A Generic Deformation Model for Dense Non-Rigid Surface Registration: a Higher-Order MRF-based Approach

Yun Zeng[1], Chaohui Wang[2], Xianfeng Gu[3], Dimitris Samaras[3], Nikos Paragios[4,5]
[1]Department of Mathematics, Harvard University, MA, USA
[2]Max Planck Institute for Intelligent Systems, Tübingen, Germany
[3]Department of Computer Science, Stony Brook University, NY, USA
[4]Center for Visual Computing, Ecole Centrale Paris, Châtenay-Malabry, France
[5]Equipe GALEN, INRIA Saclay - Ile-de-France, Orsay, France

Abstract

We propose a novel approach for dense non-rigid 3D surface registration, which brings together Riemannian geometry and graphical models. To this end, we first introduce a generic deformation model, called Canonical Distortion Coefficients (CDCs)*, by characterizing the deformation of every point on a surface using the distortions along its two principle directions. This model subsumes the deformation groups commonly used in surface registration such as isometry and conformality, and is able to handle more complex deformations. We also derive its discrete counterpart which can be computed very efficiently in a closed form. Based on these, we introduce a higher-order Markov Random Field (MRF) model which seamlessly integrates our deformation model and a geometry/texture similarity metric. Then we jointly establish the optimal correspondences for all the points via maximum a posteriori (MAP) inference. Moreover, we develop a parallel optimization algorithm to efficiently perform the inference for the proposed higher-order MRF model. The resulting registration algorithm outperforms state-of-the-art methods in both dense non-rigid 3D surface registration and tracking.*

1. Introduction

Surface registration is one of the most active research topics in 3D computer vision, due to the wide availability of 3D data acquisition techniques/devices (*e.g.*, [9, 18, 27, 33]) and in particular Microsoft Kinect [13]. It often serves as a necessary step for numerous applications, such as shape recognition/retrieval, deformation transfer, facial expression recognition and change detection [6, 23]. A main challenge in solving this problem lies in the fact that real-world deformations often have very high degrees of freedom and accurately characterizing these deformations requires sophisticated mathematical models that are generic enough to represent these deformations and whose optimal configuration can be efficiently inferred.

Most existing surface registration approaches rely on some assumption on the deformation (*e.g.*, rigid [2], isometric [5] and conformal [26]), which serves as a prior/regularization model and/or facilitates the search of optimal correspondences. Despite their success in various applications, accuracy will deteriorate drastically when the real deformation deviates from the assumed group. To overcome such a limitation, we first propose a novel deformation model that is able to represent a much wider range of deformations. According to Riemannian geometry [8], a surface can be represented in a parametrized domain (local charts) so that the deformation at any point p can be unambiguously (*i.e.*, independently of parametrization and embedding) characterized by considering a particular class of parametrizations for that surface (called *canonical parametrizations*). Based on this, we introduce *Canonical Distortion Coefficients* (CDCs), defined as the distortions along p's two principle directions and computed on the canonical parametrization domain. An intuitive explanation of CDCs is that they characterize how an infinitesimal circle is deformed into an infinitesimal ellipse at every point.

Furthermore, in the discrete setting where a surface is represented as a simplicial complex (*e.g.*, a planar or tetrahedral mesh), we show that the computation of CDCs at any point on the continuous surface corresponds to the computation of CDCs for its corresponding facet in the discrete setting, derived via the common piecewise linear assumption in finite element methods [3]. Accordingly, the canonical parametrization at a particular point, which requires the metric tensor to be Euclidean, simply corresponds to any mapping of the facet from 3D to 2D that preserves edge lengths and orientations. It follows that the CDCs for the deformation of each facet can be computed in a closed form,

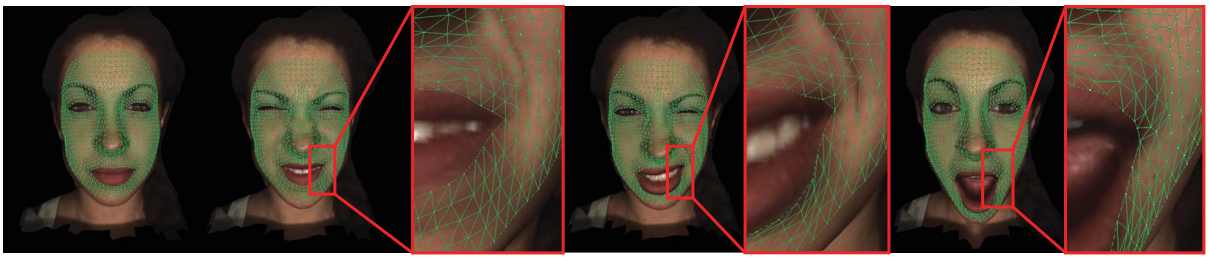


Figure 1. Surface tracking results (best viewed in color). The left-most image shows the 3D textured shape in the first frame with a template mesh (green). The following three images show the estimated configurations (with close-ups) of the template for three representative frames obtained by our method. Note that large and anisotropic deformations are correctly handled.

which only requires solving linear equations.

Finally, we introduce a higher-order MRF-based approach for dense surface registration, which integrates the above deformation model and geometry/texture similarity in a single formulation and jointly searches for the optimal correspondence for all the points via combinatorial optimization. Furthermore, we exploit the topological properties of such a class of MRFs and develop a parallel optimization algorithm based on dual-optimization techniques [28], which requires minimal memory and achieves significant speedup via an implementation in distributed hardware.

In sum, the contributions of this paper are three-fold, including the generic deformation model (CDCs), the higher-order MRF-based approach for dense surface registration and the parallel MRF inference algorithm. The whole approach brings together Riemannian geometry and graphical models. On the one hand, our deformation model is generic, leading to the fact that those deformation groups commonly used in surface registration (*e.g.*, isometry and conformality) fall into its special sub-classes. Due to this property, our surface registration method is able to effectively handle much more complex deformations such as the anisotropic or locally twisting motions (Fig. 1), which is important for addressing challenging real-world cases. On the other hand, the higher-order MRF surface registration approach takes advantage of optimality and efficiency properties of graphical models [24], which is further boosted by the developed parallel inference algorithm. Via a series of experimental comparisons with state-of-the-art methods, we demonstrate that our approach achieves significant improvement in non-rigid 3D surface registration and tracking.

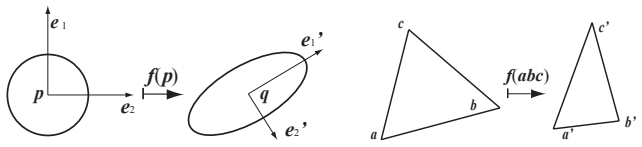
The remainder of the paper is organized as follows: the mathematical formulation of the deformation model is introduced in Sec. 2; in Sec. 3, we present our higher-order MRF model for surface registration, as well as its parallel optimization; the experimental validation of the proposed techniques in 3D surface registration and tracking is shown in Sec. 4; finally, we conclude the whole work in Sec. 5.

Related work

Accurately characterizing the deformations of an arbitrary 3D object is a very challenging task due to the high degrees of freedom exhibited in real-world deformations. Among previous approaches, certain "rigidity" assumptions have been widely made, either in extrinsic space or in intrinsic space, as a tradeoff between the accuracy in deformation representation and the simplicity in computation.

When a shape is represented in extrinsic space, a simple deformation model is the rigid deformation (*i.e.*, rotation and translation). Assuming two shapes undergo a (near) rigid deformation between them, the Iterative Closest Points (ICP) method [2] has been widely adopted for surface registration. However, global rigidity does not take into account bendable surfaces (*e.g.*, garments or rubber bands). In order to address this, the notion of local rigidity has been proposed, which assumes that the deformation between the local neighborhoods of two corresponding points is rigid [7, 16, 20, 22], leading to a higher degrees of freedom of the deformation. However, such a model does not make use of the geometric properties of surface deformation and searching for the correspondences between two surfaces with large deformations directly in the original extrinsic space would suffer from high computational complexity.

A more efficient way of handling large deformations is to consider the intrinsic representation, which assumes that each point of a surface is equipped with a metric tensor. The notion of rigidity (*i.e.*, isometry) can then be characterized by assuming that the metric tensor remains the same during the deformation of a surface. One popular approach to matching two surfaces undergoing isometric deformations is to consider the conformal mappings of the surfaces [26]. It benefits from a nice property: for surfaces undergoing only isometric deformations, their conformal parametrizations only differ by a Mobius transformation with very few degrees of freedom [11]. However, finding a globally consistent parametrization/embedding is often too restrictive. Recent works proposed searching among multiple parametrizations and/or combining multiple matching cues (*e.g.*, texture or Gaussian curvature) to improve matching accuracy [10, 11, 25, 29, 31, 32]. Another popular approach is to embed the surface into an Euclidean space such that the Euclidean distance approximates the intrinsic properties of the surface [4, 12, 14, 19]. Nevertheless, when it



(a) Continuous setting (b) Discrete setting

Figure 2. The finite element method assumes the transformation between facets to be piecewise linear and $f(\vec{ab}) = \vec{a'b'}, f(\vec{ac}) = \vec{a'c'}$. Under this assumption, the Jacobian for each mapping $\triangle abc \mapsto \triangle a'b'c'$ can be computed in a closed form.

comes to dense and anisometric deformations, the accuracy of the above methods would unavoidably deteriorate due to the isometric assumption.

2. Canonical Distortion Coefficients

In this section, we introduce the Canonical Distortion Coefficients (CDCs) for characterizing arbitrary diffeomorphisms. The continuous formulation (Sec. 2.2) is obtained based on Riemannian geometry (Sec. 2.1). Its discrete counterpart is derived via finite element analysis (Sec. 2.3).

2.1. Riemannian metric and parametrization

Let $(\mathcal{M}, g_{\mathcal{M}})$ denote a surface $\mathcal{M}$ equipped with a Riemannian metric $g_{\mathcal{M}}$. In Riemannian geometry [8], a surface is defined by its local charts $\mathcal{M} = U_\alpha \cup U_\beta \cup \ldots$, and each open subset U_α is in $1-1$ correspondences to R^2, denoted by the local parametrization $\phi_\alpha : U_\alpha \to \mathbb{R}^2$. For any point $p \in U_\alpha \subset \mathcal{M}$, a metric tensor is associated with p as a symmetric positive definite matrix:

$$g^\alpha(p) = \begin{pmatrix} g^\alpha_{11}(p) & g^\alpha_{12}(p) \\ g^\alpha_{21}(p) & g^\alpha_{22}(p) \end{pmatrix}. \quad (1)$$

Since different local parametrizations must represent the same surface, the following chain rule should be satisfied:

$$\forall p \in U_\alpha \cap U_\beta, g^\alpha(p) = J_{\alpha\beta}(p)^T g^\beta(p) J_{\alpha\beta}(p). \quad (2)$$

Here, $J_{\alpha\beta}$ is the Jacobian matrix of the transformation between the local coordinate systems of U_α and U_β. Any local representation satisfying this transformation rule is a valid parametrization of the surface. Since the metric tensor at any point p is positive definite, it is always possible to apply a proper linear transformation to its local parametrization ϕ_α such that $g^\alpha(p)$ is an identity matrix. We name this particular type of parametrization *canonical parametrization*:

Definition 1 ***(Canonical parametrization)*** *For any $p \in \mathcal{M}$, a parametrization $\phi_\alpha : \mathcal{U}_\alpha \to \mathbb{R}^2$ is called* canonical *if the metric tensor at p is the identity matrix.*

Accordingly, the Jacobian matrix J_{pq} between the two points p and q under their canonical parametrizations is called the *canonical Jacobian*. In the following, we show that considering the canonical parametrization and Jacobian leads to a representation of arbitrary deformations that are independent from both the intrinsic and extrinsic representations of the surface.

2.2. Canonical Distortion Coefficients

Let us consider arbitrary diffeomorphisms between the parametrization domains of two surfaces. For any correspondence $p \to q, p \in \mathcal{U}_\alpha \subset \mathcal{M}$ and $q \in \mathcal{U}_\beta \subset \mathcal{N}$, the change of metric $g^\alpha(p) \to J_{\alpha\beta}(p)^T g^\beta(q) J_{\alpha\beta}(p)$ reflects how an infinitesimal circle is deformed into an infinitesimal ellipse. In particular, under canonical parametrizations for points p and q (*i.e.*, both $g^\alpha(p)$ and $g^\beta(q)$ are identity matrices), the matrix $J^T_{pq} J_{pq}$ accurately characterizes such local deformation, where J_{pq} is the canonical Jacobian mapping p to q. If we only consider the change of shape (Fig. 2 (a)), *i.e.*, how a circle is deformed into an ellipse regardless of its orientation, the distortion along its two principle directions can be represented by the two eigenvalues λ_1 and λ_2 of $J^T_{pq} J_{pq}$. Therefore, the local deformation between two surfaces can be characterized by the two eigenvalues for each pair (p, q) of corresponding points. Formally, we define:

Definition 2 ***(Canonical Distortion Coefficients)*** Canonical Distortion Coefficients (CDCs) *between points p and q are defined as the eigenvalues of the matrix $J^T_{pq} J_{pq}$ where J_{pq} is the Jacobian matrix between any canonical parametrization at p and q.*

CDCs are able to characterize a wide class of deformation groups. For instance, below are two typical classes of deformations that can be characterized by CDCs:

1. In the case of the isometric deformation, a unit circle is mapped to a unit circle, *i.e.*, $\lambda_1 = \lambda_2 = 1$.

2. In the case of the conformal deformation, a unit circle can be mapped to a circle with arbitrary radius [21], *i.e.*, $\lambda_1 = \lambda_2 \neq 0$.

Our CDCs can be further connected to a general class of diffeomorphisms characterized by the *Beltrami-coefficients* (BCs) [1]. However, BCs are for conformal surface parametrization, where the scaling factor is lost. The proposed CDCs preserve the scale information which is important for surface registration. Besides, unlike the BC, the CDC is directly extendable to nD. Note that the ability of CDCs to encode scale information directly makes CDCs suitable to characterize detailed, anisometric deformation for dense surface registration.

2.3. Finite element discretization

The basic assumption in finite element analysis [3] is that continuous space can be approximated using a set of *basis*

elements (*e.g.*, polynomial functions defined on each facet) with continuity preserved at the boundaries between the basis elements. Here, we consider the most common discrete surface representation – the triangular mesh, with triangular facets as basis finite elements. In this setting, CDCs are assumed to be constant for each triangular facet. Thus, the canonical parametrization for a facet is Euclidean if its mapping to the 2D domain preserves all the edge lengths.

Next we consider the computation of the canonical Jacobian (Sec. 2.1). In the continuous setting, the Jacobian matrix at a point p represents a linear transformation that transforms tangent vectors at p to tangent vectors at q. Given a basis element $\triangle abc$, the tangent space at p is equivalent to the linear space spanned by $\triangle abc$. Hence, the linear mapping $J(\cdot)$ between two canonical domains $\triangle abc$ and $\triangle a'b'c'$ should satisfy $J(\vec{ab}) = \vec{a'b'}$ and $J(\vec{ac}) = \vec{a'c'}$. The Jacobian of a linear transformation between two triangles is a 2×2 matrix and can be computed in a closed form. Since $J(\cdot)$ is linear, it is guaranteed that $J(\vec{bc}) = \vec{b'c'}$, *i.e.*, the Jacobian for mapping $p \rightarrow q$ in the continuous case is equivalent to a linear transformation matrix for mapping $\vec{ab} \rightarrow \vec{a'b'}$, $\vec{ac} \rightarrow \vec{a'c'}$ in the discrete case (Fig. 2).

Alg. 1 summarizes the algorithm for computing CDCs. For an n-manifold surface, the computation of CDCs only requires solving n linear equations and eigenvalues. Note that although the computation looks analogous to [15, 17] for surface parametrization due to the piecewise linear assumption, Alg. 1 is derived in the context of Riemannian geometry for surface deformation.

Algorithm 1: CDC computation for each triangular facet.

Input : $\triangle abc$ and its mapping $\triangle a'b'c'$
Output : CDCs for mapping from $\triangle abc$ to $\triangle a'b'c'$.
Step One: Map the triangles $\triangle abc$ and $\triangle a'b'c'$ to 2D and keep their orientation.
Step Two: Compute the 2×2 linear transformation J mapping $\vec{ab}$ to $\vec{a'b'}$ and $\vec{ac}$ to $\vec{a'c'}$.
Step Three: Compute the eigenvalues, λ_1 and λ_2 of $J^T J$.
Step Four: Output λ_1 and λ_2

3. Surface registration framework

In this section, we first introduce our MRF formulation for surface registration. Then we present the parallel optimization algorithm for the inference in the MRF.

3.1. Higher-order MRF formulation

Given two surfaces $\mathcal{M}$ and $\mathcal{N}$ either in a continuous or a discrete (*e.g.*, point clouds) representation, we consider a triangulated set of n points $\mathcal{V} = \{p_u | p_u \in \mathcal{M}, u = 1, \ldots, n\}$, where $\mathcal{V} \subset \mathcal{M}$ are chosen as a standard template. The goal is to determine the correspondences of $\mathcal{V}$ on the other surface $\mathcal{N}$.

Our higher-order MRF model has the same topology as the graph $\mathcal{G} = (\mathcal{V}, \mathcal{F})$ corresponding to the triangulation of the set of points on the surface $\mathcal{M}$, where $\mathcal{V}$ denotes the vertex set and $\mathcal{F} \subset \mathcal{V}^3$ denotes the triangular facet set. The random variable X_u for each vertex $u \in \mathcal{V}$ represents the correspondence of the vertex u on the surface $\mathcal{N}$. Its realization[1] x_u belongs to a set of possible matching candidates indexed by $\mathcal{L}_u = \{1, \ldots, L_u\}$. We use $\mathbf{x} = (x_u)_{u \in \mathcal{V}}$ to denote the configuration of the whole MRF.

Regarding the MRF energy, we first define the unary potential function $\theta_u(x_u)$ as the difference in the feature descriptor (*e.g.*, texture or shape context) between u and its correspondence x_u:

$$\theta_u(x_u) = |\text{fea}_{\mathcal{M}}(u) - \text{fea}_{\mathcal{N}}(x_u)|^2,$$

where $\text{fea}_{\mathcal{S}}(\cdot)$ denotes the feature descriptor attached to a point on surface $\mathcal{S}$. Next, let $\lambda_{uvw}(x_u, x_v, x_w)$ denote the CDCs computed from deforming $\triangle uvw$ to $\triangle x_u x_v x_w$ (Alg. 1). We define the higher-order potential as follows:

$$\theta_{uvw}(x_u, x_v, x_w) = \rho(\lambda_{uvw}(x_u, x_v, x_w)),$$

where $\rho(\cdot)$ is a function that encodes the deformation constraints on the CDC values. Its definition in our surface registration applications will be given in Eq. 5 of Sec. 4. Finally, given the above potential functions, surface registration boils down to the search of the optimal configuration $\mathbf{x}$ that minimizes the following energy:

$$E(\mathbf{x}) = \sum_{u \in \mathcal{V}} \theta_u(x_u) + \sum_{(u,v,w) \in \mathcal{F}} \theta_{uvw}(x_u, x_v, x_w). \quad (3)$$

In the following section, we present the optimization algorithm developed for the above problem.

3.2. Efficient higher-order MRF optimization

Efficient inference in higher-order MRFs is a very active research problem and various techniques have been proposed to deal with such a challenging problem in the past decade, such as those based on order reduction (combined with graph cuts), belief propagation, and/or relaxation techniques [24]. However, the algorithms designed for general MRFs often lack efficiency in terms of computation and/or memory when solving MRFs with special topologies and/or potential energy functions. In order to efficiently perform the inference in our MRF model (Sec. 3.1), we exploit the topology property of such a class of MRFs and develop a parallel optimization algorithm, which requires minimal memory and achieves significant speedup via an implementation in distributed hardware.

[1] For the sake of clarity and simplicity, x_u will denote the corresponding label in $\mathcal{L}_u$ when describing the optimization algorithm in Sec. 3.2.

Let us first derive the dual problem for the LP relaxation of the minimization problem of the energy in Eq. 3. First, an indicator variable $\tau_{u;i}$ is introduced to any $u \in \mathcal{V}$ and $i \in \mathcal{L}$, and $\tau_{uvw;ijk}$ to any $(u,v,w) \in \mathcal{F}$ and $(i,j,k) \in \mathcal{L}\times\mathcal{L}\times\mathcal{L}$:

$$\tau_{u;i} = \begin{cases} 1 & \text{if } x_u = i \\ 0 & \text{otherwise} \end{cases} \qquad \tau_{uvw;ijk} = \begin{cases} 1 & \text{if } x_u = i, x_v = j, x_w = k \\ 0 & \text{otherwise} \end{cases}.$$

By defining $\theta_{u;i} = \theta_u(i)$ and $\theta_{uvw;ijk} = \theta_{uvw}(i,j,k)$, we obtain the following integer LP formulation for the minimization problem of the energy in Eq. 3:

$$\begin{aligned} \min_{\tau} & \sum_{u\in\mathcal{V}}\sum_{i\in\mathcal{L}} \theta_{u;i}\tau_{u;i} + \sum_{(u,v,w)\in\mathcal{F}}\sum_{(i,j,k)\in\mathcal{L}^3} \theta_{uvw;ijk}\tau_{uvw;ijk} \\ \text{s.t.} & \sum_{i} \tau_{u;i} = 1, && \forall u \in \mathcal{V} \\ & \sum_{i,j,k} \tau_{uvw;ijk} = 1, && \forall (u,v,w) \in \mathcal{F} \\ & \sum_{j,k} \tau_{uvw;ijk} = \tau_{u;i}, && \forall (u,v,w) \in \mathcal{F} \text{ and } i \in \mathcal{L} \\ & \tau_{u;i}, \tau_{uvw;ijk} \in \{0,1\}. \end{aligned}$$

By relaxing the domains of the variables $\tau_{u;i}$ and $\tau_{uvw;ijk}$ to $[0,1]$, we obtain the LP-relaxation of the above problem and then derive its dual problem as shown below:

$$\begin{aligned} \max_{M} & \sum_{u} \min_{i} \overline{\theta}_{u;i} + \sum_{(u,v,w)\in\mathcal{F}} \min_{i,j,k} \overline{\theta}_{uvw;ijk} \qquad (4) \\ \text{s.t. } & \overline{\theta}_{u;i} = \theta_{u;i} + \sum_{(u,v,w)\in\mathcal{F}} M_{uvw;u:i}, \forall u \in \mathcal{V} \text{ and } i \in \mathcal{L} \\ & \overline{\theta}_{uvw;ijk} = \theta_{uvw;ijk} - M_{uvw;u:i} - M_{uvw;v:j} - M_{uvw;w:k}, \\ & \forall (u,v,w) \in \mathcal{F} \text{ and } (i,j,k) \in \mathcal{L}\times\mathcal{L}\times\mathcal{L}. \end{aligned}$$

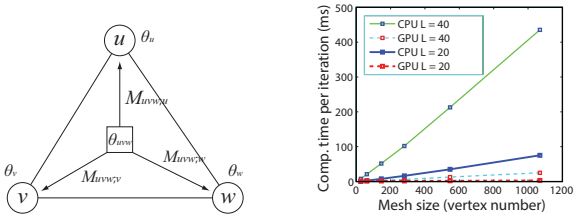
Here $M_{uvw;u:i}$ is the dual variable (message) corresponding to the constraint $\sum_{j,k} \tau_{uvw;ijk} = \tau_{u;i}$ (Fig. 3(a)).

The dual problem of Eq. 4 can be solved by *min-sum diffusion* algorithm [28] (at convergence, the J-consistency condition is satisfied) as shown in Alg. 2. Since after each update of the message, only a reparametrization of the MRF is performed, no extra memory is needed for storing all the dual variables $M_{uvw;u:i}$. Hence, the memory requirement for the Alg. 3 is only for storing primal variables, *i.e.*, $O(|V||\mathcal{L}| + |\mathcal{F}||\mathcal{L}|^3)$.

Algorithm 2: Min-sum diffusion algorithm.

repeat
  for each $M_{uvw;u:i}$ **do**
    $M_{uvw;u:i} -= \frac{1}{2}[\theta_{u;i} - \min_{j,k}\theta_{uvw;ijk}]$ and reparameterize $\theta_{u;i}$ and $\theta_{uvw;ijk}$ according to the constraints in Eq. 4.
  end for
until convergence

Each message update in Alg. 2 only involves the parameters in a triangular facet $\triangle uvw$ of the MRF. Moreover, in $\triangle uvw$, the update of the message for each label $M_{uvw;u:i}, i = \{1,\dots,L\}$ is independent from each other. Hence, the algorithm can be naturally parallelized and efficiently executed in distributed hardware. To this end, we first define the concept of *independent facet set*:



(a) (b)

Figure 3. MRF optimization algorithm. (a) illustrates the message passing (Eq. 4). (b) shows the speedup obtained by the parallel implementation of Alg. 3. L is the number of labels for each node.

Definition 3 *(**Independent facet set**) Given a graph $\mathcal{G} = (\mathcal{V},\mathcal{F})$, a subset $\mathcal{F}_k \subset \mathcal{F}$ is called* independent facet set *if for any $f_i, f_j \in \mathcal{F}_k, i \neq j$, $f_i \cap f_j = \varnothing$.*

The decomposition of a set $\mathcal{F}$ into subsets of independent facet sets $\mathcal{F} = \cup_i \mathcal{F}_i$ can be efficiently computed in polynomial time by a simple greedy algorithm. Then, we can implement Alg. 2 in parallel as shown in Alg. 3. The maximal speedup achieved in Alg. 3 is $\max_i(|\mathcal{F}_i||\mathcal{L}|)$. Fig. 3(b) shows the experimental comparison on running time between the implementations with and without GPU accelerations, and demonstrates significant speedup ($\times 100$ times with 128 CUDA cores) obtained with the parallel algorithm.

Algorithm 3: Parallel min-sum diffusion algorithm.

Decompose $\mathcal{F}$ into independent facet sets $\cup_i \mathcal{F}_i$
repeat
  for each Independent facet set $\mathcal{F}_i$, in parallel for all $(u,v,w) \in \mathcal{F}_i$ and $k \in \mathcal{L}$ **do**
    Update the message $M_{uvw;u:k}$, $M_{uvw;v:k}$ and $M_{uvw;w:k}$ and do reparametrization (Alg. 2).
  end for
until convergence

4. Experimental results

We evaluate our method in the surface registration and tracking problems. The input to our algorithm are two 3D surfaces in the case of registration (or a set of 3D surfaces in the case of tracking), and a template triangular mesh $\mathcal{G} = (\mathcal{V},\mathcal{F})$ which consists of a point set $\mathcal{V}$ sampled from the first surface $\mathcal{M}$ and whose topology is defined by a facet set $\mathcal{F}$ (*e.g.*, Fig. 5 (a)). Our goal is to find the optimal matching point x_p on the other surface (or each of the successive surfaces) $\mathcal{N}$ for each $p \in \mathcal{V}$ (*e.g.*, Fig. 5 (c)).

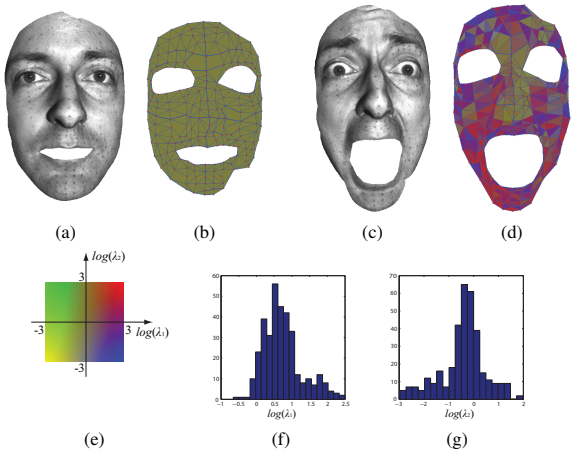


Figure 4. Expression deformation prior obtained from 3D scanned data with markers. (a) and (c) show the 3D scan of the onset and peak of a facial expression with large surface deformations respectively. (b) and (d) are the corresponding triangular meshes constructed from the markers. The color coding in (d) shows the deformation intensity as illustrated in (e). The histogram of the CDC values are shown in (f) and (g).

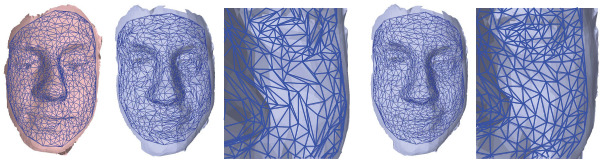
Estimation of deformation prior: To obtain the distribution of CDCs in real-world deformations, we first obtain the ground truth data from 3D scanning systems with reliable texture information (*e.g.*, markers). As shown in Fig. 4(a) and (c), the 3D dataset with markers are captured using the system introduced in [27]. To capture the maximal range of CDCs, we select two frames with the largest expression difference. Fig. 4 (b), (d), (f) and (g) visualize the distribution of CDCs. As a result, we obtain the ranges of λ_1 and λ_2 as $I_1 = [0.7, 5.66], I_2 = [0.1, 4]$, respectively, which can be used to impose priors for facial expression deformation. In our experiments, a simple uniform distribution in the allowed range was used, by defining the higher-order term in Eq. 3 as:

$$\theta_{uvw}(x_u, x_v, x_w) = \begin{cases} 0 & \text{if } \lambda_1 \in I_1 \text{ and } \lambda_2 \in I_2 \\ 10 & \text{otherwise} \end{cases}, \quad (5)$$

where λ_1 and λ_2 denote the CDCs obtained by matching $\triangle uvw$ to $\triangle x_u x_v x_w$. Note that penalty on the flip of triangles can be easily included in such higher-order terms [31]. However, in our experiments, we found that the inclusion of such terms does not improve the results.

4.1. Surface registration

For surface registration, we compare our method with two recent methods: *high-order graph matching* (HOGM) [31] and *blended intrinsic maps* (BIM) [10]. For the purpose of a fair comparison with [31], we use the same singleton term as in [31] and adopt a similar hierarchical optimization scheme to perform the registration: first establishing sparse feature correspondences based on isometric deformation and then establishing the dense correspondences based on our deformation model. Similar to [31], a set of matching candidates for each $p \in \mathcal{V}$ is computed using the candidate selection method[2] proposed in [31] and then the optimal correspondences of all points are jointly estimated through the MRF inference presented in Sec. 3. In our experiments, we set the candidate size $L = 64$. The computation of all the L^3 possible CDCs for one facet takes only $2.0ms$ on average using GPU. Accordingly, the computation of all the energy terms $\theta_{uvw;ijk}$ for a higher-order graph with 165 vertices and 272 facets takes only 0.5s.



(a) Input (b) Result by [31] and closeup (c) Our result and closeup

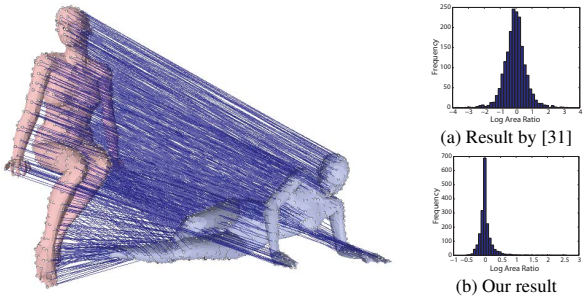
Figure 5. Surface registration result. (a) shows the input mesh with sampling points and their triangulation. The algorithm by [31] does not guarantee the quality of each triangular facet in the target matching (b). In contrast, our algorithm considers the distortion of each facet using CDCs and achieves better results (c).

The qualitative results in Fig. 5 show that the unnatural distortions of each triangular facet (Fig. 5(b)) in the result of [31] are significantly reduced in the result obtained using our method (Fig. 5(c)), which demonstrates the effectiveness of the deformation constraints encoded in our MRF model. Besides, the optimization technique in [31] requires order reduction, which introduces a large number of auxiliary variables and prevents it from searching in a large label set (due to the memory limitation). More visual results and quantitative comparisons are given in Fig. 6 and Table 1 using the same quality measure (*i.e.*, area ratios) as the one used in [31], the assumption being that most large triangle area changes are caused by wrong matches. Another quantitative comparison using the metric proposed in [10] is shown in Fig. 8. All results show that our method improves the matching quality up to an order of magnitude.

Furthermore, we have compared our approach with a recent intrinsic space based method [10] for dense surface registration. Here, we use the normalized (by the number of points evaluated) error evaluation metric proposed in [10] for the quantitative comparison (Fig. 8). In all cases, our method achieves lower errors. Note that [10] assumes the mapping between two surfaces be bijective and there is no explicit underlying deformation model in selecting the final correspondence. In contrast, our deformation model was explicitly encoded in the MRF model for selecting the optimal dense correspondence, which is a main reason for the better performance.

Last, in order to test the performance of the proposed

[2]We refer the reader to [31] for the detail of the selection method.



(a) Result by [31]

(b) Our result

Figure 6. A challenging surface registration result using our method (left). The quality of matching is measured by the ratio of area change of each triangular facet, under the assumption that most large area changes are caused by wrong matches. Our method (b) has significantly fewer triangles with large area change compared to a recent method (a) [31].

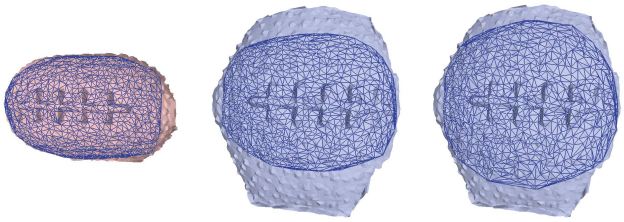
Data	Method in [31]	Our method
Face (smile)	$(2.26, 0.19, 67.83)$	$(1.24, 0.86, 4.2)$
Face (laugh)	$(1.75, 0.12, 111.11)$	$(1.36, 0.82, 11.0)$
Face (sad)	$(1.87, 0.19, 78.62)$	$(1.48, 0.87, 7.52)$

Table 1. Comparison with [31]. $(\cdot, \cdot, \cdot)$ denotes the average, minimal and maximal area ratios between the original/matched facets. The values are expected to be close to 1 for good registration.

approach in cases of significant anisometric deformations, we design the following experiment. The 3D scan of a highly deformable toy is captured by the system introduced in [27] before and after a large deformation (Fig. 7). To establish the ground truth and estimate the deformation prior, we manually select 20 facets and their matches based on texture features to calculate the average CDC values. The two surfaces are then matched *without using texture information*, *i.e.*, in Eq. 3, we only use the curvature cue for the singleton term (data likelihood) and the deformation prior for the higher-order term. Fig. 7(b) shows the result using isometric assumption $(\lambda_1, \lambda_2) = (1, 1)$ and Fig. 7(c) shows the result using the learned average CDC prior as described above, *i.e.*, $(\lambda_1, \lambda_2) = (1.028, 0.993)$. To compare the accuracy achieved in the two cases, we compute the average texture difference between the original area covered by the triangulated mesh and the matched area on the second surface (*i.e.*, the blue deformed templates in Fig. 7). Experimental results demonstrate that the use of CDCs leads to a significantly lower error (0.005 *v.s.* 0.073) compared with the use of the isometry assumption.

4.2. Template-based surface tracking

Finally, we apply our method to the challenging problem of template-base surface tracking. For the singleton term in Eq. 3, we use the robust metric proposed in [32] for a fair comparison. To impose inter-frame consistency, two consecutive frames with the largest deformation change were selected to obtain the range of CDCs between frames, *i.e.*,



(a) Original surface (b) With isometric assumption (c) With learnt CDC prior

Figure 7. The comparison between surface registration with isometric assumption (b) and with learnt CDC prior (c).

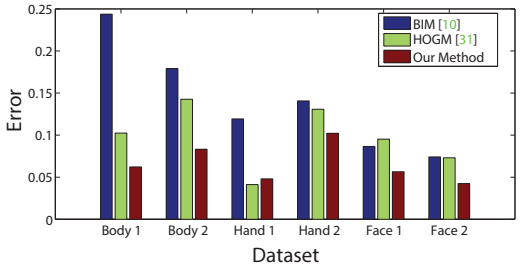


Figure 8. Comparisons based on the metric defined in [10].

$I_1 = [0.874, 1.143]$ and $I_2 = [0.846, 1.182]$ for λ_1 and λ_2 respectively. Also we handle drift error by imposing consistency between the first frame and the current frame, using the same deformation prior obtained in Fig. 4.

Figs. 1 and 9 show some tracking results on the BU-4DFE database [30], consisting of 101 different subjects each with 6 different expressions and around 100 frames/expression. A template is constructed in the first frame and tracked in the subsequent frames. Because of the temporal continuity in consecutive frames, sufficient matching candidates ($L = 64$) can be obtained by only looking at the neighborhood of each point. The tracking results demonstrate that our method is able to track the subtle expression change correctly, even in the challenging case where the deformation is highly anisometric.

We also compare with the harmonic map based method in [27] and the pairwise MRF based method in [32] (Fig. 10), based on tracking errors defined using average texture differences on 10 randomly selected subjects' videos from the BU database (six different expressions per subject). The results show that our method consistently outperforms [27] and [32].

5. Conclusion

We have presented a generic deformation model, namely CDCs, to characterize the space of deformations between two surfaces, which can be efficiently computed in a closed form in the discrete setting. Such a deformation model is applied to surface registration by combining CDCs with other geometric/photometric information within a higher-order MRF framework, whose optimal configuration is in-

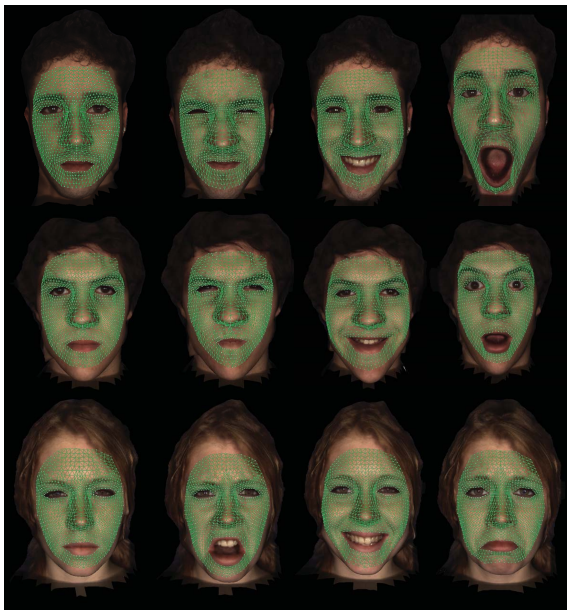


Figure 9. Surface tracking results on BU-4DFE database.

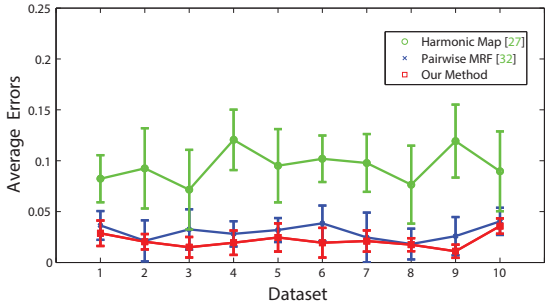


Figure 10. Quantitative comparisons on BU-4DFE database.

ferred with an efficient parallel optimization algorithm. We have demonstrated the potential of our approach in surface registration and tracking, where our approach significantly outperforms state-of-the-art methods. In the near future, we are interested in exploring the group structure of the surface deformation space represented by CDCs for facial expression recognition and deformation analysis/manipulation.

Acknowledgement The work was done while Y. Zeng was with Stony Brook University and C. Wang was with Ecole Centrale Paris and INRIA Saclay. It was partially supported from the European Research Council Starting Grant DIOCLES (ERC-STG-259112), NSF Awards IIS-1161876, IIS-1111047, IIS-0959979 and the SUBSAMPLE Project of the DIGITEO Institute, France.

References

[1] L. V. Ahlfors. *Lectures on Quasiconformal Mappings*. American Mathematical Society, 2 edition, 2006.

[2] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *TPAMI*, 14(2):239–256, 1992.

[3] S. C. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 3 edition, 2007.

[4] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *Proc. National Academy of Sciences*, 103:1168–1172, 2006.

[5] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Expression-invariant representations of faces. *TPAMI*, pages 1042–1053, 2007.

[6] R. J. Campbell and P. J. Flynn. A survey of free-form object representation and recognition techniques. *Comput. Vis. Image Underst.*, 81(2), 2001.

[7] I. Chao, U. Pinkall, P. Sanan, and P. Schröder. A simple geometric model for elastic deformations. *ACM Trans. Graph.*, 29, July 2010.

[8] M. P. do Carmo. *Riemannian Geometry*. Birkhäuser, 1992.

[9] C. Hernndez, G. Vogiatzis, G. J. Brostow, B. Stenger, and R. Cipolla. Non-rigid photometric stereo with colored lights. In *ICCV*, 2007.

[10] V. G. Kim, Y. Lipman, and T. Funkhouser. Blended intrinsic maps. *ACM Trans. Graph.*, 30(4):79:1–79:12, 2011.

[11] Y. Lipman and T. Funkhouser. Möbius voting for surface correspondence. *ACM Trans. Graph.*, 28(3):1–12, 2009.

[12] F. Memoli and G. Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Found. Comput. Math.*, 5(3):313–347, 2005.

[13] Microsoft© Kinect. 2010.

[14] M. Ovsjanikov, Q. Merigot, F. Memoli, and L. J. Guibas. One point isometric matching with the heat kernel. *Comput. Graph. Forum*, 29(5):1555–1564, 2010.

[15] U. Pinkall and K. Polthier. Computing discrete minimal surfaces and their conjugates. *Experimental Mathematics*, 2(1):15–36, 1993.

[16] M. Salzmann, J. Pilet, S. Ilic, and P. Fua. Surface deformation models for nonrigid 3D shape recovery. *TPAMI*, 29:1481–1487, 2007.

[17] P. V. Sander, J. Snyder, S. J. Gortler, and H. Hoppe. Texture mapping progressive meshes. In *SIGGRAPH*, pages 409–416, 2001.

[18] A. Shaji, A. Varol, L. Torresani, and P. Fua. Simultaneous point matching and 3D deformable surface reconstruction. In *CVPR*, 2010.

[19] A. Sharma, R. P. Horaud, J. Cech, and E. Boyer. Topologically-robust 3D shape matching based on diffusion geometry and seed growing. In *CVPR*, 2011.

[20] O. Sorkine and M. Alexa. As-rigid-as-possible surface modeling. In *SGP*, pages 109–116, 2007.

[21] K. Stephenson. *Introduction to Circle Packing: The Theory of Discrete Analytic Functions*. Cambridge University Press, 2005.

[22] R. W. Sumner, J. Schmid, and M. Pauly. Embedded deformation for shape manipulation. *ACM Trans. Graph.*, 26, July 2007.

[23] O. van Kaick, H. Zhang, G. Hamarneh, and D. Cohen-Or. A survey on shape correspondence. *Comput. Graph. Forum*, 30(6):1681–1707, 2011.

[24] C. Wang, N. Komodakis, and N. Paragios. Markov Random Field modeling, inference & learning in computer vision & image understanding: A survey. *CVIU*, 117(11):1610–1627, 2013.

[25] C. Wang, Y. Zeng, D. Samaras, and N. Paragios. Modeling shapes with higher-order graphs: Methodology and applications. In S. Dickinson and Z. Pizlo, editors, *Shape Perception in Human and Computer Vision: An Interdisciplinary Perspective*, pages 459–471. Springer, 2013.

[26] S. Wang, Y. Wang, M. Jin, X. D. Gu, and D. Samaras. Conformal geometry and its applications on 3D shape matching, recognition, and stitching. *TPAMI*, 29(7):1209–1220, 2007.

[27] 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 *ICCV*, 2005.

[28] T. Werner. Revisiting the linear programming relaxation approach to Gibbs energy minimization and weighted constraint satisfaction. *TPAMI*, 32:1474–1488, 2010.

[29] T. Windheuser, U. Schlickewei1, F. R. Schmidt, and D. Cremers. Geometrically consistent elastic matching of 3D shapes: A linear programming solution. In *ICCV*, 2011.

[30] L. Yin, X. Chen, Y. Sun, T. Worm, and M. Reale. A high-resolution 3D dynamic facial expression database. In *Automatic Face and Gesture Recognition*, 2008.

[31] Y. Zeng, C. Wang, Y. Wang, X. Gu, D. Samaras, and N. Paragios. Dense non-rigid surface registration using high-order graph matching. In *CVPR*, 2010.

[32] Y. Zeng, C. Wang, Y. Wang, X. Gu, D. Samaras, and N. Paragios. Intrinsic dense 3D surface tracking. In *CVPR*, 2011.

[33] 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.