Global Optimization for Alignment of Generalized Shapes

Hongsheng Li Tian Shen Xiaolei Huang
Department of Computer Science and Engineering
Lehigh University, Bethlehem, PA 18015
{h.li, tis207, xih206}@lehigh.edu

Abstract

In this paper, we introduce a novel algorithm to solve global shape registration problems. We use gray-scale "images" to represent source shapes, and propose a novel two-component Gaussian Mixtures (GM) distance map representation for target shapes. Based on this flexible asymmetric image-based representation, a new energy function is defined. It proves to be a more robust shape dissimilarity metric that can be computed efficiently. Such high efficiency is essential for global optimization methods. We adopt one of them, the Particle Swarm Optimization (PSO), to effectively estimate the global optimum of the new energy function. Experiments and comparison performed on generalized shape data including continuous shapes, unstructured sparse point sets, and gradient maps, demonstrate the robustness and effectiveness of the algorithm.

1. Introduction

Shape registration is a crucial problem in computer vision and medical image analysis. Global registration, also known as shape alignment, aims to recover a global transformation, such as rigid or similarity transformation, that brings the pose of a source shape as close as possible to that of a target shape. Global registration algorithms are usually different from each other in three aspects: shape representation, energy function and optimization method.

There has been a lot of prior research on global shape registration. We classify them into the following three categories based on their shape representations.

Point-based representations are widely used since they can represent shapes of arbitrary dimension and topology. The Iterative Closest Point (ICP) [2, 24] algorithm is one of the most widely used point registration methods for geometric alignment of shapes. There are a large number of variants introduced on the basic ICP concept.

In [19, 24], it was observed that longer "closest" distances tend to be between false point correspondences. Various methods were proposed to solve the problem, which can be categorized as either rejection [10, 19, 24] or weighting [6] methods. To accelerate the computation process of ICP, Fitzgibbon [5] suggested that the use of pre-computed distance map could result in faster computation. Most recently, Sandhu *et al*. [13] uses the standard ICP energy function to match points but applies Particle Filters in the iterative step to recover the best transformation.

However, energy functions of the ICP family have disadvantages such as possible false correspondences [2] or discontinuities in the energy function domain [10, 19, 24] (see Section 2.2.2 for details).

Tsin *et al*. [18] proposed a kernel correlation based point set registration approach where the cost function is proportional to the correlation of two kernel density estimates. Another significant contribution to point registration is made by [8]. Point sets are modeled as Mixtures of Gaussian (MG) distributions. L_2 distance is then used to measure dissimilarity between the distributions, which is minimized by Gradient Descent. The same MG distribution model is used in [20] but extended to group-wise point sets registration by using Jensen-Shannon Divergence for atlas creation and distance measuring.

Image-based representations can also represent shapes of arbitrary dimension and topology. Embedding shapes into image planes, this kind of representation provides more constraints and supporting information from the neighboring areas of shapes.

Different kinds of embeddings have been proposed. Paragios *et al*. [12] uses signed distance maps to represent shapes. Shape dissimilarity is then calculated by computing the sum of squared differences (SSD) between two maps and minimized by Gradient Descent. Huang *et al*. [7] adopts the same level set shape representation and optimization method as those of [12]. But Mutual Information is employed to measure the dissimilarity between two shapes. El Munim *et al*. [4] keeps a similar framework as [12], but Euclidean distance maps are replaced by Vector Distance Function (VDF)s. Along this line, Tang *et al*. [17] computes shape features, which are then organized into vectors and assigned to the nearest pixels of corresponding shape parts to create feature images; however, all descriptors mentioned are for 2D closed contours.

Continuous-contour/surface-based representations [9, 14, 22] sometimes are also based on points (acquired by

sampling the contour/surface). The main difference from point-based representation is that, for every point representing a part of a shape, its neighbors' information is also required for the computation of the representation. However, because of the requirement for additional topology information, this kind of contour/surface representations has more restrictions. For instance, they usually cannot represent unstructured sparse point sets.

In this paper, we introduce a novel global shape registration algorithm which combines *image-based* shape representations and *point-based* energy function. Inspired by [5], we propose a new asymmetric image-based representation which employs gray-scale images to represent source shapes and uses pre-computed "distance" maps to represent target shapes. The gray-scale image representation for source shape can represent shapes of arbitrary topology even including generalized shapes such as gradient maps (Section 3.3). For target shape "distance" maps, observing several drawbacks of the widely used L_2 distance model, we create a novel two-component univariate Gaussian Mixtures (GM) distance model. It is different from the GM model proposed in [8] where a n-point set is modeled as a n-component bivariate (2D) or trivariate (3D) GM distribution if no prior knowledge is given. We also propose an energy function that can be computed efficiently because of the pre-computed target "distance" maps. Such efficiency enables us to use a global optimization method, the Particle Swarm Optimization (PSO), to find the globally optimal transformation parameters.

The main contribution of our algorithm is two-fold: (1) A shape dissimilarity metric more robust than the widely used L_2 distance is created based on the new asymmetric shape representations. Moreover, the new representation can represent not only arbitrary shapes but also gradient maps, which are difficult for existing representations. (2) The high efficiency of the new energy function makes it feasible for the Particle Swarm Optimization, which has shown its effectiveness in various optimization problems, to be used to recover the best transformation. We did extensive comparison with state-of-the-art shape/point registration algorithms. Experiments on a variety of shape data, which include continuous shapes, gradient maps, and unstructured sparse point sets, in both 2D and 3D, show the advantages of our algorithm.

2. Methodology

2.1. The Standard ICP Energy Function

When point sets are used to represent shapes, let $\mathcal{C}_s$ and $\mathcal{C}_t$ be the source shape and target shape, respectively. Let $\mathbf{x}_{i,s}(i = 1, \ldots, N_s)$ be the points on the source shape $\mathcal{C}_s$, and let $\mathbf{x}_{j,t}(j = 1, \ldots, N_t)$ be the points on the target shape $\mathcal{C}_t$. The standard ICP energy function [2] based on squared Euclidean distances is given by

$$\mathcal{F}(\Theta) = \frac{1}{N_s} \sum_{i=1}^{N_s} w_i d^2(\mathcal{T}(\mathbf{x}_{i,s}, \Theta), \mathcal{C}_t) \quad (1)$$

where $d(\mathbf{x}, \mathcal{C}) = \min_{\mathbf{x}_\mathcal{C} \in \mathcal{C}} d'(\mathbf{x}, \mathbf{x}_\mathcal{C})$, and $d'(\mathbf{x}_1, \mathbf{x}_2)$ is the Euclidean distance between the two points $\mathbf{x}_1$ and $\mathbf{x}_2$. Θ are the parameters of a chosen type of transformation, and $\mathcal{T}(\mathbf{x}_{i,s}, \Theta)$ is the transformed source shape point according to Θ. w_i denotes the weight of the ith "closest" distance.

2.2. Image-based Shape Representation and Energy Function

Our goal is to build a robust, generic but computationally efficient shape dissimilarity metric because we want to employ global optimization methods other than commonly-used gradient-based local optimization methods [4, 7, 12, 24], which have shown poor performance in overcoming local optima.

2.2.1 Previous Work on Computationally Efficient Shape Dissimilarity Metric

In [5], distance transform is applied to pre-computing a target shape's "closest" distances. Given a target shape $\mathcal{C}_t$, and let $\Phi : \Omega \rightarrow R^+$ be a Lipschitz function that refers to the Euclidean distance transform of $\mathcal{C}_t$. The squared distance transform [11] of the target shape is

$$\Phi_{\mathcal{C}_t}(x, y) = d^2((x, y), \mathcal{C}_t), \quad (2)$$

where $d((x, y), \mathcal{C}_t)$ denotes the minimum L_2 distance between the pixel at location (x, y) and the shape $\mathcal{C}_t$ (Figure 1b and 1c). In [5], L_2 distance is replaced by Huber norm in distance maps. The ICP energy function based on the pre-computed distance map $\Phi_{\mathcal{C}_t}$ is then given by

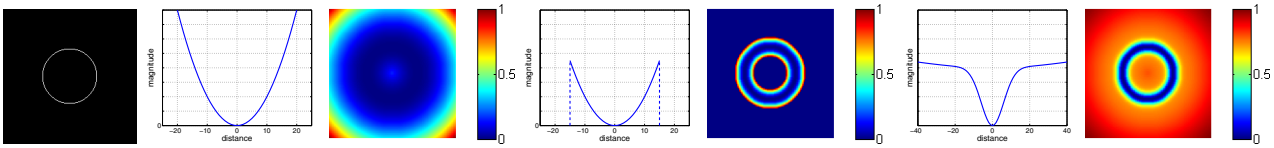
$$E(\Theta) = \sum_{i=1}^{N_s} w_i \Phi_{\mathcal{C}_t}(\mathcal{T}(\mathbf{x}_{i,s}, \Theta)). \quad (3)$$

Except for the normalization term, (3) has the same meaning as the ICP energy function (1). However, it is computationally efficient because pre-computed distance maps can circumvent the need of searching for "closest" distances for every point $\mathbf{x}_{i,s}$ in iterative steps. This technique is also related to Chamfer matching [1].

2.2.2 Mitigations of False Correspondences

Although the standard ICP energy function shows its effectiveness in various registration problems, it was observed in [19, 24] that longer "closest" distances tend to be between false correspondences, especially when outliers exist. In agreement with this observation are our experiments (Figure 3b), in which minimizing energy functions based on L_2 distance sometimes led to wrong transformations.

To alleviate this problem, there are several mitigation measures proposed which can be categorized into two classes: rejection and weighting. One naive rejection approach [19, 24] is to set a maximum distance threshold, D_{thres}. If the "closest" distance $> D_{thres}$, $w_i = 0$; otherwise $w_i = 1$ (Figure 1d and 1e). However, this approach fails to measure properly the dissimilarity between



(a) (b) (c) (d) (e) (f) (g)

Figure 1. (a) A circle ($r = 40$), (b) squared Euclidean distance model, (c) image embedding of (a) based on the model (b) (normalized), (d) squared Euclidean distance with threshold ($D_{thres} = 15$) model, (e) image embedding of (a) based on the model (d) (normalized), (f) two-component Gaussian Mixtures (GM) distance model, and (g) target shape representation of (a) based on the model (f) (normalized).

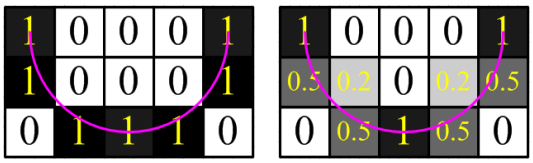


Figure 2. Source shape (purple curve) representation obtained through discretization (left) and anti-aliasing (right) techniques.

two shapes that are significantly different in scale. For instance, a source shape which has all its points' "closest" distances greater than D_{thres} would result in the rejection of all distances and thus cannot converge to the pose of the target shape. Another rejection policy, based on some multiple of the standard deviation of distances, is presented in [10]. But the deviation would be biased if strong outliers exist in the source shape. Furthermore, distance or point weights cannot easily be taken into consideration during the deviation computation.

All rejection-based methods cause possible discontinuities in the energy function domain. A linear weighting policy, which assigns longer "closest" distances lower weights, $w_i = 1 - d_i/D_{\max}$, is proposed in [6]. This scheme is sensitive to the $D_{\max}$ value however. It becomes similar to the L_2 distance model when $D_{\max}$ is large, and similar to the distance threshold model when $D_{\max}$ is small.

2.2.3 New Asymmetric Shape Representation

Although LM-ICP [5] uses Huber norm instead of L_2 distance, it still suffers from the false correspondence problem we mentioned above (see Figure 3c). Moreover, how to assign reasonable weights, w_i, for distances is not specified either.

We present a new source shape representation which is a gray-scale "image" that can represent generalized shapes including parametric curves/surfaces, sparse point sets, edge maps, and even gradient maps. Given a continuous shape or a point set, discretization and anti-aliasing (*e.g.* supersampling) techniques (Figure 2) are employed to embed the shape into the image plane. Edge maps obtained by shape extraction algorithms and even gradient maps can also be utilized directly as source shapes. Directly using gradient maps as source shapes in particular allows us to circumvent the shape extraction step and reduce the risk of misregistration because of improper shape extraction (Section 3.3).

To address the false correspondence problem, we also present a new target shape representation, which replaces the squared Euclidean distance model with a high-peak-fat-tail distance model. The high peak indicates higher weight on short distances, and the fat tail decreases the weight of longer distances while keeping the energy function smooth and differentiable within the domain. In this paper, to achieve the high-peak-fat-tail effect we choose a two-component Gaussian Mixtures (GM) model, leading to the following target shape representation,

$$\Psi_{\mathcal{C}_t}(x,y) = -e^{-\Phi_{\mathcal{C}_t}(x,y)/2\sigma_1^2} - \alpha e^{-\Phi_{\mathcal{C}_t}(x,y)/2\sigma_2^2}, \quad (4)$$

where $\alpha \in (0,1)$ is the GM weighting parameter, σ_1 and σ_2 ($\sigma_1 < \sigma_2$) are the standard deviations of the two Gaussians (Figure 1f and 1g).

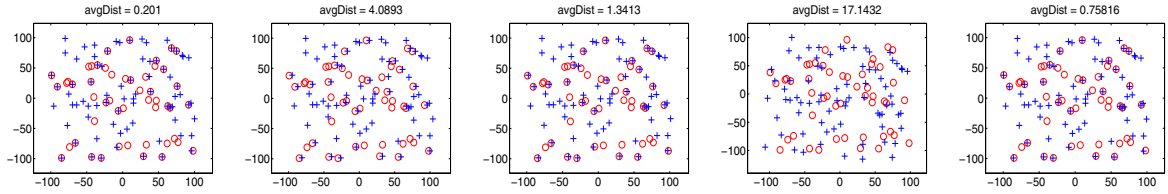
We compared our new high-peak-fat-tail GM kernel with other dissimilarity metrics including nearest L_2 distance [2], nearest Huber norm [5], L_2 distance between Gaussian mixtures [8], and Kernel Correlation [18], using extremely difficult cases of registering unstructured sparse point sets with strong outliers. In Figure 3, two point sets with certain number of corresponding points and outliers are registered using rigid transformation (see Section 3.4 for how the two point sets are generated). To compare only the effects of dissimilarity metrics, we initialize the source set's pose so that it is correctly aligned with the target set. We then observe perturbation by registering the point sets using those metrics. Results show that our GM kernel is most robust against outliers and leads to the smallest average distance error between corresponding points among all these metrics.

2.2.4 Efficient Energy Function

Based on the source and target shape representations proposed in Section 2.2.3, we define an energy function measuring the dissimilarity between source and target shapes,

$$\mathcal{M}_\Psi(\Theta) = \frac{\iint \mathcal{T}(S,\Theta)(x,y) \cdot \Psi_{\mathcal{C}_t}(x,y)dxdy}{\iint \mathcal{T}(S,\Theta)(x,y)dxdy}, \quad (5)$$

where $\mathcal{T}(S,\Theta)$ is the transformed source shape image according to transformation parameters Θ. The numerator of (5) calculates the sum of weighted "closest" GM distances between points on the transformed source shape image, $\mathcal{T}(S,\Theta)$, and the target shape, $\mathcal{C}_t$. The weights are implicitly specified by the pixel values in $\mathcal{T}(S,\Theta)$. The denominator, the sum of point weights on the transformed



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

Figure 3. Comparison between different dissimilarity metrics by registering unstructured sparse point sets with strong outliers (source as blue '+' vs. target as red 'o'). (a) The result of the proposed GM kernel, (b) the result of nearest L_2 distance [2], (c) the result of nearest Huber norm ($\sigma = 10$) [5], (d) the result of L_2 distance between Gaussian mixtures [8], and (e) the result of Kernel Correlation [18].

source shape, is the normalization term necessary to handle scaling. Note that the new target shape representation $\Psi_{\mathcal{C}_t}$ based on the GM kernel is pre-computed and most pixels in the source representation $\mathcal{T}(S, \Theta)$ have zero values. Therefore, computation of the energy function (5) is highly efficient.

The time complexity of our energy function (5) can be explained as follows: given two n-point sets, let k be the number of samples in the finite image domain, which can be assumed constant. A one-time overhead of the linear time distance transform algorithm is $O(k)$ [11]. Then it takes $O(n)$ time to evaluate our energy function. In contrast, the standard ICP energy function takes $O(n^2)$ to evaluate.

2.3. Globally Optimal Transformation Estimation using the Particle Swarm Optimization (PSO)

In previous research, local optimization methods such as Gradient Descent are widely used [4, 7, 12]. To these local methods, good initial parameter estimates are essential. To avoid this problem, some non-linear optimization methods are also introduced [13, 17]. In our algorithm, the PSO is adopted to estimate the global optimum of the energy function (5).

The PSO, which belongs to the class of swarm intelligence methods, is an effective optimization method for high-dimensional optimization problems, originally developed by Eberhart and Kennedy [3].

Let $f(\mathbf{x})$ denote the function to be minimized. The basic particle swarm model consists of a swarm of m particles moving in an n-dimensional real valued parameter space, each position of which potentially gives the global optimum of $f(\mathbf{x})$ over a given domain D. Let $\mathbf{x}_i(t) = [x_1(t), x_2(t), \ldots, x_n(t)]^T$ and $\mathbf{v}_i(t) = [v_1(t), v_2(t), \ldots, v_n(t)]^T$ be the ith particle's position and velocity at time t. Each particle knows the best position it has been at so far, $pbest_i$, and the overall best position obtained so far, $gbest$, by any particle in the swarm. Each particle updates its position and velocity according to the following equations,

$$\begin{aligned}\mathbf{v}_i(t+1) =& \omega \mathbf{v}_i(t) + c_1 \times rand() \times (pbest_i - \mathbf{x}_i(t)) \\ &+ c_2 \times rand() \times (gbest - \mathbf{x}_i(t)) \qquad (6)\end{aligned}$$

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t), \qquad (7)$$

where ω is the inertia weight representing the degree of the momentum of the particles. c_1 and c_2 are attraction terms that move the ith particle toward $pbest_i$ and $gbest$ respectively. $rand()$ generates pseudo-random numbers drawn from a uniform distribution in the range $[0, 1]$. $||\mathbf{v}_i||$ is usually limited to be within a range $[0, v_{\max}]$, where $v_{\max}$ is the maximum velocity.

Before the PSO computation begins, positions and speeds of all particles are randomly generated in the given continuous domain. Then they are iteratively renewed according to (6) and (7) until a minimum error criterion or a predefined maximum iteration is attained.

Because the objective is to find the global optimum, we choose not to use the original PSO's maximum iteration stop criterion. In our modified PSO, to decide whether the value of the energy function has globally converged or not, first we use a relative error function,

$$\Delta f_i(t) = \frac{|f_i(t) - f(gbest)|}{|\min(f_i(t), f(gbest))|}, \qquad (8)$$

proposed in [23], to determine inactive particles. If a particle's $\Delta f_i(t)$ is less than a threshold, ϵ, for more than T_c time periods, we consider this particle to be inactive. If during a certain period of time, the number of inactive particles exceeds a threshold N_c, we consider the global optimum found. During the computation and before convergence, any detected inactive particles are randomly relocated in the domain and are given random initial speeds so that they can keep searching the space.

The PSO has a strong ability to find the global optimum without any initial parameter estimate. In order to accelerate its convergence speed and to prevent premature termination, many variants have been introduced. Shi *et al.* [16] presented a strategy of time-varying inertia weight: ω linearly decreases as the number of iteration increases. The modified PSO in our algorithm also uses the strategy of time-varying inertia weight [16]. The difference from [16] is that each particle has its own inertia weight instead of all particles sharing the same one. If a particle is relocated, its inertia weight gets reinitialized to the maximum inertia weight. This enables a newly relocated particle to search more areas instead of directly traveling back to $gbest$ and becoming inactive again.

For 2D global shape registration, the search space for particles has either 4 (similarity transformation) or 6 (affine

transformation) dimensions. In 3D, we consider similarity transformation which has a 7-dimensional search space. The combination of the new energy function (5) and the modified PSO allows us to solve global shape registration problems robustly and efficiently, as we will demonstrate in our experiments.

3. Experiments

For all experiments, we set $\alpha = 0.5$, $\sigma_1 = 5$, $\sigma_2 = 50$, and let $c_1 = c_2 = 2$. ω linearly decreases from 1.0 to 0.2. We used 100 particles for 2D cases and 3000 particles for 3D. Most cases converged within 300 iterations. On an Intel E6850 CPU, single thread, a 2D case took about 1s-5s. A 3D case took 60s-200s.

3.1. Similarity and Affine Registration of 2D Shapes

The first and simplest experiment is the similarity and affine registration of 2D shapes (Figure 4). For this purpose, we use some shapes from the SIID shape database [15]. Source and target shapes in Figure 4 differ in parts, and initial poses of the sources shapes vary in a broad range. Despite such difficulties, the results show that the energy function always converges to the global optimum through our modified PSO.

3.2. Registration of Shapes with Strong Outliers and Comparison with State-of-the-art Algorithms

In practical applications, shapes as clear and sharp as those in Figure 4 are usually difficult to obtain because of complex backgrounds of interested objects. Thus, extracted shapes may contain spurious contours, which can significantly affect registration results. Some state-of-the-art algorithms [4, 7, 12, 24] have difficulties in handling such shape registration problems with strong outliers. Because our method looks for the global optimum and integrates outlier-resistant mechanisms based on the high-peak-fat-tail GM model, it performs better in solving this challenging problem. In our experiment, we used "clear" shapes as source shapes and registered them to shapes with strong artificial outliers. In Figure 5, artificial outliers "airplane" and "grids" are added to "hand" shapes to create challenging target shapes. For this "hand" shape registration problem, we compared the registration performance of our method with those in [4, 7, 12, 24][1]. In the cases of Figure 5, energy functions would have multiple local optima no matter which shape representation is used. Hence, finding the global optimum is impossible for [4, 7, 12, 24] unless an initial value close enough to the global optimum is given. On the other hand, the PSO has a strong ability to systematically search in a broad space to find the global optimum. For comparison, we assigned the same initial values to all the methods. Our method successfully found the best transform parameters. But other algorithms often got stuck at local optima.

[1]We implemented the methods according to the original papers.

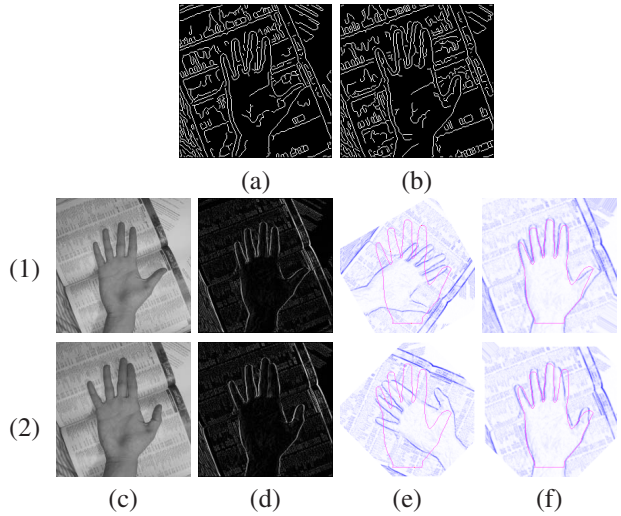


Figure 6. Registration between contour shapes and gradient maps. (a) Edges of 1c extracted by Canny detector using MATLAB default parameters, (b) edges of 2c extracted by Canny detector using MATLAB default parameters, (c) original images, (d) gradient maps of (c) as source shapes, (e) initial poses (source in blue vs. target in red), and (f) registration results.

3.3. Registration of Generalized Shapes

Extracting edges from images is a difficult task because the choice of extraction algorithm parameters has a large effect on extracted shapes (Figure 6a and 6b). Because of the natural characteristic of the gray-scale image-based shape representation, our method is capable of directly using gradient maps as source shape images. This property could reduce the risk of using shapes obtained by inappropriate edge extraction parameters or algorithms (Figure 6). So far we only tested gradient maps as source shape images. However, other gray-scale images may also be considered as candidate shape images for registration.

3.4. Similarity Registration of 2D Unstructured Sparse Point Sets

To validate our algorithm on registering unstructured sparse 2D point sets, we followed the experimental setup of [8, 20] and compared the performance of our algorithm with that of Mixtures of Gaussian (MG) [8] and Kernel Correlation (KC) [18] methods. The code of the MG[2] and KC[3] methods is available at their authors' websites.

First, we performed similarity registration experiments on two synthetic point sets without noise and outliers. The source point set is a randomly transformed version of the target point set. Figure 7(1) shows one example registration result. The target point set contains 50 points (red 'o's) randomly spread in a region $[-100, 100] \times [-100, 100]$. And the source point set (blue '+'s) is a transformed version of the target, after a 180° rotation and 1.5 scaling.

[2]`http://gmmreg.googlecode.com/svn/trunk/`
[3]`http://www.cs.cmu.edu/\~ytsin/KCReg/`

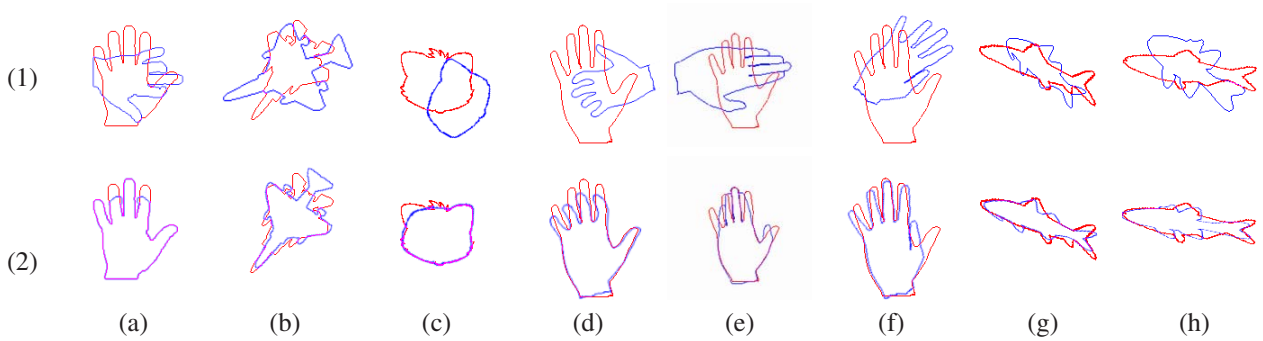


Figure 4. Similarity and affine registration results. (1) Initial poses (source in blue versus target in red), and (2) registration results. (a-f) Similarity registration, and (g-h) affine registration.

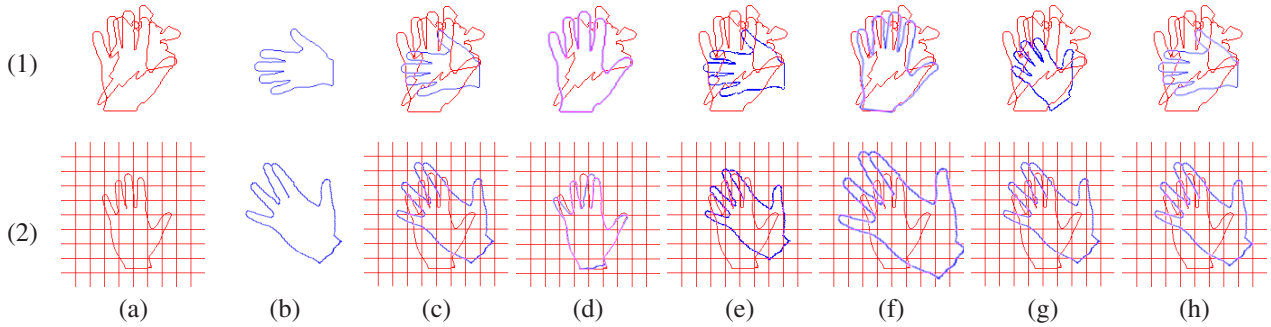


Figure 5. Registration results of "hand" shapes with strong outliers. (a) Target shapes, (b) source shapes, (c) initial poses (source in blue and target in red), (d) registration results of our method, (e) registration results of ICP with distance threshold [24], (f) registration results of Distance Functions (DF)-based method [12], (g) registration results of Mutual Information (MI)-based method [7], and (h) registration results of Vector Distance Function (VDF)-based method [4].

Next, we also did experiments in the presence of noise and outliers following the setup in [8]. The following procedures are used to generate source point sets from a target point set. For a target set with n points spread in the range $[-D, D] \times [-D, D]$ ($n = 50$ and $D = 100$ for 2D point sets), we control the degree of corruption by (1) discarding a subset of $(1 - \rho)n$ points from the target set, (2) adding uniformly distributed noise in the range of $[-\epsilon, \epsilon] \times [-\epsilon, \epsilon]$ to all points in the target set, (3) applying a similarity transformation (s, θ, t_x, and t_y) to the target set, and (4) adding $(\tau - \rho)n$ spurious, uniformly distributed points to the target set. Therefore, after the corruption, a source set would have a total number of τn points, in which only ρn have corresponding points in the target set. We quantify the strength of outliers and noise as $S_o = (\tau - \rho)/\rho$ and $S_n = \epsilon/D$ respectively. In Figure 7, the average L_2 distance between known correspondences is computed as the error measure and displayed on top of each registered frame. Our extensive experiments give us empirical convergence ranges for 2D unstructured sparse point set registration (Table 1).

We compared our method with two state-of-the-art point registration methods, MG [8] and KC [18]. The comparison was done using rigid transformation (no scaling). At each outlier or noise-strength level, we generated 50 target sets and 50 corrupted source sets. For each pair, we used the proposed, MG and KC methods for registration and computed the resulting average distance between known correspondences. Results show that when outlier and noise levels are low, both MG and KC methods register two sets accurately. But as the outlier and noise levels rose, our method performed more robustly than the MG and KC methods (Figure 8).

3.5. Similarity Registration of 3D Point Sets

Experiments similar to those with 2D unstructured sparse point sets were done on 3D point sets. In 3D, we set $n = 100$, $D = 50$. Empirical convergence ranges show the robustness of our algorithm for registration of 3D point sets (Table 2).

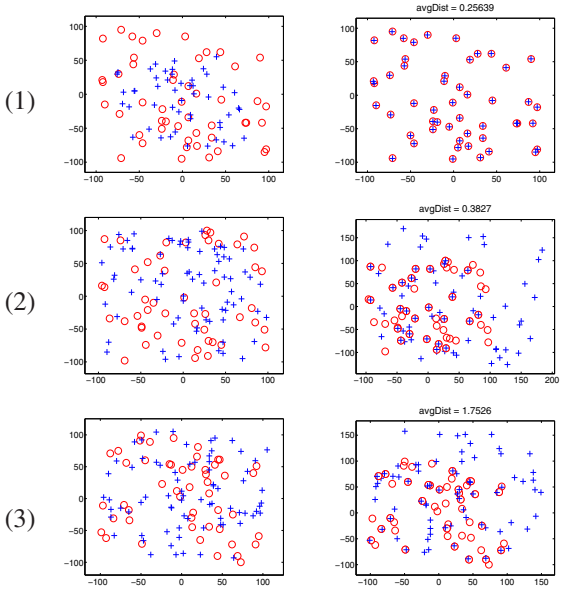
For performance evaluation, we followed the experimental setup in [13]. 1000 points are randomly chosen from the "bunny" model [19]. We then generated translation $\vec{t} = [t_x, t_y, t_z]$ from a normal distribution with each component having a standard deviation of 70; this deviation is chosen based on the range of the target set ($[-126, 127] \times [-125, 125] \times [-98, 98]$). The rotation angle θ is generated randomly along the z axis, from a normal distribution $\mathcal{N}(0, (\pi/3)^2)$. The scaling factor is chosen randomly from a uniform distribution $\mathcal{U}(0.7, 1.3)$. Before the transformation is applied to the target set, a certain percentage of points is replaced with zero-mean Gaussian noise $\mathcal{N}(0, 60^2)$. We consider three noise levels, with 5, 20, and 35 percent of the

Method	Outlier and noise level	Empirical convergence range (scale×rotation angle×translation in x and y directions)
MG method	$S_o = 0, S_n = 0$	$[1, 1] \times [-120°, 120°] \times [-40, 40] \times [-40, 40]$ [8]
Proposed method	$S_o = 0, S_n = 0$	$[0.667, 1.5] \times [-180, 180] \times [-80, 80] \times [-80, 80]$
	$S_o = 2, S_n = 0$	$[0.75, 1.333] \times [-180°, 180°] \times [-40, 40] \times [-40, 40]$
	$S_o = 2, S_n = 2\%$	$[0.8, 1.25] \times [-180°, 180°] \times [-40, 40] \times [-40, 40]$

Table 1. Convergence ranges of the proposed and MG methods in the 2D unstructured sparse point set registration tasks.

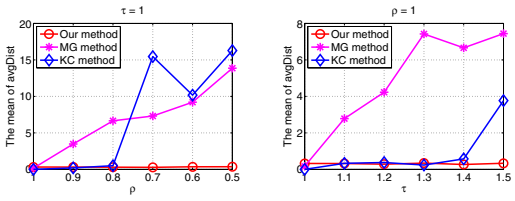
Outlier and noise level	Empirical convergence range (scale×rotation angle×rotation axis×translation in x, y, and z directions)
$S_o = 0, S_n = 0$	$[0.667, 1.5] \times [-180°, 180°] \times$ *all possible axes* $\times [-40, 40] \times [-40, 40] \times [-40, 40]$
$S_o = 2, S_n = 0$	$[0.8, 1.25] \times [-180°, 180°] \times$ *all possible axes* $\times [-20, 20] \times [-20, 20] \times [-20, 20]$
$S_o = 2, S_n = 2\%$	$[0.8, 1.25] \times [-180°, 180°] \times$ *all possible axes* $\times [-20, 20] \times [-20, 20] \times [-20, 20]$

Table 2. Convergence ranges of the proposed method in the 3D unstructured sparse point set registration tasks.



(a) (b)

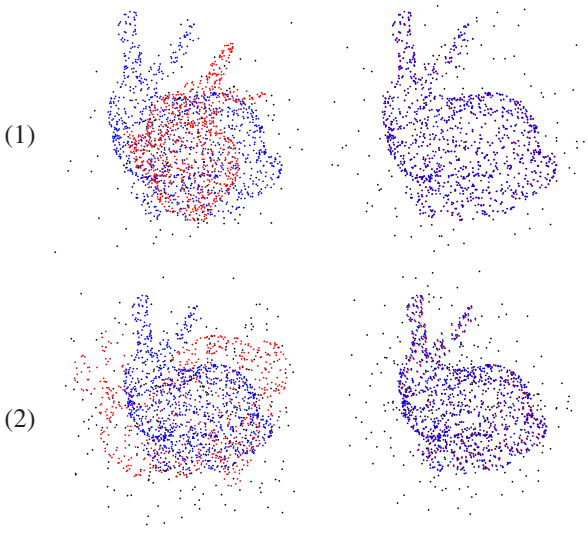
Figure 7. Similarity registration results of 2D point sets (source as blue '+'s versus target as red 'o's). (a) Initial poses, and (b) registration results. (1) Registration of point sets without outliers and noise ($S_o = 0$, $S_n = 0$) , (2) registration of point sets with outliers ($\rho = 0.5$, $\tau = 1.5$, $S_n = 0$), and (3) registration of point sets with outliers and higher noise ($\rho = 0.5$, $\tau = 1.5$, $S_n = 2\%$).



(a) (b)

Figure 8. Comparison between the proposed, MG and KC methods. The mean of 50 average distances at each outlier level versus varying ρ or τ. (a) $\tau = 1, \rho \in [0.5, 1]$, and (b) $\rho = 1, \tau \in [1, 1.5]$.

points substituted, respectively. We performed 100 experiments in each noise level. Two example registrations are shown in Figure 9. The statistics of errors with respect to ground truth are given in Table 3.



(a) (b)

Figure 9. Testing algorithm robustness to noise and initialization with the 3D "bunny" model (source in red, target in blue, and noise points in black). (1) 20% Gaussian zero mean noise level, and (2) 35% Gaussian zero mean noise level. (a) Initial poses, and (b) registration results.

Noise	5%	20%	35%
Δs	$\mu = 0.00165$ $\sigma = 0.00112$ $\max = 0.00382$	$\mu = 0.00431$ $\sigma = 0.00251$ $\max = 0.00836$	$\mu = 0.0122$ $\sigma = 0.0132$ $\max = 0.0798$
$\Delta\vec{R}$	$\mu = 0.00258$ $\sigma = 0.00127$ $\max = 0.00549$	$\mu = 0.00900$ $\sigma = 0.0113$ $\max = 0.0580$	$\mu = 0.00893$ $\sigma = 0.00845$ $\max = 0.0340$
$\Delta\theta$	$\mu = 0.124°$ $\sigma = 0.0986°$ $\max = 0.324°$	$\mu = 0.182°$ $\sigma = 0.155°$ $\max = 0.669°$	$\mu = 0.394°$ $\sigma = 1.10°$ $\max = 6.10°$
$\Delta\vec{t}$	$\mu = 0.670$ $\sigma = 0.274$ $\max = 1.25$	$\mu = 0.951$ $\sigma = 0.710$ $\max = 2.92$	$\mu = 1.20$ $\sigma = 0.705$ $\max = 2.70$

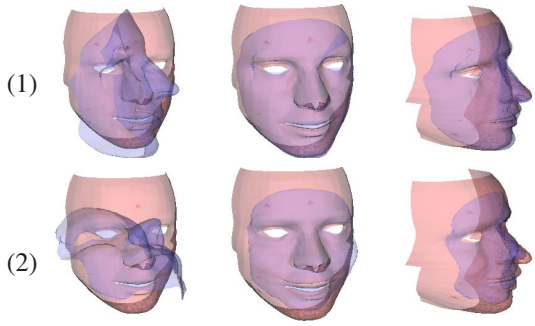
Table 3. Mean, standard deviation, and maximum of errors of scaling factor Δs, rotation axis $\Delta\vec{R}$, rotation angle $\Delta\theta$, and translation $\Delta\vec{t}$ for each noise level.

3.6. Similarity Registration of 3D Surfaces

We also did experiments on 3D surface registration. A male's face (Figure 10a) is used as the target surface and two other persons' faces with different facial expressions

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

Figure 10. Face models used in 3D surface registration. (a) The male target model, (b-c) a male subject, and (d-e) a female subject.



(1)

(2)

(a) (b) (c)

Figure 11. 3D surface registration experiments. (1) Registration of models in Figure 10a and 10c, and (2) registration of models in Figure 10a and 10e. (a) Initial poses (source in blue versus target in red), (b) registration results (front view), and (c) registration results (side view).

(smiling and non-smiling) are registered to it (Figure 10b-e) [21]. Figure 11 shows two examples of this experiment. Our method successfully recovered the best transformation even when the initialization is far from the correct pose.

4. Conclusion

In this paper, we present a novel global shape registration algorithm. Flexible gray-scale "images" and GM "distance" maps were proposed to represent source and target shapes respectively. GM "distance" maps are pre-computed by the novel two-component GM distance transform which provides a more robust dissimilarity metric for shapes. Because of the high efficiency of our new energy function, a global optimum estimation algorithm, the Particle Swarm Optimization (PSO), was introduced to optimize the energy function. The proposed method showed its robustness and effectiveness in solving various challenging registration cases using generalized shape data.

References

[1] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf. Parametric correspondence and chamfer matching: two techniques for image matching. *IJCAI*, pages 1175–1177, 1977.

[2] P. Besl and H. McKay. A method for registration of 3-d shapes. *TPAMI*, pages 239–256, 1992.

[3] R. Eberhart and J. Kennedy. A new optimizer using particle swarm theory. *Int'l Symposium on Micromachine and Human Science*, pages 39–43, 1995.

[4] H. El Munim and A. Farag. Shape representation and registration using vector distance functions. *CVPR*, 2007.

[5] A. Fitzgibbon. Robust registration of 2d and 3d point sets. *IVC*, 21:1145–1153, 2003.

[6] G. Godin, M. Rioux, and R. Baribeau. Three-dimensional registration using range and intensity information. *SPIE: Videometrics III*, 2350, 1994.

[7] X. Huang, N. Paragios, and D. N. Metaxas. Shape registration in implicit spaces using information theory and free form deformations. *TPAMI*, 28:1303–1318, 2006.

[8] B. Jian and B. Vemuri. A robust algorithm for point set registration using mixture of gaussians. *ICCV*, 2:1246–1251, 2005.

[9] S. Manay, D. Cremers, B.-W. Hong, A. Yezzi, and S. Soatto. Integral invariants for shape matching. *TPAMI*, 28:1602–1618, 2006.

[10] T. Masuda, K. Sakaue, and N. Yokoya. Registration and integration of multiple range images for 3-d model construction. *ICPR*, 1:879–883, 1996.

[11] C. R. Maurer, R. Qi, and V. Raghavan. A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. *TPAMI*, 25, 2003.

[12] N. Paragios, M. Rousson, and V. Ramesh. Non-rigid registration using distance functions. *CVIU*, 89:142–165, 2003.

[13] R. Sandhu, S. Dambreville, and A. Tannenbaum. Particle filtering for registration of 2d and 3d point sets with stochastic dynamics. *CVPR*, 2008.

[14] F. Schmidt, D. Farin, and D. Cremers. Fast matching of planar shapes in sub-cubic runtime. *ICCV*, 2007.

[15] D. Sharvit, J. Chan, H. Tek, and B. Kimia. Symmetry-based indexing of image databases. *IEEE Workshop on Content-Based Access of Image and Video Libraries*, 1998.

[16] Y. Shi and R. Eberhart. A modified particle swarm optimizer. *IEEE World Congress on Computational Intelligence*, pages 69–73, 1998.

[17] L. Tang and G. Hamarneh. SMRFI: Shape matching via registration of vector-valued feature images. *CVPR*, 2008.

[18] Y. Tsin and T. Kanade. A correlation-based approach to robust point set registration. *ECCV*, 3:558–569, 2004.

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

[20] F. Wang, B. Vemuri, A. Rangarajan, and S. Eisenschenk. Simultaneous nonrigid registration of multiple point sets and atlas construction. *TPAMI*, pages 2011–2022, 2008.

[21] Y. Wang, X. Huang, C. S. Lee, S. Zhang, Z. Li, D. Samaras, D. Metaxas, A. Elgammal, and P. Huang. High resolution acquisition, learning and transfer of dynamic 3-D facial expressions. *Eurographics*, 23, 2004.

[22] P. Xiao, N. Barnes, T. Caetano, and P. Lieby. An mrf and gaussian curvature based shape representation for shape matching. *CVPR*, 2007.

[23] X. Xie, W. Zhang, and Z. Yang. Adaptive particle swarm optimization on individual level. *Int'l Conf. on Signal Processing*, 2:1215–1218, 2002.

[24] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *IJCV*, 13:119–152, 1994.