Robust Point Matching for Nonrigid Shapes by Preserving Local Neighborhood Structures

Yefeng Zheng, *Member*, *IEEE*, and
David Doermann, *Member*, *IEEE*

Abstract—In previous work on point matching, a set of points is often treated as an instance of a joint distribution to exploit global relationships in the point set. For nonrigid shapes, however, the local relationship among neighboring points is stronger and more stable than the global one. In this paper, we introduce the notion of a neighborhood structure for the general point matching problem. We formulate point matching as an optimization problem to preserve local neighborhood structures during matching. Our approach has a simple graph matching interpretation, where each point is a node in the graph, and two nodes are connected by an edge if they are neighbors. The optimal match between two graphs is the one that maximizes the number of matched edges. Existing techniques are leveraged to search for an optimal solution with the shape context distance used to initialize the graph matching, followed by relaxation labeling updates for refinement. Extensive experiments show the robustness of our approach under deformation, noise in point locations, outliers, occlusion, and rotation. It outperforms the shape context and TPS-RPM algorithms on most scenarios.

Index Terms—Point matching, shape matching, image registration, nonrigid shapes, relaxation labeling.

1 Introduction

SHAPE matching or image registration is often encountered in image analysis, computer vision, and pattern recognition. It is typically formulated as a point matching problem since point representations are general and easy to extract. In this paper, we focus on point-pattern-based shape matching. Shapes can be roughly categorized as rigid or nonrigid, and they may undergo deformation in captured images. With a small number of transformation parameters, rigid shape matching is relatively easy. Rigid shape matching under the affine or projective transformation has been widely studied [1]. Recently, point matching for nonrigid shapes has received a great deal of attention. There are two unknowns in a shape matching problem: the correspondence and the transformation [2]. Since solving for either without information regarding the other is quite difficult, most approaches to nonrigid shape matching use an iterated estimation framework. Given an estimate of the correspondence, the transformation may be estimated and used to update the correspondence. The iterated closest point (ICP) algorithm, a well-known heuristic approach proposed by Besl and McKay, is one example [3]. Assuming two shapes are roughly aligned, for each point in one shape, the closest point in the other shape is taken as the current estimate of the correspondence. Recently, Chui and Rangarajan [2] proposed an optimization-based approach, the TPS-RPM algorithm, in which two unknown variables (transformation and correspondence) are combined into an objective function. The soft assignment technique and deterministic annealing algorithm are used to search for an optimal solution. Belongie et al. [4] proposed another method for nonrigid point matching. In this approach, the shape context is assigned to each point, which describes the distribution of the remaining points relative to this point. The solution that minimizes the overall shape context distances is the optimal match between two point sets. One drawback of this approach is that neighboring points in one shape may be matched to two points far apart in the other shape.

We observe that although the absolute distance between two points may change significantly under nonrigid deformation, the neighborhood structure of a point is generally well preserved due to physical constraints. For example, although a human face is a nonrigid shape, the relative position of the chin, nose, mouth, and eyes, etc., cannot deform independently due to underlying constraints of bones and muscles. These physical constraints restrict the deformation of the point set sampled from a face. Furthermore, the rough structure of a shape is typically preserved; otherwise, even people cannot match shapes reliably under arbitrary large deformation. Such constraints may be represented as the ordering of points on a curve. Sebastian et al. [5] demonstrated the effectiveness of point ordering in matching curves, but for general shapes other than curves, local point ordering is hard to describe, and is ignored in many point matching algorithms [4]. As a major contribution of this paper, we formulate point matching as an optimization problem to preserve local neighborhood structures. In addition to the physical constraint explanation, our approach gets support from cognitive experiments of human shape perception. There is strong evidence that the early stages of human visual processing is local, parallel, and bottom-up, though feedback may be necessary in later stages. Preserving local neighborhood structures is important for people to detect and recognize shapes efficiently and reliably [6], [7].

Our formulation has a simple graph matching interpretation, where each point is a node in the graph, and two nodes are connected by an edge if they are neighbors. The optimal match between two graphs is the one that maximizes the number of matched edges (i.e., the number of neighborhood relations). Graph matching is an NP-hard problem. Exhaustive or branch-and-bound search for a global optimal solution is only realistic for graphs with few nodes. As an alternative approach, a discrete optimization problem can be converted to a continuous one, allowing many continuous optimization techniques to be applied [8], [9]. In this paper, we use the shape context distance to initialize graph matching, followed by a relaxation labeling process to refine the match.

The remainder of this paper is organized as follows: In Section 2, we formulate point matching as an optimization problem. Section 3 describes our approach based on relaxation labeling to search for an optimal solution of the objective function. In Section 4, we demonstrate the robustness of our approach by comparing it with two state-of-the-art algorithms: shape context [4] and TPS-RPM [2]. The paper concludes with a discussion of the limitations and future work in Section 5.

2 Problem Formulation

In this section, we formulate point matching as an optimization problem. Suppose a template shape T is composed of M points, $S_T = \{T_1, T_2, \cdots, T_M\}$, and a deformed shape D is composed of N points, $S_D = \{D_1, D_2, \cdots, D_N\}$. It is a common practice to enforce

• *The authors are with the Language and Media Processing Laboratory, Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. E-mail: {zhengyf, doermann}@cfar.umd.edu.*

Manuscript received 10 Dec. 2004; revised 30 Aug. 2005; accepted 12 Sept. 2005; published online 14 Feb. 2006.
Recommended for acceptance by C. Kambhamettu.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0659-1204.

the one-to-one matching constraint in point matching, so the point sets S_T and S_D are augmented to $S'_T = \{T_1, T_2, \cdots, T_M, nil\}$ and $S'_D = \{D_1, D_2, \cdots, D_N, nil\}$, respectively, by introducing a dummy or nil point. A match between shapes T and D is $f : S'_T \Leftrightarrow S'_D$, where the match of normal points is one-to-one, but multiple points may be matched to a dummy point.

Under a rigid transformation (translation and rotation), the distance between any pair of points is preserved. Therefore, the optimal match $\hat{f}$ is

$$\hat{f} = arg\min_f C(T, D, f), \tag{1}$$

where

$$\begin{aligned} C(T, D, f) = &\sum_{m=1}^{M}\sum_{i=1}^{M}\big(\|T_m - T_i\| - \|D_{f(m)} - D_{f(i)}\|\big)^2 \\ &+ \sum_{n=1}^{N}\sum_{j=1}^{N}\big(\|D_n - D_j\| - \|T_{f^{-1}(n)} - T_{f^{-1}(j)}\|\big)^2. \end{aligned} \tag{2}$$

In this cost function, we penalize any matching error which does not preserve the distance of a point pair. If $M = N$ and no points are matched to dummy points, the first term and the second term in (2) should be equal, and the optimal match should achieve zero penalty, $C(T, D, \hat{f}) = 0$. Points matching a dummy point need special treatment, however, to simplify the representation, we do not explicitly describe such treatment here. We will come back to this issue later.

If nonrigid deformation is present, the distance between a pair of points will, in general, not be preserved, especially for points which are far apart. On the other hand, due to physical constraints and in order to preserve the rough structure, the local neighborhood of a point may not change freely. We, therefore, define a neighborhood for point i as $\mathcal{N}_i$. The neighborhood relationship is symmetric, meaning if $j \in \mathcal{N}_i$, then $i \in \mathcal{N}_j$. Since we only want to preserve the distances of neighboring point pairs under deformation, (2) becomes

$$\begin{aligned} C(T, D, f) = &\sum_{m=1}^{M}\sum_{i\in\mathcal{N}_m}\big(\|T_m - T_i\| - \|D_{f(m)} - D_{f(i)}\|\big)^2 \\ &+ \sum_{n=1}^{N}\sum_{j\in\mathcal{N}_n}\big(\|D_n - D_j\| - \|T_{f^{-1}(n)} - T_{f^{-1}(j)}\|\big)^2. \end{aligned} \tag{3}$$

The absolute distance of a pair of points is not preserved well under scale changes. Therefore, we quantize the distance to two levels as

$$\|T_m - T_i\| = \begin{cases} 0 & m \in \mathcal{N}_i \\ 1 & m \notin \mathcal{N}_i \end{cases} \quad \text{and} \quad \|D_n - D_j\| = \begin{cases} 0 & n \in \mathcal{N}_j \\ 1 & n \notin \mathcal{N}_j. \end{cases} \tag{4}$$

Equation (3) then is simplified to

$$C(T, D, f) = \sum_{m=1}^{M}\sum_{i\in\mathcal{N}_m} d(f(m), f(i)) + \sum_{n=1}^{N}\sum_{j\in\mathcal{N}_n} d(f^{-1}(n), f^{-1}(j)), \tag{5}$$

where

$$d(i, j) = \begin{cases} 0 & j \in \mathcal{N}_i \\ 1 & j \notin \mathcal{N}_i. \end{cases} \tag{6}$$

To deal with points matched to a dummy point, we let $d(., nil) = d(nil, .) = d(nil, nil) = 1$ to discourage the match.

Simple mathematical deduction will convert the above minimization problem to a maximization problem [10]

$$\hat{f} = arg\ \max_f S(T, D, f), \tag{7}$$

where

$$S(T, D, f) = \sum_{m=1}^{M}\sum_{i\in\mathcal{N}_m}\delta(f(m), f(i)) + \sum_{n=1}^{N}\sum_{j\in\mathcal{N}_n}\delta(f^{-1}(n), f^{-1}(j)), \tag{8}$$

$$\delta(i, j) = 1 - d(i, j). \tag{9}$$

This formulation has a simple graph matching interpretation. Each point is a node in the graph and two nodes are connected by an edge if they are neighbors. The dummy node is not connected to other nodes in the graph. If connected nodes m and i in one graph are matched to connected nodes $f(m)$ and $f(i)$ in the other graph, $\delta(f(m), f(i)) = 1$. The optimal solution of (7) is the one that maximizes the number of matched edges of two graphs.

There is no obvious neighborhood definition for a point set. In the following, we present a simple neighborhood definition.[1] Initially, the graph is fully connected, and we then remove long edges until a predefined number of edges are preserved. Suppose there are M nodes in the graph, on average, each point has E_{ave} neighbors, then the number of preserved edges is $M \times E_{ave}/2$ ($E_{ave} = 5$ in default). With this neighborhood definition, the graph representation of a point set is translation, rotation, and scale invariant.

We can represent the matching function f in (7) with a set of supplemental variables, which are organized as a matrix P with dimension $(M + 1) \times (N + 1)$.

If point T_m in the template shape T is matched to point D_n in the deformed shape D, then $P_{mn} = 1$; otherwise, $P_{mn} = 0$. The last row and column of P represent the case that a point may be matched to a dummy point. Matrix P satisfies the following normalization conditions

$$\sum_{n=1}^{N+1} P_{mn} = 1 \quad \text{for } m = 1, 2, \cdots, M, \tag{11}$$

$$\sum_{m=1}^{M+1} P_{mn} = 1 \quad \text{for } n = 1, 2, \cdots, N. \tag{12}$$

Using matrix P, the objective function (8) can be written as

$$S(T, D, P) = 2\sum_{m=1}^{M}\sum_{i\in\mathcal{N}_m}\sum_{n=1}^{N}\sum_{j\in\mathcal{N}_n} P_{mn}P_{ij}. \tag{13}$$

3 Searching for an Optimal Solution

Since $P_{mn} \in \{0, 1\}$, searching for an optimal P that maximizes $S(T, D, P)$ is a hard discrete combinatorial problem. In this paper, we use relaxation labeling to solve the optimization problem, where the condition $P_{mn} \in \{0, 1\}$ is relaxed as $P_{mn} \in [0, 1]$ [9]. After relaxation, P_{mn} is a real number, and the problem is converted to a constrained optimization problem with continuous variables.

1. Our framework is general enough to incorporate other neighborhood definitions. Please refer to Section 4 for a definition, which is robust under nonuniform scale changes for different parts of a shape.

3.1 Matching Initialization

Relaxation labeling only converges to a local optimal solution, so a good initialization is crucial for getting a good solution. In this paper, we use the shape context distance [4] to initialize the matching probability matrix P. The shape context of a point is a measure of the distribution of other points relative to it. Consider two points, m in one shape and n in the other shape. Their shape contexts are $h_m(k)$ and $h_n(k)$, for $k = 1, 2, , K$, respectively. Let C_{mn} denote the matching cost of these two points. As shape contexts are distributions represented as histograms, it is natural to use the χ^2 test statistic as a cost measure [4]

$$C_{mn} = \frac{1}{2}\sum_{k=1}^{K}\frac{[h_m(k) - h_n(k)]^2}{h_m(k) + h_n(k)}. \quad (14)$$

The Gibbs distribution is widely used in statistical physics and image analysis to relate the energy of a state to its probability. Taking cost C_{mn} as the energy of the state that points m and n are matched, the probability of the match is

$$P_{mn} \propto e^{-C_{mn}/T_{init}}. \quad (15)$$

Parameter T_{init} ($T_{init} = 0.1$ in default) is used to adjust the reliability of the initial probability measures. We set the probability for a point matching a dummy point ($P_{m,nil}$ or $P_{nil,n}$) to 0.2. Experiments show that our approach is not very sensitive to this parameter.

It is easy to make shape context translation and scale invariant, and in some applications, rotation invariance is also required. Our graph representation is rotation invariant, so we need a rotation invariant initialization scheme. Without referring to original images from which the point set are sampled, the rotation invariant shape context based on the tangent direction cannot be used [4]. Instead, we use the mass center of a point set as a reference point, and use the direction from a point to the mass center as the positive x-axis for the local coordinate system [10].

3.2 Relaxation Labeling

Relaxation labeling was first proposed in a seminal paper by Rosenfeld et al. in mid-1970s [9]. The basic idea is to use iterated local context updates to achieve a globally consistent result. The contextual constraints are expressed in the compatibility function $R_{mn}(i, j)$, which measures, in our case, the strength of compatibility between T_m matching D_n and T_i matching D_j. The support function S_{mn} measures the overall support the match between points T_m and D_n gets from its neighbors.

$$S_{mn} = \sum_{i\in\mathcal{N}_m}\sum_{j\in\mathcal{N}_n} R_{mn}(i, j)P_{ij}. \quad (16)$$

The original updating rule is [9]

$$P_{mn} := \frac{P_{mn}S_{mn}}{\sum_{j=1}^{N} P_{mj}S_{mj}}. \quad (17)$$

The denominator is used to enforce the normalization constraint. Traditionally, only the one-way normalization constraint, (11), is enforced in relaxation labeling.

Theoretic analysis [11] shows the updating process will converge to a local optimum if 1) the objective function is a polynomial with positive coefficients and 2) S_{mn} is defined as the derivative of the objective function. With the objective function of (13), S_{mn} is defined as

$$S_{mn} = 4\sum_{i\in\mathcal{N}_m}\sum_{j\in\mathcal{N}_n} P_{ij}. \quad (18)$$

Interpreted in the relaxation labeling theory, our compatibility function is $R_{mn}(i, j) = 1$ if a pair of neighbors T_m and T_i are matched to a pair of neighbors D_n and D_j; otherwise, $R_{mn}(i, j) = 0$. We can show that our objective function corresponds to the average local consistency measure of the Hummel-Zucker consistency theory [12]. According to the consistency theory, each update step will increase the overall consistency of the system. Each unambiguous consistent solution is a local attractor. Starting from a nearby point, the relaxation labeling process will converge to it [12], [11]. Although there is no guarantee that the updating process will converge to an unambiguous solution starting from an *arbitrary* initialization, our experiments show that most elements of matrix P do converge to zero or one.

If one-to-one matching is necessary, we perform alternated row and column normalizations to the matching probability matrix P, after each step of relaxation update. It is well-known that this normalization process will convert a square matrix with positive elements to a doubly stochastic matrix [2]. We can show that alternated row and column normalizations (except the last row and column) of a positive matrix will result in a generalized doubly stochastic matrix.[2] This technique is also used in the soft assignment approach without proof [2]. Though the original relaxation labeling is theoretically well founded [11], it is not clear whether the updating process will converge to a local optimum and increase the consistency after imposing the one-to-one matching constraint. We found the updating process still converges by experiments, but we cannot prove it theoretically.

It is difficult to achieve a good match for shapes under nonrigid distortions with a single approach. In this paper, we also use the common framework of iterated correspondence and transformation estimations. Similar to [4], the thin plate spline (TPS) deformation model is exploited to bring two shapes closer in each iteration.

3.3 Relationship to Previous Work

One difference from the previous applications of relaxation labeling on point matching [13] is that we use it to solve a constrained optimization problem, so the relaxation labeling process is guaranteed to converge to a local optimal solution [11]. In the previous work, relaxation labeling is used in an ad hoc way without an objective function to be optimized, so there is no clear indication of the quality of the solution. Furthermore, unlike previous work, we can enforce one-to-one matching in our approach if necessary.

The relaxation labeling method used in this paper is similar to the well-known soft assignment technique [8], [2]. Both of them convert the discrete combinatorial optimization problem to one with continuous variables by assigning a probability measure to a match. The procedure is called "relaxation" or "soft" in these two techniques, respectively. Generally, deterministic annealing is used to solve the soft assignment problem. It first starts with a high temperature where the matching probabilities are uniform. Gradually decreasing the temperature, the matching probabilities will converge to a local optimal solution of the objective function. An appropriate choice of the initial temperature and temperature reduction ratio is necessary to achieve good results [2]. On the contrary, the relaxation labeling-based approach is parameter free. Another advantage is that we can easily incorporate a meaningful

2. Refer to our companion technical report for a proof [10].

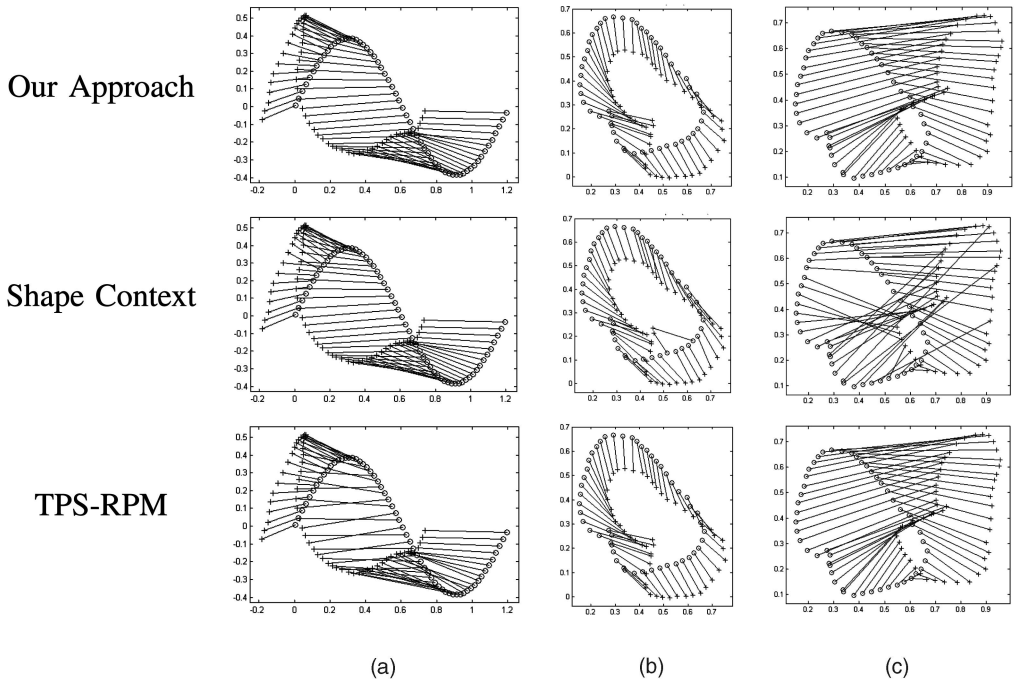

Fig. 1. Point matching results on (a) open curves and (b) and (c) closed contours. Top row: our approach. Middle row: the shape context algorithm. Bottom row: the TPS-RPM algorithm.

initialization in relaxation labeling. Since the distribution of local optima is complex, a good initialization is crucial to achieve a good result. Unfortunately, it is hard to incorporate an initialization method into the deterministic annealing framework. It is also a drawback of another continuous optimization technique for graph matching proposed by Pelillo [14]. We tested the soft assignment based graph matching method [8] and found the results were much worse than the relaxation labeling-based approach.

4 Experiments

In this section, we first show that our approach preserves sequential ordering of points (a degenerated neighborhood structure) on open curves and closed contours during matching. We then quantitatively compare it with two state-of-the-art algorithms for robustness under deformation, noise in point locations, outliers, occlusion, and rotation.

4.1 Examples

We have tested our algorithm on the samples used in [2], and compare our results with two other algorithms: shape context [4] and TPS-RPM [2]. In these examples, the template and deformed shapes have the same number of points. In order to achieve a direct and fair comparison, we force all algorithms to match as many point pairs as possible without outlier rejection. Fig. 1 shows the point matching results of three algorithms on a pair of open curves and two pairs of closed contours. As shown in Fig. 1a, all three algorithms achieve good results for the pair of open curves even though the deformation is large. Neighboring points may swap their matches in TPS-RPM, however. For the first pair of closed contours, all algorithms achieve reasonable results, but the shape context algorithm introduces a few mismatches as shown in Fig. 1b. Since the rotation between two shapes is large for the second pair of closed contours, the rotation invariance shape context is used for initialization in our approach and the shape context algorithm. Both our approach and TPS-RPM achieve good results and preserve the sequential ordering of points. The result of the shape context algorithm is not as good: Neighboring points in one shape may be matched to points far apart in the other shape.

Our approach is general enough to incorporate other neighborhood definitions. In this experiment, we use another neighborhood definition which is robust when different parts of a shape have significantly different scales. For two points in a shape, they are neighbors if and only if they are both in each other's top E_{ave} nearest points. It is easy to verify that this neighborhood relationship is still symmetric and scale invariant. A fish shape is used to synthesize test samples. We first applied a moderate amount of nonrigid deformation to the template shape, and then enlarged the fish tail. Figs. 2b and 2c show the graphs of the template and deformed shapes when $E_{ave} = 5$ and the fish tail is enlarged to four times of the original size. This neighborhood definition is robust: 91 percent of the edges in the deformed graph have a correspondence in the template graph, and a good point match is achieved as shown in Fig. 2d.

4.2 Quantitative Evaluation

Synthetic data is easy to obtain and can be designed to test specific aspects of an algorithm. We test our algorithm on the same synthesized data as in [2] and [4]. There are three sets of data designed to measure the robustness of an algorithm under **deformation**, **noise** in point locations, and **outliers**. In each test, the template point set is subjected to one of the above distortions to create a *target* point set (for the latter two test sets, a moderate amount of deformation is present). Two shapes (a fish and a Chinese character) are used, and 100 samples are generated for each degradation level. We then run our algorithm to find the correspondence between these two sets of points and use the estimated correspondence to warp the template shape. The accuracy of the match is quantified as the average Euclidean distance between a point in the warped template and the corresponding point in the target. Alternative evaluation metrics

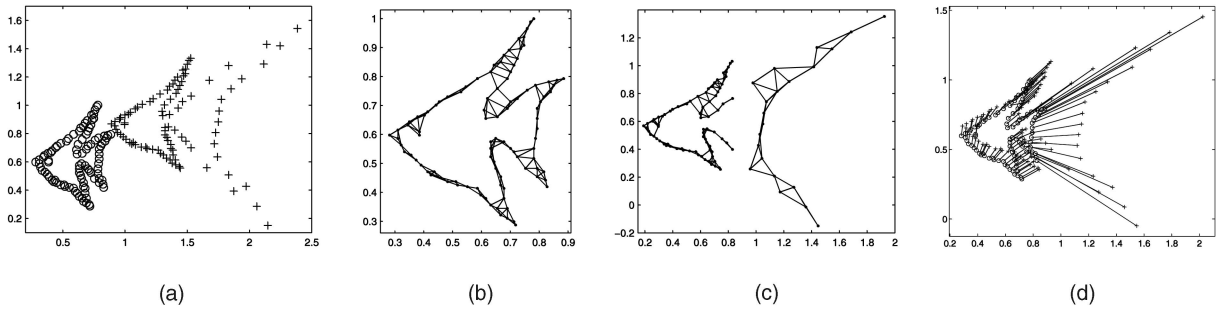

Fig. 2. A neighborhood definition which is robust under large nonuniform scale changes for different parts of a shape. (a) Point sets of the template (o) and deformed (+) shapes. (b) Template graph with 210 edges. (c) Deformed graph with 196 edges. Among them, 178 (91 percent) edges also present in the template graph. (d) Point matching result.

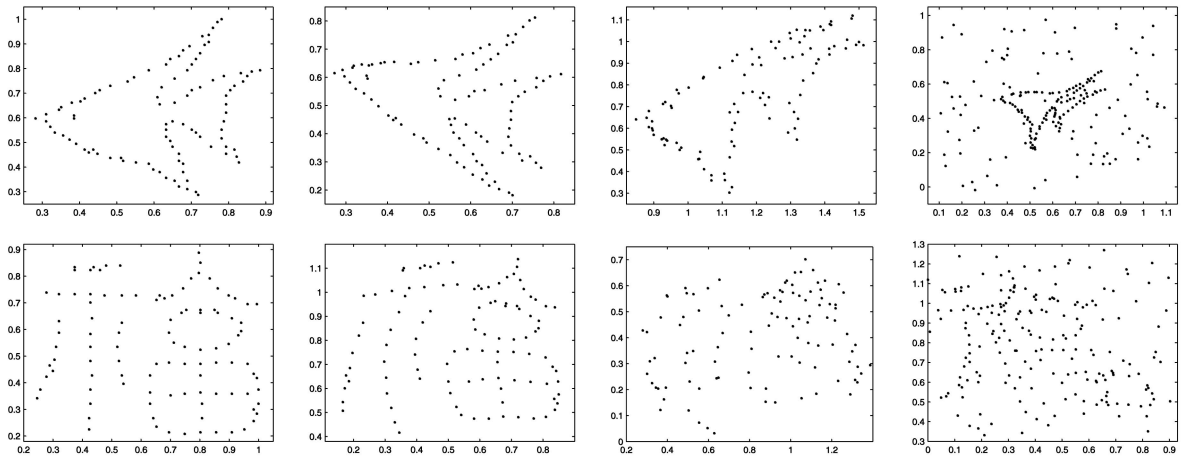

Fig. 3. Chui-Rangarajan synthesized data sets. The template point sets are shown in the first column. Columns 2-4 show examples of target point sets for the deformation, noise, and outlier tests, respectively.

are possible (e.g., the number of correctly matched point pairs), but in order to compare our results directly with two other algorithms, we use the same evaluation metric as in [2] and [4]. Several examples from the synthesized data sets are shown in Fig. 3. The quantitative evaluation results are shown in Fig. 4. Since the new neighborhood definition presented above is not robust to outliers, the original version is used throughout this experiment. Our algorithm performs best on the deformation and noise sets. For the outlier test set, however, there is no clear winner. The TPS-RPM algorithm outperforms our algorithm on the Chinese character shape under large outlier ratios. Since points are spread out on the Chinese character shape, when a large number of outliers are present, the neighborhood of a point changes significantly (as shown in the last column of Fig. 3), which violates our assumption. On the contrary, points on the fish shape are clustered, and the neighborhood of a point is preserved well even with a large outlier ratio. Therefore, better results are achieved by our algorithm on this shape.

Occlusion is often present in real applications and is a challenge for many algorithms. In the following experiments, we test the three algorithms under occlusion using synthesized data. A moderate amount of nonlinear deformation is applied to a shape. We then randomly select a point and remove it together with some of its closest points. Six occlusion levels are used: 0 percent, 10 percent, 20 percent, 30 percent, 40 percent, and 50 percent, and 100 samples are generated for each level. The left two columns of the top row of Fig. 5 show two synthesized samples. Quantitative evaluation results are shown in the left two columns of the bottom row of Fig. 5. The TPS-RPM algorithm treats all extra points as outliers, which are assumed to be independently distributed. Since it does not model the distribution of occlusions well, the performance of TPS-RPM deteriorates quickly. In our approach, the change of neighborhood structure is restricted to points close to the occlusion. As shown in Fig. 5, our approach achieves the best results for up to 40 percent occlusion. When the occlusion ratio is large, a shape is likely to be broken into several parts and the neighborhood structure of almost all remaining points may be changed. Therefore, when the occlusion ratio is 50 percent, the difference between our approach and the shape context algorithm is small.

In some applications, rotation invariance is a critical property of a shape matching algorithm. We test our algorithm under rotations using synthesized data of the same fish and Chinese character shapes. A moderate amount of nonlinear deformation is applied to a shape, and the ground-truthed correspondences are used to correct the rotation introduced in the deformation. We then rotate the deformed shape. The probability of selecting a clockwise or counterclockwise rotation is equal. Six rotations are used: 0, 30, 60, 90, 120, and 180 degrees. One hundred samples are generated for each rotation. The right two columns of the top row of Fig. 5 show two synthesized samples. At the first iteration, the rotation invariant shape context distance is used to initialize the graph matching in our approach. The rotation between two shapes is corrected by the affine transformation in the first iteration. After

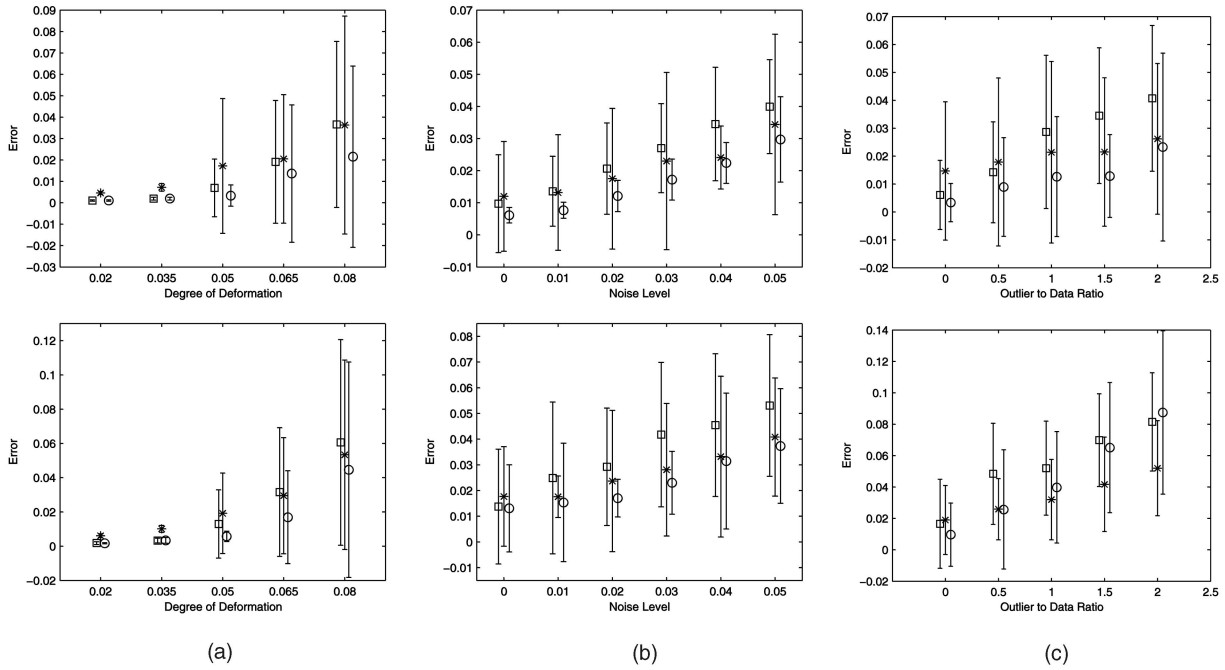

Fig. 4. Comparison of our results (○) with the TPS-RPM (∗), and shape context (□) algorithms on the Chui-Rangarajan synthesized data sets. The error bars indicate the standard deviation of the error over 100 random trials. Top row: the shape of fish. Bottom row: the shape of a Chinese character.

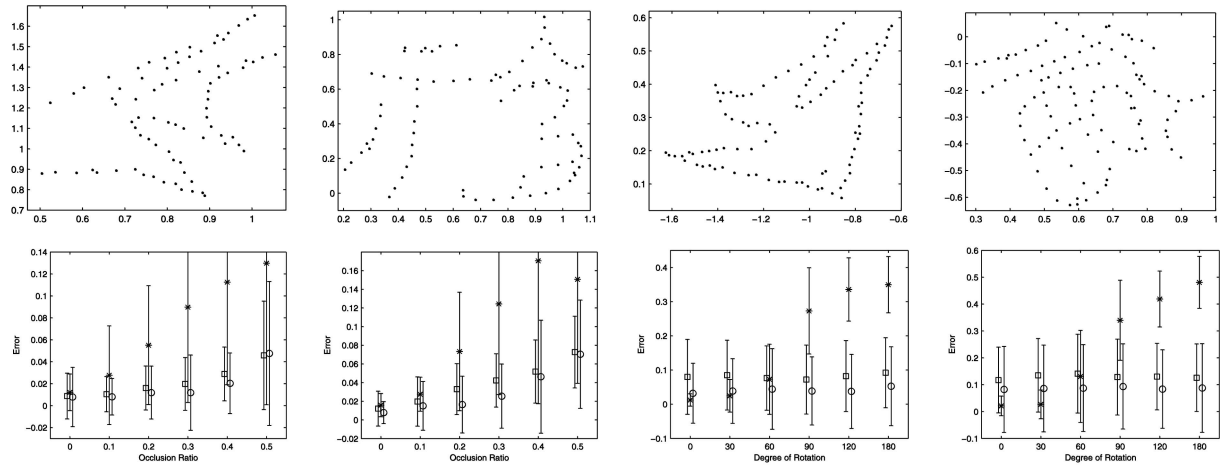

Fig. 5. Comparison of our results (○) with the TPS-RPM (∗) and shape context (□) algorithms under occlusion and rotation. Left two columns: the fish and Chinese character shapes under occlusion. Right two columns: the fish and Chinese character shapes under rotation. Top row: synthesized samples. Bottom row: mean and variance of errors.

that, the normal shape context distance is used. Quantitative evaluation results are shown in the right two columns of the bottom row of Fig. 5. We can see that our method is truly rotation invariant, and it consistently outperforms the shape context algorithm. TPS-RPM, however, can only tolerate rotations up to 60 degrees.

In Figs. 4 and 5, the variance of all algorithms is large. Therefore, a statistical analysis must be applied to ascertain whether the performance difference between these algorithms is significant. Since the distribution of errors is not Gaussian, we use the Wilcoxon paired signed rank test, which is distribution free and powerful [15]. Table 1 shows the statistical analysis (with two-sided significance level of 0.01) of the performance of our approach compared with two other algorithms. Here, $+$ $(-)$ means the improvement (deterioration) of our approach is statistically significant compared with the other algorithm, and $=$ means there is no significant difference between two algorithms. The statistical test verifies that the improvement of our approach on most data sets is significant.

5 Conclusions and Future Work

In this paper, we have introduced the notion of a neighborhood structure for the general point matching problem. We formulated point matching as an optimization problem to preserve local neighborhood structures during matching. Extensive experiments were presented to demonstrate the robustness of our approach. Compared with the other two state-of-the-art algorithms, our approach performs as well or better under nonrigid deformation, noise in point locations, outliers, occlusion, and rotation.

TABLE 1
Wilcoxon Paired Signed Rank Test

	Fish					Chinese character				
	Deformation	Noise	Outlier	Occlusion	Rotation	Deformation	Noise	Outlier	Occlusion	Rotation
Ours vs. shape context	= = + + +	= + + + + +	+ + + + +	= + + + + =	+ + + + + +	= = + + +	+ = + + + +	+ + + = =	+ + + + + =	= + + = = =
Ours vs. TPS-RPM	+ + + + +	+ + + + = +	+ + + + +	+ + + + + +	- - + + + +	+ + + + +	+ + + + + +	+ = = - -	+ + + + + +	= = + + + +

+, −, and = mean the former algorithm is better, worse, or no difference than the latter, respectively.

Large outlier or occlusion ratios (especially if the occlusion breaks a shape into several isolated parts) can significantly change the local neighborhood structure. A combination of different sources of degradation, such as large rotation, noise, and occlusion, is also a challenge, which should be addressed in our future research. In this work, the relaxation labeling method is used to solve the constrained optimization problem. Only converging to a local optimum, it is by no means the best approach. We are testing other optimization methods such as simulated annealing, genetic algorithms, and graduated nonconvexity methods. Our graph matching formulation is applicable for both 2D and 3D shapes. Using the shape context distance for initialization, we only demonstrate it on 2D shapes, since the original shape context is only defined for 2D point sets. We will test the effectiveness of our approach for 3D shape matching by extending the shape context to 3D point sets.

ACKNOWLEDGMENTS

The authors would like to thank Haili Chui, Anand Rangarajan, Serge Belongie, and Jitendra Malik for providing their implementation source code and test data sets. This makes the comparison experiments much easier. The authors also thank the anonymous reviewers for their constructive comments. The support of this research by the US Department of Defense under contract MDA-9040-2C-0406 is gratefully acknowledged.

REFERENCES

[1] L.G. Brown, "A Survey of Image Registration Techniques," *ACM Computing Survey*, vol. 24, no. 4, pp. 325-376, 1992.
[2] H. Chui and A. Rangarajan, "A New Point Matching Algorithm for Non-Rigid Registration," *Computer Vision and Image Understanding*, vol. 89, nos. 2-3, pp. 114-141, 2003.
[3] P.J. Besl and N.D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Trans. Pattern Analysis and Machine Inteligence*, vol. 14, no. 2, pp. 239-256, Feb. 1992.
[4] S. Belongie, J. Malik, and J. Puzicha, "Shape Matching and Object Recognition Using Shape Contexts," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509-522, Apr. 2002.
[5] T.B. Sebastian, P.N. Klein, and B.B. Kimia, "On Aligning Curves," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 1, pp. 116-124, Jan. 2003.
[6] I. Biederman, "Recognition by Components: A Theory of Human Image Understanding," *Psychological Rev.*, vol. 94, no. 2, pp. 115-147, 1987.
[7] D.J. Field, A. Hayes, and R.E. Hess, "Contour Integration by the Human Visual System: Evidence for a Local 'Association Field,'" *Vision Research*, vol. 33, no. 2, pp. 173-193, 1993.
[8] S. Gold and A. Rangarajan, "A Graduated Assignment Algorithm for Graph Matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 4, pp. 377-388, Apr. 1996.
[9] A. Rosenfeld, R.A. Hummel, and S.W. Zucker, "Scene Labeling by Relaxation Operations," *IEEE Trans. System, Man, and Cybernetics*, vol. 6, no. 6, pp. 420-433, 1976.
[10] Y. Zheng and D. Doermann, "Robust Point Matching for Non-Rigid Shapes: A Relaxation Labeling Based Approach," Technical Report, LAMP-TR-117, Univ. of Maryland, College Park, 2004.
[11] M. Pelillo, "The Dynamics of Nonlinear Relaxation Labeling Processes," *J. Math. Imaging and Vision*, vol. 7, no. 4, pp. 309-323, 1997.
[12] R.A. Hummel and S.W. Zucker, "On the Foundations of Relaxation Labeling Processes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 5, no. 3, pp. 267-287, 1983.
[13] J. Ton and A.K. Jain, "Registering Landsat Images by Point Matching," *IEEE Trans. Geoscience and Remote Sensing*, vol. 27, no. 5, pp. 642-651, 1989.
[14] M. Pelillo, "Replicator Equations, Maximal Cliques, and Graph Isomorphism," *Neural Computation*, vol. 11, no. 8, pp. 1933-1955, 1999.
[15] G.K. Kanji, *100 Statistical Tests*. Sage Publications, 1999.