

Robust Point Matching using Mixture of Asymmetric Gaussians for Nonrigid Transformation

Gang Wang, Zhicheng Wang, Weidong Zhao, and Qiangqiang Zhou

CAD Research Center, Tongji University, Shanghai 201804, China
gwang.cv@gmail.com, {zhichengwang,wd}@tongji.edu.cn, zqqcsu@gmail.com

Abstract. In this paper, we present a novel robust method for point matching under noise, deformation, occlusion and outliers. We introduce a new probability model to represent point sets, namely asymmetric Gaussian (AG), which can capture spatially asymmetric distributions. Firstly, we use a mixture of AGs to represent the point set. Secondly, we use L_2 -minimizing estimate (L_2E), a robust estimator to estimate densities between two point sets, to estimate the transformation function in reproducing kernel Hilbert space (RKHS) with regularization theory. Thirdly, we use low-rank kernel matrix approximation to reduce the computational complexity. Experimental results show that our method outperforms the comparative state-of-the-art methods on most scenarios, and it is quite robust to noise, deformation, occlusion and outliers.

1 Introduction

The point matching problem can be categorized into rigid and nonrigid matching depends on the transformation pattern. Generally, rigid transformation, containing translation, rotation and scaling, is relatively easy to estimate. By contrast, nonrigid transformation is hard to resolve since the transformation model is often unknown and difficult to model. Nonrigid transformation exists in numerous applications, including hand-written character recognition, facial-expression recognition and medical image registration.

However, there are many parameters in nonrigid transformation, causing several problems: 1) sensitive to noise, deformation, occlusion and outliers; 2) the trap of local minima; 3) high computational complexity. The nonrigid matching problem, in this sense, remains unsolved. Therefore, a point matching method should construct the complex transformation model with low computational complexity.

The Iterative Closest Point (ICP) algorithm [3] is one of the best known algorithms for point matching, because of its simplicity and low computational complexity. However, ICP requires a good initial position, *i.e.*, adequately close distance between the *Model* and the *Scene* point sets.

In order to address the limitations of ICP and improve the performance of matching, many interesting methods are proposed recently. Chui et al. [4]

proposed a robust point matching algorithm named TPS-RPM. TPS-RPM is more robust than ICP to noise, deformation, occlusion and outliers.

Tsin et al. [20] proposed a correlation-based named kernel correlation (KC) point set registration method where the correlation of two kernel density estimates is used to formulate the cost function. Zheng et al. [21] proposed a robust point matching method for nonrigid shapes by preserving local neighborhood structures. Myronenko et al. [13, 14] proposed another algorithm, namely the Coherence Point Drift (CPD), based on the motion coherence theory (MCT). CPD can get good results in a very short time when handling a large number of points.

Moreover, Jian et al. [7] proposed a robust point set registration approach using Gaussian mixture models (GMM), they leverage the closed-form expression for the L_2 distance between two Gaussian mixtures which represent the given point sets. Alternatively, Ma et al. [10] introduced L_2 -minimizing estimate (L_2E) [17], a robust estimator in statistics, to the nonrigid transformation estimation problem. Then they proposed a robust point matching method named RPM- L_2E .

In this paper, we are present a novel robust point matching method. Briefly, the core of our method is using a mixture of asymmetric Gaussians (AG) to represent the density of the given point set. Then we use L_2E [17] to estimate the transformation parameters.

The rest of this paper organized as follows: A novel robust point matching method using mixture of asymmetric Gaussians and L_2 -minimizing estimate for nonrigid transformation is presented in Section 2. Section 3 shows an optimal solution of our proposed method. The experiments and performance evaluation of our proposed method is shown in Section 4. In Section 5, we present a conclusion.

2 Method

2.1 Point Set Representation Using Mixture of Asymmetric Gaussians

We introduce a new probability model named Asymmetric Gaussian (AG) [8] which can capture spatially asymmetric distributions. AG is another form extending from Gaussian. It is shown that Gaussian has a symmetric distribution while AG has an asymmetric distribution by Fig. 1 where the density functions are plotted. Thus the distribution of AG is given by:

$$\mathcal{A}(x|\mu, \sigma^2, r) = \frac{1}{(2\pi\sigma^2)^{D/2}((r+1)/2)^D} \begin{cases} \exp\left(-\frac{|x-\mu|^2}{2\sigma^2}\right) & x \leq \mu \\ \exp\left(-\frac{|x-\mu|^2}{2r^2\sigma^2}\right) & x > \mu \end{cases} \quad (1)$$

where D is the dimension of data, μ , σ^2 and r are parameters of AG where $r = 1$ means AG equaling to Gaussian.

Since the definition of the density model, it is easy to construct a mixture of AG which may be well approximate almost any density with a linear combination

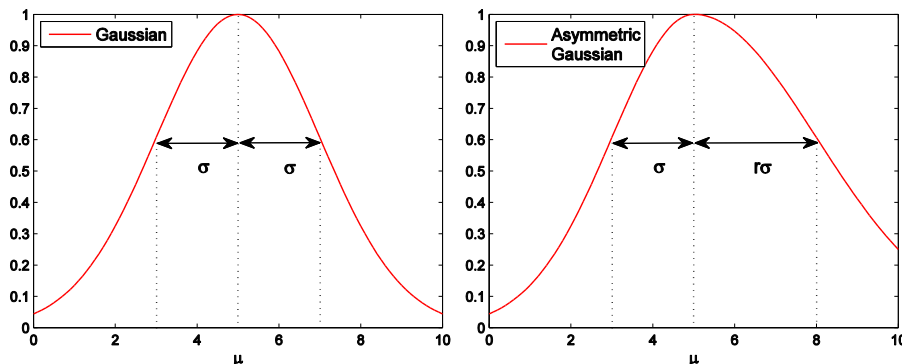


Fig. 1. Density functions of Gaussian and asymmetric Gaussian (AG).

of local AGs. Then the overall density of the J -component mixture is given by:

$$p(x) = \sum_{j=1}^J w_j \mathcal{A}(x|\mu, \sigma^2, r) \quad (2)$$

where w_j is the weight of each component, and $\{w_j\}_{j=1}^J$ are mixing proportions satisfying $0 \leq w_j \leq 1$ and $\sum_{j=1}^J w_j = 1$.

Given two point sets: 1) the *Model* set $X_{M \times D} = (x_1, \dots, x_M)^T$ which needs to be moved; 2) the *Scene* set $Y_{N \times D} = (y_1, \dots, y_N)^T$ which is fixed. In this paper, we represent a discrete point set by a mixture of AGs where the number of AG components is equivalent to the number of points. Note that all AG components are weighted equally, and all point sets are normalized as distributions with zero mean and unit variance.

2.2 The Robust Point Matching Method

In this article, we select L_2E [17] to estimate the unknown parameters of the transformation between two AG mixtures. The estimation error of L_2E which maximizes the sum of the densities is less than the estimation error of maximum likelihood estimation (MLE) which maximizes the product of the densities. In nonrigid transformation, the transformation f can be solved by minimizing the following cost function:

$$F_{AG}(f, \sigma^2, r) = \frac{1}{(2\pi\sigma^2)^D ((r+1)/2)^{2D}} - \frac{2}{MN} \sum_{i=1}^N \sum_{j=1}^M \mathcal{A}(y_i - f(x_j) | 0, \sigma^2 I, rI) \quad (3)$$

where f is the transformation model, $M \leq N$, and I is an identity matrix of size $D \times D$.

Following the idea of TPS-RPM [4], we use a slightly simpler form to estimate the transformation without considering outliers:

$$F_{AG}(f, \sigma^2, r) = \frac{1}{(2\pi\sigma^2)^D((r+1)/2)^{2D}} - \frac{2}{L} \sum_{l=1}^L \mathcal{A}(y_l - f(x_l) | 0, \sigma^2 I, rI) \quad (4)$$

where L is the number of correspondences and $L \leq M$. y_l , which is recovered from Eq. 3 or other fitting models (*e.g.*, Soft-assignment [4], Shape Context [2]), denotes the correspondence to x_l .

Here we introduce a special space, reproducing kernel Hilbert space (RKHS), and then finding the functional form of the transformation model f using calculus of variation. In RKHS, it is given the *Model* set $X \in \mathbb{R}^D$, the *Scene* set $Y \in \mathbb{R}^D$, and their correspondence set $S = \{(x_1, y_1), \dots, (x_L, y_L)\}$. Then we define an RKHS \mathcal{H} with a positive definite kernel function k . In this paper, we use the Gaussian kernel: $k(x_i, x_j) = \exp(-\beta\|x_i - x_j\|^2)$, where β is a constant. Thus we can define the kernel matrix K :

$$K = \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_L) \\ \vdots & \ddots & \vdots \\ k(x_L, x_1) & \dots & k(x_L, x_L) \end{bmatrix} \quad (5)$$

The transformation function $f \in \mathcal{H}$ can be found by minimizing the following regularized least-squares [1, 16]:

$$\varepsilon(f) = \min_{f \in \mathcal{H}} F_{AG}(f, \sigma^2, r) + \frac{\lambda}{2} \|f\|_K^2 \quad (6)$$

where the first term is the empirical risk and the second term is the Tikhonov regularization [18], $\lambda > 0$ is a trade-off parameter, $\|\cdot\|_K$ denotes a norm in the RKHS. Tikhonov regularization form smoothly trades-off $\|f\|_K^2$ and the empirical risk and solves the ill-posed problem in point matching.

According to the representation theorem [12] and related study in [1, 14], the solution of Eq. (6) to the Tikhonov regularization can be written in the following form:

$$f^*(\cdot) = \sum_{i=1}^L h_i K(x_i, \cdot) \quad (7)$$

for some $h_i \in \mathbb{R}^L$.

Substituting Eq. (7) into the cost function (4), we can therefore rewrite it with the Tikhonov regularization as

$$F_{AG}(H, \sigma^2, r) = \frac{2^D}{(\pi\sigma^2(r+1)^2)^D} - \frac{2^{(2+D)/2}}{L(\pi\sigma^2(r+1)^2)^{D/2}} \Gamma + \frac{\lambda}{2} \text{tr}(H^T K H) \quad (8)$$

where

$$\Gamma = \begin{cases} \exp\left(-\frac{\|Y - KH\|^2}{2\sigma^2}\right) & \text{if } y_i \leq (KH)_i \\ \exp\left(-\frac{\|Y - KH\|^2}{2r^2\sigma^2}\right) & \text{otherwise} \end{cases} \quad (9)$$

and $tr(\cdot)$ denotes the trace, $H = (h_1, \dots, h_L)^T$ is an coefficient matrix of size $L \times D$.

2.3 Low-rank Kernel Matrix Approximation

The matrix-valued kernel [12, 19] plays an important role in the regularization theory, it provides an easy way to choose an RKHS. However, in this paper, the computational complexity of our method is $O(N^3)$, hopefully, low-rank kernel matrix approximation [11] can yield a large increase in speed with little loss in accuracy. Low-rank kernel matrix approximation \hat{K} is the closest τ -rank matrix approximation to K and satisfying L_2 and Frobenius norms.

Using eigenvalue decomposition of K , the approximation matrix can be rewritten as:

$$\hat{K} = Q\Lambda Q^T \quad (10)$$

where Λ is a diagonal matrix of size $\tau \times \tau$ with τ largest eigenvalues and Q is an $L \times \tau$ matrix with the corresponding eigenvectors. The object function of our method therefore can be rewritten as:

$$F_{AG}(\hat{H}, \sigma^2, r) = \frac{2^D}{(\pi\sigma^2(r+1)^2)^D} - \frac{2^{(2+D)/2}}{L(\pi\sigma^2(r+1)^2)^{D/2}} \hat{\Gamma} + \frac{\lambda}{2} tr(\hat{H}^T \hat{K} \hat{H}) \quad (11)$$

where

$$\hat{\Gamma} = \begin{cases} \exp\left(-\frac{\|Y-U\hat{H}\|^2}{2\sigma^2}\right) & \text{if } y_i \leq (U\hat{H})_i \\ \exp\left(-\frac{\|Y-U\hat{H}\|^2}{2r^2\sigma^2}\right) & \text{otherwise} \end{cases} \quad (12)$$

and $U_{L \times \tau} = Q\Lambda$, parameter matrix \hat{H} of size $\tau \times D$ instead of the original matrix H .

3 Searching for An Optimal Solution

In this paper, the aforementioned cost function is convex in the neighborhood of the optimal position and, most importantly, always differentiable. Thus, the numerical optimization problem can be solved by employing some gradient-based optimization methods, such as quasi-Newton method [15]. The derivative of Eq. (11) with respect to the coefficient matrix \hat{H} is given by:

$$\frac{\partial F_{AG}}{\partial \hat{H}} = \lambda U \hat{H} + \frac{2\hat{K}}{L\sigma^2(2\pi\sigma^2)^{D/2}((r+1)/2)^D} \begin{cases} V \circ (C \otimes 1) & \text{if } y_i \leq (U\hat{H})_i \\ \frac{1}{r^2} V \circ (\hat{C} \otimes 1) & \text{otherwise} \end{cases} \quad (13)$$

where $V = U\hat{H} - Y$, $C = \exp(\text{diag}(VV^T)/2\sigma^2)$, $\hat{C} = \exp(\text{diag}(VV^T)/2r^2\sigma^2)$ and 1 is an $1 \times D$ row vector of all ones. \circ denotes the Hadamard product, \otimes denotes the tensor product.

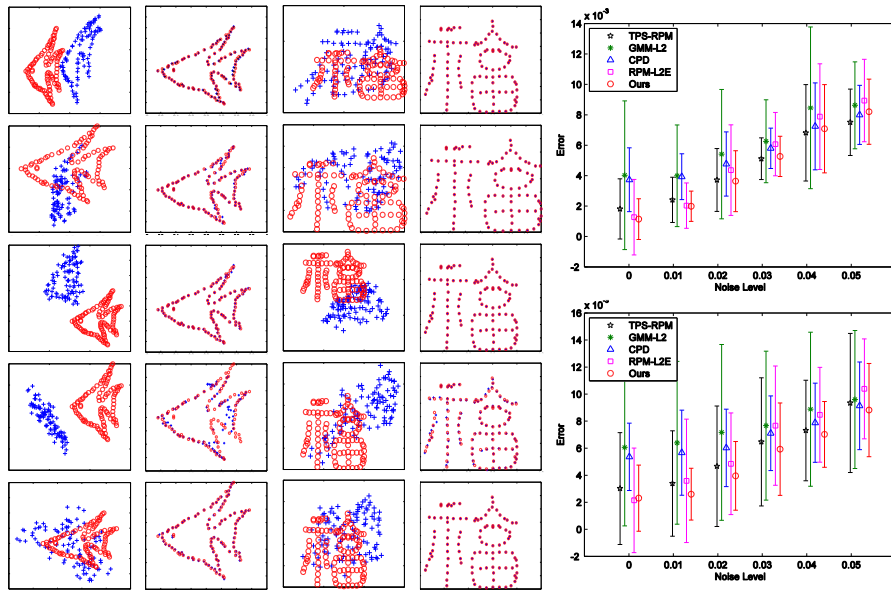


Fig. 2. Experiments on noise. The left figure of each group denotes initial point sets (the *Model* set: blue crosses, the *Scene* set: red circles), and the right figure denotes the registration result of our method. Note that increasing degrees of degradation from top to bottom. The rightmost figure (top: *fish*, bottom: *Chinese character*) is a performance comparison of our results (red circle) with the TPS-RPM (black pentagram), GMM- L_2 (green star), CPD (blue triangle) and RPM- L_2E (magenta square) methods. The error bars indicate the registration error means and standard deviations over 100 random trials.

Finally, we use deterministic annealing introduced by [4, 7] which is a useful heuristic method to escape from the trap of local minima. The initialization value of σ^2 , r and \hat{H} are 0.05, 9 and 0, respectively, and we set $\alpha = 0.95$, $\beta = 0.8$, $\lambda = 0.1$, and $\tau = 15$ for our method throughout this paper. Note that the termination condition of iteration is $\sigma^2 < 0.005$.

4 Experiments

In order to evaluate the performance of our method, we implemented it in Matlab and tested it on a laptop with Pentium CPU 2.4GHz and 4GB RAM. In this section, we first present the results of point sets qualitatively evaluate our method. Then we quantitatively evaluate our method via registration error which is the average Euclidean distance between the *Model* set and the *Scene* set, and compare our results with several comparative methods: TPS-RPM [4], GMM- L_2 [7], CPD [13, 14] and RPM- L_2E [10], which are implemented using publicly available codes.

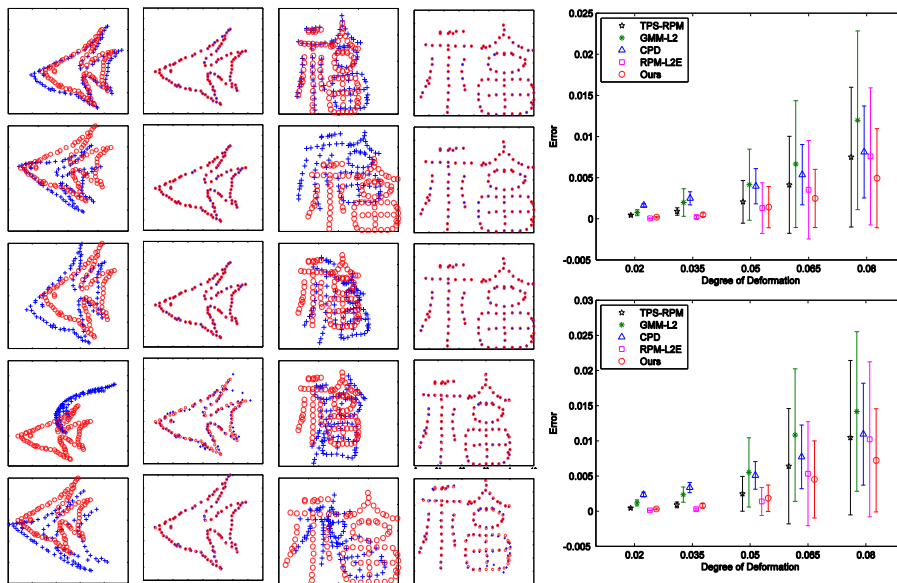


Fig. 3. Experiments on deformation. The left figure of each group denotes initial point sets (the *Model* set: blue crosses, the *Scene* set: red circles), and the right figure denotes the registration result of our method. Note that increasing degrees of degradation from top to bottom. The rightmost figure (top: *fish*, bottom: *Chinese character*) is a performance comparison of our results (red circle) with the TPS-RPM (black pentagram), GMM- L_2 (green star), CPD (blue triangle) and RPM- L_2E (magenta square) methods. The error bars indicate the registration error means and standard deviations over 100 random trials.

4.1 Synthetic Data

We have tested our method on the same data as in [4, 21, 10] named Chui-Rangarajan synthesized data sets. We chose four sets of the aforementioned data and designed them to evaluate the robustness of a method under 4 degradations: noise, deformation, occlusion and outliers. *Fish* and *Chinese character* shapes of data are used for point set registration, and there are 100 samples in each degradation level.

Noise: The noise, due to the processes of image acquisition and feature extraction are not accurate completely, arising from these processes and leading to the resulting feature points cannot be exactly matched [4]. The second and fourth columns of Fig. 2 show results of two shapes under the noise where its level from 0.01 to 0.05. Observing that when the noise level is large, two point sets are not aligned together perfectly. The rightmost column of Fig. 2 denotes the registration error of ours and the comparative methods on two shapes respectively. The registration results show that errors are becoming larger gradually as increasing

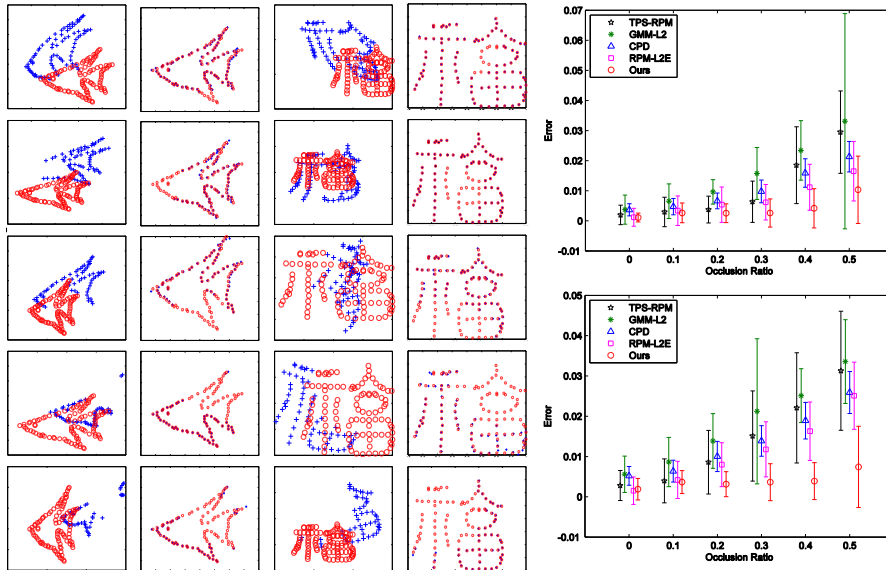


Fig. 4. Experiments on occlusion. The left figure of each group denotes initial point sets (the *Model* set: blue crosses, the *Scene* set: red circles), and the right figure denotes the registration result of our method. Note that increasing degrees of degradation from top to bottom. The rightmost figure (top: *fish*, bottom: *Chinese character*) is a performance comparison of our results (red circle) with the TPS-RPM (black pentagram), GMM- L_2 (green star), CPD (blue triangle) and RPM- L_2E (magenta square) methods. The error bars indicate the registration error means and standard deviations over 100 random trials.

the noise level. Our method and the other four methods have nearly means of registration errors, but our method is slightly more robust than the others by comparing their standard deviations.

Deformation: Nonrigid transformation is quite difficult than rigid. In our method, we use the Gaussian radial basis function (GRBF) to model the transformation. Observing that the registration results of our method are quite well, as shown in the top three rows of Fig. 3, but the fourth and fifth rows show some points drifted because of the large degree of deformation of the given *Model* set. Comparing our results with the other methods, as shown in the rightmost column of Fig. 3, means of errors of our method are less than TPS-RPM’s, GMM- L_2 ’s and CPD’s clearly. When the degree of deformation is large, such as 0.065 and 0.08, our method outperforms RPM- L_2E .

Occlusion: Occlusion *a.k.a.*, missing points, some point features have no corresponding points in the other point set. In this paper, we follow the idea of TPS-RPM that the missing points are treated as outliers, where the outliers

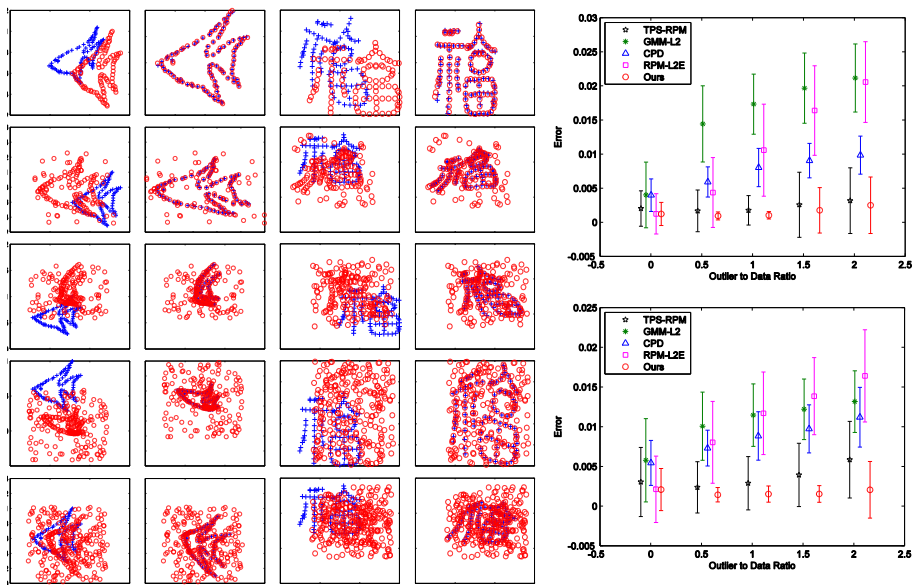


Fig. 5. Experiments on outliers. The left figure of each group denotes initial point sets (the *Model* set: blue crosses, the *Scene* set: red circles), and the right figure denotes the registration result of our method. Note that increasing degrees of degradation from top to bottom. The rightmost figure (top: *fish*, bottom: *Chinese character*) is a performance comparison of our results (red circle) with the TPS-RPM (black pentagram), GMM- L_2 (green star), CPD (blue triangle) and RPM- L_2E (magenta square) methods. The error bars indicate the registration error means and standard deviations over 100 random trials.

satisfy the normal distribution. Observing that input *Model* point set not only contains missing points, but also is deformed in several degrees, as shown in the first and third columns of Fig. 4. Results on the *fish* data, as shown in the second column of Fig. 4, show that almost extra points are aligned correctly. But on the *Chinese character* data, as shown in the fourth column of Fig. 4, the results are not aligned very well, because points on the *Chinese character* shape are not clustered. The rightmost column of Fig. 4 shows the comparison of our results with the comparative methods. The difference between our method and other methods becomes larger as increasing the occlusion ratio. Most importantly, the errors of our method increase much slower than the others.

Outliers: The existence of outliers means many points in one point set that have no corresponding points (homologies) in the other and affects matching results significantly [4]. In this paper, five outlier to data ratios are used: 0, 0.5, 1.0, 1.5, 2.0, and the corresponding results are shown in the second and fourth columns of Fig. 5 respectively. Excitingly, correspondence points are aligned perfectly using our method even in the largest outlier ratio. The comparison of our results

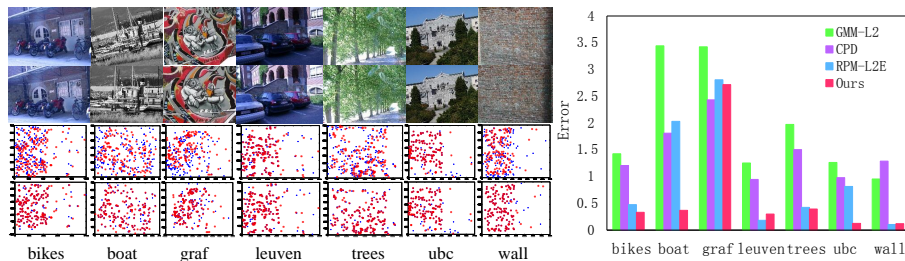


Fig. 6. Experiments on 2D real images. The third figure of each group denotes initial point sets (the *Model* set: blue crosses, the *Scene* set: red circles), and the last figure denotes the registration result of our method. The rightmost figure is a performance comparison of our results with GMM- L_2 , CPD and RPM- L_2E methods. The height of each bar indicates the registration error means over 20 random trials.

with other methods, as shown in the rightmost column of Fig. 5, shows that our method outperforms GMM- L_2 , CPD and RPM- L_2E significantly and is more robust than TPS-RPM.

4.2 Real Image Data

In this experiment, we select 7 image pairs from the Oxford affine covariant regions data set ¹, as shown in the foremost two rows in Fig. 6. There are five different imaging conditions: image blur (bikes and trees), scale changes (boat), multi-viewpoint (graf and wall), JPEG compression (ubc), and illumination (leuven). We extract the SIFT [9] features of those image pairs, and construct initial correspondences using BBF (Best Bin First) method [9], the sparse initialization point sets as shown in the third row of Fig. 6. Note that the point sets do not satisfy arbitrary shapes. The TPS-RPM failed to match and register any point sets, so we did not draw its result in the rightmost figure of Fig. 6.

We repeated the process 20 times, and obtained standard deviations of the methods. GMM- L_2 and CPD are stable without deviations. The standard deviations of RPM- L_2E are 0.21, 0.39, 0.34, 0.03, 0.05, 0.04, 0.02, from left (bikes) to right (wall). However, the standard deviations of our method are almost zeros except the image pairs of graf (std: 0.52). The comparison of our results with other methods shows that our method outperforms GMM- L_2 , CPD and RPM- L_2E significantly in most cases.

5 Conclusion

In this paper, we focus on the case of degradations in point matching problem. Under the previous work of literatures [3–7, 10, 14, 21], we introduce a novel robust method for point matching. There are many registration methods based

¹ <http://www.robots.ox.ac.uk/~vgg/data/>

the Gaussian model, while the asymmetric Gaussian model experimented more accurate than the former in this paper, such as comparing the proposed method with GMM- L_2 and RPM- L_2E which use Gaussian. Moreover, as pointed out in [8], asymmetric Gaussian can capture more accurate spatially asymmetric distributions than Gaussian. In addition, we choose L_2E [17], a robust estimator between two densities, to estimate the similarity between two input point sets. We use low-rank kernel matrix approximation to speed up our method. Note that our method is different with graph matching method [22], because the focus of our method is on the estimation robustness to noise, deformation, occlusion and outliers, while the latter one is mainly used to recover the correspondences accurately. Experimental results on synthetic and real image data illustrate that our proposed method outperforms the comparative methods. Future work includes validating our method on large number of data sets and applying it to many applications, *e.g.*, image retrieval, and 3D image registration.

Acknowledgement. This work was supported by National Natural Science Foundation of China (NSFC, No. 61103070).

References

1. Baldassarre, L., Rosasco, L., Barla, A., Verri, A.: Vector field learning via spectral filtering. *Machine Learning and Knowledge Discovery in Databases*, Springer (2010) 56–71
2. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24** (2002) 509–522
3. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14** (1992) 239–256
4. Chui, H., Rangarajan, A.: A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding* **89** (2003) 114–141
5. Granger, S., Pennec, X.: Multi-scale EM-ICP: A fast and robust approach for surface registration. *Computer Vision–ECCV 2002*, Springer (2002) 418–432
6. Huang, X., Paragios, N., Metaxas, D.N.: Shape registration in implicit spaces using information theory and free form deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28** (2006) 1303–1318
7. Jian, B., Vemuri, B.C.: Robust point set registration using gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33** (2011) 1633–1645
8. Kato, T., Omachi, S., Aso, H.: Asymmetric gaussian and its application to pattern recognition. *Structural, Syntactic, and Statistical Pattern Recognition*, Springer (2002) 405–413
9. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* **60** (2004) 91–110
10. Ma, J., Zhao, J., Tian, J., Tu, Z., Yuille, A.L.: Robust estimation of nonrigid transformation for point set registration. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE (2013) 2147–2154
11. Markovsky, I., USEVICH, K.: *Low Rank Approximation*. Springer (2012)

12. Micchelli, C.A., Pontil, M.: On learning vector-valued functions. *Neural Computation* **17** (2005) 177–204
13. Myronenko, A., Song, X., Carreira-Perpinn, M.A.: Non-rigid point set registration: Coherent point drift. *Advances in Neural Information Processing Systems* **19** (2007) 1009–1016
14. Myronenko, A., Song, X.: Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32** (2010) 2262–2275
15. Nocedal, J., Wright, S.J.: *Conjugate gradient methods*. Springer (2006)
16. Chui, H., Rangarajan, A.: Regularized least-squares classification. *Nato Science Series Sub Series III Computer and Systems Sciences* **190** (2003) 131–154
17. Scott, D.W.: Parametric statistical modeling by minimum integrated square error. *Technometrics* **43** (2001) 274–285
18. Tikhonov, A.N., Arsenin, V.I., John, F.: *Solutions of ill-posed problems*. Winston Washington, DC (1977)
19. Tschumperle, D., Deriche, R.: Vector-valued image regularization with PDEs: A common framework for different applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27** (2005) 506–517
20. Tsin, Y., Kanade, T. A correlation-based approach to robust point set registration. *Computer Vision—ECCV 2004*, Springer (2004) 558–569
21. Zheng, Y., Doermann, D.: Robust point matching for nonrigid shapes by preserving local neighborhood structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28** (2006) 643–649
22. Scott, D.W.: Deformable graph matching. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE* (2013) 2922–2929