Proceedings of the 2016 IEEE
International Conference on Robotics and Biomimetics
Qingdao, China, December 3-7, 2016

Non-Rigid Point Set Registration via Mixture of Asymmetric Gaussians with Integrated Local Structures*

Mingliang Fu** and Weijia Zhou

Abstract—**As a fundamental problem in computer vision community, non-rigid point set registration is a challenging topic since the corresponding transformation model is often unknown and difficult to model. In this paper, we present a robust method for non-rigid point set registration. Firstly, a mixture of asymmetric Gaussian model (MoAG) is employed to capture spatially asymmetric distributions which the Gaussian mixture model (GMM) based methods neglect instinctively. Secondly, local structures among adjacent points are integrated into the MoAG-based point set registration framework to improve the correspondence estimation. Thirdly, Expectation-Maximization (EM) algorithm which provides a numerical method for finding maximum likelihood estimators is utilized to estimate parameters of the latent variable model in our proposed method. The transformation model is solved under regularization theory in the Reproducing Kernel Hilbert Space (RKHS). Conveniently, a fast implementation is achieved by the sparse approximation. Finally, experiments results show that the robustness of our algorithm outperforms the comparative state of art methods under various types of distortions, such as deformation, noise, outliers and occlusion.**

I. Introduction

Registration of point cloud data plays a key role in numerous computer vision applications, including hand-written character recognition [1], facial expression recognition [2] and medical image registration [3]. In general, the process of registration of rigid point cloud data can be described like this: the model point set is performed with transformations include such action as translation, rotation, scaling, or any combination of these to align with the scene point set. The transformation process can be modeling directly. Non-rigid transformation allows anisotropic scaling and skews [3], which makes it difficult to give the corresponding mathematical model.

In this paper, the degradation case, such as deformation, noise, occlusion and outlier is our major focus. The non-rigid transformation is very difficult to estimate since the real transformation is usually unknown and difficult to approximate. The solution of non-rigid point set registration generally consists of two main steps: correspondence estimation and transformation updating.

To achieve approximations of the true non-rigid transformation, many probabilistic models like Gaussian mixture model (GMM) based methods [3, 4, 5, 6, 7] are proposed for the non-rigid point set registration. Gaussian mixture model is a simple liner superposition of Gaussian components, aimed at providing a richer class of density models than the single Gaussian [8]. GMM-based methods makes the reasonable assumption that points form the model point set are normally distributed around points in the scene point set. However, they do not always fit any distribution of data. A new probability model asymmetric Gaussian (AG) [9] is originally proposed to capture spatially asymmetric distribution. Recently, method [10] utilizes MoAG to represent point sets in their framework. However, the correspondence between model and scene point sets is estimated only with global relationship in these probabilistic methods. The local neighborhood information is ignored, which is quite important for establishing a comprehensive correspondence [7, 11, 12].

In this paper, a non-rigid point set registration algorithm based on MoAG is presented. Different from previous MoAG based method [10], we conduct the correspondence estimation with integrated local structures. Besides, the optimal model parameters are updated by the well-known EM algorithm iteratively. To reduce the high computational complexity, we follow the rule used in [13]. The transformation is estimated in reproducing kernel Hilbert space (RKHS) with regularization theory. Thus, a fast implement can be achieved with the sparse approximation. Experimental results demonstrate that the proposed method outperforms several state of art methods in most tested scenarios.

The rest of the paper is organized as follows: In Section 2, the mixture of AG model is described. An outlier term is integrated into MoAG model to approximate the real data model as much as possible. The experiments and performance evaluation are given in Section 4. Our conclusion is given in Section 5.

II. Method

Existing GMM-based approaches are based on a hypothesis that the model point set are normally distributed relative to points belonging to the scene point set. In general, this hypothesis is reasonable. However, the GMM cannot represent the asymmetric distribution properly. Thus, a mixture of AG model is chosen instead of mixture Gaussian model to represent distribution of the model and scene point set in the proposed method. The MoAG can capture spatially

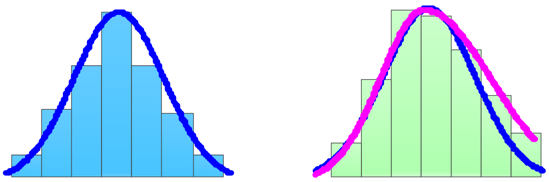
* This research is supported by Grant Y4K3170301, provided by Department of Space Automation Technologies and Systems (No. 3 Department at SIA), Chinese Academy of Sciences.

Mingliang Fu is a Ph.D. student in University of Chinese Academy of Sciences, Beijing, 100049, China and works in Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, Liaoning province, 110016, China. E-mail: fumingliang@sia.cn.

Weijia Zhou is a professor and doctoral supervisor in Shenyang Institute of Automation, Shenyang, Liaoning province, 110016, China. E-mail: zwj@sia.cn.

**Corresponding author: fumingliang@sia.cn (Mingliang Fu).

asymmetric distribution [9], which approaches the real data distribution. Fig. 1 gives a clear delineation about the fact.



(a) (b)

Figure 1. Fitting of univariate Gaussian (blue curve) and AG (violet curve) to data. (a) The histogram represents data obeys a normal distribution. (b) The histogram represents data obey an asymmetric distribution.

A. The mixture of AG model

In this section, we introduce a view of MoAG. Among lots of methods, GMM has received great success in solving the issue of correspondence estimation. Points in model set and scene set are treated as the centroid of the GMM. Unfortunately, GMM do not always fit any distribution of data, so MoAG is chosen instead of mixture Gaussian model. AG model extends from Gaussian model substantially. Let $X=(x_1,\cdots,x_N)\in\mathbb{R}^{D\times N}$ and $Y=(y_1,\cdots,y_N)\in\mathbb{R}^{D\times M}$ be model points and scene points, respectively. In the similar manner as Gaussian, the distribution of dimensional AG can be expressed as [10]

$$AG\left(x\middle|\mu,\sigma^2,r\right)=\frac{1}{\left(2\pi\sigma^2\right)^{D/2}\left(\left(r+1\right)/2\right)^D}\left[\gamma\exp\left(-\frac{|x-\mu|^2}{2\sigma^2}\right)+\left(1-\gamma\right)\exp\left(-\frac{|x-\mu|^2}{2r^2\sigma^2}\right)\right], \quad (1)$$

where μ,σ^2 and r are parameters of $AG\left(x\middle|\mu,\sigma^2,r\right)$. $\gamma=\begin{cases}1, & x\le\mu\\ 0, & x>\mu\end{cases}$ represents the weighting between two Gaussian components in single AG model. AG degenerates to a standard Gaussian model when $r=1$.

Recall from the definition of the density model, it's easy to construct MoAG. Similar to the rule as [7], the membership probability of the MoAG π_{mn} satisfying $0\le\pi_{mn}\le 1$ and $\sum_{n=1}^{N}\pi_{mn}=1$. If model point sets contain N elements, MoAG can be expressed as a linear superstition of AGs in the form

$$P\left(x\middle|\theta\right)=\sum_{n=1}^{N}\pi_{mn}AG\left(x\middle|\mu,\sigma^2,r\right). \quad (2)$$

To account for noise and outliers in point sets, an additional uniform distribution $p\left(outlier\right)$ is added to MoAG. Denoting the weight of the MoAG and the uniform distribution as ω and $1-\omega\left(\omega\in[0,1]\right)$, respectively. The overall data distribution is

$$P_{AG}\left(x\middle|\theta\right)=\left(1-\omega\right)\sum_{n=1}^{N}\pi_{mn}AG\left(x\middle|\mu,\sigma^2,r\right)+\omega p\left(outlier\right). \quad (3)$$

Based on the assumption that the structural similarity between two pixels depends on the similarity between the geometrical configurations in their whole neighbourhoods, local structures are utilized to improve the correspondence estimation. Moreover, we add a membership probability to MoAG

$$P_{AG}\left(y_m\middle|\theta\right)=\omega p\left(outlier\right)+\left(1-\omega\right)\sum_{n=1}^{N}\frac{\pi_{mn}}{\left(2\pi\sigma^2\right)^{D/2}\left(\left(r+1\right)/2\right)^D}\left[\gamma\exp\left(-\frac{\left|y_m-\Gamma\left(x_n\right)\right|^2}{2\sigma^2}\right)+\left(1-\gamma\right)\exp\left(-\frac{\left|y_m-\Gamma\left(x_n\right)\right|^2}{2r^2\sigma^2}\right)\right]. \quad (4)$$

Shape context [1] and the fast point feature histograms (FPFH) [14] are utilized to measure the similarity in 2D and 3D case, respectively. The basic idea of the non-uniform weight is to encode the feature similarity into the mixture coefficients. It's reasonable to assign a higher weight to AG component so that a pair of points with similar features has more chance to form a correspondence. An adaptive weight factor is designed to integrate the local structures to the solution process. After the initial correspondences are obtain, we use them to specialize the membership probability as follows

$$\pi_{mn}=\begin{cases}\dfrac{vol}{|\Phi|}, & if\ x_n\in\Phi\\ \dfrac{1-vol}{N-|\Phi|}, & if\ x_n\notin\Phi\end{cases} \quad (5)$$

where Φ denotes the set which a data point y_m corresponds to, *vol* represents the confidence of a correspondence calculated with shape context. $|\cdot|$ describes the number of elements in a set. Note that if a data point do not match a corresponding point, π_{mn} is assigned to a constant *1/N.*

B. EM Algorithm for model parameters updating

For GMM-based methods, L2 distance [6] and L2E [13] have been proven the popular approaches to perform similarity measures between Gaussian mixtures. However, the correspondences and the transformation cannot be solved simultaneously in these methods. In this paper, we choose EM [15] algorithm to solve the optimal transformation iteratively. Based on the Bayes' theorem, the MAP solution can be written as

$$\theta^*=\arg\max_{\theta}P\left(\theta\middle|Y\right)=\arg\max_{\theta}P\left(Y\middle|\theta\right)P\left(\Gamma\right). \quad (6)$$

Equivalently, the optimal transformation can be determined by the negative log-likelihood function

$$L\left(\theta\middle|Y\right)=-\sum_{m=1}^{M}\ln P_{AG}\left(y_m\middle|\theta\right)-\ln P\left(\Gamma\right). \quad (7)$$

According to the Jensen's inequality, (7) can be re-expressed as

$$L(\theta|Y) = -\sum_{m=1}^{M}\ln P_{AG}(y_m|\theta) - \ln P(\Gamma)$$
$$\geq -\sum_{m=1}^{M}\sum_{n=1}^{N} P(z_m = n|y_m,\theta^{old})\left[\ln\left[P_{AG}(y_m|z_m = n,\theta)P(z_m = n|\theta)\right]\right.$$
$$\left. -\ln P(z_m = n|y_m,\theta^{old})\right), \tag{8}$$

where z_m denotes the latent variable and θ^{old} represent the current parameter values.

The Q function, which we call the objective function, is also an upper bound of the negative log-likelihood function (7). To simplify things, terms not including θ in (8) are omitted. The Q function can be expressed as

$$Q(\theta,\theta^{old}) = \sum_{m,n=1}^{M,N}\left(\frac{\gamma}{2\sigma^2} + \frac{1-\gamma}{2r^2\sigma^2}\right)P(z_m = n|y_m,\theta^{old})\|y_m - \Gamma(x_n)\|^2$$
$$+\frac{M_p D}{2}\ln\sigma^2 + M_p D\ln(r+1)$$
$$+M_p\ln(1-\omega) - (M - M_p)\ln\omega + \frac{\lambda}{2}\phi(\Gamma). \tag{9}$$

Where $M_p = \sum_{m,n=1}^{M,N} P(z_m = n|y_m,\theta^{old}) \leq M$ (with $M_p = M$ only if $\omega = 0$).

Next, the EM algorithm is employed to update model parameters. After some initial values for the GMM are chosen, we alternate between the following updates that we shall call E-step and M-step. In E-step, the current values for model parameters are utilized to evaluate the responsibilities. In the subsequent M-step, the expectation is maximized. A detailed course of these two steps is given as follows.

E-Step: Evaluate the responsibilities using the current parameter values. The posterior probability of MoAG can be written as:

$$P(z_m = n|y_m,\theta^{old}) = \frac{P(y_m|z_m = n,\theta^{old})P(z_m = n|\theta^{old})}{P(y_m|\theta^{old})}$$
$$= \frac{\pi_{mn}\left[\gamma\exp\left(-\frac{|y_m - \Gamma(x_n)|^2}{2\sigma^2}\right) + (1-\gamma)\exp\left(-\frac{|y_m - \Gamma(x_n)|^2}{2r^2\sigma^2}\right)\right]}{\sum_{k=1}^{N}\pi_{mk}\left[\gamma\exp\left(-\frac{|y_m - \Gamma(x_n)|^2}{2\sigma^2}\right) + (1-\gamma)\exp\left(-\frac{|y_m - \Gamma(x_n)|^2}{2r^2\sigma^2}\right)\right] + \frac{\omega(2\pi\sigma^2)^{D/2}((r+1)/2)^D}{(1-\omega)a}}. \tag{10}$$

M-Step: Re-estimate the parameters using the current responsibilities.

To maximize (10), take $\frac{\partial Q}{\partial\sigma^2} = 0$, and we obtain

$$\sigma^2 = \frac{1}{M_p D}\sum_{m,n=1}^{M,N}\left(\gamma + \frac{1-\gamma}{r^2}\right)P(z_m = n|y_m,\theta^{old})\|y_m - \Gamma(x_n)\|^2. \tag{11}$$

Then, take $\frac{\partial Q}{\partial\omega} = 0$ and ω have the form as follow:

$$\omega = \frac{M - M_p}{M}. \tag{12}$$

Similarity, take $\frac{\partial Q}{\partial r} = 0$ and we can get r from

$$\frac{1}{r^3} + \frac{1}{r^2} = \frac{\sigma^2 M_p D}{\sum_{m,n=1}^{M,N}(1-\gamma)P(z_m = n|y_m,\theta^{old})\|y_m - \Gamma(x_n)\|^2}. \tag{13}$$

Note that r cannot be solved directly, and the *solve* function in Matlab is employed to solve the issue. Equation (13) has three solutions, two complex solutions and one real solution. Here the real solution is selected for the final result.

The pseudo code of the proposed method is shown in Algorithm 1.

Algorithm 1 Summary of our proposed algorithm

Input: Two point sets X and Y, parameters γ, β, λ
Output: Optimal transformation
1 Construct the Gram matrix Γ;
2 Initialization $C = 0, \sigma^2 = \frac{1}{DMN}\sum_{m,n=1}^{M,N}\|y_m - x_n\|^2$
3 Compute feature descriptors for point set Y;
4 While not converge do
E-step:
Compute the feature descriptors for point set $\Gamma(X)$
Establish correspondence between $\Gamma(X)$ and Y;
Initialize π_{mn} based on the feature correspondence;
Update posterior probability P by (10);
M-step:
Update σ^2 by (11).
Update ω by (12).
Update r by (13).
Until a termination condition is reached;
5 End while
End

The solving process of non-rigid transformation is similar with [3]. In every iteration step, the transformed position of point set Y becomes

$$\Gamma(Y,v)=Y+v(Y), \tag{14}$$

where v denotes a displacement function and has the form $v(x)=\sum_{n=1}^{N}\Gamma(x,x_n)c_n$. c_n represents the coefficient set.

Note that the regularization term comes from a displacement field:

$$P(v)\propto \exp\left(-\frac{\lambda}{2}\phi(v)\right). \tag{15}$$

Adding a regularization term to (11) and omitting terms are independent of transformation Γ , we obtain

$$\ell(C)=\sum_{m=1}^{M}\left(\frac{1-\gamma}{2r^2\sigma^2}+\frac{\gamma}{2\sigma^2}\right)\left\|\left(Y_m^N-X-C\Gamma\right)d\left(P_m\right)^{1/2}\right\|_F^2$$

$$+\frac{\lambda}{2}C\Gamma C^T. \tag{16}$$

Let $\frac{\partial \ell}{\partial \Gamma}=0$ and C can be determined by the following linear equation

$$C\left(\Gamma d\left(1^T P\right)+\lambda\frac{I}{k}\right)=YP-Xd\left(1^T P\right), \tag{17}$$

where $k=\left(\frac{1-\gamma}{r^2\sigma^2}+\frac{\gamma}{\sigma^2}\right)$, $\Gamma\in\mathbb{R}^{N\times N}$ is an RKHS with kernel element $\Gamma_{ij}=\exp\left(-\frac{\left\|x_i-x_j\right\|^2}{2\beta^2}\right)$.

By (16) and (17), the displacement function in (14) is updated in every iteration step.

In general, the data volume of point cloud captured by sensors is huge. Quick and effective registration algorithm is of great significance in real applications. Existing fast implement technology mainly include fast gauss transform [16], low-rank matrix approximation [3, 7, 17, 18], mixture decoupling [19]. In this paper, sparse approximation [13] is utilized to accelerate the proposed algorithm. The subset we pick contains 30 points in our experiments.

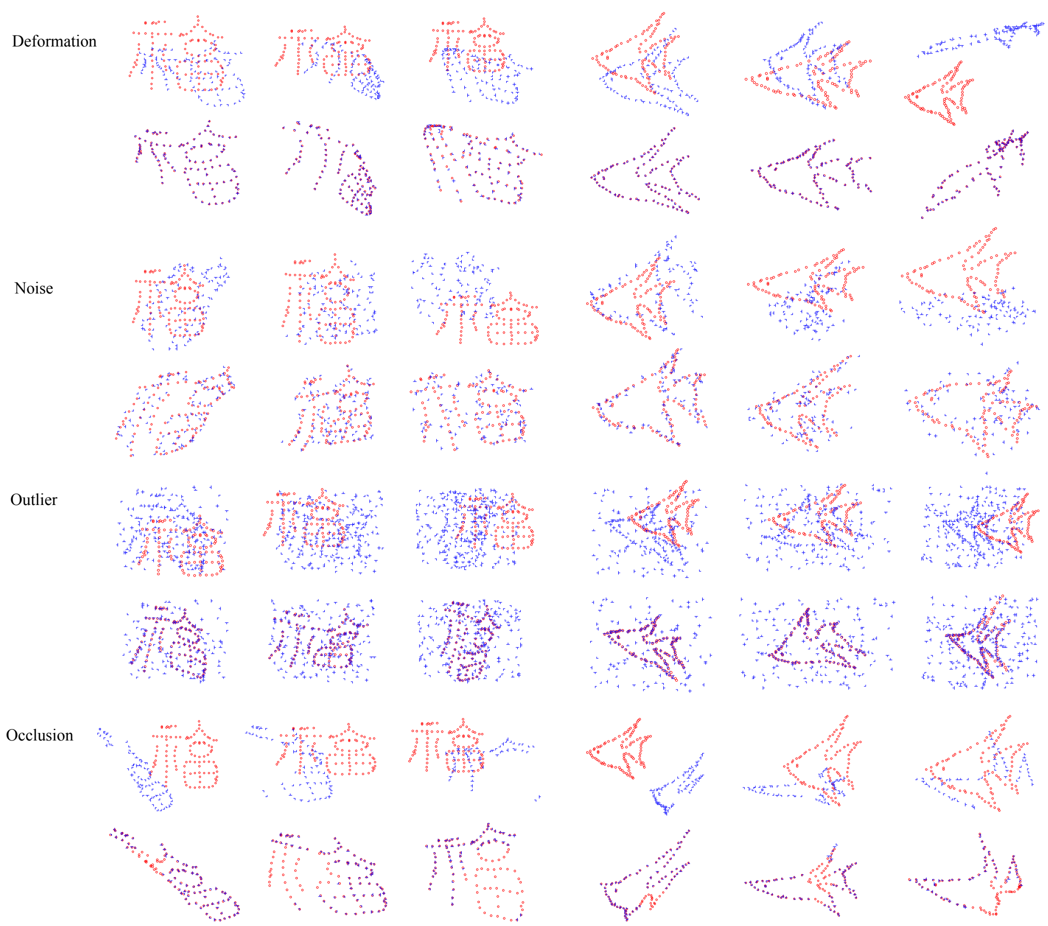


Figure 2. Registration results with our method on the Chinese character (left) and fish (right) patterns. From left to right are degradations of two types of in three different levels. The mission is to align the model point set (red circles) onto the scene point set (blue pluses).

III. Experimental Results

In order to evaluate the performance of the proposed algorithm, we conducted different degradation types of experiments on synthetic data, e.g., deformation, noise, occlusion and outlier. The test data sets are available at http://www.csee.ogi.edu/~myron. All the experiments were performed using Matlab R2014a on a PC with an Intel Core i3 3.0 GHz processor and 6.0 Gbytes of physical memory. We compare the proposed method on the above data with other two state of art algorithms namely TPS-RPM [20] and CPD [3].

The test data consists of two models with different shapes, where the first model consists of 98 points representing a fish and the second model consists of 105 points, representing a Chinese character. For each model, there are five sets of data designed to measure the robustness of registration algorithms with respect to different degrees of deformation, noise, outliers and occlusion. More experimental details can be found in [20]. In our proposed algorithm, there are mainly three parameters. We set $\gamma = 0.1$, $\beta = 2$ and $\lambda = 3$ throughout the paper. Part of the registration results with the proposed method are shown in Fig. 2.

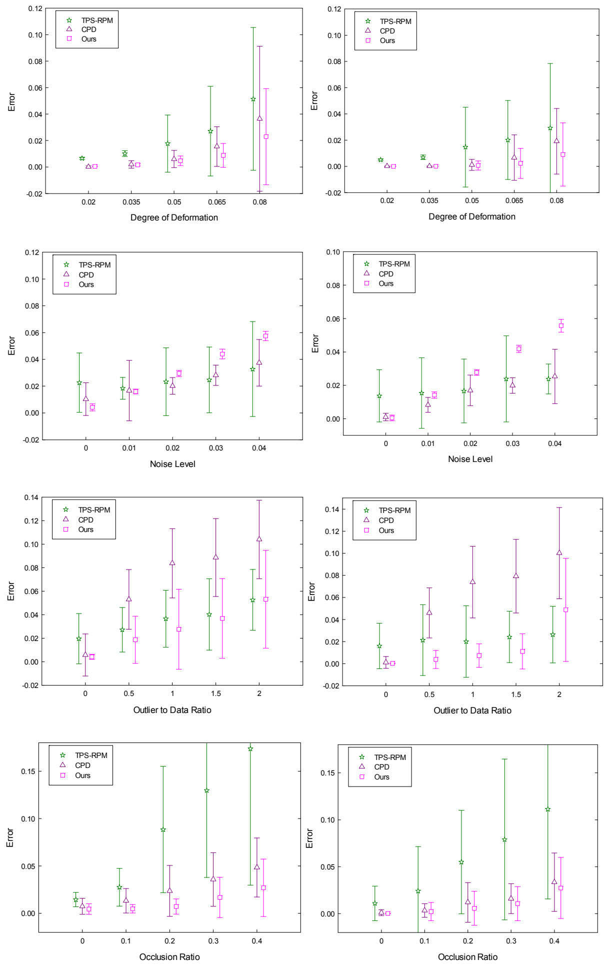


Figure 3. Quantification of the proposed algorithm compared toTPS-RPM [20] and CPD [3] on the corresponding datasets. Right column: the shape of fish. Left column: the shape of a Chinese character. Each error bar indicates the registration error mean and the standard deviation over 100 runs.

The registration error of the model and scene point set is quantified as the average Euclidean distance based on the ground truth correspondences. In the deformation test, the number of the model point set and the scene point set are the same. The average Euclidean distance is calculated with all the point pairs. For the *fish* patterns corrupted with outliers, the number of scene point set N is fixed at 98, M is varied from 98 to 294. The error is computed using the point pairs with the ground truth correspondences. The performance statistics for different degradation types are summarized in Fig. 3. All three methods are compared by the error mean and the standard deviation of the registration error of 100 trials. The proposed method is more robust compared with the other two algorithms in most scenes.

IV. Conclusion

GMM-based methods for non-rigid point set registration can achieve a good result. However, these algorithms neglect the asymmetric distribution among data inevitably. To solve this issue, MoAG is utilized to capture spatially asymmetric distribution in our method. Meanwhile, we employ the EM algorithm to estimate the correspondence and transformation. The experiment results show that our method is robust to different degradations and outperforms other two state of the art registration algorithms.

Acknowledgment

Sincerely appreciate that Andriy Myronenko and Haili Chui for providing their implementation source codes and test datasets.

References

[1] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509-522, Apr 2002.

[2] K.W. Bowyer, K. Chang, and P. Flynn, "*A Survey of Approaches and Challenges in 3D and Multi-Modal 3D + 2D Face Recognition*", Computer Vision and Image Understanding, vol. 101, pp.1-15, 2006.

[3] A. Myronenko and X. Song, "Point Set Registration: Coherent Point Drift," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2262-2275, Dec. 2010.

[4] V. Panaganti and R. Aravind, "Robust Nonrigid Point Set Registration Using Graph-Laplacian Regularization," *2015 IEEE Winter Conference on Applications of Computer Vision*, Waikoloa, HI, 2015, pp. 1137-1144.

[5] P. Wang, P. Wang, Z. G. Qu, et al, "A refined coherent point drift (CPD) algorithm for point set registration," *Science China Information Sciences*, vol. 54, pp. 2639-2646, Dec. 2011.

[6] B. Jian and B. C. Vemuri, "Robust Point Set Registration Using Gaussian Mixture Models," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1633-1645, Aug. 2011.

[7] J. Ma, J. Zhao, and A. L. Yuille, "Non-Rigid Point Set Registration by Preserving Global and Local Structures," in *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 53-64, Jan. 2016.

[8] C.M. Bishop, *Pattern Recognition and Machine Learning.* New York, CA: Springer, 2007, pp. 423-443.

[9] T. Kato, S. Omachi, and H. Aso, "Asymmetric Gaussian and its application to pattern recognition," Proc. Joint IAPR Int. Workshops SSPR 2002 and SPR 2002, pp.404-413.

[10] G. Wang, Z. Wang, W. Zhao and Q. Zhou, "Robust point matching using mixture of asymmetric gaussians for nonrigid transformation," *Computer Vision-ACCV 2014,* pp. 433-444, 2015, Springer.

[11] Y. Yang, S. H. Ong, and K. W. C. Foong, "A robust global and local mixture distance based non-rigid point set registration," Pattern Recognition, vol. 48, no. 1, pp. 156-173, 2015.

[12] Y. Zheng and D. Doermann, “Robust point matching for nonrigid shapes by preserving local neighborhood structures,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 643-649, April 2006.

[13] J. Ma, J. Zhao, J. Tian, Z. Tu, and A. L. Yuille, “Robust Estimation of Nonrigid Transformation for Point Set Registration,” *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, Portland, OR, 2013, pp. 2147-2154.

[14] R. B. Rusu, N. Blodow, and M. Beetz, “Fast Point Feature Histograms (FPFH) for 3D registration,” *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, Kobe, 2009, pp. 3212-3217.

[15] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *J. R. Stat. Soc.,* vol. 39, no. 1, pp. 1-38, 1977.

[16] L. Greengard and J. Strain, “The Fast Gauss Transform,” *SIAM J. Scientific and Statistical Computing*, vol. 12, no.1, pp.79-94, 1991.

[17] R. Rifkin, G. Yeo, and T. Poggio, “Regularized least-squares classification,” *Nato Science Series Sub Series III Computer and Systems Sciences*, vol. 190, pp. 131-154, 2003.

[18] J. Ma, J. Zhao, J. Tian, X. Bai, and Z. Tu, “Regularized vector field learning with spare approximation for mismatch removal,” *Pattern Recognit.*, vol. 46, no.12, pp.3519-3532, 2013.

[19] B. Eckart, K. Kim, A. Troccoli, A. Kelly and J. Kautz, “MLMD: Maximum Likelihood Mixture Decoupling for Fast and Accurate Point Cloud Registration,” *3D Vision (3DV), 2015 International Conference on*, Lyon, 2015, pp. 241-249.

[20] H. Chui, A. Rangarajan, “A New Point Matching Algorithm for Non-Rigid Registration,” *Computer Vision and Image Understanding*, pp.114-141, Feb., 2003.