

Sparse Non-rigid Registration of 3D Shapes

Jingyu Yang¹, Ke Li¹, Kun Li^{2†} and Yu-Kun Lai³

¹School of Electronic Information Engineering, Tianjin University, Tianjin, China

²School of Computer Science and Technology, Tianjin University, Tianjin, China

³School of Computer Science and Informatics, Cardiff University, Wales, UK

Abstract

Non-rigid registration of 3D shapes is an essential task of increasing importance as commodity depth sensors become more widely available for scanning dynamic scenes. Non-rigid registration is much more challenging than rigid registration as it estimates a set of local transformations instead of a single global transformation, and hence is prone to the overfitting issue due to underdetermination. The common wisdom in previous methods is to impose an ℓ_2 -norm regularization on the local transformation differences. However, the ℓ_2 -norm regularization tends to bias the solution towards outliers and noise with heavy-tailed distribution, which is verified by the poor goodness-of-fit of the Gaussian distribution over transformation differences. On the contrary, Laplacian distribution fits well with the transformation differences, suggesting the use of a sparsity prior. We propose a sparse non-rigid registration (SNR) method with an ℓ_1 -norm regularized model for transformation estimation, which is effectively solved by an alternate direction method (ADM) under the augmented Lagrangian framework. We also devise a multi-resolution scheme for robust and progressive registration. Results on both public datasets and our scanned datasets show the superiority of our method, particularly in handling large-scale deformations as well as outliers and noise.

Categories and Subject Descriptors (according to ACM CCS): I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Hierarchy and geometric transformations

1. Introduction

3D registration is an active research topic in computer graphics and computer vision [BTP13, ARV07, ZNI*14]. Given a template shape (represented either as a point cloud or a mesh) and a target shape, the aim of 3D registration is to find a mapping between them that optimally transforms the template onto the target. 3D registration is an essential technique for 3D scanning systems, where several partial scans from different viewpoints are fused to form complete 3D models. The scanned data often contains noise and outliers; this is particularly problematic for cheap depth cameras such as Microsoft Kinects, which have now seen wider use. Registration methods robust to noise and outliers are thus highly desirable.

Rigid registration aims to find a *global* rigid-body transformation that aligns two shapes. Various methods have been

developed over the past decades [RKGB12]. Iterative Closest Point (ICP) and its variants [BM92] are the most classic ones. This kind of methods alternates between finding closest points and solving the optimal transformation. The major problem of such methods is its sensitivity to noise and outliers which are often observed in 3D scans. To address this, Bouaziz *et al.* [BTP13] propose a new variant of the ICP algorithm with sparsity-inducing norms, achieving superior results for the data with noise and outliers.

For dynamic scenes containing deformable objects, non-rigid registration is required. Such methods need to find a set of *local* transformations instead of a single global transformation. It is still a difficult, under-constrained problem due to high degrees of freedom and lack of prior knowledge. Most work in literature imposes the smoothness prior as the total energy of transformation differences over all the local neighbors in the ℓ_2 norm to make the problem well-posed [SSP07, LSP08]. However, the smoothness term in the ℓ_2 -norm tends to penalize large transformation differences. Therefore it is not well suited to deformations with

† Corresponding author: lik@tju.edu.cn.

piecewise near rigid parts (common for articulated models) where large changes happen in local areas, and is more substantially affected by noise and outliers of geometry and/or correspondences.

We propose a sparse non-rigid registration (SNR) method. Our method stands on the iterative framework consisting of two steps: 1) correspondence computation and 2) transformation estimation. We investigate the characteristics of non-rigid transformations of 3D shapes, and observe that the transformation functional on the graph is piecewise smooth, and can be well described by the Laplacian distribution on its local differences. Inspired by this, we propose a transformation estimation model based on the ℓ_1 -norm regularization of transformation differences. The model is solved by the alternating direction method under the augmented Lagrangian multiplier (ADM-ALM) framework. Our method is extended to a multi-resolution approach to achieve robust registration for complex deformations. The proposed method is evaluated on public datasets [BBK08, VBMP08] and real datasets captured using a Kinect depth sensor. The results demonstrate that the proposed SNR method obtains better results than the conventional non-rigid registration with the ℓ_2 -norm regularization. In particular, results on datasets with noise and outliers (both geometry and correspondences) show that our ℓ_1 -norm regularization is less sensitive to such problems typical in the input.

The main contributions of this work are summarized as:

- We propose a sparse non-rigid registration method. The SNR model is constructed based on the verified observation that non-rigid transformations are piecewise smooth on the underlying graph, and is able to handle flexible deformations of local geometries. The SNR model is transferred into a series of alternating optimization subproblems with exact solutions and guaranteed convergence.
- We establish a multi-resolution non-rigid registration scheme. The template and target shapes are downsampled into low resolution versions of several scales. The non-rigid registration at the full resolution is obtained in a coarse-to-fine manner. This strategy is not only more efficient, but also prevents the method from trapping into poor local minimums, providing robust registration for complicated deformations.

2. Related Work

Over the last two decades, non-rigid registration has been an intensively studied problem. A complete survey is beyond the scope of this paper. Please refer to [TCL*13] for a recent survey. In this section, we provide a brief summary of most relevant work from two aspects:

Model Selection: Non-rigid registration requires some assumptions about the underlying deformation. Some methods use a piecewise rigid transformation model, in which global

rigid transformations are computed for bones and local non-rigid deformations occur near joints. Allen *et al.* [ACP02] place markers on the object to help reconstruct the pose of a scan, which is then used as a basis for modeling deformation. Pekelny *et al.* [PG08] use predefined bone information to find and track transformations of rigid components. Further techniques take more generic deformations into consideration. Chui *et al.* [CR03] adopt the thin-plate spline (TPS) as the model to represent non-rigid transformations. Myronenko *et al.* [MS10] regularize the displacement field to ensure smoothness. Papazovet *et al.* [PB11] model as-rigid-as-possible shape deformations using an ordinary differential equation.

Recently, local affine transformations are frequently used in non-rigid registration [ACP03]. Amberg *et al.* [ARV07] use a stiffness term to impose similarity between neighboring affine matrices. Liao *et al.* [LZW*09] use differential coordinates, which can be considered as local affine transformations with smoothness constraints. Rouhani *et al.* [RBS14] model non-rigid deformation as an integration of locally rigid transformations. The template shape is clustered into a collection of small patches that can deform rigidly, and the target is represented by an implicit function. The implicit function induces a distance field that helps align the template to the target without explicit correspondences. In our work, we use local affine transformation as it allows more flexibility to capture fine surface details.

Regularization and Optimization: Non-rigid registration is often formulated as an optimization problem. Most methods formulate some energy functional with both data (measuring fitting error) and regularization terms. The latter helps to preserve smoothness, avoid overfitting and thus make the optimization more robust to noise and outliers.

Amberg *et al.* [ARV07] use a stiffness term to penalize differences between transformation matrices of neighboring vertices. Süßmuth *et al.* [SWG08] and Liao *et al.* [LZW*09] use a generalized as-rigid-as-possible energy [SA07] and thin-plate splines (TPS) [CR03] to promote smoothness. Wand *et al.* [WAO*09] take a set of time-varying point data as input, and reconstruct a single shape and a deformation field that fit the data. The algorithm encourages the normals of adjacent points to vary as smoothly as possible. To improve robustness, Li *et al.* [LSP08] solve correspondences, confidence weights, and a deformation field within a single optimization framework. Their work [LSP08, LAGP09] uses regularizers to improve local rigidity and transformation smoothness.

Regularization using the ℓ_1 -norm: The ℓ_1 -norm has been widely used to solve the registration problem. Hontani *et al.* [HMS12] propose a statistical shape model built based on some training surfaces, which is then incorporated into the non-rigid ICP framework. Outliers can be detected based on their sparseness. Based on the ℓ_1 -norm, Flöry *et al.* [FH10] propose a technique for rigid registration

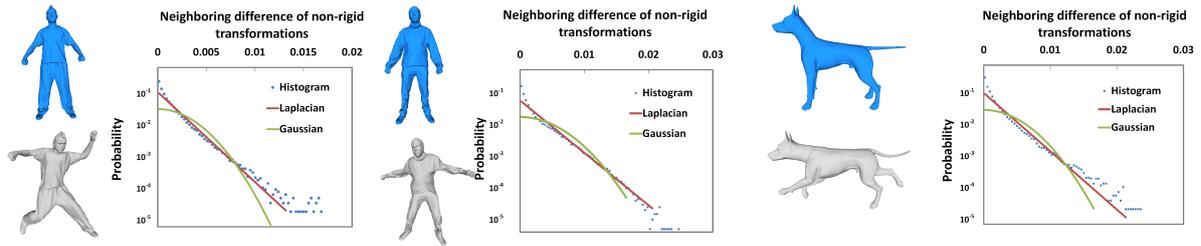


Figure 1: Normalized histograms of transformation differences and the associated fitted Laplacian distributions and Gaussian distributions for datasets (a) Bouncing, (b) Jumping, and (c) Dog. The Laplacian distribution (corresponding to the ℓ_1 -norm regularizer) fits the histograms significantly better than the Gaussian distribution (corresponding to the ℓ_2 -norm regularizer), suggesting the use of sparsity — promoting ℓ_1 norm regularization.

of surfaces. Their technique performs better than ℓ_2 -method in the presence of outliers. Bouaziz *et al.* [BTP13] propose an ICP-based algorithm for rigid registration with a general ℓ_p sparse data term ($p \leq 1$), which shows improved robustness to noise and outliers. Yu *et al.* [YZL*14] employ sparsity constraints to handle the outliers and gross errors when dealing with noisy medical images.

In this paper, we propose a new non-rigid registration method based on sparse representation. We formulate the smoothness term with the ℓ_1 norm instead of the conventional ℓ_2 norm to ensure that adjacent transformations are similar. We will show that our method is more robust to noise and outliers, consistent with theoretical discussion [KK03].

3. Proposed Method

3.1. Motivation

Non-rigid registration has much higher degrees of freedom, and is more challenging than rigid registration. It is prone to the over-fitting problem as one can optimize a transformation for each vertex with very small matching error, but the deformation of the surface may be unnatural or even violate the underlying physical model. Hence, the key is to impose proper priors to make the problem well-posed.

Most work in the literature assumes smoothness of transformations as deformations between neighboring vertices are usually similar. Following this, the smoothness is defined as the total energy of transformation differences over all the local neighbors in the ℓ_2 norm [SSP07,LSP08]. While it helps to make the problem well-posed, the ℓ_2 -norm based smoothness tends to penalize large transformation differences, causing incorrect estimation around vertices with intensive deformations. From the statistic point of view, the ℓ_2 -norm regularization is to assume the Gaussian distribution of transformation differences [Bis06]. However, deformations of most interested 3D surfaces, e.g., humans, animals, and robots, are articulated. In other words, the deformations vary smoothly over the vertices except for flexible

joints with intensive deformations, and hence, the transformations are piecewise-smooth signals residing on 3D surfaces. This suggests that the transformation differences are sparse, and should be modeled by a heavy-tailed distribution, rather than being dense and modeled by the rapidly-vanishing Gaussian distribution. This is analogous to the case of piecewise smooth images with sparse representations under some transforms/dictionaries [EFM10].

This intuition is verified in Fig. 1. These 3D shapes come with ground truth dense correspondences that describe the non-rigid deformations. Their normalized histograms of local rigid transform differences between adjacent vertices are presented, together with fitted Laplacian distributions and Gaussian distributions. The Laplacian distributions tightly fit the histograms while the Gaussian distributions show significant departures. This motivates us to model transformation differences with the Laplacian distribution, or equivalently to use the ℓ_1 norm in measuring the smoothness of non-rigid transformations.

3.2. Sparse Non-rigid Registration Method

3.2.1. Iterative Framework

Our non-rigid registration takes two shapes (a template shape and a target shape) and finds a collection of local transformations to deform the template shape to the target. Like most non-rigid registration methods, our method operates iteratively. Each iteration consists of two steps, i.e., correspondence estimation and transformation estimation. In the first step, the correspondences between the template and the target are estimated given the alignment with the transformations obtained from the last iteration. Accurate correspondence estimation is highly desirable for rapid convergence. Similar to ICP, we also use closest points between template and target shapes to suggest potential correspondences. However, this is only effective when the template and the target are in reasonably good alignment. To address this, in particular at the beginning of the iteration where both shapes can be arbitrarily placed, we incorporate correspon-

dences based on intrinsic geometric features. We use a state-of-the-art technique based on local geometric similarity and diffusion pruning of inconsistent correspondence [TMRL14] which often provides reliable correspondences. Alternative correspondence techniques or manual specification of a few correspondences may instead be used (see Fig. 8). After the first iteration, the template mesh is generally close to the target so we add ICP-based correspondences for the following iterations. In the second step (Sec. 3.2.2), non-rigid transformations are estimated by a proposed energy-minimization approach based on sparse representation (SR) given the correspondences from the first step.

3.2.2. SR-based Transformation Estimation

We start by giving a formal definition of the problem. Denote a template set of 3D points by $\mathcal{V} \triangleq \{\mathbf{v}_1, \dots, \mathbf{v}_N\}$, where $\mathbf{v}_i \triangleq [x_i, y_i, z_i, 1]^\top$ is a 3D point in the homogeneous coordinate and N is the number of points. Similarly, we define a target set of 3D points as $\mathcal{U} \triangleq \{\mathbf{u}_1, \dots, \mathbf{u}_M\}$. Let $f: \{1, \dots, N\} \mapsto \{1, \dots, M\}$ be the index mapping from the template points to the target points established by correspondence computation: $\mathbf{u}_{f(i)} \in \mathcal{U}$ is the correspondence point of $\mathbf{v}_i \in \mathcal{V}$. For non-rigid registration, we allow an affine transformation for each point in the template to cover a wide range of non-rigid deformations. Denote the set of non-rigid transformations by $\mathcal{X} \triangleq \{\mathbf{X}_1, \dots, \mathbf{X}_N\}$, where \mathbf{X}_i is the 3×4 transformation matrix for point \mathbf{v}_i . For convenience, denote by $\mathbf{X} \triangleq [\mathbf{X}_1, \dots, \mathbf{X}_N]^\top$ of size $4N \times 3$ the ensemble matrix containing N transformation matrices to be estimated. Given a correspondence mapping f , the aim of the proposed method is to find non-rigid transformations \mathcal{X} , or equivalently \mathbf{X} , that transforms the template \mathcal{V} into the target \mathcal{U} as accurately as possible while ensuring the validity of 3D deformations (e.g., subjecting to physical constraints of the 3D objects) by imposing a sparse prior of transformations.

To this end, we design an energy function to evaluate the goodness of transformations:

$$E(\mathbf{X}; f) = E_{data}(\mathbf{X}; f) + \alpha E_{smooth}(\mathbf{X}), \quad (1)$$

where $E_{data}(\mathbf{X})$ is the data term to measure the registration accuracy, $E_{smooth}(\mathbf{X})$ is the smoothness term to measure the smoothness of local transformations, and α is a weight to twist the importance of the two terms. The data term and the smoothness term are defined as follows.

Data term: The accuracy of deformation can be measured by the closeness of the transformed points to their corresponding target points. For some points in \mathcal{V} , no reliable correspondences from \mathcal{U} can be found, which is common in most matching algorithms such as SHOT [STDS14], diffusion pruning [TMRL14] or closest points (using a standard correspondence rejection strategy). We introduce a weight denoted by w_i for each point: w_i is set at zero if \mathbf{v}_i does not have a corresponding point in \mathcal{U} , and one otherwise. Given

the correspondence mapping f , the data term is defined as

$$E_{data}(\mathbf{X}; f) = \sum_{\mathbf{v}_i \in \mathcal{V}} w_i \|\mathbf{X}_i \mathbf{v}_i - \tilde{\mathbf{u}}_{f(i)}\|_2^2 \quad (2)$$

where $\tilde{\mathbf{u}}_{f(i)}$ is the Cartesian coordinate of $\mathbf{u}_{f(i)}$.

We define the following matrix/vector form of the variables to reformulate the data term (2) for the compact representation in algorithm derivation.

$$\begin{aligned} \mathbf{W} &= \text{diag}(\sqrt{w_1}, \dots, \sqrt{w_N}), \\ \mathbf{V} &= \text{diag}(\mathbf{v}_1^\top, \dots, \mathbf{v}_N^\top), \\ \tilde{\mathbf{U}}_f &= [\tilde{\mathbf{u}}_{f(1)} \quad \dots \quad \tilde{\mathbf{u}}_{f(N)}]^\top, \end{aligned} \quad (3)$$

where $\text{diag}(\cdot)$ returns the block-wise diagonal matrix of the input vectors. Then, the data term can be rewritten as

$$E_{data}(\mathbf{X}; f) = \|\mathbf{W}(\mathbf{V}\mathbf{X} - \tilde{\mathbf{U}}_f)\|_F^2, \quad (4)$$

where $\|\cdot\|_F^2$ denotes Frobenius norm of a matrix.

Smoothness term: To impose a sparse prior on transformation differences, we define a graph $\mathcal{G} \triangleq (\mathcal{V}, \mathcal{E})$, where the vertices of the graph are the 3D points in \mathcal{V} , and the edges of the graph are denoted by \mathcal{E} . For a 3D mesh, edges of the graph are simply defined by the edges of the mesh; for 3D point clouds, edges can be defined by connecting each vertex with its K -nearest neighbors (K is typically set to 6).

For both cases, the edge set can be defined with a neighboring system. Denote by \mathcal{N}_i the neighborhood of vertex \mathbf{v}_i , and an edge e_{ij} is defined between each neighboring vertex \mathbf{v}_j and \mathbf{v}_i . So, we have $\mathcal{E} = \{e_{ij} \mid \mathbf{v}_j \in \mathcal{N}_i, \mathbf{v}_i \in \mathcal{V}\}$. The analysis in Sec. 3.1 reveals that transformation differences on 3D surfaces are sparse, and are well fitted by a Laplacian distribution. Therefore, the cost of transformation smoothness is measured by the ℓ_1 norm of transformation differences over the neighboring system \mathcal{E} :

$$E_{smooth}(\mathbf{X}) = \sum_{e_{ij} \in \mathcal{E}} \|\mathbf{X}_i - \mathbf{X}_j\|_1, \quad (5)$$

where $\|\cdot\|_1$ here represents the ℓ_1 norm of the matrix considered as a long vector.

Similar to the data term, we define a differential matrix $\mathbf{K} \in \{-1, 1\}^{|\mathcal{E}| \times |\mathcal{V}|}$ on the graph \mathcal{G} for concise presentation. Concretely, each row of \mathbf{K} corresponds to an edge in \mathcal{E} and each column corresponds to a vertex in \mathcal{V} . Each row in \mathbf{K} has only two nonzero entries. For example, assuming the r^{th} row in \mathbf{K} associates with edge e_{ij} , then the entry linked to the reference vertex \mathbf{v}_i is set at 1, while the one linked to the neighboring vertex \mathbf{v}_j is set at -1, i.e. $k_{ri} = 1$ and $k_{rj} = -1$. Let \mathbf{I}_4 be a 4×4 identity matrix to expand the differential matrix \mathbf{K} for differentiating 4×3 matrices (transpose of \mathbf{X}_i). The smoothness term can be rewritten as

$$E_{smooth} = \|\mathbf{K} \otimes \mathbf{I}_4 \mathbf{X}\|_1, \quad (6)$$

where \otimes denotes the operator of Kronecker product. The

matrix-vector form of the energy function is obtained by substituting (4) and (6) into (1), and the transformation matrices are obtained by minimizing the following function:

$$\min_{\mathbf{X}} \|\mathbf{W}(\mathbf{V}\mathbf{X} - \tilde{\mathbf{U}}_f)\|_F^2 + \alpha \|\mathbf{K} \otimes \mathbf{I}_4 \mathbf{X}\|_1, \quad (7)$$

whose algorithm is detailed in the following section.

3.3. Proposed ADM-ALM Algorithm

Minimization (7) is essentially an unconstrained optimization of \mathbf{X} . The $\ell_1 - \ell_2$ structure of energy functions has been intensively studied over the past decade in line with sparse representation [YPXD09]. However, the differential matrix $\mathbf{K} \otimes \mathbf{I}_4$ impedes the direct use of efficient proximal operators to attack the non-differential ℓ_1 term. We first transform the minimization (7) into the following form with an auxiliary variable \mathbf{A} .

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{A}} \|\mathbf{W}(\mathbf{V}\mathbf{X} - \tilde{\mathbf{U}}_f)\|_F^2 + \alpha \|\mathbf{A}\|_1, \\ \text{s.t. } \mathbf{A} = \mathbf{B}\mathbf{X}, \end{aligned} \quad (8)$$

where $\mathbf{B} \equiv \mathbf{K} \otimes \mathbf{I}_4$ is introduced for concise notations.

We solve the constrained minimization (8) using the augmented Lagrangian method [Ber82]. The ALM method converts the original problem (8) to iterative minimization of its augmented Lagrangian function:

$$\begin{aligned} L(\mathbf{X}, \mathbf{A}, \mathbf{Y}, \mu) = \|\mathbf{W}(\mathbf{V}\mathbf{X} - \tilde{\mathbf{U}}_f)\|_F^2 + \alpha \|\mathbf{A}\|_1 \\ + \langle \mathbf{Y}, \mathbf{A} - \mathbf{B}\mathbf{X} \rangle + \frac{\mu}{2} \|\mathbf{A} - \mathbf{B}\mathbf{X}\|_F^2, \end{aligned} \quad (9)$$

where μ is a positive constant, \mathbf{Y} is the Lagrangian multiplier, and $\langle \cdot, \cdot \rangle$ denotes the inner product of two matrices considered as long vectors. Under the standard ALM framework, \mathbf{Y} and μ can be efficiently updated. However, each iteration has to solve \mathbf{X} and \mathbf{A} simultaneously, which is difficult and computationally demanding. Hence, we resort to the alternate direction method (ADM) [BPC*11] to optimize \mathbf{A} and \mathbf{X} separately at each iteration:

$$\begin{cases} \mathbf{A}^{(k+1)} = \arg \min_{\mathbf{A}} \alpha \|\mathbf{A}\|_1 + \langle \mathbf{Y}^{(k)}, \mathbf{A} - \mathbf{B}\mathbf{X}^k \rangle \\ \quad + \frac{\mu^{(k)}}{2} \|\mathbf{A} - \mathbf{B}\mathbf{X}^k\|_F^2, \\ \mathbf{X}^{(k+1)} = \arg \min_{\mathbf{X}} \|\mathbf{W}(\mathbf{V}\mathbf{X} - \tilde{\mathbf{U}}_f)\|_F^2 + \langle \mathbf{Y}^{(k)}, \\ \quad \mathbf{A}^{(k+1)} - \mathbf{B}\mathbf{X} \rangle + \frac{\mu^{(k)}}{2} \|\mathbf{A}^{(k+1)} - \mathbf{B}\mathbf{X}\|_F^2, \\ \mathbf{Y}^{(k+1)} = \mathbf{Y}^{(k)} + \left(\mathbf{A}^{(k+1)} - \mathbf{B}\mathbf{X}^{(k+1)} \right), \\ \mu^{(k+1)} = \rho \mu^{(k)}, \quad \rho > 1 \end{cases} \quad (10)$$

The \mathbf{A} -subproblem has the following closed solution:

$$\mathbf{A}^{(k+1)} = \text{shrink} \left(\mathbf{B}\mathbf{X}^{(k)} - \frac{1}{\mu^{(k)}} \mathbf{Y}^{(k)}, \frac{\alpha}{\mu^{(k)}} \right), \quad (11)$$

where $\text{shrink}(\cdot, \cdot)$ is the shrinkage function applied on the matrix element-wise:

$$\text{shrink}(x, \tau) = \text{sign}(x) \max(|x| - \tau, 0). \quad (12)$$

Algorithm 1. Non-rigid registration

1. Input: template \mathcal{V} , target \mathcal{U} .
 2. While not converged do
 3. Find correspondence mapping $f^{(l)} : \mathcal{V} \mapsto \mathcal{U}$;
 4. Solve transformations $\mathbf{X}^{(l)}$ via Algorithm 2;
 5. End while
 6. Output: \mathbf{X}
-

The \mathbf{X} -subproblem is a quadratic optimization, which can be solved by setting its derivatives with respect to \mathbf{X} to zeros, yielding the following linear equations:

$$\begin{aligned} \left(\mathbf{V}^\top \mathbf{W}^\top \mathbf{W} \mathbf{V} + \mu^{(k)} \mathbf{B}^\top \mathbf{B} \right) \mathbf{X} = \mathbf{V}^\top \mathbf{W}^\top \mathbf{W} \tilde{\mathbf{U}}_f \\ + \mathbf{B}^\top \mathbf{Y}^{(k)} + \mu^{(k)} \mathbf{B}^\top \mathbf{A}^{(k+1)}. \end{aligned} \quad (13)$$

$\mathbf{X}^{(k+1)}$ can be obtained by multiplying the left-side with the inverse of $\mathbf{V}^\top \mathbf{W}^\top \mathbf{W} \mathbf{V} + \mu^{(k)} \mathbf{B}^\top \mathbf{B}$. However, the inversion can be problematic when the scale of the problem is large, e.g., when registration point clouds with tens of thousands of points. We avoid this problem by using the LDL decomposition, a variant of Cholesky factorization.

$$(\mathbf{L}, \mathbf{D}) = \text{ldl} \left(\mathbf{V}^\top \mathbf{W}^\top \mathbf{W} \mathbf{V} + \mu^{(k)} \mathbf{B}^\top \mathbf{B} \right). \quad (14)$$

where \mathbf{L} and \mathbf{D} are the lower triangular matrix and the diagonal matrix returned by the LDL decomposition. Then, we introduce two auxiliary variables \mathbf{Q} and \mathbf{Z} for clear presentation, and $\mathbf{X}^{(k+1)}$ is obtained by solving the following one diagonal and two triangular systems:

$$\begin{aligned} \mathbf{L}\mathbf{Q} &= \mathbf{D}^\top \mathbf{V}^\top \mathbf{W}^\top \mathbf{W} \tilde{\mathbf{U}}_f + \mathbf{B}^\top \mathbf{Y}^{(k)} + \mu^{(k)} \mathbf{B}^\top \mathbf{A}^{(k+1)}, \\ \mathbf{D}\mathbf{Z} &= \mathbf{Q}, \\ \mathbf{L}^\top \mathbf{X} &= \mathbf{Z}. \end{aligned} \quad (15)$$

We choose the ADM-ALM algorithm as it is much faster than many other ℓ_1 -minimization solvers such as interior-point algorithms and proximal point methods, and is less sensitive to problem sizes [YZB*13]. The iterative non-rigid registration framework is summarized in Algorithm 1. The ADM-ALM algorithm (an inner iteration of Algorithm 1) to estimate non-rigid transformations \mathbf{X} is summarized in Algorithm 2. Fig. 2 illustrates the typical convergence behavior of both algorithms. We observe that the total energy decreases at each iteration for both the inner and outer loops. The outer loop (Algorithm 1) converges after about 15 iterations. As the registration gets more accurate with the increasing iteration of Algorithm 1, fewer iterations are needed for Algorithm 2 to converge.

3.4. Multi-resolution Approach

The transformation estimation model (7) is convex w.r.t. \mathbf{X} given a correspondence mapping $f : \mathcal{V} \mapsto \mathcal{U}$, and the ADM-ALM algorithm is guaranteed to converge to the optimum [BPC*11]. However, the ideal minimization should actually

Algorithm 2. ADM-ALM algorithm to solve (7)

1. Input: $\tilde{\mathbf{U}}_{f^{(l)}} \in \mathbf{R}^{N \times 3}$, $\mathbf{V} \in \mathbf{R}^{N \times 4N}$, $\mathbf{B} \in \mathbf{R}^{|\mathcal{E}| \times 4|\mathcal{V}|}$;
2. Initialize: $\mathbf{X}^{(l,0)} = \mathbf{0}$, $\mathbf{Y}^{(0)} = \mathbf{0}$, $\mu > 0$, $\rho > 1$;
3. While not converged do
4. Solve $\mathbf{A}^{(l,k+1)}$ by (11);
5. Solve $\mathbf{X}^{(l,k+1)}$ by (14)~(15);
6. $\mathbf{Y}^{(k+1)} = \mathbf{Y}^{(k)} + \mu^{(k)} (\mathbf{V}^{(k+1)} - \mathbf{B}\mathbf{X}^{(k+1)})$;
7. $\mu^{(k+1)} = \rho\mu^{(k)}$;
8. End while
9. Output: $\mathbf{X}^{(l)}$.

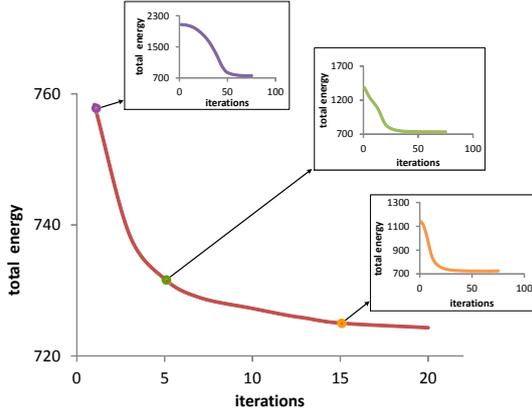


Figure 2: Total energy of (7) vs. the number of iterations for Algorithm 1 (curve in the large subplot) and Algorithm 2 (curves in three smaller subplots) on dataset Bouncing. The number of iterations for Algorithm 2 to converge decreases with increasing number of iterations of Algorithm 1.

be w.r.t. both \mathbf{X} and f , and possibly has many local optima. The separate treatment in Algorithm 1 may be trapped to a local minimum that significantly departs from the desired one. We propose a continuation approach that operates on multi-resolution template and target shapes from coarse to fine to reach a promising solution. Furthermore, the multi-resolution approach also helps to attack large scale problems that would otherwise be prohibitively expensive (in terms of computation and storage costs).

Denote by $\mathcal{V}^{(s)}$ the s^{th} scale of the template shape \mathcal{V} , obtained by downsampling using a standard approach [GH97, PGK02]. Assume that we totally have S scales, we have multi-resolution shapes, $\mathcal{V}^{(S)}, \mathcal{V}^{(S-1)}, \dots, \mathcal{V}^{(0)}$, where $\mathcal{V}^{(S)}$ is the shape at the coarsest resolution and $\mathcal{V}^{(0)} \equiv \mathcal{V}$ is that at the full resolution. Similarly, we define the multi-resolution series for the target: $\mathcal{U}^{(S)}, \mathcal{U}^{(S-1)}, \dots, \mathcal{U}^{(0)}$.

As illustrated in Fig. 3, we first obtain non-rigid transformation $\mathbf{X}^{(S)}$ and corresponding mapping $f^{(S)}$ from $\mathcal{V}^{(S)}$ and $\mathcal{U}^{(S)}$ at the coarsest scale using Algorithms 1 and 2. The non-rigid registration at the s^{th} scale is solved by warm-starting

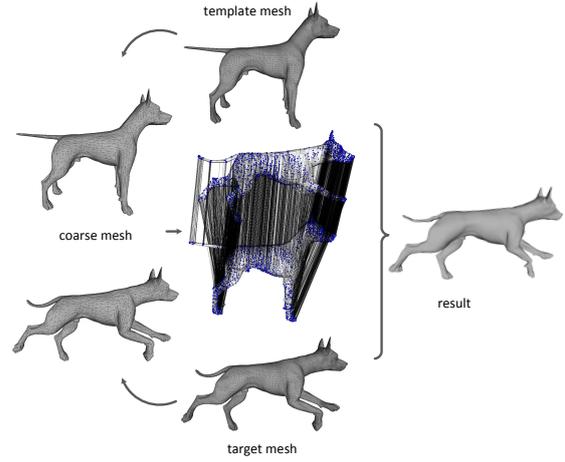


Figure 3: Illustration of multi-resolution non-rigid registration. The deformation of high-resolution template is predicted by the deformation of low-resolution template.

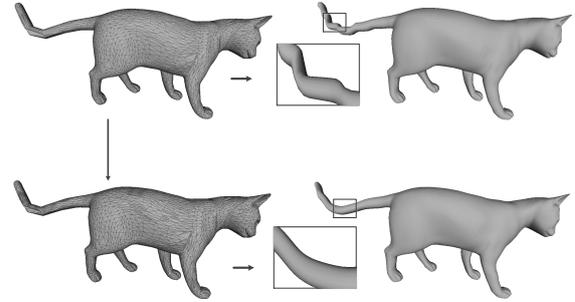


Figure 4: Example of multi-resolution registration. When the resolution is too low, the transformations cannot be effectively represented, causing artifacts. This is resolved by proceeding to a finer level of shapes.

the algorithm with the solution at the $(s+1)^{\text{th}}$ scale. Let $\mathbf{v}_i^{(s)}$ be a vertex at the s^{th} scale. The deformation of $\mathbf{v}_i^{(s)}$ is predicted by the weighted average of the deformations of vertices within a spherical neighborhood of radius r at the $(s+1)^{\text{th}}$ scale [LAGP09]. Denoted by $\tilde{\mathbf{v}}_i^{(s)}$ the predicted position of $\mathbf{v}_i^{(s)}$, and by $\Gamma_i^{(s+1)}$ the index set of vertices within the spherical neighborhood.

$$\tilde{\mathbf{v}}_i^{(s)} = \sum_{j \in \Gamma_i^{(s+1)}} a_j(\mathbf{v}_i^{(s)}, \mathbf{v}_j^{(s+1)}, r) \mathbf{X}_j^{(s+1)} \mathbf{v}_j^{(s+1)}, \quad (16)$$

The weight is defined as

$$a_j(\mathbf{v}_i, \mathbf{v}_j, r) = \max\left(0, \left(1 - \frac{d^2(\mathbf{v}_i, \mathbf{v}_j)}{r^2}\right)\right), \quad (17)$$

where r is the effective radius (set based on the average edge length) and $d(\mathbf{v}_i, \mathbf{v}_j)$ is the Euclidean distance between \mathbf{v}_i

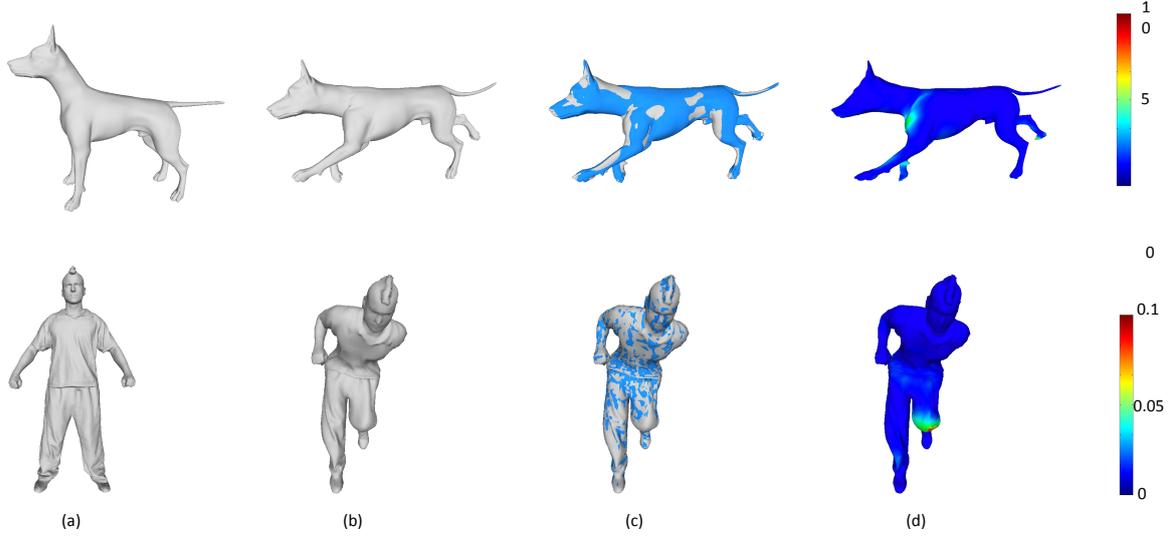


Figure 5: Results on Dog (top) and Bouncing (bottom): (a) template, (b) target, (c) results, and (d) fitting errors.

and \mathbf{v}_j . The weight drops when the distance between these two vertices increases.

By predicting the deformation of each vertex in $\mathcal{V}^{(s)}$ via (16), we obtain a predicted deformation, denoted by $\tilde{\mathcal{V}}^{(s)}$. Then, the correspondence mapping $f^{(s)}: \mathcal{V}^{(s)} \mapsto \mathcal{U}^{(s)}$ is estimated as in the single resolution ICP situation, using a combination of intrinsic-based and ICP-based correspondences. This procedure proceeds from a coarser level to a finer level and terminates if either the highest level is reached, or the transformation obtained is sufficiently accurate when applied to the input template and target shapes using (16). Fig. 4 demonstrates that coarse template and target shapes are not sufficient to represent the transformations effectively, leading to the wavy artifact on the tail. Higher resolution shapes effectively solve this problem. By using a coarse-to-fine strategy, our method has less chance of getting stuck at poor local minima, and handles high resolution shapes effectively as often coarser shapes are sufficient to represent the transformations. In practice, the actual number of scales and the resolution of the simplified meshes are not critical. The meshes can be obtained by repeated simplification up to a required number of vertices at the coarsest scale. In our implementation, 1000 vertices are used at the coarsest scale.

4. Experimental Results

4.1. Results on Clean Datasets

Fig. 5 gives registration results for two datasets: *Dog* from the TOSCA high-resolution dataset [BBK08] and *Bouncing* from the human motion dataset [VBMP08]. Final registration results are shown as the overlap of the deformed template (blue) and the target (ground truth), and fitting errors

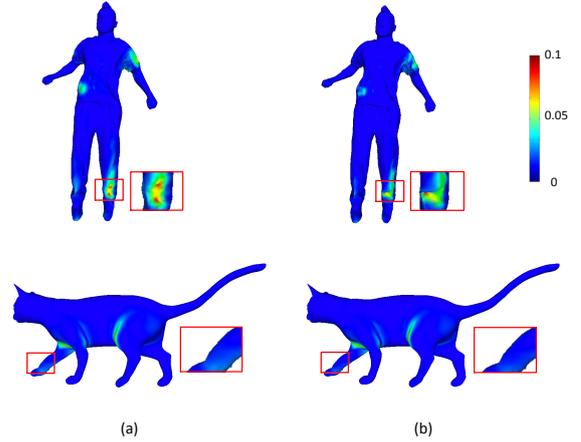


Figure 6: Comparison results of ℓ_2 -norm regularization (left) and our ℓ_1 -norm regularization (right) on datasets Bouncing (top) and Cat (bottom).

errors are color-coded on the reconstructed mesh for visual inspection. For vertex \mathbf{v}_i , registration error is defined as $\|\mathbf{X}_i \mathbf{v}_i - \mathbf{g}_i\|_2^2$, where \mathbf{g}_i is the ground-truth correspondence of \mathbf{v}_i . The results show that our method has promising registration results on the test datasets.

We compare our method with the following classic ℓ_2 -norm regularized non-rigid ICP method:

$$\min_{\mathbf{X}} \alpha \|\mathbf{B}\mathbf{X}\|_F^2 + \|\mathbf{W}(\mathbf{V}\mathbf{X} - \tilde{\mathbf{U}}_f)\|_F^2. \quad (18)$$

This is equivalent to solving the following linear equations:

$$\left(\mathbf{V}^\top \mathbf{W}^\top \mathbf{W} \mathbf{V} + \alpha \mathbf{B}^\top \mathbf{B}\right) \mathbf{X} = \mathbf{V}^\top \mathbf{W}^\top \tilde{\mathbf{U}}_f, \quad (19)$$

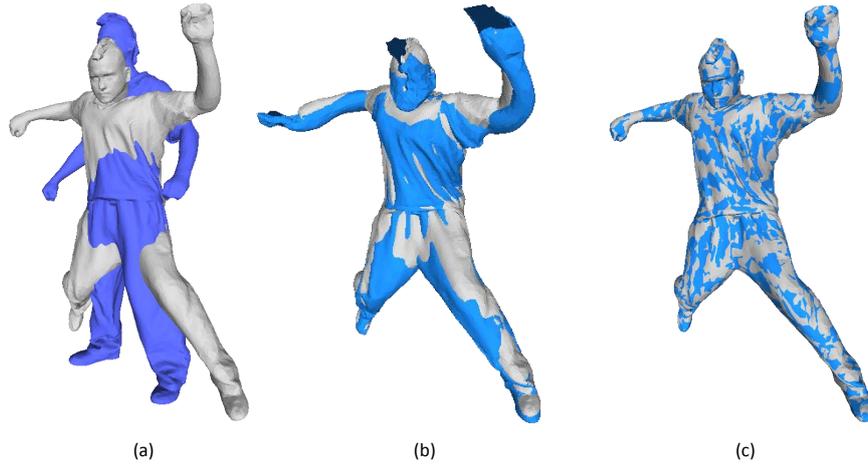


Figure 7: Comparison results on Bouncing dataset: (a) template and target, (b) the method in [LSP08] and (c) our method.

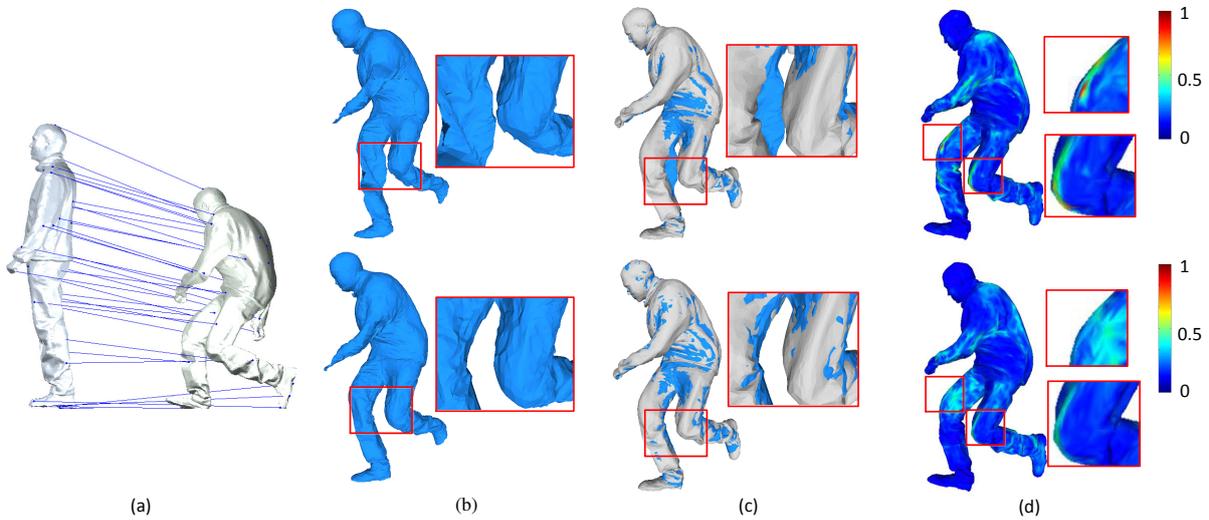


Figure 8: Comparison results on Crane dataset with 35 manually-specified correspondences: (a) given correspondences, (b) deformed meshes, (c) overlap results, and (d) color-coded local transformation differences (in Frobenius norm). For (b)~(d), the top row place the ℓ_2 -norm registration results while the bottom row presents our results.

by the LDL decomposition, similar to Eq. (13). In each registration, we adjust the value of α until we get the best result (accurate without loss of smoothness). As shown in Fig. 6, our method has better registration results with less fitting errors in the areas with intensive deformations, such as the wrinkles around the pant leg of the person and the paw of the cat highlighted in rectangles.

Our method is also compared with a state-of-the-art method [LSP08] in Fig. 7. Obvious errors on the head and hands can be seen on the result of the method in [LSP08], while our method achieves accurate non-rigid registration for the whole body. [LSP08] iteratively optimizes correspondences and transformations, with the initial correspondences established using a closet point approach. This approach

works effectively when the template and target shapes are close such that good initial correspondences can be obtained. In this example however, the substantial pose change means that it is difficult to obtain a good initialization.

We further compare our method with the ℓ_2 -regularized method by manually giving 35 correspondences on Crane dataset in Fig. 8. In this case, there is no correspondence around the right knee of the actor. As shown in Fig. 8 (b) and (c), with ℓ_1 regularization, registration tends to follow significant deformation (e.g. at joints) better than ℓ_2 regularization. This is further illustrated by the the color-coded local transformation differences in Fig. 8 (d), showing that the ℓ_2 regularization presents significant discontinuities due to incorrect deformation. The results demonstrate that our

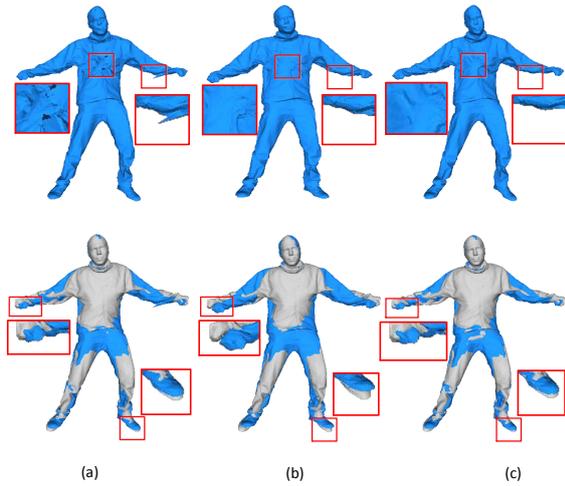


Figure 9: Registration results with partially incorrect correspondences on Jumping dataset shown as the deformed template and target shapes (top row) and overlaps of the deformed template and target shapes (bottom row): (a)(b) ℓ_2 -norm registration results with $\alpha = 1.5$ and 3 respectively, (c) ℓ_1 -norm registration results.

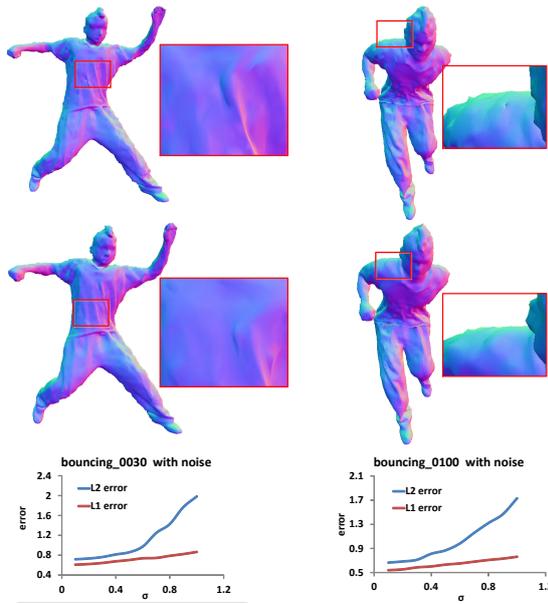


Figure 10: Registration results of ℓ_2 -norm regularization (top) and our ℓ_1 -norm regularization (middle) on datasets with noise. The curves of average fitting errors vs. normalized noise levels (bottom) show that our method is more robust than the ℓ_1 -norm regularized method.

method is particularly effective when the deformation is substantial and nearly piecewise rigid as the ℓ_1 regularization allows large deviations without compensating others. As a

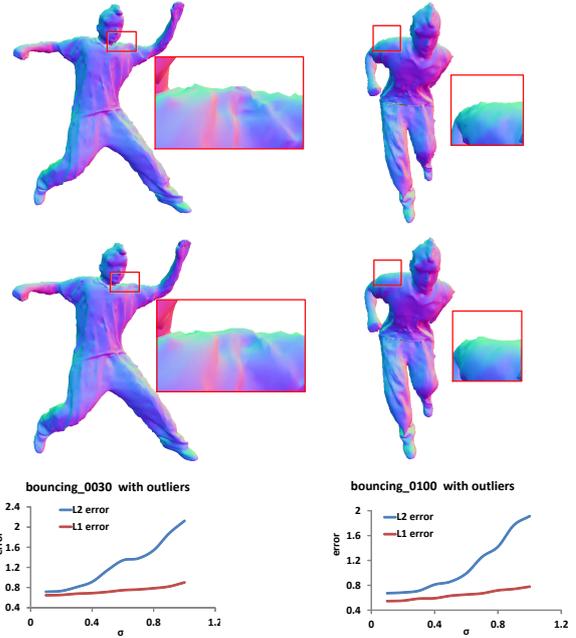


Figure 11: Registration results of ℓ_2 -norm regularization (top) and our ℓ_1 -norm regularization (middle) on datasets with outliers. The curves of average fitting errors vs. normalized outliers levels (bottom) show that our method is more robust than the ℓ_1 -norm regularized method.

result, ℓ_1 regularization helps converge to a much more accurate local minimum than ℓ_2 regularization.

4.2. Results on Noisy Datasets

This section evaluates the robustness of our method in the following noisy cases:

1) Correspondences with partially incorrect matchings:

It is common to include incorrect correspondences using established methods. We simulate this in a controlled manner. We obtain two thirds of correspondences using diffusion pruning [TMRL14] and the remaining one third using local geometric feature matching based on SHOT signatures [STDS14]. The majority of correspondences from the former are correct while many correspondences from the latter are incorrect due to the ambiguity of local features. The correspondences computed by SHOT signatures are shown with red lines in our supplementary document.

Fig. 9 shows that the sparseness regularizer in our model (7) is significantly more robust than the classic quadratic regularizer in (18). The ℓ_2 norm based registration presents holes and outliers in Fig. 9 (a) when the weight on the smoothness term is small ($\alpha = 1.5$). We try to increase the weight, but the registration result will become over-smoothed, which is inferior to our result. Fig. 9 (b) presents the result of the ℓ_2 norm based registration with $\alpha = 3$: severe misalignments are observed at the hands and feet.

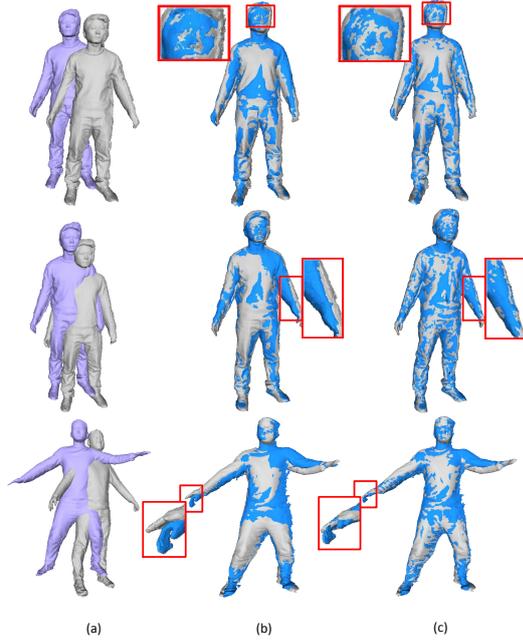


Figure 12: Results on real data captured by Kinect: (a) template and target, (b) ℓ_2 -norm regularization, (c) our ℓ_1 -norm regularization.

2) Target shapes with noise or outliers: 3D shapes of targets are polluted with dense noise or sparse outliers along the normal directions of the associated vertices. For dense noise, all the target vertices are perturbed with Gaussian noise; while, for the outliers, 1% of target vertices are contaminated with Gaussian noise. Fig. 10 and Fig. 11 present registration results for datasets with noise and outliers, respectively. The variance of the noise σ is normalized by \bar{l}^2 , where \bar{l} is the average length of triangle edges on the associated target mesh, and chosen in the range of $[0, 1]$. The datasets with noise or outliers are shown in our supplementary document. The results show that our method is more robust than the classical non-rigid method with ℓ_2 -norm regularization, particularly for high noise levels.

4.3. Results on Real Scans

Fig. 12 presents results on real scans generated by Kinect Fusion [NIH*11] using Kinect V2. The real scans are challenging due to 1) noise and outliers, 2) inconsistent topology between the template and the target, and 3) uncertain correspondences. The results show that the ℓ_2 -norm regularization presents misalignments around the hands while our ℓ_1 -norm regularization provides promising registration results.

4.4. Discussion

We compare the running times of the proposed method with the ℓ_2 -norm regularized method on *Bouncing* dataset by

downsampling the meshes into smaller ones with 1K to 10K vertices. The number of iterations for non-rigid ICP is set as 20, and ℓ_1 has extra 20 inner iterations for each outer iteration. All the experiments are performed on a desktop computer with an Intel i5 3.2GHz CPU and 8GB RAM. Table 1 gives the comparison results. The proposed method needs about $4\times$ running times compared with the ℓ_2 -regularized method due to the ADM-ALM algorithm for each outer iteration in sparse non-rigid registration (Algorithm 1). This is the main limitation of the proposed method, and further numerical optimization would be beneficial.

While our method presents promising non-rigid registration results, some issues that could potentially improve performance are still open for future work: e.g., 1) The data consistency can also be measured by a sparsity-promoting function [BTP13] for more robust registration; 2) The rotation and translation parameters are equally treated in the regularization. However, rotation and translation with the same deviation could lead to different amount of registration errors, which suggests to assign different weights for more effective regularization.

Table 1: Comparison on running times.

Num. vertexes	1000	2000	5000	10000
ℓ_2 -norm	1.67s	3.60s	12.41s	26.73s
ℓ_1 -norm	8.05s	17.36s	52.48s	119.06s

5. Conclusions

This paper proposes a new non-rigid registration method with sparse modeling of non-rigid transformations. We reveal and characterize the piecewise smoothness of transformation functional of non-rigid 3D shapes, which inspires the incorporation of a sparseness regularization on the transformation differences. An ADM-ALM algorithm is devised to solve the ℓ_1 -norm regularized transformation estimation. Non-rigid 3D shapes are registered by alternating between correspondence computation and transformation estimation. A multi-resolution scheme is developed to register 3D shapes in a coarse-to-fine manner for efficient and robust registration. The results on various datasets demonstrate the superiority of the proposed method over the classic ℓ_2 -norm regularization on non-rigid registration.

Acknowledgements

We would like to thank Shuai Lin for help with comparative experiments with [LSP08].

References

- [ACP02] ALLEN B., CURLESS B., POPOVIĆ Z.: Articulated body deformation from range scan data. *ACM Trans. Graph.* 21, 3 (2002), 612–619. 2

- [ACP03] ALLEN B., CURLESS B., POPOVIĆ Z.: The space of human body shapes: reconstruction and parameterization from range scans. *ACM Trans. Graph.* 22, 3 (2003), 587–594. 2
- [ARV07] AMBERG B., ROMDHANI S., VETTER T.: Optimal step nonrigid icp algorithms for surface registration. In *Proc. CVPR* (2007), pp. 1–8. 1, 2
- [BBK08] BRONSTEIN A. M., BRONSTEIN M. M., KIMMEL R.: *Numerical geometry of non-rigid shapes*. Springer Science & Business Media, 2008. 2, 7
- [Ber82] BERTSEKAS D. P.: Constrained optimization and Lagrange multiplier methods. *Computer Science and Applied Mathematics, Boston: Academic Press 1* (1982). 5
- [Bis06] BISHOP C. M.: Linear models for regression. In *Pattern recognition and machine learning*, Bishop C. M., (Ed.). Springer New York, 2006. 3
- [BM92] BESL P. J., MCKAY N. D.: Method for registration of 3-D shapes. In *Robotics-DL tentative* (1992), pp. 586–606. 1
- [BPC*11] BOYD S., PARIKH N., CHU E., PELEATO B., ECKSTEIN J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning* 3, 1 (2011), 1–122. 5
- [BTP13] BOUAZIZ S., TAGLIASACCHI A., PAULY M.: Sparse iterative closest point. *Computer Graphics Forum* 32, 5 (2013), 113–123. 1, 3, 10
- [CR03] CHUI H., RANGARAJAN A.: A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding* 89, 2 (2003), 114–141. 2
- [EFM10] ELAD M., FIGUEIREDO M. A., MA Y.: On the role of sparse and redundant representations in image processing. *Proceedings of the IEEE* 98, 6 (2010), 972–982. 3
- [FH10] FLÖRY S., HOFER M.: Surface fitting and registration of point clouds using approximations of the unsigned distance function. *Comput. Aided Geom. Des.* 27, 1 (2010), 60–77. 2
- [GH97] GARLAND M., HECKBERT P. S.: Surface simplification using quadric error metrics. In *Proc. ACM SIGGRAPH* (1997), pp. 209–216. 6
- [HMS12] HONTANI H., MATSUNO T., SAWADA Y.: Robust non-rigid icp using outlier-sparsity regularization. In *CVPR* (2012), pp. 174–181. 2
- [KK03] KE Q., KANADE T.: *Robust subspace computation using L_1 norm*. Tech. rep., 2003. 3
- [LAGP09] LI H., ADAMS B., GUIBAS L. J., PAULY M.: Robust single-view geometry and motion reconstruction. *ACM Trans. Graph.* 28, 5 (2009), 175. 2, 6
- [LSP08] LI H., SUMNER R. W., PAULY M.: Global correspondence optimization for non-rigid registration of depth scans. *Computer graphics forum* 27, 5 (2008), 1421–1430. 1, 2, 3, 8, 10
- [LZW*09] LIAO M., ZHANG Q., WANG H., YANG R., GONG M.: Modeling deformable objects from a single depth camera. In *Proc. ICCV* (2009), pp. 167–174. 2
- [MS10] MYRONENKO A., SONG X.: Point set registration: Coherent point drift. *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 12 (2010), 2262–2275. 2
- [NIH*11] NEWCOMBE R. A., IZADI S., HILLIGES O., MOLYNEUX D., KIM D., DAVISON A. J., KOHI P., SHOTTON J., HODGES S., FITZGIBBON A.: Kinectfusion: Real-time dense surface mapping and tracking. In *IEEE Int. sym. Mixed and augmented reality* (2011), pp. 127–136. 10
- [PB11] PAPAZOV C., BURSCHKA D.: Deformable 3D shape registration based on local similarity transforms. In *Computer Graphics Forum* (2011), vol. 30, pp. 1493–1502. 2
- [PG08] PEKELNY Y., GOTSMAN C.: Articulated object reconstruction and markerless motion capture from depth video. *Computer Graphics Forum* 27, 2 (2008), 399–408. 2
- [PGK02] PAULY M., GROSS M., KOBELT L. P.: Efficient simplification of point-sampled surfaces. In *Proc. Visualization* (2002), pp. 163–170. 6
- [RBS14] ROUHANI M., BOYER E., SAPPA A. D.: Non-rigid registration meets surface reconstruction. In *Proc. 3D Vision* (2014). 2
- [RKGB12] RUHNKE M., KUMMERLE R., GRISETTI G., BURGARD W.: Highly accurate 3D surface models by sparse surface adjustment. In *Proc. IEEE Intl. Conf. on Robotics and Automation* (2012), pp. 751–757. 1
- [SA07] SORKINE O., ALEXA M.: As-rigid-as-possible surface modeling. In *Proc. Symposium on Geometry Processing* (2007), pp. 109–116. 2
- [SSP07] SUMNER R. W., SCHMID J., PAULY M.: Embedded deformation for shape manipulation. *ACM Trans. Graph.* 26, 3 (2007), 80. 1, 3
- [STDS14] SALTI S., TOMBARI F., DI STEFANO L.: SHOT: unique signatures of histograms for surface and texture description. *Computer Vision and Image Understanding* 125 (2014), 251–264. 4, 9
- [SWG08] SÜSSMUTH J., WINTER M., GREINER G.: Reconstructing animated meshes from time-varying point clouds. *Computer Graphics Forum* 27, 5 (2008), 1469–1476. 2
- [TCL*13] TAM G. K., CHENG Z.-Q., LAI Y.-K., LANGBEIN F. C., LIU Y., MARSHALL D., MARTIN R. R., SUN X.-F., ROSIN P. L.: Registration of 3D point clouds and meshes: a survey from rigid to nonrigid. *IEEE Trans. Vis. Comp. Graph.* 19, 7 (2013), 1199–1217. 2
- [TMRL14] TAM G. K., MARTIN R. R., ROSIN P. L., LAI Y.-K.: Diffusion pruning for rapidly and robustly selecting global correspondences using local isometry. *ACM Trans. Graph.* 33, 1 (2014), 4. 4, 9
- [VBMP08] VLASIC D., BARAN I., MATUSIK W., POPOVIĆ J.: Articulated mesh animation from multi-view silhouettes. *ACM Trans. Graph.* 27, 3 (2008), 97. 2, 7
- [WAO*09] WANG M., ADAMS B., OVSIANIKOV M., BERNER A., BOKELOH M., JENKE P., GUIBAS L., SEIDEL H.-P., SCHILLING A.: Efficient reconstruction of nonrigid shape and motion from real-time 3D scanner data. *ACM Trans. Graph.* 28, 2 (2009), 15. 2
- [YXP09] YANG J., PENG Y., XU W., DAI Q.: Ways to sparse representation: an overview. *Science in China series F: information sciences* 52, 4 (2009), 695–703. 5
- [YB*13] YANG A. Y., ZHOU Z., BALASUBRAMANIAN A. G., SASTRY S. S., MA Y.: Fast ℓ_1 -minimization algorithms for robust face recognition. *IEEE Trans. Image Processing* 22, 8 (2013), 3234–3246. 5
- [YZL*14] YU Y., ZHANG S., LI K., METAXAS D., AXEL L.: Deformable models with sparsity constraints for cardiac motion analysis. *Medical Image Analysis* 18, 6 (2014), 927–937. 3
- [ZNI*14] ZOLLHÖFER M., NIESSNER M., IZADI S., RHEMANN C., ZACH C., FISHER M., WU C., FITZGIBBON A., LOOP C., THEOBALT C., STAMMINGER M.: Real-time non-rigid reconstruction using an RGB-D camera. *ACM Trans. Graph.* 33, 4 (2014). 1