Φ represent the radius function of circles on each vertex and the intersection angle between two circles at one edge, respectively. Each edge length is determined by using the cosine law with the radii of two circles and the intersection angle on the edge. The vertex curvature $k(v)$ measures the flatness of its neighborhood, which is defined as $2\pi - \sum_i \alpha_i$ for an interior vertex and $\pi - \sum_i \alpha_i$ for a boundary vertex, where α_i's are the corner angles surrounding a vertex v. As in the smooth case, the discrete version of Gauss-Bonnet formulae holds, $\sum_v k(v) = 2\pi\chi(M)$. Two circle packing metrics (M,Γ_1,Φ_1) and (M,Γ_2,Φ_2) are conformal, if and only if $\Phi_1 \equiv \Phi_2$.

Therefore, the discrete Ricci flow can be defined in the same way

$$\frac{d\gamma_i(t)}{dt} = -2k_i\gamma_i(t), \tag{17}$$

which converges to constant curvature under the constraint that the total area of the mesh is fixed. If we define the *discrete conformal factor* as $\mathbf{u} = \{\ln\gamma_1, \ln\gamma_2, \cdots, \ln\gamma_n\}$ and the prescribed target curvature as $\bar{\mathbf{k}} = \{\bar{k_1}, \bar{k_2}, \cdots, \bar{k_n}\}$, then the general discrete Ricci flow is

$$\frac{du_i(t)}{dt} = \bar{k_i} - k_i(t), \tag{18}$$

which will lead to the desired conformal metric satisfying the desired target curvature.

Actually, Ricci flow is the gradient flow of a specific energy form, *Ricci energy*:

$$E(\mathbf{u}) = \sum_i (\bar{k_i} - k_i)du_i. \tag{19}$$

Ricci energy is convex, and therefore it has a unique global optima $\bar{\mathbf{u}}$, which induces the target curvature $\bar{\mathbf{k}}$. In practice, optimizing Ricci energy is more efficient than computing Ricci flow. The convex Ricci energy can be stably optimized using Newton's method. The Hessian matrix of the energy $E(\mathbf{u})$ can be computed explicitly as follows.

Suppose $f_{ijk} \in F$ is a face on the mesh, there exists a unique circle orthogonal to all three circles at the vertices, shown as the red circle in Figure 21. We denote the center of that circle as o_{ijk}. The distance from o_{ijk} to the edge e_{ij} is denoted as h_{ij}^k. If an interior edge $e_{ij} \notin \partial M$ is shared by two faces f_{ijk} and f_{jkl}, its weight is $w_{ij} = h_{ij}^k + h_{ji}^l$. If the edge is on the boundary, and only adjacent to face f_{ijk}, its weight is $w_{ij} = h_{ij}^k$.

The Hessian matrix of $E(\mathbf{u})$ has the formulae as

$$\frac{\partial^2 E(\mathbf{u})}{\partial u_i \partial u_j} = \begin{cases} -w_{ij} & i \neq j, e_{ij} \in E \\ \sum_k w_{ik} & i = j, e_{ik} \in E \\ 0 & otherwise \end{cases} \tag{20}$$

which is positive definite on the hyperplane $\sum u_i = const$.

Now, we will describe how to obtain the conformal map of a given surface from the conformal metric obtained from Ricci flow. Alg. 1 shows the procedures of computing the desired conformal metric by prescribed curvature by minimizing the Ricci energy. For surfaces with boundaries, we introduce a novel metric computed by Alg. 2, such that all interior points have zero curvature, and the vertices on the same boundary component share the same curvature. Intuitively, this metric will flatten the surface and map all the boundaries to circles, as shown in Figure 23.

Once the metric is obtained, the mesh can be isometrically embedded onto the plane. The isometric embedding is denoted as $\tau : V \to \mathbb{C}$. This embedding in fact is harmonic, namely, it minimizes the harmonic energy $E(\tau) = \sum_{e_{ij} \in E} w_{ij} |\tau(v_i) - \tau(v_j)|^2$, i.e., $\Delta(\mathbf{u})\tau = 0$.

4.3 Feature Based Canonical Domain Decomposition

In practice, it is often useful to add feature constraints, such as point and curve correspondences when comparing 3D shapes. Hence we propose the incorporation

Algorithm 1 Compute Conformal Metric by Prescribed Curvature Using Ricci Flow

while $|\bar{\mathbf{k}} - \mathbf{k}| > \varepsilon$ **do**
  Compute dual circles
  Compute edge weight w_{ij}
  Form the Hessian matrix $\Delta(\mathbf{u})$
  Solve $\Delta(\mathbf{u})d\mathbf{u} = (\bar{\mathbf{k}} - \mathbf{k})$ constrained at $\sum_i du_i = 0$
  Update $\mathbf{u} = \mathbf{u} + d\mathbf{u}$
  Update $\mathbf{k}$
end while

Algorithm 2 Compute Uniform Flat Metric

Compute the boundary components, $\partial M = C_1 \cup C_2 \cup \cdots \cup C_k$.
$\forall v \notin \partial M$, set $\bar{k} \Leftarrow 0$.
$\forall v \in C_j$, $\bar{k}(v) \Leftarrow s_j \frac{2\pi}{|C_j|}$, where $s_1 = 1$, $s_j = -1$ for $j \neq 1$.
while $|\bar{\mathbf{k}} - \mathbf{k}| > \varepsilon$ **do**
  Compute $\bar{\mathbf{u}}$ by $\bar{\mathbf{k}}$ using the Ricci energy algorithm.
  $\forall v \in C_j$, $\bar{k}(v) \Leftarrow \frac{s_j \pi (\bar{l}(e_-) + \bar{l}(e_+))}{\sum_{e \in C_j} \bar{l}(e)}$,
  where $\bar{l}(e)$ is the edge length under $\bar{u}$, and
  e_- and e_+ represent two boundary edges incident to v.
end while

of such constraints to the energy minimization and formulate the main framework of surface matching using Ricci flow in the following commutative diagram,

$$\begin{array}{ccc} S_1 & \xrightarrow{\phi} & S_2 \\ {\scriptstyle\tau_1}\downarrow & & \downarrow{\scriptstyle\tau_2} \\ D_1 & \xrightarrow{\bar{\phi}} & D_2 \end{array} \tag{21}$$

S_1, S_2 are two given original surfaces, $\phi : S_1 \to S_2$ is the desired 3D surface matching. We use Ricci flow to compute $\tau_i : S_i \to D_i$ which maps S_i conformally onto the canonical domain D_i. D_1 and D_2 are simple planar domains. The topology and the curvature of D_1 and D_2 incorporate the major feature information of the original surfaces S_1 and S_2. If there are certain feature constraints, we can further incorporate them using the method described below and compute a map $\bar{\phi} : D_1 \to D_2$. The final map ϕ is induced by $\phi = \tau_2^{-1} \circ \bar{\phi} \circ \tau_1$.

For surfaces with significant point features, we design the target curvature such that those features are transformed to the branch points of the Riemann surfaces of in the target domains. Alg. 3 uses features to design such target domains.

After computing the metric incorporating all the major features using Alg. 4, the surface is decomposed to canonical patches, each of which is mapped to a rectangle or a trapezoid as shown in Figure 22.

As described in Section 4.2.2, each patch is embedded onto the plane by minimizing the harmonic energy with the feature point position constraints as described in Alg. 5.

Algorithm 3 Computing Feature Based Flat Metric

Specify the feature curves
Slice the surface open along the feature curves.
Specify the feature points $\{v_1, v_2, \cdots, v_m\}$.
Compute the boundary components, $\partial M = C_1 \cup C_2 \cup \cdots \cup C_k$.
for each vertex $v \in M$, set $\bar{k} \Leftarrow 0$.
Allocate curvature on feature points, $\bar{k_i} = 2m_i\pi, m_i \in \mathcal{Z}$, $\sum_i m_i = \chi(M)$.
Use Ricci energy optimization to compute the metric.

An example result on human face scan is demonstrated in Figure 22. The original surface is a 2-holed annulus. We select the nose tip as the only feature point,

Algorithm 4 Computing Feature Based Domain Decomposition

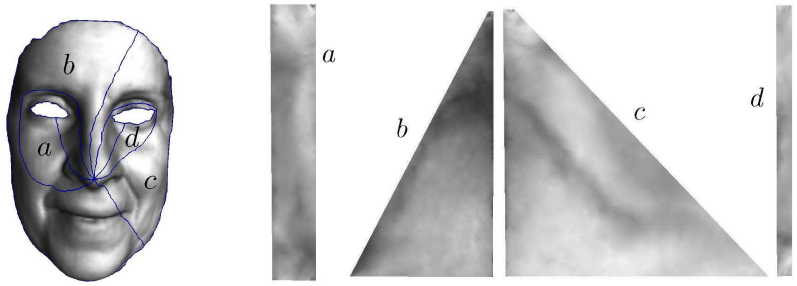
Compute the feature based target metric, such that all boundaries become straight lines under the target metric.
Compute the straight lines starting from the feature points and are perpendicular to the boundaries under the new metric.
Compute the straight lines parallel to boundaries under the new metric.
Slice the surface open along the straight lines to decompose the surface to patches, each of which is conformally mapped to a rectangle or a trapezoid.

Algorithm 5 Computing the Isometric Embedding

On each face, compute the dual circle which is orthogonal to all three vertex circles.
Compute the distance from the center of the dual circle to three edges.
Embed a seed triangle f on the mesh.
Minimize the harmonic energy with the constraints of the embedded triangle f.

and set the target curvature to be zero everywhere (including both the interior points and the boundary points) except for the noise tip, whose curvature equals to -2π. Then, we use Ricci flow to compute the target metric of the Riemann surface, which is a flat surface with a single branch point. Because the target surface can not be embedded in $\mathbb{R}^3$ directly, we decompose it to canonical patches, shown as blue curves in the figure intersecting the boundaries, where all the boundaries are straight lines in the target domain. The decomposition includes three steps: First, under the target metric we find straight lines from the branch point to the boundaries, each of which is perpendicular to a boundary; Next, we trace the straight lines which are parallel to the boundaries under the target metric, shown as curves circling around the eye contours in the figure; Finally, all the straight lines partition the surface to patches and each patch is conformally embedded onto the plane either as a rectangle or a trapezoid. Thanks to the conformal deformation, this decomposition is solely determined by the geometry of the original surface and the choice of features. Therefore, surface matching and registration can be carried out by matching the decomposed patches on the planar domain, while the features are guaranteed to match as they become patch corners or boundaries.

The main reason for the decomposition is to improve the efficiency and accuracy of the method. We convert surface matching to matching between rectangles and trapezoids, as simpler process. Because we incorporate feature constraints to



(a) A face surface with segmenting curves

(b) Planar domain for each patch

Figure 22: Canonical surface decomposition using Ricci flow. The nose tip is selected as a feature point. A flat metric is computed using Ricci flow, such that all interior points and all boundary points are with zero curvatures, except the feature point where the Gaussian curvature equals to -2π. Straight lines under the new metric, which are either parallel or perpendicular to the boundaries, result in the blue curves on the original surface that pass through the feature point in (a). Then the surface is decomposed to patches, each patch is conformally equivalent to a rectangle or a trapezoid on the plane, shown in (b).

the mapping by minimizing harmonic energy, which requires the domain to be convex, the decomposition is also necessary to ensure the convexity and to guarantee the globally optimum solution.

4.4 Ricci Flow Based Shape Representation

In this section, we present a new shape representation for 3D surface analysis, such as shape matching and registration based on Ricci flow, which can handle surfaces with varying boundaries and arbitrary topologies. Moreover, it also allows multiple types of feature constraints, such as feature point constraints, feature curve constraints, and target curvature constraints. Therefore, it provides a unified framework for non-rigid 3D surface analysis.

Ricci Flow Shape and Texture Images The main advantage of the Ricci flow method is that it can convert all 3D problems into 2D domains. By computing conformal maps using the Ricci flow method, each 3D surface, even with a complicated topology (e.g. having multiple holes) can be mapped to a 2D domain through a global optimization. Therefore, we can generate the Ricci flow shape images

by associating a shape attribute with each vertex in the Ricci flow conformal maps. Among the shape attributes, we use mean curvatures to obtain Ricci flow conformal images since mean curvature depends only on surface geometry. In our method, the mean curvature is computed in the same way as in [32]. Moreover, it is also possible to generate other Ricci flow conformal images by associating other attributes such as textures.

Surface Matching with Ricci Flow Representation Given two general surfaces S_1 and S_2, we first compute the Ricci flow shape or texture images. Because the resulting maps do not have any singularities and are a diffeomorphism, i.e., one-to-one and onto, we can register these two 3D surfaces by simply matching and registering with the aligned Ricci flow shape or texture images. We evaluate the accuracy of surface matching by using the error distance between the two resulting maps, as follows

$$normalized\ error_{S_1,S_2} = \frac{\sum_{i=1}^{N} \| p_i^{S_1} - p_i^{S_2} \|}{\sum_{i=1}^{N} \| p_i^{S_1} \|}, \tag{22}$$

where N is the number of overlapping points in the Ricci flow conformal shape or texture images of 3D surface S_1 and S_2, and $p_i^{S_k}$ is the value of point i in the 2D image of surface $S_k (k = 1, 2)$. This is the matching method used in our experiments.

4.5 Experimental Results

In this section, we demonstrate the performance of our framework by several experiments on real 3D data, such as isometrically deformed surfaces, dynamic facial expression with complex topology, and human heart surfaces undergoing complex non-rigid motion deformations.

4.5.1 Experiments on Isometrically Deformed Surfaces

Firstly, as a simple experiment, we test our method on isometrically deforming data. We scanned a flexible (but non-stretchable) toy mask two times, one for the original and another for its deformed version. Since there is no stretching the

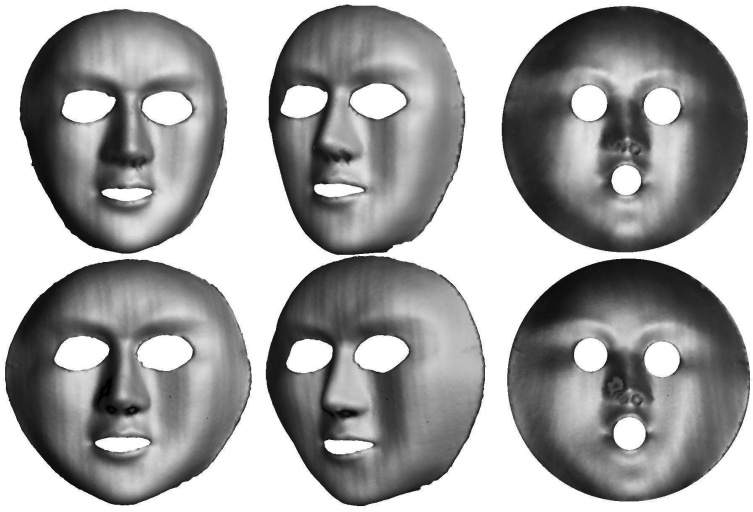


Figure 23: Surface matching under isometric deformation using a toy mask. The first row shows two views of the original surface and its conformal image; the second row shows two views of the deformed surface and its conformal image. Pixel intensities in the conformal images are copied from the corresponding points in the 3D scans. Under isometric deformation, the conformal images are identical. The normalized registration error is 0.018 computed using Equation 22.

deformation is isometric which can be easily handled by Ricci flow, as shown in Figure 23. The conformal images of the two scans are practically identical with normalized error 0.018.

4.5.2 Experiments on Complex Topology

As described in Section 4.2, the major advantage of our method against to the existing 3D surface-matching methods based on conformal geometric maps [49, 79, 83, 89], is that our framework can handle surfaces with arbitrary topology directly. Therefore, for surfaces with multiple holes our method does not require additional pre-processing steps such as hole fillings. Figure 24 shows a comparison between our method and the LSCM-based method [79] and the harmonic map based [83] methods. Figure 24 (a, f) show the original 3D surfaces of the same subject with different expressions, and Figure 24 (b, g) depict the resulting Ricci flow texture images computed by our method. Since the LSCM-based method and harmonic maps can only handle disk topology, the holes in the eye and mouth on the original

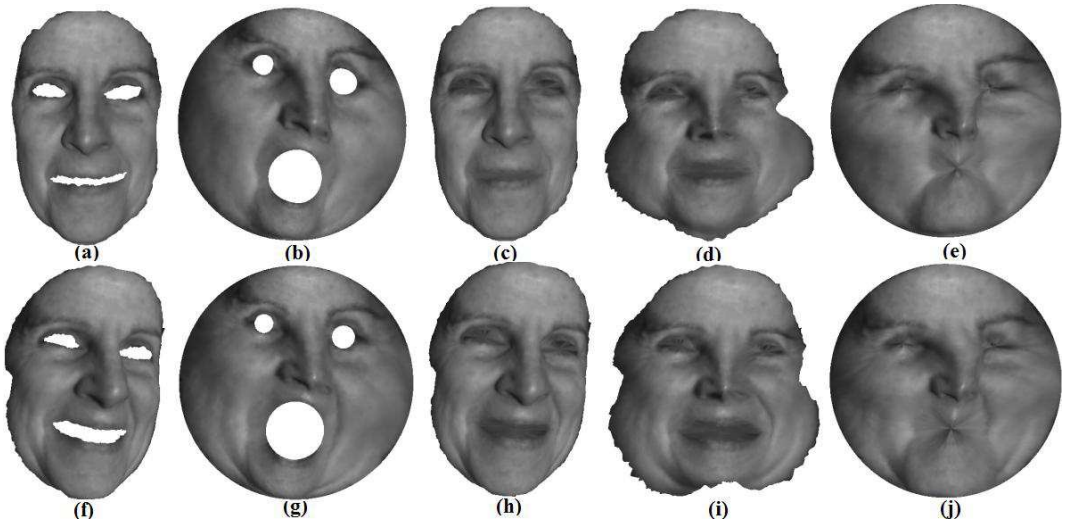


Figure 24: Comparison of Ricci flow with LSCM and Harmonic maps. (a) and (f) are two surfaces to be registered. (b) and (g) are their Ricci flow maps. (c) and (h) are these two surfaces after hole-filling. (d) and (i) are their LSCMs. (e) and (j) are their harmonic maps. The registration error of Ricci flow using Equation 22 is 0.058, while, the registration errors (without including hole area) of LSCMs and Harmonics are 0.072 and 0.081, respectively.

3D scan data need to be filled before computing the 2D conformal map, as shown in Figure 24(c, h). As shown in Figure 24(d, i) and Figure 24(e, j), the introduction of fake geometry to fill the holes leads the large distortion errors around the eye and mouth areas in both of the least-squares conformal maps and the harmonic maps. Notice that our texture images obtained from Ricci flow in have no signification distortions as shown in Figure 24 (b, g), although we leave the holes as they are in the raw data.

The normalized matching error of Ricci flow is 0.058, compared to 0.072 for LSCM and 0.081 for harmonic maps. All errors were computed using Equation 22, where hole areas were not included. For each 3D facial scan with around 100K vertices, The process time of LSCM and harmonic maps is approximately 40 seconds on a Pentium4 2.4 GHz PC, while the process time of Ricci flow is around 400 seconds. Moreover, Our method is robust enough to handle limited amounts of non-isometric deformations (which violate the Ricci flow definition). However large deformations can be handled with incorporation of feature constraints as described in the following experiment.

4.5.3 Registration with Large Non-Rigid Deformation

In order to demonstrate the performance of our method on surfaces with large non-rigid deformation, we captured 3D facial expression data using a phase-shifting structured light ranging system [91] with large non-rigid deformations. Since our method allows feature curve correspondence constraints, we detect the contour of the lips and the eyes and integrate them into the computation of the Ricci flow maps (these curves can be detected by methods such as [53, 88]). Based on the resulting 2D maps, we can perform the registration between two scanned faces with different expressions, as shown in Figure 25. Since the deformation between two scanned faces is non-rigid, the surface matching with single maps is problematic as shown in the 2nd column in Figure 25. In this case, we apply the decomposition method described in Section 4.3. The 3rd-5th columns in Figure 25 illustrate the pairs of the Ricci flow images which corresponds to a part of the original surfaces as shown in the 1st column in Figure 25. Each pair of patches are registered with the corresponding patch boundaries, and we measure the errors between the patches by Equation 22.

The original registration error between two faces in Figure 25 is 0.0447. In order to demonstrate the robustness of the feature detection accuracy in our decomposition method, we randomly perturb the feature point around the nose tips in Figure 25. The average error of three different perturbations within a 3mm (resp. 6mm) radius is 0.045 (resp. 0.048).

Although our method is not limited to handle face data, it is worth to compare our method with the face registration method based on multi-dimensional scaling [7]. Compared to Bronstein et al. [7], which cannot guarantee to obtain global optima in isometric embedding, our method reaches global optima in handling anisometric data with arbitrary topologies, as shown the following experiments.

For medical data application, we use a 3D deforming heart sequence[1]. The original tagging data were acquired using a 3T MRI machine. The data are image sequences from end diastole to end systole. The reconstruction was done based on methods developed by the authors of [35], who made the data available to us. The output from the analyzed data result are 3D corresponding points over time from

[1]This experiment was carried out by Yun Zeng

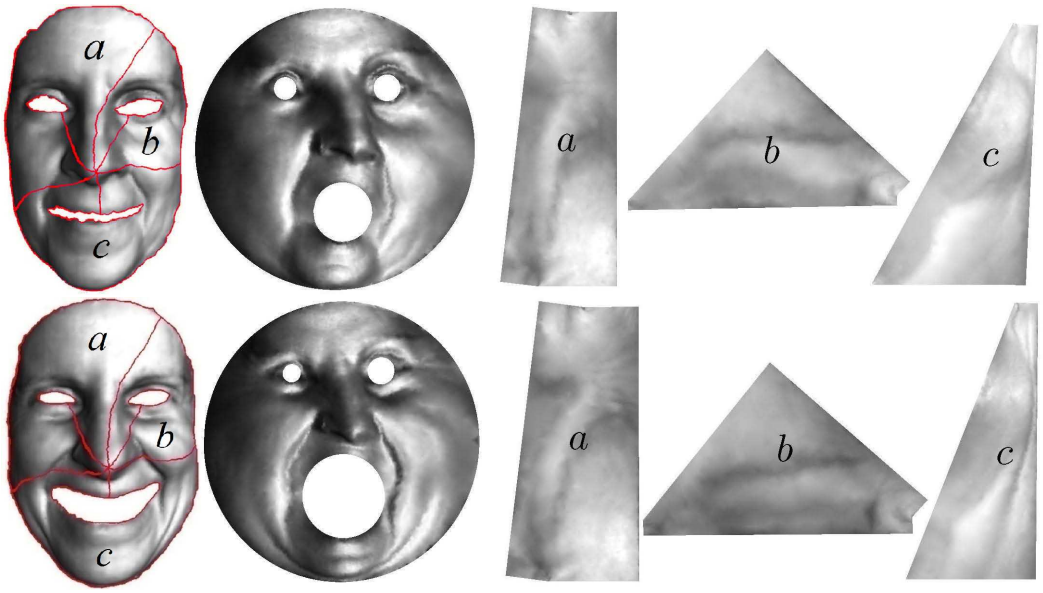


Figure 25: Registration of facial expression data using feature based domain decomposition. The first column shows two face scans with very large non-rigid deformation. The second column shows the planar domains computed using uniform flat metrics. Because of the large deformation, there is significant difference between the planar domains. Selecting the nose tip as the feature points, the surfaces are decomposed to canonical planar domains using the method described in Section 4.3. The surfaces are registered by matching the corresponding planar domains.

end diastole to end systole. We experimented using a sequence of 21 frames of 3D corresponding points. Experiments were performed on the deforming 3D surface. The given 3D correspondences were not used in the experiments, but only as ground truth.

In order to test the robustness of our method to initial surface segmentation, we experiment using only the left ventricle data. We first detect and segment along the boundary between the surface of the left ventricle and the rest of the heart. After segmenting the heart data for each frame, we apply the Ricci flow algorithm to map each heart into its canonical planar domain, and register each adjacent frame by mapping the corresponding planar domains. In a first experiment, we manually defined a boundary on the first frame and consistently kept these points as the boundary points throughout the sequence. Even though there are large interior deformations, the boundary is sufficient in establishing almost perfect surface

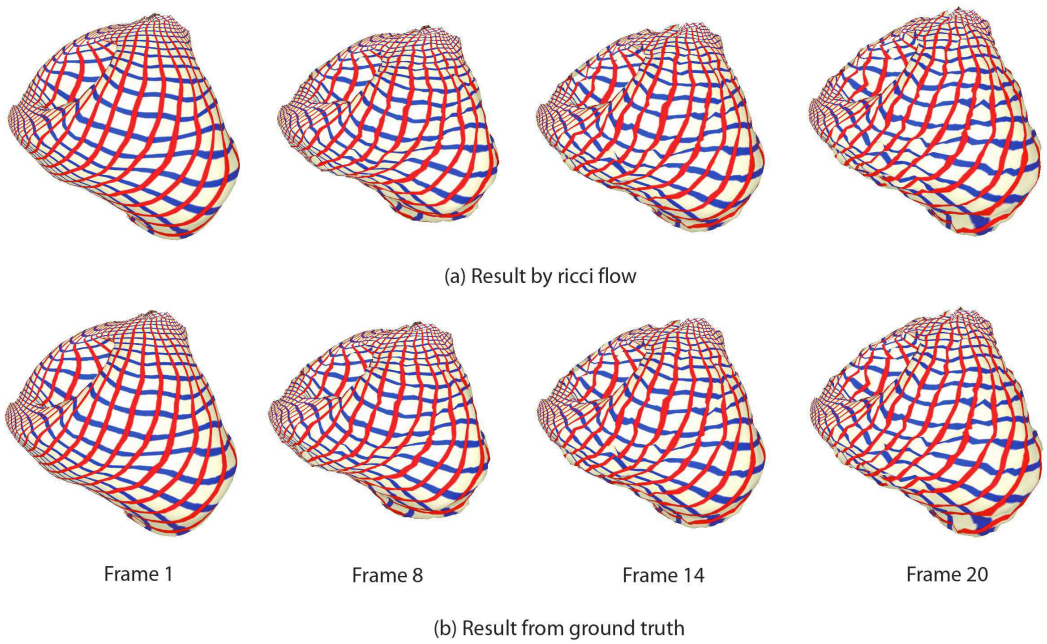


Figure 26: Registration of 3D dynamic heart data. Registration results using Ricci flow for 4 different frames are shown in the top row. The original heart data for the same frames are shown in the bottom row. The data on frame 1 were texture mapped with a grid pattern, that helps to visualize the subsequent deformations. *(Heart dataset courtesy of Professor Dimitris N. Metaxas at Rutgers University)*

correspondences, with an average registration error of 0.006. In the second experiment, the boundary was automatically determined based on curvature, using the VTK software package. These boundary points are not guaranteed to be consistent across frames. The method is still very robust with an average registration error of 0.03.

Figure 26 illustrates the effectiveness of registration using Ricci flow. The first frame is texture-mapped with a grid pattern both in the experimental and ground truth data, in order to better visualize the deformation. Although the non-rigid deformation of the heart is significant between different frames, our method captures the deformation almost indistinguishably from the ground truth.

4.6 Discussion

In this chapter, we proposed a 3D shape analysis method based on surface Ricci flow. Since Ricci flow is a powerful tool to handle geometries with arbitrary topologies our method can unify conventional methods based on conformal geometry. It also allows different types of feature constraints, such as feature point and curve constraints, to handle large deformations and to further improve the accuracy of surface matching and registration. A series of algorithms was introduced to map the 3D surfaces onto canonical 2D domains, and a new surface representation is proposed to combine multiple features for 3D shape analysis. Finally, the generality and flexibility of Ricci flow were demonstrated by various experiments on human face scans and dynamic heart surface data.

Chapter 5

Dynamic Non-Rigid Registration for Facial Expression Analysis

In this chapter, we present a novel framework for automatic non-rigid registration of 3D dynamic facial data using least squares conformal maps with additional feature correspondences detected by employing active appearance models (AAMs). Based on this registration method, we also develop a new system of facial expression synthesis and transfer. We perform a series of experiments to evaluate our non-rigid registration method and demonstrate its efficacy and efficiency in the applications of facial expression synthesis and transfer.

This work has been published in the proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition 2008 [78].

5.1 Introduction

Automatic non-rigid registration of 3D time-varying densely-sampled data is a fundamental and critical issue in 3D vision and graphics which has widespread applications. As 3D scanning technologies continue to improve, 3D dynamic densely-sampled data is becoming more and more prevalent for analysis and synthesis. To study and analyze such huge data, an efficient non-rigid registration algorithm is necessary to establish one-to-one inter-frame correspondences automatically. However, automatic 3D non-rigid registration still remains a challenging task, especially

for dynamic densely-sampled facial expression data with many degrees of freedom. Most of existing 3D non-rigid registration methods rely on recovering low dimensional parameters of face model or register 3D faces with local optimization that may not establish accurate one-to-one inter-frame correspondences successfully. In this chapter, an automatic non-rigid registration algorithm of 3D dynamic densely-sampled facial data is developed using least squares conformal maps with additional interior feature correspondences detected by active appearance model (AAM) [17, 28, 29]. The least squares conformal maps between two 3D surfaces are globally optimized with less angle distortion and the resulting 2D map is stable, one-to-one, insensitive to resolution changes and robust in the existence of noise. Through the way of mapping 3D surfaces to a 2D common domain, it simplifies the original 3D surface-registration problem to a 2D registration problem. Thus, more accurate and efficient non-rigid registration algorithms could be achieved by using least squares conformal maps. In sharp contrast to previous works on 3D non-rigid registration, especially the methods using attached markers, which unavoidably require much laborious human intervention and also are more invasive to human subjects, our new method can register non-rigid 3D dynamic data automatically and efficiently with minimum manual work.

Conformal maps have already been employed in many vision and graphics applications most recently. A surface matching method based on harmonic maps was proposed in [89]. Sharon et al. [68] use conformal maps to analyze similarities of 2D shapes. Moreover, conformal maps are used for 3D face and brain surface matching in [32,81,82]. Least squares conformal maps are introduced by Levy et al. [49] for texture atlas generation and used by Wang et al. [79] to conduct 3D surface matching with feature detection using the technique of spin-image. Because spin-image can only detect features on surfaces with rigid transformation, their method can not guarantee to successfully match surfaces with non-rigid deformation. For non-rigid 3D surface registration, Wang et al. [83] use a modified harmonic map to track 3D high resolution facial motion data. In order to calculate these harmonic maps, the surface boundary must be identified and a boundary mapping from 3D surfaces to the 2D domain must be properly created, which can be a difficult task especially when parts of the surface are occluded. In contrast, conformal maps and least squares conformal maps do not necessarily require boundary information to be

aligned and so give rise to a natural choice to combat this difficulty. Moreover, least squares conformal maps enable users to enforce more interior feature constraints which will guarantee to achieve more accurate registration results in an automatic way.

Facial expression undergos complicated global and local nonlinear deformation between frames and is represented by a high dimensional vector (a collection of 3D vertices). It is impossible to analyze and synthesize facial expression in high dimensional space. In this chapter, we describe a dynamic facial expression synthesis system using isomap [74] which can embed facial expression manifolds in high dimension into low dimensional space. Finally, we present a facial expression transfer framework based on our non-rigid registration method using least-squares conformal maps and our approaches lead to more accurate results with minimum human intervention.

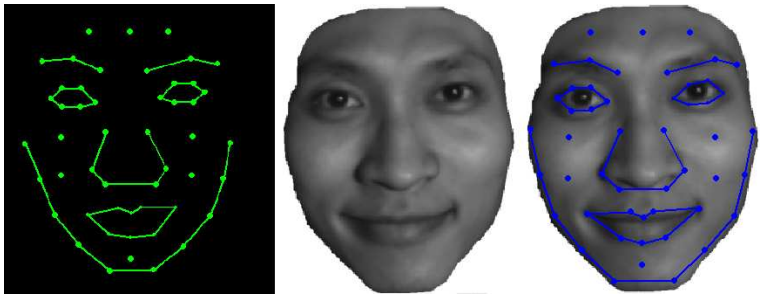
5.2 Non-Rigid Registration Algorithm for 3D Dynamic Facial Data

We now introduce an automatic non-rigid registration algorithm by using least-squares conformal maps which can map 3D surfaces to a 2D common domain with global optimization. Therefore, they can simplify the original 3D surface-registration problem to a 2D registration problem. In particular, our registration algorithm includes two steps: First, interior feature correspondences are detected by using Active Appearance Model (AAM); After that, by generating and registering the 2D least-squares conformal maps of 3D faces in two frames, we compute their dense one-to-one correspondences to register these two frames.

5.2.1 Feature Tracking

There are many features in the human face such as corners of eyes, nose and mouth. Detecting and tracking these features accurately and efficiently in 3D dynamic facial data still presents difficulties. Active Appearance Model (AAM) [17, 28, 29] is successfully used to track facial features in video sequences. AAM is a face detection technique that combines shape and texture information into one

PCA space. The model iteratively searches a new image by using the texture residual to update the model parameters. To use AAM to detect features in 3D dynamic facial data, firstly, we use a projection matrix P to project the 3D faces onto a 2D image plane. Then we use AAM to detect the features in each 2D video frame. After that, with the known projection and depth information of these 3D data, we can project the features detected by AAM back to 3D face surfaces. Finally, we can automatically get the initial inter-frame feature correspondences in these 3D dynamic data. In experiments, we select 200 frames in training data containing different facial expressions to build the AAM and the facial feature template contains 50 vertices, as shown in Figure 27.



(a) (b) (c)

Figure 27: AAM feature detection. (a) The feature template of AAM. (b) A 3D face projected onto an image plane. (c) The detected features on the face.

5.2.2 Dynamic Non-Rigid Registration

After detecting the initial corresponding features in two frames S_i and S_{i+1}, we can compute their least squares conformal maps (LSCMs) using the method described in Section 2.1.1.3. As the LSCMs are driven by representative motion features between the two frames, they capture the inter-frame non-rigid deformation. Furthermore, because this mapping is *one-to-one and onto*, by registering their 2D LSCMs, we can recover the inter-frame registration on these 3D face surfaces.

As an example, Figure 28(c,f) show the LSCMs of the inter-frame 3D faces in Figure 28(a,d). Figure 28(a,d) are the original faces with texture information and Figure 28(c,f) are their registered 2D LSCMs. The similarity of these two LSCMs

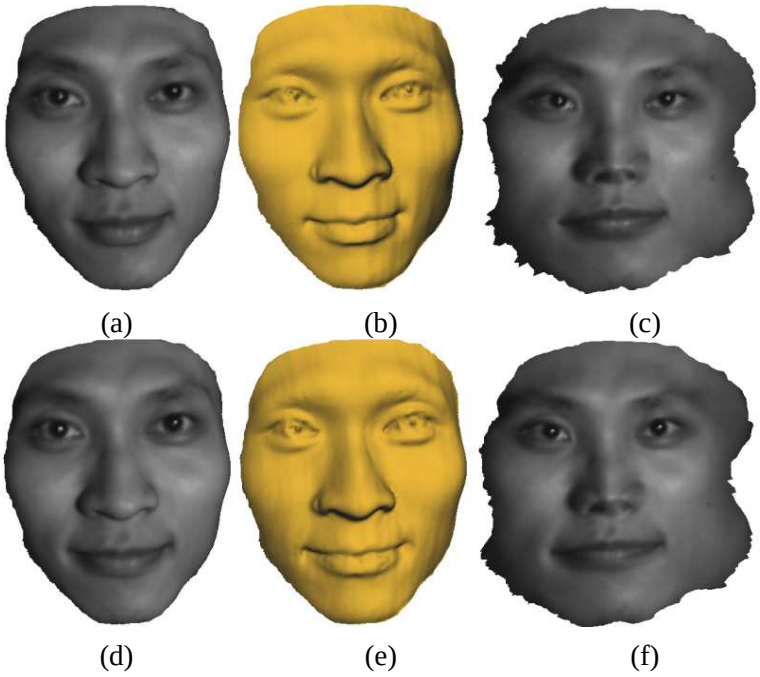


Figure 28: Registration using least squares conformal maps (LSCMs). (a) and (d) are two original inter-frame 3D face surfaces with texture information. (b) and (e) are these faces without texture. (c) and (f) are their registered LSCMs.

in Figure 28(c,f) shows that we can successfully register two inter-frame non-rigid 3D faces by just registering their 2D LSCMs.

5.3 A Framework of Facial Expression Synthesis and Transfer

We now present the new framework of dynamic facial expression synthesis and transfer based on our non-rigid registration method.

5.3.1 Facial Expression Synthesis

Expression synthesis generates new facial animations using existing expression data. Our expression synthesis framework includes two steps: The first step

analyzes existing expression data by embedding them into a low-dimensional manifold using Isomap [74] after registering these data using our 3D non-rigid registration method described in section 5.2. The second step synthesizes new expression by selecting parameters of these expression data analyzed in the first step.

5.3.1.1 Facial Expression Manifold Embedding

Facial expression undergos complicated global and local nonlinear deformation between frames. In order to analyze expression data easily and efficiently, we need embed facial expression manifolds non-linearly into a low dimensional space. We adapt Isomap framework [74] to achieve a low dimensional manifold embedding for individual facial expressions that provides a good representation of facial motion. Isomap finds the best embedding manifold with nonlinear dimensionality reduction by preserving the proportion of distance in the embedding space and the original facial motion space. Figure 29 shows the embedding of smile motion to a 3D space. It is an elliptical one dimensional manifold in 3-dimensional space. In embedding space, the expression manifolds are elliptical curves with distortions according to face geometry, expression types. To analyze these expression manifolds, we need align these one dimensional manifolds in embedding space. For each manifold, correspondences are initially established using the points with high curvatures. Then, multiple manifolds are aligned using an approach similar to [16]. Thus, we can align the original expression sequences in temporal space by aligning expression manifolds in the embedding space.

5.3.1.2 Dynamic Expression Synthesis

After we align N expression styles $s_1, s_2, ..., s_n$ of the same person using method described above, we then generate a new style vector s_{new} by linear interpolation of these N styles using control parameters $w_1, w_2, ..., w_n$ as follows:

$$s_{new} = w_1 s_1 + w_2 s_2 + ... + w_n s_n, \tag{23}$$

where $\sum_{i=1}^{N} w_i = 1$. For example, if we want to generate new expression as style with 50% of the first style and 30% of the second style and 20% of the third style, then we generate new style as $s_{new} = 0.5s_1 + 0.3s_2 + 0.2s_3$.

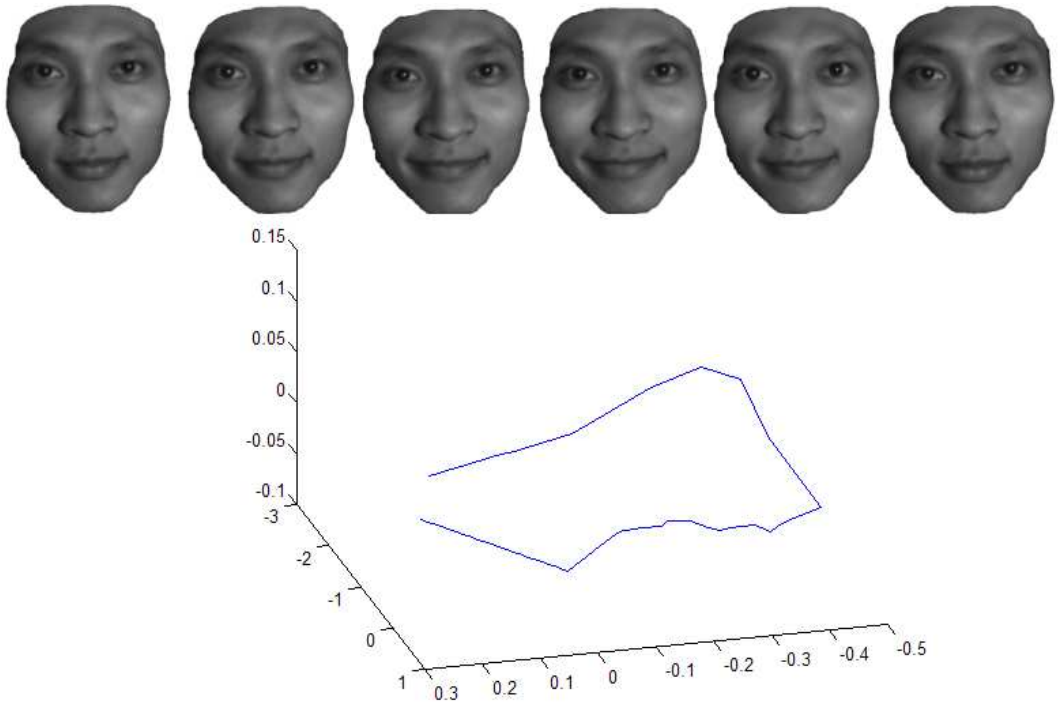


Figure 29: Facial expression manifold. The curve is Isomap for 3D registered facial expression sequence(some frames are shown in upper row).

5.3.2 Facial Expression Transfer

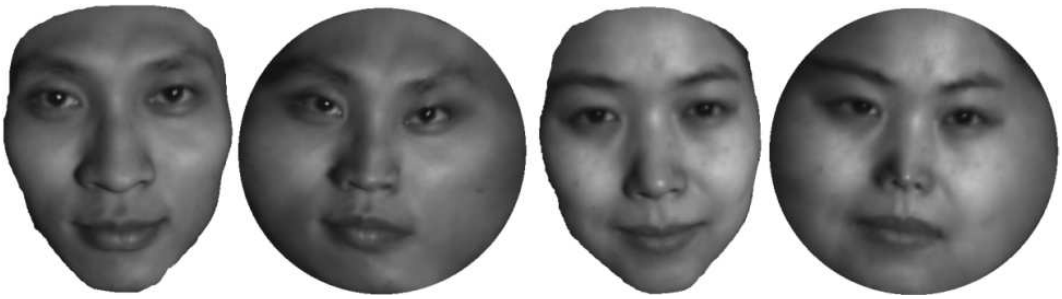
Expression transfer directly maps expressions of the source model to the target model. In particular, our expression transfer framework includes two steps: The first step determines temporal correspondences between every two adjacent frames of the source model and spatial correspondences between the source and target models; The second step transfers the adjusted motion vectors from source model vertices to target model vertices.

5.3.2.1 Dense Surface Correspondences

Source models at each frame do not have temporal inter-frame correspondences. In addition, source model and target model do not have spatial correspondences as they may have different structures. However, we can establish both temporal and spatial correspondences by using parameterization methods [22,25,36,49] to map 3D source and target models to a 2D domain. Therefore, we can compute 3D dense surface correspondences by just detecting correspondences in their 2D maps.

Temporal Correspondences: In our experiments, we use fine facial motion data which are captured by a structured lighting method [91] with 30 frame per second. A 3D face in each frame has approximately 70K points with both shape and texture information. To utilize this 3D dynamic data, we use our 3D non-rigid registration method described in Section 5.2 to obtain the one-to-one inter-frame correspondences, as shown in Figure 28.

Spatial Correspondences: For expression transfer, it is crucial to find spatial correspondences between the source and target models. Harmonic mapping is a popular approach for recovering dense surface correspondences [22, 46]. However, difficulties arise when specific points need to be matched exactly between models. Our approach to finding spatial correspondences starts with initial corresponding feature points which the user specifies [46] between the source and target models. After that, we simplify the source and target models and map them to a 2D plane by minimizing the harmonic energy [22, 32, 89] with user-specified corresponding feature points as interior constraints. By detecting and interpolating the one-to-one



(a) (b) (c) (d)

Figure 30: Spatial feature correspondence detection using harmonic maps. (a) and (c) are source and target faces. (b) and (d) are their harmonic maps computed by our method. After detecting the one-to-one correspondences in their 2D harmonic maps, we can obtain the spatial feature correspondences between 3D source and target faces.

correspondences in the 2D harmonic maps, we can obtain the spatial correspondences between the source and target models, as shown in Figure 30.

5.3.2.2 Expression Transfer with Motion Vectors

A transferred expression animation displaces each target vertex to match the motion of a corresponding surface point in source model. Since facial geometry and aspect ratios are different between the scans of source models and the target face, source displacement vectors can not be simply transferred without adjusting the direction and magnitude of each motion vector. In our experiments, we adjust both the scale and orientation of motion vectors before transferring the source motion to target model by using the method described in [56]. An example of motion vector transfer is shown in Figure 31.

5.4 Experimental Results

The performances of our approaches on non-rigid registration of 3D time-varying data and facial expression synthesis and transfer are evaluated in a number of experiments. First, we analyze the accuracy of our 3D non-rigid registration method and compare results with two previous methods. Second, we evaluate the

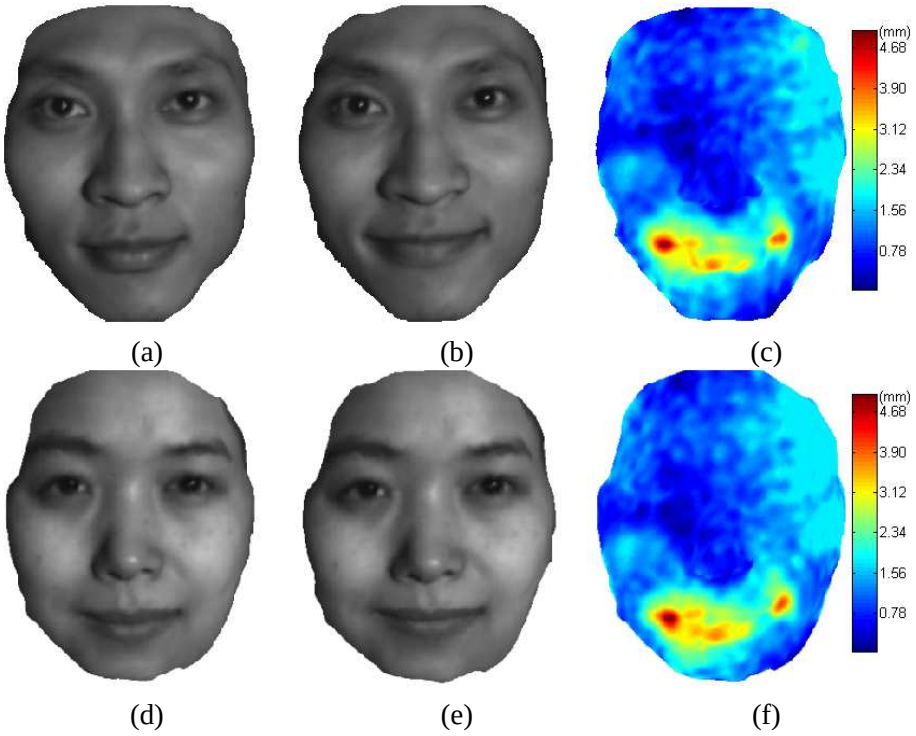


Figure 31: An example of motion vector transfer. (a) and (b) are source faces with different expressions. (c) is the color-coded magnitude of motion vectors in the source model. (d) is the target face model. (e) is transferred expression on the target face. (f) is the color-coded magnitude of motion vectors to be transferred to the target face model (d).

performance of facial expression synthesis and transfer based on our non-rigid registration method.

5.4.1 Evaluation of 3D Non-Rigid Registration

We apply our non-rigid registration method on 3D dynamic facial data and compare results with the tracking method based on modified harmonic maps [83] and Iterative Closest Point (ICP) method [65] which have been widely used for 3D registration. In order to evaluate their accuracy, we compute the registration error by approximately using the difference in the intensity values of vertices of registered 3D face surfaces between two frames as:

$$RegistrationError = \frac{\sum_{i=1}^{N} \|t_j^i - t_{j+1}^i\|}{\sum_{i=1}^{N} t_j^i}, \tag{24}$$

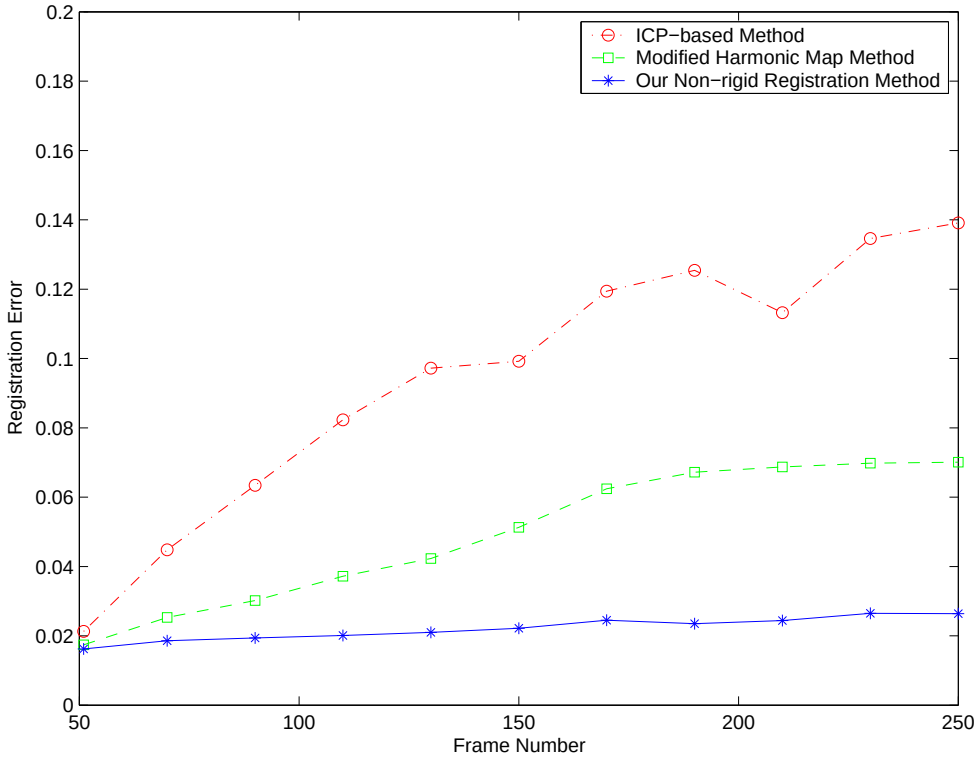


Figure 32: Comparison of the three registration methods.

where t_j^i is intensity value of the ith vertex of 3D face surface in the jth frame and N is the number of registered vertices. If the registration is perfect, the only difference in the intensity values of vertices of registered two 3D faces will result from the change of shadowing and shading effects due to geometric deformation.

We present the comparison of these three techniques in Figure 32 by plotting the registration errors according to different frames. From the results, we can see that our method performs considerably better than the other two methods. The ICP method can not achieve good results in 3D non-rigid shape registration. The modified harmonic map method uses optical flow to track very few feature points which are very sensitive to noise. Moreover, their method will have larger registration errors in the 3D face data with varying boundary, because of the limitation of harmonic maps.

5.4.2 Evaluation of Facial Expression Synthesis and Transfer

Firstly, We apply our facial expression synthesis framework on 3D dynamic facial data to synthesize new facial expressions. Actors perform four different type of expressions: smile, surprise, sad and angry. The expressions were captured using our structured lighting ranger scanner. We then registered and analyzed the captured ranger data using our facial expression synthesis framework described in section 5.3.1. Figure 33 shows the generation of two expressions: smile and surprise and the synthesis of a new in-between expression by changing the weight of these two original input expressions. With our method we can generate a convincing combination of two different expressions without loss of details. The generated in-between expressions are shown in the second and third rows.

Next, we apply our facial expression transfer framework on facial data with different expressions and transfer these expression styles and details to target face models. We perform two group experiments to evaluate the accuracy and robustness of our facial expression transfer method both qualitatively and quantitatively. Our first group experiments are intended to qualitatively show the effectiveness of our expression transfer approach. Figure 34 shows the expression transfer results with various exaggerated expressions and Figure 35 shows the results with different kinds of expressions which are neutral, happy, surprise, sad and angry. We also perform expression transfer from the source model to a topologically different target face model caused by missing data in eye regions during data acquisition under different resolutions and the results are shown in Figure 36. As shown in these results, the expressions of the source model are reproduced in the target model with convincingly better effects.

The second group experiments are intended to quantitatively measure the effectiveness of our expression transfer approach. In the third experiment, we use two different 3D scans of the male subject in Figure 35 as source and target models, respectively, that is, transferring expressions from a person to himself. In the last experiment, we transfer expressions of the male subject to the female subject in Figure 35 and then transfer intermediate results back to another 3D scan of the male subject. By using Equation (24), the average errors of intensities are measured between the original and final face models in all frames as shown in Table 3. Figure 37 exhibits some of these expression transfer results in different frames.

Figure 33: Synthesis of new facial expression by weighting two different expression type: smile in the first row and surprise in fourth row. Second row: $70\%smile + 30\%surprise$. Third row: $30\%smile + 70\%surprise$.

Figure 34: Exaggerated expression transfer. Source face model with exaggerated expressions are shown in the first row. Transferred expressions on two target faces are showed in the second and third row, respectively. The target faces have different shapes and textures but the expressions are proportionally scaled to fit each model well.

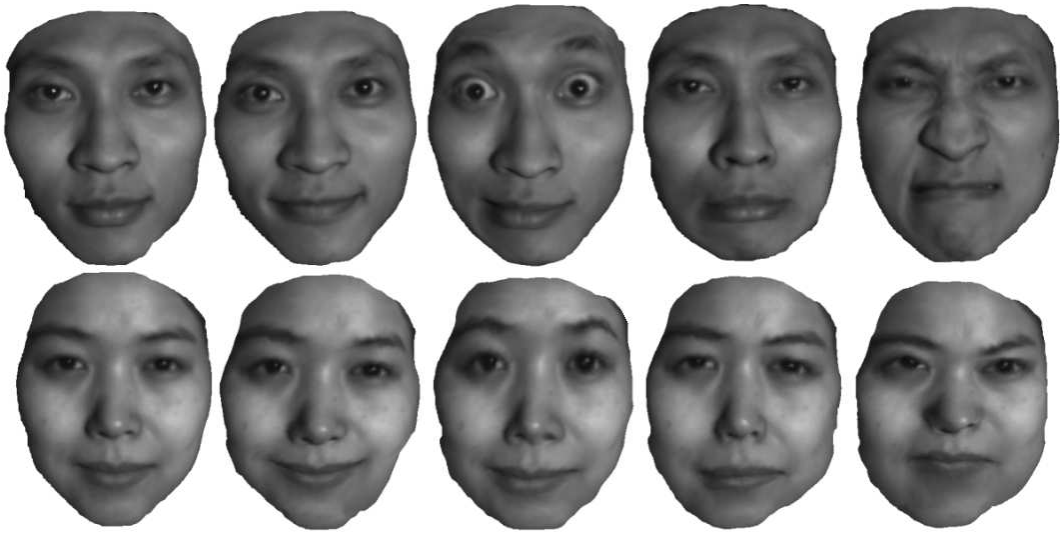


Figure 35: Expression transfer. Source face model with different expressions are shown in the first row. Transferred expressions on the target face are shown in the second row. From left to right, emotional expressions are neutral, happy, surprised, sad and angry, respectively.

Table 3: Average errors of expression transfer.

	Man ⇒ Man	Man ⇒ Woman ⇒ Man
Average RegistrationError	2.312%	2.379%

From the results, we can see that in each frame, the finial faces after expression transfer are very similar to the original source face data and the only difference results from the change of the shadowing and shading effects due to face geometry deformation. The overall processing time including 3D non-rigid registration and expression transfer is approximately 1 minute per frame on a Pentium4 2.4 GHz PC. From all of these results, comparing with the previous research on expression transfer which typically require many manual labors, our method can transfer expression from one person to another efficiently and automatically.

Figure 36: Expression transfer from a male subject to a topologically different face model under different resolutions. Source face model with different expressions are shown in the first row. Transferred expressions on the target face which has different topology due to missing data (missing in eye region during data acquisition) are shown in the second row. Expression transfer results of the target face with only 1/4 of the original resolution are shown in the third row.

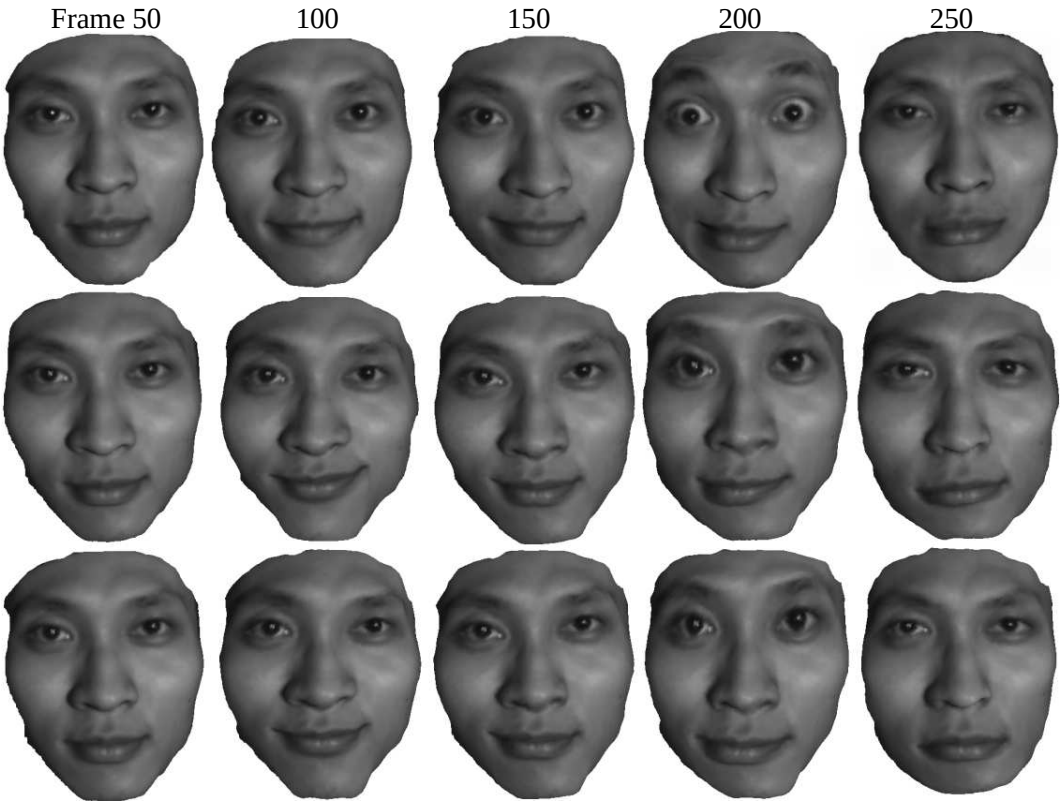


Figure 37: Expression Transfer results (Man ⇒ Man and Man ⇒ Woman ⇒ Man). Source face models in different frames are shown in the first row. Expression transfer results (Man ⇒ Man) are shown in the second row. Expression transfer results (Man ⇒ Woman ⇒ Man) are shown in the third row.

5.5 Discussion

We have developed a novel method for non-rigid registration using least squares conformal maps to automatically compute one-to-one inter-frame correspondences for 3D time-varying facial data. Moreover, based on this registration method, we have also implemented a new visual modeling framework of expression synthesis and transfer for 3D dynamic facial data. Our experimental results demonstrate that our novel facial modeling framework leads to better registration for 3D dynamic facial data and subsequent applications such as dynamic facial expression analysis, synthesis and transfer.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

In this dissertation, I present our research results and future research directions within our generalized framework of shape registration and analysis. Our proposed framework takes the advantages of new shape representations by using different parameterization technology which can map 3D surfaces to a 2D common domain, and thus can simplify all 3D problems to 2D image problems.

In particular, We analyze a family of quasi-conformal maps including harmonic maps, conformal maps and least squares conformal maps with regards to 3D shape matching. As a result, we propose a novel and computationally efficient shape matching framework by using least squares conformal maps. The robustness of least square conformal maps is evaluated and analyzed comprehensively in 3D shape matching with occlusion, noise and resolution variation. In order to further demonstrate the performance of our proposed method, we also conduct a series of experiments on two computer vision applications, i.e., 3D face recognition and 3D non-rigid surface alignment and stitching. We show that previous methods based on conformal geometries, such as harmonic maps and least squares conformal maps, which can only handle 3D shapes with simple topology are subsumed by our Ricci flow based method which can handle surfaces with arbitrary topology. The solution to Ricci flow is unique and its computation is robust to noise. Large non-rigid deformations can be registered with feature constraints, hence we introduce a method that

constrains Ricci flow computation using feature points and feature curves. We also demonstrate the applicability of this intrinsic shape representation through standard shape analysis problems, such as 3D shape matching and registration. Moreover, we present a new method for automatic non-rigid registration of 3D dynamic facial data using least squares conformal maps, and based on this registration method, we also develop a new framework of facial expression synthesis and transfer. A non-rigid registration algorithm of 3D dynamic facial data is developed by using least squares conformal maps with additional feature correspondences detected by employing active appearance models. We also perform a series of experiments to evaluate our non-rigid registration method and demonstrate its efficacy and efficiency in the applications of facial expression synthesis and transfer.

6.2 Future Work

There are many avenues for possible future work, including exploring the theoretical foundation of shape representation, designing efficient and accurate algorithms for shape registration, developing new functionalities, improving shape analysis in wide range of applications such as human facial and body motion analysis, medical imaging, volumetric data reconstruction and visualization, 3D surveillance, 3D robotics, and etc.

6.2.1 Novel Shape Representations

With development of 3D scanning technologies, large 3D shape databases require automated methods for matching and registration. However, matching surfaces undergoing non-rigid deformation is still a challenging problem, especially when data is noisy and with complicated topology. The conventional conformal geometric methods can only handle surfaces with simple topologies or compute simple maps. As a result, most existing algorithms are limited to surfaces with simple topology such as genus zero with/without a single boundary [31]. Thus, novel and effective representations for 3D shapes registration need to be discovered using new parameterization methods such as hyperbolic Ricci flow, which can handle surfaces with arbitrary topologies for shape analysis and registration.

6.2.2 New Shape Retrieval Frameworks

Shape matching and retrieval is a fundamental issue and still a challenging task in computer vision and graphics. By discovering new shape representations, we can develop novel and efficient shape matching and retrieval frameworks with many applications, such as shape registration, partial scan alignment and stitching, 3D object recognition and classification. Moreover, our shape matching and retrieval frameworks also can be applied in biometrics and homeland security, such as 3D human face matching, retrieval, and recognition. Our new shape representations and matching technique will definitely benefit the registration and recognition in these areas.

6.2.3 Non-Rigid Shape Registration with Manifold Learning

Trackers based on template matching work well for objects whose shape and appearance change only slightly during motion. However, in most real-world applications, the objects being registered or tracked can undergo a variety of complex non-rigid deformations that are difficult to parameterize using traditional methods. After employing the non-rigid registration method to get dense one-to-one correspondences in each frame [78, 84], it will be possible to construct a manifold of transitions to deformed objects. Therefore, these manifolds will provide us with a great opportunity to analyze object motion and deformation. There are many applications of this methodology such as human face and body motion analysis and animation, 3D surveillance and robotics.

6.2.4 Dynamic Shape Analysis for New Medical Imaging Modalities

Ultrasound, electron microscopy, and diffusion tensor MRI are imaging modalities which require different assumptions about the underlying physical processes of image formation. The 2D or 3D time-varying images should be registered because of image variation. Recovering registrations for each modality which accurately model the object transformation while ignoring other effects, such as imaging noise and no-rigid deformation, will provide methods both for merging the output

of disparate imaging techniques and allowing for new patient protocols, such as high-resolution gated-MRI without patient breath-holding, which is difficult with current technology.

In the long term, my research will focus on theoretical shape representation, registration, and analysis in computer vision and graphics with applications to shape retrieval, human facial and body motion analysis, animation, medical imaging. This work incorporates ideas from image processing, machine learning, human-computer interface and biomedical engineering.

6.3 Concluding Remarks

These directions for future work, and the many other open problems that exist, are sure to encourage interesting and exciting research in shape registration and analysis for years to come. As technical difficulties are overcome, and existing computational algorithms are improved, the applications of shape registration and analysis will increase in variety and number. We are pleased to have taken the first step in uncovering the heretofore untapped potential of shape registration and analysis by presenting our framework to the computer vision and graphics. It is our hope that this integrated approach and demonstrated applications will foster continued interest and research in this area. We are looking forward to the continued exploration of shape registration and analysis and predict a successful future for it.

Bibliography

[1] P. Alliez, D. Cohen-Steiner, O. Devillers, B. Levy, and M. Desbrun. Anisotropic polygonal remeshing. *ACM Transactions on Graphics*, pages 485–493, 2003.

[2] P. Alliez, M. Meyer, and M. Desbrun. Interactive geometry remeshing. In *SIGGRAPH '02*, pages 347–354, 2002.

[3] G. D. Anderson, M. K. Vamanamurthy, and M. K. Vuorinen. *Conformal Invariants, Inequalities, and Quasiconformal Mappings*. Wiley-Interscience, March 1997.

[4] S. Angenent, S. Haker, A. Tannenbaum, and R. Kikinis. On the laplace-beltrami operator and brain surface flattening. *IEEE Trans. Med. Img.*, 18(4):700–711, 1999.

[5] V. Athitsos, J. Alon, S. Sclaroff, and G. Kollios. Boostmap: a method for efficient approximate similarity rankings. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages II: 268–275, 2004.

[6] M. Botsch, C. Rössl, and L. Kobbelt. Feature sensitive sampling for interactive remeshing. In *Proceedings of the 2000 Conference on Vision Modeling and Visualization*, pages 129–136, 2000.

[7] A. Bronstein, M. Bronstein, and R. Kimmel. Three dimensional face recognition. *International Journal on Computer Vision*, 64(1):5–30, 2005.

[8] B. Brown and S. Rusinkiewicz. Non-rigid range-scan alignment using thin-plate splines. In *Symposium on 3D Data Processing, Visualization, and Transmission*, 2004.

[9] V. Camion and L. Younes. Geodesic interpolating splines. In *EMMCVPR01*, pages 513–527, 2001.

[10] R. Campbell and P. Flynn. A survey of free-form object representation and recognition techniques. *Computer Vision and Image Understanding*, 81:166–210, 2001.

[11] M. D. Carmo. *Differential Geometry of Curves and Surfaces*. Prentice Hall, 1976.

[12] J. Chai, J. Xiao, and J. Hodgins. Vision-based control of 3d facial animation. In *Symposium on Computer Animation*, pages 193–206, 2003.

[13] C. Chen, Y. Hung, and J. Cheng. RANSAC-based DARCES: A new approach to fast automatic registration of partially overlapping range images. *IEEE Transection on Pattern Analysis and Machine Intelligence*, 21(11):1229–1234, 1999.

[14] B. Chow and F. Luo. Combinatorial ricci flows on surfaces. *Journal Differential Geometry*, 63(1):97–129, 2003.

[15] C. Chua and R. Jarvis. 3d free-form surface registration and object recognition. *International Journal on Computer Vision*, 17:77–99, 1996.

[16] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Comput. Vis. Image Underst.*, 89(2-3):114–141, 2003.

[17] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. In *ECCV98*, pages 484–498, 1998.

[18] D. DeCarlo and D. Metaxas. The integration of optical flow and deformable models with applications to human face shape and motion estimation. In *IEEE International Conference on Computer Vision and Pattern Recognition*, page 231, 1996.

[19] M. Desbrun, M. Meyer, and P. Alliez. Intrinsic parameterizations of surface meshes. In *Eurographics02*, pages 209–218, 2002.

[20] C. Dorai, J. Weng, and A. K. Jain. Optimal registration of object views using range data. *IEEE Transection on Pattern Analysis and Machine Intelligence*, 19(10):1131–1138, 1997.

[21] M. Eck, T. DeRose, T. Duchamp, H. Hoppe, M. Lounsbery, and W. Stuetzle. Multiresolution analysis of arbitrary meshes. In *SIGGRAPH'95*, pages 173–182, 1995.

[22] M. Eck, T. DeRose, T. Duchamp, H. Hoppe, M. Lounsbery, and W. Stuetzle. Multiresolution analysis of arbitrary meshes. In *SIGGRAPH95*, pages 173–182, 1995.

[23] J. Eells and J. H. Sampson. Harmonic mappings of riemannian manifolds. *Amer. J. Math.*, 86:109–160, 1964.

[24] O. Faugeras and M. Hebert. The representation, recognition and locating of 3d object. *Int. Journal of Robotics Research*, 5(3):27–51, 1986.

[25] M. S. Floater and K. Hormann. Surface parameterization: a tutorial and survey. In *Advances in Multiresolution for Geometric Modelling*, pages 157–186. Springer, 2004.

[26] A. Frome, D. Huber, R. Kolluri, T. Bulow, and J. Malik. Recognizing objects in range data using regional point descriptors. In *ECCV04*, May 2004.

[27] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs. A search engine for 3d models. In *ACM Transactions on Graphics*, pages 83–105, 2003.

[28] R. Gross, I. Matthews, and S. Baker. Generic vs. person specific active appearance models. *Image and Vision Computing*, 23(11):1080–1093, November 2005.

[29] R. Gross, I. Matthews, and S. Baker. Active appearance models with occlusion. *Image and Vision Computing*, 24(6):593–604, 2006.

[30] X. Gu and B. C. Vemuri. Matching 3d shapes using 2d conformal representations. *MICCAI, Lectures Notes in Computer Science*, 3216:771–780, 2004.

[31] X. Gu, S. Wang, J. Kim, Y. Zeng, Y. Wang, H. Qin, and D. Samaras. Ricci flow for 3d shape analysis. In *IEEE International Conference on Computer Vision*, 2007.

[32] X. Gu, Y. Wang, T. F. Chan, P. M. Thompson, and S. Yaun. Genus zero surface conformal mapping and its application to brain surface mapping. *IEEE Transaction on Medical Imaging*, 23(7), 2004.

[33] X. Gu and S.-T. Yau. Global conformal parameterization. *SGP 2003*, pages 127–137, 2003.

[34] B. Guenter, C. Grimm, D. Wood, H. Malvar, and F. Pighin. Making faces. In *SIGGRAPH98*, pages 55–66, 1998.

[35] E. Haber, D. N. Metaxas, and L. Axel. Motion analysis of the right ventricle from mri images. In *MICCAI '98*, 1998.

[36] S. Haker, S. Angenent, A. Tannenbaum, R. Kikinis, G. Sapiro, and M. Halle. Conformal surface parameterization for texture mapping. *IEEE Transactions on Visualization and Computer Graphics*, 6:181–189, 2000.

[37] R. S. Hamilton. The ricci flow on surfaces. *Mathematics and general relativity*, 71:237–262, 1988.

[38] K. Hormann and G. Greiner. MIPS: An efficient global parametrization method. In *Curve and Surface Design: Saint-Malo 1999*, pages 153–162. 2000.

[39] X. Huang, N. Paragios, and D. Metaxas. Establishing local correspondences towards compact representations of anatomical structures. *MICCAI03*, 2:926–934, 2003.

[40] D. Huber, A. Kapuria, R. Donamukkala, and M. Hebert. Parts-based 3d object classification. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages II: 82–89, June 2004.

[41] M. Hurdal, K. Stephenson, P. Bowers, D. Sumners, and D. Rottenberg. Coordinate systems for conformal cerebellar flat maps. *NeuroImage*, 11:S467, 2000.

[42] A. Johnson. *Spin-Images: A Representation for 3D Surface Matching*. PhD thesis, Robotics Institute, CMU, 1997.

[43] A. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21:433–449, 1999.

[44] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz. Rotation invariant spherical harmonic representation of 3d shape descriptors. In *Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 156–164, 2003.

[45] L. Kharevych, B. Springborn, and P. Schröder. Discrete conformal mappings via circle patterns. *ACM Trans. Graph.*, 25(2):412–438, 2006.

[46] A. W. F. Lee, D. Dobkin, W. Sweldens, and P. Schroder. Multiresolution mesh morphing. In *SIGGRAPH99*, pages 343–350, 1999.

[47] Y. Lee, D. Terzopoulos, and K. Walters. Realistic modeling for facial animation. In *SIGGRAPH95*, pages 55–62, 1995.

[48] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk. The digital michelangelo project: 3d scanning of large statues. In *ACM SIGGRAPH 2000*, pages 131–144, 2000.

[49] B. Levy, S. Petitjean, N. Ray, and J. Maillot. Least squares conformal maps for automatic texture atlas generation. In *SIGGRAPH'02*, pages 362–371, 2002.

[50] P. Liepa. Filling holes in meshes. In *SGP '03: Proceedings of the symposium on Geometry processing*, pages 200–205, 2003.

[51] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal on Computer Vision*, 60(2):91–110, 2004.

[52] R. Malladi, J. A. Sethian, and B. C. Vemuri. A fast level set based algorithm for topology-independent shape modeling. *J. Math. Imaging and Vision*, 6(2/3):269–290, 1996.

[53] I. Matthews, T. Cootes, J. Bangham, S. Cox, and R. Harvey. Extraction of visual features for lipreading. *IEEE Transection on Pattern Analysis and Machine Intelligence*, 24(2):198–213, 2002.

[54] G. Mori, S. Belongie, and J. Malik. Efficient shape matching using shape contexts. *IEEE Transection on Pattern Analysis and Machine Intelligence*, 27(11):1832–1837, 2005.

[55] K. Na and M. Jung. Hierarchical retargetting of fine facial motions. In *EUROGRAPHICS04*, pages 687–695, 2004.

[56] J. Y. Noh and U. Neumann. Expression cloning. In *SIGGRAPH01*, pages 277–288, 2001.

[57] B. O'Neill. *Elementary Differential Geometry*. 1997.

[58] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Shape distributions. In *ACM Transactions on Graphics*, volume 21, pages 807–832, 2002.

[59] G. Perelman. Finite extinction time for the solutions to the ricci flow on certain three-manifolds. Technical Report arXiv.org, July 17 2003.

[60] F. Pighin, R. Szeliski, and D. Salesin. Resynthesizing facial animation through 3d model-based tracking. In *IEEE International Conference on Computer Vision*, pages 143–150, 1999.

[61] E. Praun and H. Hoppe. Spherical parametrization and remeshing. *ACM Trans. Graph.*, 22(3):340–349, 2003.

[62] K. Pulli. Multiview registration for large data sets. In *3DIM99*, pages 160–168, 1999.

[63] H. Pyun, Y. Kim, W. Chae, H. W. Kang, and S. Y. Shin. An example-based approach for facial expression cloning. In *SCA03*, pages 167–176, 2003.

[64] S. Ruiz-Correa, L. Shapiro, and M. Meila. A new paradigm for recognizing 3d object shapes from range data. In *IEEE International Conference on Computer Vision*, pages 1126–1133, 2003.

[65] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy. Real-time 3d model acquisition. In *SIGGRAPH02*, pages 438–446, 2002.

[66] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proc. the 3rd Intl. Conf. on 3D Digital Imaging and Modeling*, pages 145–152, 2001.

[67] R. Schoen and S. T. Yau. *Lectures on Harmonic Maps*. 1997.

[68] E. Sharon and D. Mumford. 2d-shape analysis using conformal mapping. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages II: 350–357, 2004.

[69] A. Sheffer and E. de Sturler. Parameterization of faceted surfaces for meshing using angle-based flattening. *Engineering with Computers*, 17(3):326–337, 2001.

[70] F. Stein and G. Medioni. Structural indexing: Efficient 3d object recognition. *IEEE Transection on Pattern Analysis and Machine Intelligence*, 14(2):125–145, 1992.

[71] R. W. Sumner and J. Popovic. Deformation transfer for triangle meshes. In *SIGGRAPH04*, pages 399–405, 2004.

[72] Y. Sun and M. Abidi. Surface matching by 3d point's fingerprint. In *IEEE International Conference on Computer Vision*, pages II: 263–269, 2001.

[73] H. Tao and T. Huang. Explanation-based facial motion tracking using a piecewise bezier volume deformation model. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages I: 611–617, 1999.

[74] J. Tenenbaum, V. de Silva, and J. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, Dec 2000.

[75] D. Terzopoulos, A. Witkin, and M. Kass. Constraints on deformable models: Recovering 3d shape and nonrigid motion. *Artificial Intelligence*, 35:91–123, 1988.

[76] G. Turk and M. Levoy. Zippered polygon meshes from range images. In *SIGGRAPH '94*, pages 311–318, 1994.

[77] B. Vemuri, A. Mitiche, and J. Aggarwal. Curvature-based representation of objects from range data. *Image and Vision Computing*, 4:107–114, 1986.

[78] S. Wang, X. Gu, and H. Qin. Automatic non-rigid registration of 3d dynamic data for facial expression synthesis and transfer. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[79] S. Wang, Y. Wang, M. Jin, X. Gu, and D. Samaras. 3d surface matching and recognition using conformal geometry. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages II: 2453– 2460, 2006.

[80] S. Wang, Y. Wang, M. Jin, X. Gu, and D. Samaras. Conformal geometry and its applications on 3d shape matching, recognition, and stitching. *IEEE Transection on Pattern Analysis and Machine Intelligence*, 29(7):1209–1220, 2007.

[81] Y. Wang, M. Chiang, and P. M. Thompson. Mutual information-based 3d surface matching with applications to face recognition and brain mapping. *IEEE International Conference on Computer Vision*, 1:527–534, 2005.

[82] Y. Wang, X. Gu, K. Hayashi, T. Chan, P. Thompson, and S. Yau. Surface parameterization using riemann surface structure. In *IEEE International Conference on Computer Vision*, pages II: 1061–1066, 2005.

[83] Y. Wang, M. Gupta, S. Zhang, S. Wang, X. Gu, D. Samaras, and P. Huang. High resolution tracking of non-rigid 3d motion of densely sampled data using harmonic maps. In *IEEE International Conference on Computer Vision*, pages I: 388–395, 2005.

[84] Y. Wang, M. Gupta, S. Zhang, S. Wang, X. Gu, D. Samaras, and P. Huang. High resolution tracking of non-rigid motion of densely sampled 3d data using harmonic maps. *International Journal of Computer Vision*, 76(3):283–300, 2008.

[85] Y. Wang, X. Huang, C. Lee, S. Zhang, Z. Li, D. Samaras, D. Metaxas, A. Elgammal, and P. Huang. High resolution acquisition, learning and transfer of dyanmic 3-d facial expressions. In *Computer Graphics Forum*, pages III: 677–686, 2004.

[86] K. Waters. A muscle model for animation three-dimensional facial expression. In *SIGGRAPH87*, pages 17–24, 1987.

[87] J. Wyngaerd, L. Gool, R. Koch, and M. Proesmans. Invariant-based registration of surface patches. In *IEEE International Conference on Computer Vision*, pages 301–306, 1999.

[88] A. Yuille, P. Hallinan, and D. Cohen. Feature extraction from faces using deformable templates. *International Journal on Computer Vision*, 8(2):99–111, 1992.

[89] D. Zhang and M. Hebert. Harmonic maps and their applications in surface matching. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages II: 524–530, 1999.

[90] L. Zhang, N. Snavely, B. Curless, and S. M. Seitz. Spacetime faces: high resolution capture for modeling and animation. *ACM Trans. Graph.*, 23(3):548–558, 2004.

[91] S. Zhang and P. Huang. High resolution, real time 3d shape acquisition. In *CVPR04 Workshop on Real-time 3D Sensors and Their Use*, page 28, 2004.