# Registration of Multi-View Point Sets Under the Perspective of Expectation-Maximization

Jihua Zhu, *Member, IEEE*, Rui Guo, Zhongyu Li, Jing Zhang, and Shanmin Pang, *Member, IEEE*

***Abstract*— Multi-view registration plays a critical role in 3D model reconstruction. To solve this problem, most previous methods align point sets by either partially exploring available information or blindly utilizing unnecessary information, which may lead to undesired results or extra computation complexity. Accordingly, we propose a novel solution for the multi-view registration under the perspective of Expectation-Maximization (EM). The proposed method assumes that each data point is generated from one unique Gaussian Mixture Model (GMM), where its corresponding points in other point sets are regarded as Gaussian centroids with equal covariance and membership probabilities. As it is difficult to obtain real corresponding points in the registration problem, they are approximated by the nearest neighbor in each other aligned point sets. Based on this assumption, it is reasonable to define the likelihood function including all rigid transformations, which require to be estimated for multi-view registration. Subsequently, the EM algorithm is derived to estimate rigid transformations with one Gaussian covariance by maximizing the likelihood function. Since the GMM component number is automatically determined by the number of point sets, there is no trade-off between registration accuracy and efficiency in the proposed method. Finally, the proposed method is tested on several benchmark data sets and compared with state-of-the-art algorithms. Experimental results demonstrate its superior performance on the accuracy, efficiency, and robustness for multi-view registration.**

***Index Terms*— Gaussian distribution, Gaussian mixture model, expectation-maximization, point set registration.**

## I. Introduction

POINT set registration is a fundamental task in many domains, such as computer vision [1], [2], robotics [3], [4], and computer graphics [5], [6]. Recent frontiers of point scanning devices make it possible to reconstruct accurate 3D objects or scene models. Due to the limited fields of view, most scanning devices can only scan a part of the object or scene from a single viewpoint. For the 3D model reconstruction, multiple point sets should be acquired from different viewpoints to cover the entire object or scene surface, and then unified into the common reference frame by the multi-view registration. Accordingly, the multi-view registration is a prerequisite for 3D model reconstruction. Given multiple point sets, the goal of multi-view registration is to estimate the optimal rigid transformation for each point set and transform them from a set-centered frame into the same coordinate frame.

The point set registration problem has attracted immense attention, and many effective methods have been proposed to solve this problem. Among these methods, one of the most popular solutions is the iterative closest point (ICP) algorithm [7], [8], which can achieve the pair-wise registration with good efficiency and accuracy. However, including the ICP algorithm, most of them cannot be directly employed to solve the multi-view registration problem. Compared with the pair-wise registration, the multi-view registration is comparatively more difficult with less attention. Although some methods have been proposed to tackle this challenging problem, most existing methods are unable to appropriately explore available information for the accurate multi-view registration. For the multi-view registration, some methods only establish point correspondences between one and some other point sets, which cannot fully explore available information for accurate registration. While, other methods may blindly establish point correspondences between one and all other point sets, which may lead to the amount of extra computation complexity.

Recently, the expectation-maximization (EM) algorithm has been introduced as an effective means to solve the multi-view registration problem [9]–[13]. These methods suppose that all data points are generated from a central Gaussian Mixture Model (GMM), which contains many components required to be estimated. As all points are utilized to estimate each Gaussian component and rigid transformation, these methods are time-consuming to achieve multi-view registration. Besides, there is a trade-off between registration accuracy and efficiency. What's more, they are likely trapped into the local minimum due to many parameters required to be estimated.

To this end, we propose a novel method under the EM perspective. As data collected by sensors inevitably contains noise, we assume that each data point is generated from one unique GMM, where its corresponding points in other point sets are regarded as Gaussian centroids with equal covariance and membership probabilities. Since it is difficult to find the real corresponding point in the registration problem, we utilize the nearest neighbor (NN) in each other aligned point set to approximate the corresponding point. Based on this assumption, we formulate the multi-view registration problem

Manuscript received November 29, 2019; revised June 10, 2020 and August 14, 2020; accepted September 6, 2020. Date of publication September 21, 2020; date of current version September 29, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61573273 and in part by the Fundamental Research Funds for Central Universities under Grant xzy012019045. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Joao M. Ascenso. *(Corresponding author: Zhongyu Li.)*

Jihua Zhu, Rui Guo, Zhongyu Li, and Shanmin Pang are with the School of Software Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: zhongyuli@xjtu.edu.cn).

Jing Zhang is with the School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, China.

Digital Object Identifier 10.1109/TIP.2020.3024096

into the framework of maximum likelihood estimation, which can be optimized by the EM algorithm. In our method, the GMM component number is automatically determined by the number of point sets. Therefore, there is no trade-off between registration accuracy and efficiency. Besides, this method defines all Gaussian components by establishing the relationship between one data point and several NNs, so it only requires to estimate rigid transformations and one covariance. Since this method requires to estimate fewer parameters and considers data noise in registration, it is more likely to obtain promising registration results. What's more, each data point is only related to a small number of Gaussian components, so it is efficient to achieve multi-view registration.

The remainder of this paper is organized as follows. Section 2 briefly review related works on point set registration. Section 3 formulates the multi-view registration problem under the EM perspective. Following that is section 4, in which the proposed method is derived to solve the multi-view registration problem. In Section 5, the proposed method is tested and evaluated on benchmark data sets. Then, it is applied to robot mapping and scene reconstruction. Finally, some conclusions are drawn in Section 6.

## II. Related Work

This section only surveys existing works related to the proposed method for multi-view registration. For convenience, we use the terms point set and scan interchangeably throughout this paper.

Due to the number of point sets, the registration problem can be divided into two sub-problems, the pair-wise registration and the multi-view registration. For the pair-wise registration, one of the most popular methods is the ICP algorithm, which can achieve pair-wise registration with good performance. But it cannot deal with non-overlapping point sets. Besides, it belongs to the local convergent algorithms. To improve its performance, many ICP variants have been proposed for pair-wise registration [14]. For non-overlapping point sets, Chetverikov *et al.* [15] proposed the trimmed ICP (TrICP) algorithm, which introduces the overlap percentage to automatically trim non-overlapping regions for accurate registration. To address local convergence, the Genetic algorithm [16] or the particle filter [17] is integrated with the TrICP algorithm to search the desired results. For the efficiency, many point feature methods [2], [18] are proposed to provide good initial parameters for the TrICP algorithm or its variants. What's more, some deep learning based methods have been proposed to estimate the initial transformations for point set registration [19], [20].

Recently, some GMM-based methods, such as CPD [21], GMMReg [22] and FilterReg [23] were also proposed to solve the pair-wise registration problem. Both CPD and FilterReg represent one point set as a GMM, then cast the pair-wise registration problem as a maximum likelihood estimation problem. While, GMMReg utilizes two Gaussian mixture models to represent both point sets and reformulate the pair-wise registration as the problem of aligning two Gaussian mixtures, where a statistical discrepancy measure between two GMMs is minimized by the EM algorithm. Although these methods can achieve pair-wise registration with good accuracy and robustness, they are time-consuming due to the large number of point correspondences required to be established. Besides, they are unable to directly solve the multi-view registration problem.

For the multi-view registration, the intuitive method is the alignment-and-integration method [8], which sequentially aligns and integrates two point sets until all point sets are integrated into one model. This approach is simple but suffers from the error accumulation problem due to many point sets in multi-view registration. Then, Bergevin *et al.* [24] proposed the first solution for multi-view registration. It organizes all point sets by a star-network and sequentially puts one point set in the center of star-network. For the center point set, point correspondences are searched from each other point sets to estimate the rigid transformation by the ICP algorithm. As the ICP algorithm is unable to deal with non-overlapping regions, this approach is difficult to obtain promising results. Considering non-overlapping regions, Williams and Bennamoun [25] proposed a method to calculate one weight for each point correspondence by its residual error, where point pairs in non-overlapping regions are assigned with a small weight. Since this method also requires to build point correspondences between all point set pairs, it is time-consuming too. To this end, Zhu *et al.* [26] proposed a coarse-to-fine method, which sequentially traverses and refines the rigid transformation of each point set. More specifically, this method takes the traversed point set and all other coarse aligned point sets as data set and model set, respectively. Then the TrICP algorithm is utilized to refine the rigid transformation of each traversed point set. As these methods sequentially estimate one rigid transformation, they become more and more easily to be trapped into a local optimum with the increase of point set number.

To avoid sequential estimation, Krishnan *et al.* [27] proposed an optimization-on-a-manifold method for multi-view registration. This method can simultaneously estimate all rigid transformation from established correspondences between all point sets. However, it is difficult to establish accurate point correspondences in practical applications. Accordingly, Mateo *et al.* [10] treat pair-wise correspondences as missing data and proposed the method for multi-view registration under the Bayesian perspective. Although this method may be accurate, it requires to calculate a large number of latent variables. Further, Tang and Feng [28] proposed a hierarchical method for multi-view registration under the graph perspective, where each node and edge denotes a single point set and a connection between two point sets with large overlapping regions, respectively. Based on the graph, hierarchical optimization is implemented on the edges, the loops, and the entire graph. As this method require to detect all small loops, it turns to be difficult or impossible if none loop exists. Before this method, some other graph-based approaches [29], [30] were also proposed for multi-view registration. The difference is that each edge denotes the pair-wise registration of two connected nodes. Then the graph optimization approach is performed to diffuse the registration error over a graph of adjacent point sets. Without the update of point correspondences, these methods

only transfer registration errors among graph nodes and are unable to really reduce the total registration errors.

Meanwhile, some information theoretic measures were utilized to solve the multi-view registration problem. Given multi-view point sets, each of them can be represented by a probability distribution function (PDF). Accordingly, Wang *et al.* [31] proposed multi-view approach called CDF-JS, which minimizes the Jensen-Shannon divergence among CDFs. Although the CDFs are robust to noises, their estimation is time-consuming and their updates have no closed-form solutions. To this end, Chen *et al.* [32] generalizes the Jensen-Shannon divergence and proposed the Havrda-Charvát divergence for CDFs. Compared with CDF-JS, CDF-HC is more efficient and much simpler in implementation for multi-view registration. Further, Giraldo *et al.* [33] proposed multi-view registration method based on Renyi's quadratic entropy. Compared with two previous methods, this method provides a closed-form solution for the updates of CDFs. What's more, it is more efficient than two previous CDF-based methods. However, this method is still time-consuming for multi-view registration.

As multi-view registration problem is more difficult than pair-wise registration problem, Govindu and Pooja [34] proposed the motion averaging algorithm to solve it. This algorithm can directly estimate global motions (multi-view registration) from a set of relative motions (pair-wise registration ) by the motion averaging algorithm. Meanwhile, Arrigoni *et al.* [35], [36] introduced the low-rank and sparse (LRS) matrix decomposition algorithm to recover global motions from a block matrix, which is filled with all available relative motions and zero elements. In comparison, the LRS method is more robust to some unreliable relative motions. However, to obtain desired registration results, it requires more relative motions. What's more, the reliability of each relative motions is different, but these two methods consider that their contributions to the motion averaging are equal, which inevitably decreases the registration performance. To address this issue, Guo *et al.* [37] proposed the weighted motion averaging algorithm to solve the multi-view registration problem. Meanwhile, Jin *et al.* [38] proposed the weighted LRS algorithm for multi-view registration. As these two methods can pay more attention to reliable relative motions, they can achieve more accurate and robust registration results than their original methods.

Recently, Evangelidis and Horaud [11] proposed the JRMPC method, which assumes that all points are the realization of one central GMM and then casts multi-view registration into a clustering problem. Therefore, the EM algorithm is derived to simultaneously achieve clustering and multi-view registration. Actually, JRMPC is similar to the previous work [9] proposed to solve the non-rigid registration problem in medical image processing. Based on JRMPC, Ravikumar *et al.* [12] replaced the Gaussian distribution by the Student's t-distribution and proposed scaling registration method for multi-view point sets. In addition, Min *et al.* [13], Min and Meng [39], and Min [40] extent JRMPC to align high-dimensional points consisting of both positional vectors and orientational information. As these methods require to estimate many components in the mixture model, it is time-consuming and easy to be trapped into local optimum. Accordingly, Zhu *et al.* [41] derived the K-means algorithm to simultaneously achieve clustering and registration. Compared with previous works, this method requires to estimate fewer parameters, so it is more efficient and robust. However, there is a trade-off between registration accuracy and efficiency in these clustering-based methods. What's more, the clustering of points leads to the information loss, which seriously reduces registration accuracy. To address these issues, we propose a novel GMM-based registration method under the EM perspective.

## III. Problem Formulation

Given the model centered frame, the goal of multi-view registration is to estimate the rigid transformation including a rotation matrix $\mathbf{R}_i$ and a translation vector $t_i$ for each point set, so as to transform them from a set-centered frame into the model centered frame. For the same position in object or scene, only one point may be acquired in each point set. Therefore, a data point has at most one corresponding point in each other point set. To achieve multi-view registration, it usually requires to establish point correspondences between different point sets. However, all corresponding points inevitably deviated from the original position due to noise or sensor resolution. Even for well-aligned point sets, it is difficult to directly establish accurate point correspondences. Fortunately, it is easy to search the NN of one data point from each other aligned point set. These NNs can be utilized to approximate corresponding points of the data point. Considering data noise, it is reasonable to assume that one data point is generated from the Gaussian distribution, which is adhered to its NN with covariance. To achieve promising registration, it requires to explore more available information from all the other point sets. Therefore, we suppose that each data point is generated from the unique GMM, where its NNs in other aligned data sets are regarded as Gaussian centroids with equal covariance and membership probabilities.

Let $V = \{V_i\}_{i=1}^{M}$ be the union of $M$ point sets and $V_i = [v_{i,1} \cdots v_{i,l} \cdots v_{i,N_i}] \in \mathbb{R}^{3\times N_i}$ be $N_i$ points that belong to the $i$th point set. Fig. 1 illustrates the principle of the proposed method. For one data point $v_{i,l}$ in the $i$th point set, it is easy to search its NN $\{v_{j,c(j,l)}\}_{j=1,j\neq i}^{M}$ in each other aligned point sets. Assigned with a covariance, each NN can be regarded as one centroid of the unique GMM, which contains $(M-1)$ Gaussian components. Besides, we use equal covariance and equal membership probabilities for all GMM components. Under this assumption, it is reasonable to formulate the joint probability of data point $v_{i,l}$ as follows:

$$P(v_{i,l}) = \sum_{j\neq i}^{M} \frac{1}{M'} \mathcal{N}(\mathbf{R}_i v_{i,l} + t_i; \mathbf{R}_j v_{j,c(j,l)} + t_j, \Sigma), \quad (1)$$

where $M' = (M-1)$, $\mathbf{R}_i \in \mathbb{R}^{3\times 3}$ and $t_i \in \mathbb{R}^3$ denote the rotation matrix and the translation vector of the $i$th rigid transformation, respectively. For simplicity, we define the function $\phi(v_{i,l}) = \mathbf{R}_i v_{i,l} + t_i$ for the rigid transformation $\{\mathbf{R}_i, t_i\}$ imposed on the data point $v_{i,l}$.

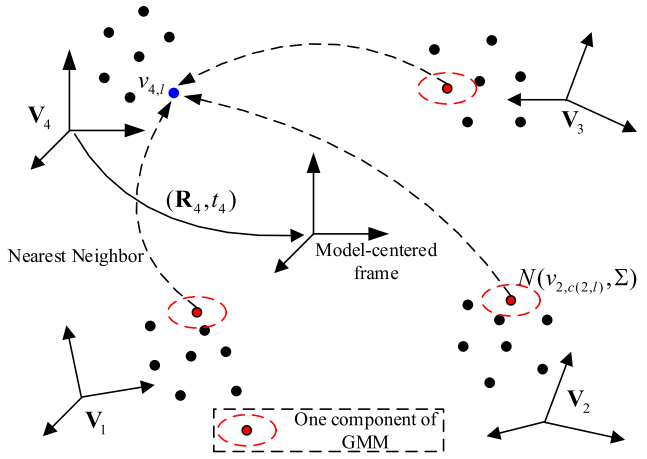


Fig. 1. Illustration of our method. It assumes that each data point is generated from one unique GMM, where its corresponding points in other point sets are regarded as Gaussian centroids with equal covariance and membership probabilities. For one data point, e.g. $v_{4,l} \in V_4$, once rotated and translated from the set-centered frame to the model-centered frame, has one NN in each other aligned point set. These NNs can approximate corresponding points and are regarded as all centroids of the unique GMM to generate the data point itself. Accordingly, the multi-view registration problem is cast into the framework of maximum likelihood estimation solved by the EM algorithm.

Considering data noise and outliers, it is essential to add an extra uniform distribution into the probability function:

$$P(v_{i,l}) = wU(M) + (1-w)\sum_{j\neq i}^{M}\frac{1}{M'}\mathcal{N}(\phi(v_{i,l}); \phi(v_{j,c(j,l)}), \Sigma), \tag{2}$$

where $w$ is the parameter representing the ratio of outliers and $\mathcal{U}(M) = 1/M$ denotes the uniform distribution parameterized by the number of point sets involved in the multi-view registration. As shown in Eq. (2), the probability function contains all rigid transformations for the multi-view registration. Accordingly, it requires to estimate these model parameters $\Theta = \{\{\mathbf{R}_i, \mathbf{t}_i\}_{i=1}^{M}, \Sigma\}$, which can be achieved by maximizing the corresponding likelihood function.

## IV. Multi-View Registration Approach Under the Perspective of EM

To estimate model parameters, it is reasonable to maximize the likelihood function by utilizing the EM algorithm. Therefore, it requires to define a set of hidden variables $Z = \{Z_{i,l} | i \in [1, 2, .., M], l \in [1, 2, \ldots, N_i]\}$, where $Z_{i,l} = c(j, l)$ means the observation $v_{i,l}$ is drawn from the Gaussian distribution $\mathcal{N}(v_{j,c(j,l)}, \sum)$. Given all point sets $V$, model parameters can be estimated by maximizing the expected completed data log-likelihood function as follows:

$$\begin{aligned}\varepsilon(\Theta|V, Z) &= E_Z[\log P(V, Z; \Theta)] \\ &= \sum_Z P(Z|V; \Theta)\log P(V, Z; \Theta) \\ &= \sum_Z P(Z|V; \Theta)\log P(V|Z; \Theta)P(Z; \Theta)\end{aligned} \tag{3}$$

As we utilize equal membership probabilities for all GMM components, $P(Z; \Theta)$ denotes the constant term. Therefore, $\varepsilon(\Theta|V, Z)$ is reformulated as:

$$\varepsilon(\Theta|V, Z) = \sum_Z P(Z|V; \Theta)\log P(V|Z; \Theta) \tag{4}$$

For simplicity, it is reasonable to assume that all data points are independent and identically distributed. Accordingly, Eq. (4) can be straightforwardly rewritten as:

$$\varepsilon(\Theta|V, Z) = \sum_{i,l,j}\alpha_{i,l,j}\log P(v_{i,l}|Z_{i,l} = c(j, l); \Theta), \tag{5}$$

where $\alpha_{i,l,j} = P(Z_{i,l} = c(j, l)|v_{i,l}; \Theta)$ denotes the posterior. By replacing the probability density of Gaussian distribution and ignoring constant terms, the objective function can be reformulated as follows:

$$f(\Theta) = -\sum_{i,l,j}\alpha_{i,l,j}(\left\|\phi(v_{i,l}) - \phi(v_{j,c(j,l)})\right\|_{\Sigma}^{2} + \log|\Sigma|), \tag{6}$$

where $\|v\|_{\Sigma}^{2} = v^T\Sigma^{-1}v$ and $|\Sigma|$ denotes the determinant of matrix $\Sigma$. For simplicity, we restrict each Gaussian distribution to the isotropic covariance, i.e., $\Sigma = \sigma^2\mathbf{I}_3$, where $\mathbf{I}_3$ denotes the $3 \times 3$ identity matrix. Therefore, Eq. (6) can be reformulated as:

$$f(\Theta) = -\sum_{i,l,j}\alpha_{i,l,j}(\frac{\left\|\phi(v_{i,l}) - \phi(v_{j,c(j,l)})\right\|_2^2}{\sigma^2} + d\log\sigma^2) \tag{7}$$

where $d$ denotes point dimension, e.g. $d = 3$ for range point.

As these parameters $\{\mathbf{R}_i\}_{i=1}^{M}$ are of the Special Orthogonal $SO(3)$, particular care should pay attention to their estimation. Therefore, the multi-view registration can be formulated as the constrained optimization problem:

$$\begin{cases}\arg\max\limits_{\Theta} f(\Theta) \\ \text{s.t.}\ \ \mathrm{R}_i{}^T\mathrm{R}_i = \mathrm{I} \text{ and } |\mathrm{R}_i| = 1,\ \forall i \in [1, \ldots, M].\end{cases} \tag{8}$$

This problem depicted in Eq. (8) can be optimized by the EM algorithm, which is augmented with the establishment of point correspondences in E-step. The proposed method can maximize the likelihood function to estimate all rigid transformations. It is called as the *expectation-maximization perspective for multi-view registration* (EMPMR), which achieves multi-view registration by iterations. In each iteration, both E-step and M-step are included in the EMPMR.

### A. E-Step

This step requires to calculate the posterior probability of one data point $v_{i,l}$ generated from each component of the corresponding GMM. Before calculation, all centroids should be assigned to each GMM.

*1) E-Corresponding-Step:* Given the parameter set $\Theta^{k-1}$ obtained from the previous iteration, it is easy to transform all point sets into the same coordinate frame. Then, for one data point $v_{i,l}$ in the $i$th point set, we require to find its corresponding point in each other point set. Although we are difficult to find real corresponding points, they can be approximated by the principle of the minimum Euclidean distance:

$$c(j, l) = \arg\min_{h\in[1,2,..,N_j]}\left\|(\mathbf{R}_i v_{i,l} + t_i) - \phi(v_{j,h})\right\|_2. \tag{9}$$

Eq. (9) denotes the NN search problem, which can be efficiently solved by the $k$-d tree based method [42]. For each data point in one point set, $M'$ NNs can be searched from other point sets and they are regarded as the centroids of GMM components to generate the data point itself.

*2) E-Probability-Step:* Given the centroid $v_{j,c(j,l)}$ and covariance $\Sigma$, it is easy to calculate the posterior probability of data point $v_{i,l}$ generated from the Gaussian distribution $\mathcal{N}(v_{j,c(j,l)}, \Sigma)$ as follows:

$$\alpha_{i,l,j} = \frac{\beta_{i,l,j}}{\sum_{j=1,j\neq i}^{M} \beta_{i,l,j} + \lambda}, \tag{10}$$

where $\lambda = \frac{wM'}{(1-w)M}$ accounts for the outlier term and the notation $\beta_{i,l,j}$ denotes the probability density of Gaussian distribution defined as:

$$\beta_{i,l,j} = \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{\left\|\phi(v_{i,l}) - \phi(v_{j,c(j,l)})\right\|_2^2}{2\sigma^2}\right). \tag{11}$$

Therefore, $a_{i,l,i} = 1 - \sum_{j=1,j\neq i}^{M} a_{i,l,j}$ represents the posterior probability of the $v_{i,l}$ being an outlier. Obviously, the posterior probability indicates the credibility of one corresponding point approximated by the NN.

### B. M-Step

In M-step, it requires to estimate all model parameters $\Theta = \{\{\mathbf{R}_i, \mathbf{t}_i\}_{i=1}^{M}, \Sigma\}$ by maximizing the function $f(\Theta)$. Although it is difficult to simultaneously estimate these model parameters, their estimation can be carried out sequentially and independently.

Given current values of $\alpha_{i,l,j}$, $c(j,l)$, and $\Sigma$, we can alternatively estimate one rigid transformation by setting other rigid transformations and the standard deviation $\sigma$ to their current values. Accordingly, the second term of $f(\Theta)$ is a constant and the $i$th rigid transformation can be estimated from the follows:

$$\begin{cases} \arg\min\limits_{\mathbf{R}_i, t_i} \left(\sum\limits_{l=1}^{N_i} \sum\limits_{j=1,j\neq i}^{M} \alpha_{i,l,j} \left\|(\mathbf{R}_i v_{i,l} + t_i) - \phi(v_{j,c(j,l)})\right\|_2^2\right) \\ \text{s.t.} \quad {\mathbf{R}_i}^T \mathbf{R}_i = \mathrm{I} \text{ and } |\mathrm{R}_i| = 1, \quad \forall i \in [1, \ldots, M]. \end{cases} \tag{12}$$

Eq. (12) denotes a weighted least square (LS) problem, which can be solved by the Singular Value Decomposition (SVD) based method [43].

Finally, when all rigid transformations have been updated, it requires updating the covariance matrix $\Sigma$ for the GMM. Take the derivative of $f(\Theta)$ with respective to $\sigma^2$ and set it to 0, $\Sigma$ can be updated as:

$$\Sigma = \sigma^2 \mathbf{I}_3, \tag{13}$$

where

$$\sigma^2 = \frac{\sum\limits_{i,l,j} \alpha_{i,l,j} \left\|(\mathbf{R}_i v_{i,l} + t_i) - (\mathbf{R}_j v_{j,c(j,l)} + t_j)\right\|_2^2}{d \sum\limits_{i,l,j} \alpha_{i,l,j}}. \tag{14}$$

**Algorithm 1** EMPMR Method

**Input**: Point sets $\mathbf{V} = \{\mathbf{V}_i\}_{i=1}^{M}$, maximum iteration $K$, initial guesses $\Theta^0$.
**Output**: $\{\mathbf{R}_i, \mathbf{t}_i\}_{i=1}^{M}$.

```
h = 0;
repeat
    h = h + 1;
    for (i = 1 : M) do
        E-step
        Build the correspondence {v_{i,l}, v_{i,c^h(j,l)}} by Eq. (9);
        Estimate the posterior probability α^h_{i,l,j} by Eq. (10);
        M-step
        Update R^h_i and t^h_i by solve Eq. (12);
        Update Σ^h by Eq. (13).
    end for
until ((1/M) |f(Θ^h) − f(Θ^{h−1})| < ε)) or (h > H)
```

TABLE I
THE TOTAL COMPUTATION COMPLEXITY OF THE PROPOSED METHOD

| Operation | Complexity |
|---|---|
| $k$-d tree building | $O(\sum_{i=1}^{M}(N_i \lg N_i))$ |
| E-Corresponding-step | $O(HM^{'}N \lg N_j)$ |
| E-Probability-Step | $O(HNM^{'})$ |
| M-Step | $O(HNM^{'})$ |

Obviously, the proposed method utilizes Eqs. (9) and (13) to specify the GMM to generate the data point $v_{i,l}$. Therefore, the number of Gaussian components is automatically determined by the number of point sets involved in multi-view registration. Compared with previous GMM-based methods, there is no trade-off between accuracy and efficiency in the proposed method. What's more, it only requires to search NNs for the definition of Gaussian components. Therefore, it is efficiently to estimate rigid transformations with the isotropic covariance.

### C. Implementation

Similar to most registration methods, EMPMR is a local convergent method and it requires to be provided with an initial guess of model parameters $\Theta$. Different from some methods, EMPMR requires to be additionally provided with an initial guess of $\Sigma$. Intuitively, the value of $\Sigma$ is related to the resolution of each range scan. Therefore, we set $\sigma = d_r$, where $d_r$ indicates the average point resolution of all point sets. To obtain desired registration results, the proposed method should alternately operate E-step and M-step until model parameters have no significant change or $k$ reaches the maximum iterations.

Based on the above description, the proposed EMPMR approach is summarized in Algorithm 1, where we set $\varepsilon = 0.0005$ and $H = 300$.

### D. Complexity Analysis

This section analyzes the computation complexity of EMPMR. As EMPMR is proposed for registration of

TABLE II
DETAILS OF SIX OBJECT DATA SETS UTILIZED IN THE EXPERIMENT

| | Angel | Armadillo | Bunny | Buddha | Dragon | Hand |
|---|---|---|---|---|---|---|
| Scans | 36 | 12 | 10 | 15 | 15 | 36 |
| Points | 2347854 | 307625 | 362272 | 1099005 | 469193 | 1605575 |

multi-view point sets, the point set number $M$ and the number of points $N_i$ or $N_j$ in each point set is the central quantity. For ease analysis, we suppose the iteration number of this method is $H$. Before iteration, it requires building $k-d$ tree to accelerate the NN search and the complexity is $O(N_i \lg N_i)$ for each point set. At each iteration, three operations are implemented to estimate one rigid transformation.

**E-Corresponding-step**. For each point $v_{i,l}$ in the $i$th point set, it is required to search its NN in each other point sets and the complexity is $O(M' \lg N_j)$. As there are $N_i$ in the $i$th point set and $M$ point sets involved in multi-view registration, the total complexity is $O(HM' \lg N_j \Sigma_{i=1}^{M} N_i)$ for $H$ iterations.

**E-Probability-Step**. For each point $v_{i,l}$ in the $i$th point set, there are $M'$ hidden variables. Accordingly, the proposed method requires to calculate $M'N_i$ hidden variables for the $i$th point set. Given $M$ point sets, the total complexity is $O(HM' \Sigma_{i=1}^{M} N_i)$ for $H$ iterations.

**M-Step**. Our method utilizes $M'N_i$ point pairs with corresponding hidden variables to estimate the $i$th rigid transformation. To estimate $M$ rigid transformations with one isotropic covariance, the total complexity is $O(HM' \Sigma_{i=1}^{M} N_i)$ for $H$ iterations.

Therefore, Table I lists the total computation complexity of the proposed method, where $N = \Sigma_{i=1}^{M} N_i$ denotes the total point number in multi-view registration.

## V. EXPERIMENTS

In this section, EMPMR is tested on several public available data sets, including objects and environmental data sets. To illustrate its performance, EMPMR is compared with five state-of-the-art methods: the motion averaging based method with the TrICP [34], the joint registration of multiple point clouds method [11], the low-rank and sparse decomposition based method with the TrICP [36], the multi-view registration method based on Student's t-mixture model [12], and the K-means based method [41], which are abbreviated MATrICP, JRMPC, TMM, LRS, and K-means, respectively. For different data sets, all compared methods utilize the same setting for each parameter. As these methods are local convergent, initial rigid transformations are provided by the feature match method [2], which achieves coarse multi-view registration by the alignment-and-integration strategy. To reduce the run time of registration, all data sets were uniformly down-sampled to around 2000 points per scan.

Experimental results are reported in the form of the run time and registration errors, includes rotation error [44] and translation error. Denote $\{\mathbf{R}_{g,i}, t_{g,i}\}$ and $\{\mathbf{R}_{m,i}, t_{m,i}\}$ as the ground truth and the estimated one of the $i$th rigid transformation, registration errors are defined as $e_{\mathbf{R}} = \frac{1}{M} \sum_{i=1}^{M} \arccos(\frac{\mathrm{tr}(\mathbf{R}_{m,i}(\mathbf{R}_{g,i})^T)-1}{2})$ and $e_t = \frac{1}{M} \Sigma_{i=1}^{M} \|t_{m,i} - t_{g,i}\|_2$, respectively. Accordingly, the unit of rotation error is radians and the unit of translation error depends on the coordinate unit of data points in each data set. All compared approaches are implemented on Matlab2014a without any extra library and utilize the $k$-d tree method to search the NN. Experiments are performed on a four-core 3.6 GHz computer with 8 GB of memory. Without any clarification, the ratio of outliers $w$ is set to be 0.01 in EMPMR.

### A. Object Data Sets

These compared methods are tested and evaluated on six data sets, where four data sets are taken from the Stanford 3D Scanning Repository [45] and the other two data sets were provided by Torsello *et al.* [30]. Each of these data sets was acquired from one object model in different views and the multi-view point sets were provided along with the ground truth of rigid transformations for their registration. Table II lists some details of these data sets, including the number of scans and total number of points for each data set.

*1) Parameter Sensitivity and Convergence:* In EMPRM, there is a free parameter $w$, which requires to be set empirically. Thus, one question arising here is whether its performance is sensitive to $w$ or not. To answer this question, we conduct experiments on six data sets to observe its effect on registration performance with different values. Experimental results are reported in the form of registration errors. Fig. 2 displays experimental results with different values of $w$ on six data sets. Obviously, the setting of $w \in [10^{-5}, 0.01]$ is more likely to obtain accurate results for multi-view registration. And the performance of EMPRM only has small variations as long as $w$ is chosen in a suitable range, i.e., from $10^{-5}$ to 0.01. Summarily, EMPRM is relatively insensitive to its parameter $w$ as long as it is chosen from a suitable range. This makes it easy to apply EMPRM without much effort for parameter tuning.

In EMPRM, the number of GMM components is automatically determined by the scan number. To illustrate its performance with respect to component number, we test EMPRM on Stanford Bunny with different scan configurations. Table III displays averaging registration errors and run time over 10 times under scan configurations listed in Table IV. As shown in Table III, the scan configuration has no significant effect on the registration accuracy of EMPRM. While, the run time increases with the increase of scan number. But a small number of scans may lead to registration failure. For example, we test EMPRM on Stanford Bunny with the scan configuration [1 4 10], where the 1st scan was acquired from front side of Bunny model and two other scans were acquired from back side of Bunny model. Accordingly, most points in the 1st scan have no real corresponding point in the other two scans. Fig 3 displays the corresponding registration

TABLE III

PERFORMANCE OF EMPRM TESTED ON STANFORD BUNNY WITH DIFFERENT SCAN CONFIGURATIONS, WHERE THE SCAN CONFIGURATION HAS NO SIGNIFICANT EFFECT ON THE REGISTRATION ACCURACY OF EMPRM

| $M$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| $e_\theta$(rad.) | 0.0039 | 0.0031 | 0.0046 | 0.0041 | 0.0040 | 0.0040 | 0.0042 | 0.0033 | 0.0035 |
| $e_t$ (mm) | 0.4981 | 0.3049 | 0.3496 | 0.3258 | 0.3098 | 0.2924 | 0.3492 | 0.3250 | 0.3439 |
| Time (min.) | 0.0005 | 0.0022 | 0.0103 | 0.0161 | 0.0280 | 0.0519 | 0.0711 | 0.0951 | 0.1456 |

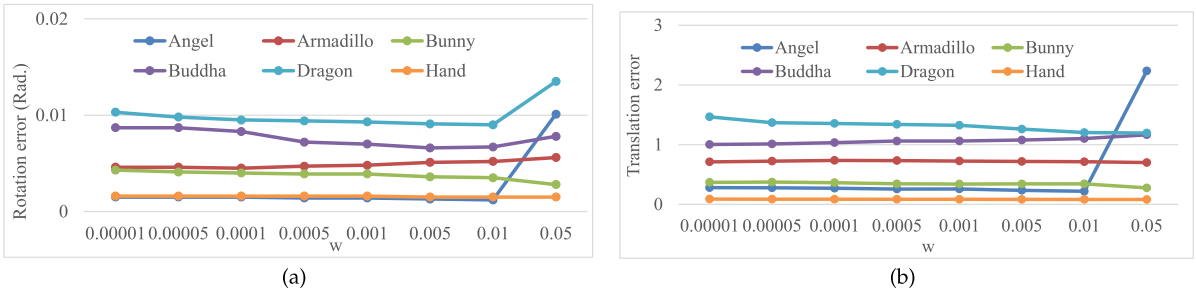


Fig. 2. Multi-view registration error of EMPMR under varied $w$. These results illustrate that EMPMR is relatively insensitive to $w$ as long as it is chosen from a suitable range. (a) Rotation errors. (b) Translation errors.

TABLE IV

THE SCAN CONFIGURATIONS OF STANFORD BUNNY FOR EMPRM TEST

| Number | Configuration |
|---|---|
| 2 | [1 3] |
| 3 | [1 2 3] |
| 4 | [1 2 3 6] |
| 5 | [1 2 3 6 8] |
| 6 | [1 2 3 6 7 8] |
| 7 | [12 3 4 5 6 7] |
| 8 | [1 2 3 4 5 7 8 ] |
| 9 | [1 2 3 4 5 6 7 8 9 ] |
| 10 | [1 2 3 4 5 6 7 8 9 10] |

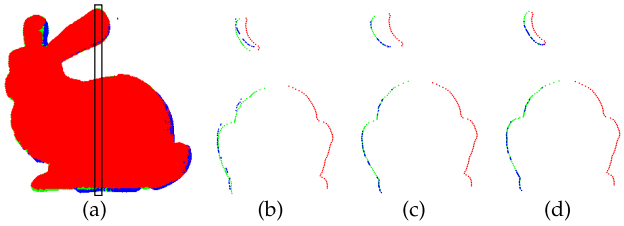


Fig. 3. Failure result of EMPRM tested on Stanford Bunny with the scan configuration [1 4 10], where points in different colors are belong to different scans. (a) Well-aligned scans. (b) Initial cross-section. (c) EMPRM's cross-section. (d) Ground truth cross-section.

results in the form of cross-section. As shown in 3, EMPRM fails to achieve accurate registration. To obtain desired results, most points should have at least one real corresponding point in another scan. Otherwise, EMPRM may be failed to achieve promising registration. This is the main limitation of EMPRM. However, it should be noted that most methods for multi-view registration proposed so far share this limitation as well.

What's more, we also explore the convergence properties of EMPRM on six data sets. Fig. 4 displays the convergence of EMPRM with respect to iterations. As shown in Fig. 4, the log-likelihood function of EMPRM increases with the increase of iterations. For most data sets, this method converges within 100 iterations, except for Angel and Hand data sets, which contains more scans than other data sets. Given initial

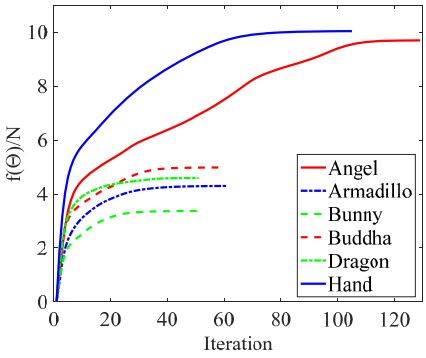


Fig. 4. Convergence results of EMPRM tested on six data sets. For view convenience, the vertical axis displays the mean log-likelihood value subtracted by the initial value.

model parameters, EMPRM has the closed-form solution in each iteration and can obtain desired results within acceptable iterations.

*2) Effect of Initial Rigid Transformations:* To illustrate the effect of initial rigid transformations, we test EMPRM on Stanford Bunny with different initial rigid transformations and compare it with other methods. As one rigid transformation includes the rotation and translation, we draw rotation angles (or translation variables) from different uniform distributions with varying intervals and assign translation variables (or rotation angles) with random typical values, so as to generate rotation matrix $\Delta\mathbf{R}_i$ and translation vector $\Delta t_i$, respectively. Then initial rigid transformations can be estimated as $\mathbf{R}_i^0 = \Delta\mathbf{R}_i \cdot \mathbf{R}_{g,i}$ and $t_i^0 = \Delta t_i + t_{g,i}$. Subsequently, $\{\mathbf{R}_i^0, t_i^0\}_{i=1}^M$ are taken as the inputs of different multi-view registration methods. To eliminate randomness, each group of experiments is carried out by 20 independent tests. Table V and Table VI illustrate the performance of all compared methods under different rotations with typical translations and different translations with typical rotations, respectively.

TABLE V

COMPARISON RESULTS (MEAN±STD.) ON STANFORD BUNNY WITH DIFFERENT INITIAL ROTATIONS, WHERE BOLD NUMBERS DENOTE THE BEST PERFORMANCE IN THE SAME SETTING

| **Method** | | $[-0.005, 0.005]$ rad. | $[-0.01, 0.01]$ rad. | $[-0.015, 0.015]$ rad. | $[-0.02, 0.02]$ rad. | $[-0.025, 0.025]$ rad. |
|---|---|---|---|---|---|---|
| MATrICP | $e_R$ | 0.0084 ±0.0002 | 0.0088 ±0.0003 | 0.0088 ±0.0004 | **0.0088 ±0.0005** | 0.0200 ±0.0353 |
| | $e_t$ | 0.7071 ±0.0155 | 0.7227 ±0.0255 | 0.7237 ±0.0280 | **0.7270 ±0.0290** | 2.1742 ±4.5691 |
| JRMPC | $e_R$ | 0.0157 ±0.0016 | 0.0224 ±0.0035 | 0.0265 ±0.0086 | 0.0328 ±0.0079 | 0.0444 ±0.0123 |
| | $e_t$ | 1.5831 ±0.2726 | 2.1815 ±0.4712 | 2.5631 ±0.7148 | 2.9610 ±0.7313 | 3.6737 ±1.1629 |
| TMM | $e_R$ | 0.0108±0.0009 | 0.0145±0.0035 | 0.0173±0.0031 | 0.0191±0.0032 | 0.0271±0.0099 |
| | $e_t$ | 0.8652±0.0600 | 1.1782±0.2118 | 1.2962±0.2253 | 1.6296±0.3543 | 2.0992±0.6654 |
| LRS | $e_R$ | 0.0115±0.0015 | 0.0137±0.0027 | 0.0166±0.0076 | 0.0168±0.0032 | 0.0270±0.0095 |
| | $e_t$ | 1.0543±0.0825 | 1.1635±0.1626 | 1.3072±0.4540 | 1.7662±0.6641 | 5.0631±3.8756 |
| K-means | $e_R$ | 0.0140 ±0.0021 | 0.0173 ±0.0028 | 0.0196 ±0.0029 | 0.0235 ±0.0036 | 0.0256 ±0.0063 |
| | $e_t$ | 1.8172 ±0.1708 | 1.9558 ±0.1823 | 2.1607 ±0.2569 | 2.3297 ±0.2945 | 2.6266±0.3697 |
| EMPMR | $e_R$ | **0.0037±0.0003** | **0.0039±0.0004** | **0.0039±0.0004** | 0.0091±0.0100 | **0.0181±0.0212** |
| | $e_t$ | **0.3638±0.0250** | **0.3735±0.0278** | **0.4853±0.0309** | 0.7536±0.6602 | **1.2096±1.1074** |

TABLE VI

COMPARISON RESULTS (MEAN±STD.) ON STANFORD BUNNY WITH DIFFERENT INITIAL TRANSLATIONS, WHERE BOLD NUMBERS DENOTE THE BEST PERFORMANCE IN THE SAME SETTING AND $d_r$ DENOTES THE MEAN RESOLUTION OF POINT SETS

| **Method** | | $[-0.8, 0.8] \times d_r$ | $[-1.6, 1.6] \times d_r$ | $[-2.4, 2.4] \times d_r$ | $[-3.2, 3.2] \times d_r$ | $[-4, 4] \times d_r$ |
|---|---|---|---|---|---|---|
| MATrICP | $e_R$ | 0.0086 ±0.0003 | 0.0088 ±0.0005 | 0.0088 ±0.0005 | 0.0151 ±0.0281 | 0.0217 ±0.0397 |
| | $e_t$ | 0.7174 ±0.0212 | 0.7276 ±0.0098 | 0.7225 ±0.0058 | 1.5691 ±3.7862 | 2.5039 ±5.3224 |
| JRMPC | $e_R$ | 0.0148 ±0.0016 | 0.0191 ±0.0030 | 0.0272 ±0.0059 | 0.0299 ±0.0059 | 0.0346 ±0.0072 |
| | $e_t$ | 1.7641 ±0.3219 | 2.4267 ±0.5660 | 3.5482 ±0.8480 | 4.5224 ±1.3969 | 4.8404 ±0.8692 |
| TMM | $e_R$ | 0.0088±0.0011 | 0.0131±0.0020 | 0.0164±0.0038 | 0.0170±0.0041 | 0.0214±0.0047 |
| | $e_t$ | 0.8185±0.1055 | 1.4051±0.1813 | 1.5058±0.3450 | **1.7415±0.4582** | **1.7933±0.3290** |
| LRS | $e_R$ | 0.0126±0.0011 | 0.0465±0.0523 | 0.1801±0.0630 | 0.2415±0.0721 | 0.2730±0.0914 |
| | $e_t$ | 1.0752±0.1366 | 3.4058±3.4232 | 6.1969±5.5621 | 8.7087±7.8585 | 9.3293±8.5560 |
| K-means | $e_R$ | 0.0128 ±0.0014 | 0.0158 ±0.0024 | 0.0182 ±0.0025 | 0.0209 ±0.0031 | 0.0238 ±0.0045 |
| | $e_t$ | 1.8116 ±0.1468 | 1.9361 ±0.2898 | 2.1542 ±0.2637 | 2.4217 ±0.4704 | 2.6961±0.5208 |
| EMPMR | $e_R$ | **0.0036±0.0003** | **0.0036±0.0004** | **0.0046±0.0032** | **0.0112±0.0106** | **0.0174±0.0136** |
| | $e_t$ | **0.3561±0.0237** | **0.3667±0.0317** | **0.4997±0.4340** | 2.1392±1.3644 | 2.4468±1.6291 |

TABLE VII

COMPARISON REGISTRATION ERRORS OF ALL COMPARED METHODS, WHERE BOLD NUMBERS DENOTE THE BEST PERFORMANCE OF EACH DATA SET

| Method | | Angel (dm) | Armadillo (mm) | Bunny (mm) | Buddha (mm) | Dragon (mm) | Hand (dm) |
|---|---|---|---|---|---|---|---|
| Initial | $e_\theta$ | 0.0221 | 0.0234 | 0.0239 | 0.0262 | 0.0251 | 0.0282 |
| | $e_t$ | 2.0388 | 2.5333 | 2.1260 | 1.6535 | 1.5216 | 0.4945 |
| MATrICP | $e_\theta$ | 0.0103 | 0.0237 | 0.0088 | 0.0085 | 0.0116 | 0.0262 |
| | $e_t$ | 2.8651 | 4.2915 | 0.7181 | **0.8581** | 1.1037 | 1.3390 |
| JRMPC | $e_\theta$ | 0.0065 | 0.0121 | 0.0166 | 0.0161 | 0.0154 | 0.0044 |
| | $e_t$ | 1.9816 | 3.0651 | 1.7571 | 0.9629 | 1.7466 | 0.7069 |
| TMM | $e_\theta$ | 0.0062 | 0.0061 | 0.0080 | 0.0106 | 0.0154 | 0.0072 |
| | $e_t$ | 2.5197 | 1.9646 | 0.7031 | 1.2137 | **1.0689** | 0.7682 |
| LRS | $e_\theta$ | 0.0087 | 0.0410 | 0.0125 | 0.0181 | 0.0162 | 0.0265 |
| | $e_t$ | 2.0633 | 9.4194 | 0.9586 | 1.0597 | 1.5358 | 1.3825 |
| K-means | $e_\theta$ | 0.0056 | 0.0066 | 0.0100 | 0.0147 | 0.0134 | 0.0059 |
| | $e_t$ | 0.9789 | 2.3207 | 1.6088 | 1.0276 | 1.6140 | 0.4615 |
| EMPMR | $e_\theta$ | **0.0012** | **0.0052** | **0.0035** | **0.0067** | **0.0090** | **0.0014** |
| | $e_t$ | **0.2161** | **0.7145** | **0.3439** | 1.1023 | 1.1433 | **0.0887** |

As shown in Table V and Table VI, the performance of all compared methods is reduced with the increase of rotation and translation. Under small rotations and translations, both EMPMR and MATrICP are accurate and robust than other compared methods. However, when rotations or translations exceed a certain threshold, all compared methods are unable to obtain desired registration results. This is because all these compared methods are local convergent algorithms. To obtain desired results, EMPMR as well as other compared methods require to be provided with approximate initial rigid transformations.

*3) Accuracy:* Table VII illustrates registration errors of all compared methods tested on six data sets. To view results in a more intuitive way, Fig. 5 illustrates all multi-view registration results in the form of cross-section. As shown in Fig. 5 and Table VII, EMPMR can always achieve the most accurate

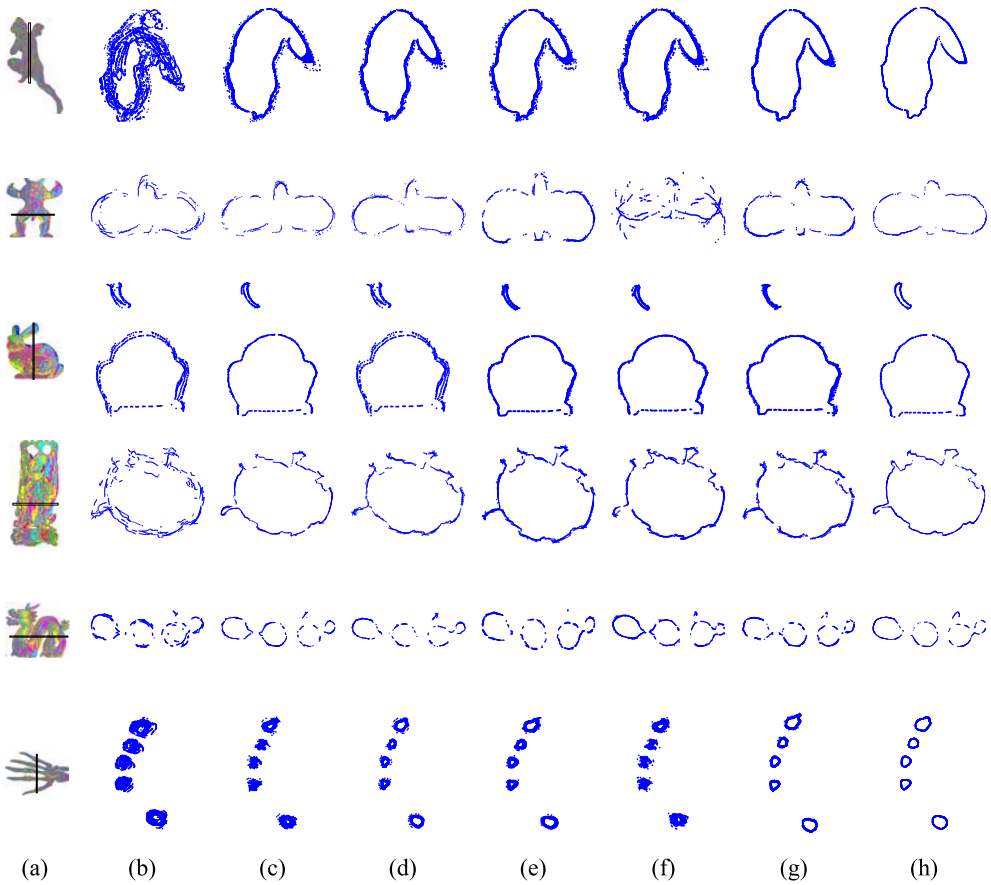


Fig. 5. Multi-view registration results in the form of cross-section, where corresponding regions are illustrated on the aligned 3D models and thin curves denote good registration results. (a) Aligned 3D models. (b) Initial results. (c) Results MATrICP. (d) Results of JRMPC. (e) Results of TMM. (f) Results of LRS. (g) Results of K-means. (h) Results of EMPMR.

registration, except for the Stanford Buddha and Dragon. For these two data sets, EMPMR obtains the smallest rotation errors but a little larger translation errors than some other methods. As shown in Fig. 5, considering both rotation and translation errors, EMPMR also obtains the most accurate registration results for these two data sets.

Actually, both MATrICP and LRS recover all global motions from a set of relative motions estimated by the pair-wise registration. These two methods can transform the multi-view registration problem into the pair-wise registration problem, which is easier than the multi-view registration. To refine global motions, MATrICP implements the motion averaging algorithm on a set of available relative motions. Given a set of reliable relative motions, it can achieve good multi-view registration, such as Stanford Bunny, Buddha, and Dragon data sets. However, this method ignores the noises in data and only partially explore available information to estimate one rigid transformation. Accordingly, it may difficult to obtain very accurate registration results. Besides, the motion averaging algorithm is sensitive to unreliable relative motions, even one relative motion will lead to the failure of multi-view registration, such as Hand and Stanford Armadillo data sets. While, LRS concatenates all available relative motions into one matrix, where the missing elements are filled with zero. Then, the complete matrix containing all global motions is recovered from the concatenated one by low-rank and sparse matrix decomposition algorithm. LRS is able to ignore unreliable relative motions, but it requires more relative motions than that of the motion averaging algorithm to recover global motions. Accordingly, it is difficult to obtain promising registration results. Moreover, it may be failed to achieve multi-view registration due to insufficient relative motions, such as Stanford Armadillo data set.

While, JRMPC and TMM assume that all data points are generated from a central GMM and a central TMM, respectively. These two methods contain many Gaussian distributions or Student's t-distributions required to be estimated. To achieve multi-view registration, they utilize the EM algorithm to estimate all model parameters including rigid transformations. Compared with Gaussian distribution, the Student's t-distribution has more good properties. Therefore, TMM can obtain more accurate results than JRMPC. This conclusion is verified by experimental results illustrated in Fig. 5 and Table VII. But neither of them can obtain satisfactory results. Theoretically, they are both probabilistic methods and expected to obtain more accurate results. As mentioned before, both of them require to estimate many model parameters, whose initial values should be fine-tuned and provided in advance. Without good initial parameters, these methods are more likely trapped into local optimum. However, it is difficult to provide good initial parameters for different data sets without enough prior information. Besides, the oper-

ation of clustering inevitably causes the loss of information, which inevitably reduces registration accuracy. What's more, the trade-off between accuracy and efficiency in clustering-based methods make it difficult to obtain promising results for multi-view registration.

Similar to JRMPC and TMM, K-means also casts the multi-view registration into the clustering problem. But it only requires to estimate clustering centroids with rigid transformations, so it is more likely to obtain accurate registration results. To achieve multi-view registration, this method sequentially traverses each point set, then alternatively implements the operations of clustering and rigid transformation estimation. For clustering, it uses the K-means algorithm to divide all aligned points into the preset number of clusters, where one clustering centroid represent all points in this cluster. The clustering operation inevitably causes the loss of information. For registration, it aligns the scan to clustering centroids, where the correspondence of each data point is searched from all cluster centroids. Since cluster centroids are far less than raw points in all scans, this method is very efficient but far from accuracy for multi-view registration due to the information loss.

Different from JRMPC and TMM, EMPMR assumes that each data point is generated from one corresponding GMM, whose centroids are specified by the NN search. As EMPMR utilizes equal covariance and equal membership probabilities for all GMM components, it only requires to estimate one covariance as well as rigid transformations for multi-view registration. For each data point, this approach fully explores available information from all other opposite point sets by the NN search. What's more, it also considers noises in data points to estimate rigid transformations. Given initial parameters, it is more likely to obtain accurate registration results than other compared methods. Unfortunately, its translation error is a little larger than other methods for some data sets. The reason may be that the estimation of rigid transformations involves false corresponding points in non-overlapping regions. Although low posterior probabilities are assigned to these corresponding points, they also may increase translation errors of EMPMR.

*4) Efficiency:* For efficiency comparison, all compared methods are tested on six data sets. To eliminate randomness, each group of experiment is carried out by 10 independent tests. Fig. 6 displays the averaging runtime of all compared methods tested on six data sets. As displayed in Fig. 6, K-means is the most efficient method. EMPMR is comparable with LRS, they are more efficient than the other three compared methods. Besides, TMM is the most time-consuming one. In these methods, two operations are time-consuming, i.e., the establishment of point correspondences and update of rigid transformations. For facilitate analysis, we restate that there are $M$ point sets in multi-view registration, where the $i$-th (or $j$-th) point set has $N_i$ (or $N_j$) points and the total number of points is $N = \sum_{i=1}^{M} N_i$. Besides, the number of clusters is $K$ in clustering-based methods.

In LRS and MATrICP, the main computation is to estimate all relative motions. Specifically, each data point requires to find one nearest neighbor from some other point sets by the

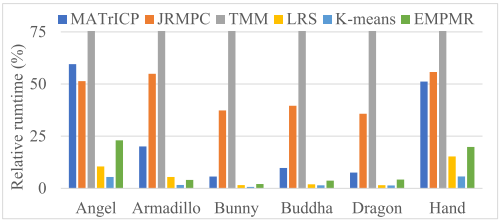


Fig. 6. Comparison of averaging runtime over 10 independent tests on six data sets. For view convenience, we display relative run time for different methods, where TMM is most time-consuming method with (17.2717, 5.4772, 6.5912, 9.276, 10.4413, 15.9024) minutes corresponding to 100% run time of each data set.

$k$-d tree method and $\bar{M}N_i$ point pairs are utilized to update relative motions related to the $i$th point set, where $\bar{M}(1 \leq \bar{M} \leq M)$ denotes the number of point sets containing high overlap percentages to the $i$th point set. For $M$ point sets, the total computation of these two steps are $O(N\bar{M}\lg N_j)$ and $O(N\bar{M})$ at each iteration. It should be noted, that the efficiency of MATrICP may be seriously reduced for the failed registration, such as Angel and Hand data sets. This is because the failed registration requires more iterations than the normal registration.

To achieve multi-view registration, clustering-based methods sequentially traverses each point set, then alternatively implements the operations of clustering and rigid transformation estimation. In both JRMPC and TMM, all points are utilized to estimate the distribution of each component and all components are involved in the estimation of one rigid transformation. Therefore, each data point requires to establish the relationships with all $K$ components by the calculation of posterior probability and $KN_i$ point pairs are utilized to update the $i$th rigid transformation. For $M$ point sets, the total computation of these two steps are $O(NK)$ and $O(NK)$ at each iteration. However, the Student's t-distribution is more complex than Gaussian distribution, which costs less time to be estimated. Therefore, TMM is more time-consuming than JRMPC. In K-means, the clustering operation divides all aligned points into the preset number of clusters and utilizes one centroid to represent all points in this cluster. For registration, each data point only requires to find one NN from $K$ cluster centers by the $k$-d tree method and only $N_i$ point pairs are utilized to update the $i$th rigid transformation. For $M$ point sets, the total computation of these two steps are $O(N\lg K)$ and $O(N)$ at each iteration. Obviously, K-means is more efficient than JRMPC and TMM.

For multi-view registration, EMPMR only requires to establish the relationship between one data point and $M'$ NNs for the definition of Gaussian components. Further, $M'N_i$ point pairs are utilized to estimate the $i$th rigid transformation and $O(NM')$ point pairs assigned with posterior probability are used to estimate the only isotropic covariance, so the total computation complexity of these two steps are $O(NM'\lg N_j)$ and $O(NM')$ at each iteration.

Table VIII lists the computation complexity of these compared methods in each iteration. Usually, the cluster number $K$ is larger than the scan number $M$ and the term $M'\lg N_j$

TABLE VIII
COMPUTATIONAL COMPLEXITY OF ALL COMPARED METHODS IN ONE ITERATION

| | MATrICP | JRMPC | TMM | LRS | K-means | EMPMR |
|---|---|---|---|---|---|---|
| Build correspondence | $O(N\bar{M}\lg N_j)$ | $O(NK)$ | $O(NK)$ | $O(N\bar{M}\lg N_j)$ | $O(N\lg K)$ | $O(NM^{'}\lg N_j)$ |
| Update transformation | $O(N\bar{M}\lg N_j)$ | $O(NK)$ | $O(NK)$ | $O(N\bar{M})$ | $O(N)$ | $O(NM^{'})$ |

TABLE IX
COMPARISON RESULTS (MEAN±STD.) ON SIX DATA SETS PERTURBED BY GAUSSIAN NOISES (SNR=50DB), WHERE BOLD NUMBERS DENOTE THE BEST PERFORMANCE

| Method | | Angel (dm) | Armadillo (cm) | Bunny (cm) | Buddha (cm) | Dragon (cm) | Hand (dm) |
|---|---|---|---|---|---|---|---|
| MATrICP | $e_R$ | 0.0102 ±0.0001 | 0.0204 ±0.0062 | 0.0089 ±0.0004 | 0.0086 ±0.0002 | 0.0114 ±0.0003 | 0.0261 ±0.0002 |
| | $e_t$ | 2.8710 ±0.0228 | 4.0146 ±0.4889 | 0.6996 ±0.0265 | **0.8424 ±0.0168** | **1.0464 ±0.0449** | 1.3417 ±0.0065 |
| JRMPC | $e_R$ | 0.0062 ±0.0002 | 0.0119 ±0.0002 | 0.0146 ±0.0002 | 0.0162 ±0.0001 | 0.0153 ±0.0004 | 0.0044 ±0.0001 |
| | $e_t$ | 1.9069 ±0.0137 | 3.0249 ±0.0162 | 1.7402 ±0.0125 | 0.9690 ±0.0042 | 1.7823 ±0.0048 | 0.7075 ±0.0024 |
| TMM | $e_R$ | 0.0061±0.0003 | 0.0060±0.0003 | 0.0084±0.0005 | 0.0102±0.0004 | 0.0153±0.0004 | 0.0072±0.0002 |
| | $e_t$ | 2.4665±0.0842 | 1.9844±0.0427 | 0.7487±0.0377 | 1.1904±0.0374 | 1.0662±0.0201 | 0.7792±0.0071 |
| LRS | $e_R$ | 0.0086±0.0000 | 0.0397±0.0007 | 0.0127±0.0008 | 0.0180±0.0001 | 0.0150±0.0007 | 0.0265±0.0001 |
| | $e_t$ | 2.0599±0.0055 | 9.1893±0.1191 | 0.9882±0.0776 | 1.0443±0.0035 | 1.4367±0.0562 | 1.3803±0.0012 |
| K-means | $e_R$ | 0.0056±0.0003 | 0.0064±0.0004 | 0.0104±0.0004 | 0.0140±0.0005 | 0.0141±0.0006 | 0.0055±0.0002 |
| | $e_t$ | 1.0117 ±0.0636 | 2.3448 ±0.1086 | 1.6936 ±0.0628 | 1.0684 ±0.0559 | 1.3913 ±0.0724 | 0.4674 ±0.0107 |
| EMPMR | $e_R$ | **0.0012 ±0.0000** | **0.0052 ± 0.0000** | **0.0035 ± 0.0000** | **0.0068 ±0.0001** | **0.0090 ± 0.0000** | **0.0014 ±0.0000** |
| | $e_t$ | **0.2191± 0.0043** | **0.7170 ±0.0018** | **0.3439 ±0.0024** | 1.0972 ±0.0036 | **1.1179±0.0052** | **0.0839±0.0000** |

TABLE X
COMPARISON RESULTS (MEAN±STD.) ON SIX DATA SETS PERTURBED BY GAUSSIAN NOISES (SNR=25DB), WHERE BOLD NUMBERS DENOTE THE BEST PERFORMANCE

| **Method** | | Angel (dm) | Armadillo (cm) | Bunny (cm) | Buddha (cm) | Dragon (cm) | Hand (dm) |
|---|---|---|---|---|---|---|---|
| MATrICP | $e_R$ | 0.0103 ±0.0001 | 0.0204 ±0.0073 | 0.0119 ±0.0006 | 0.0123 ±0.0004 | 0.0162 ±0.0005 | 0.0369 ±0.0003 |
| | $e_t$ | 2.8726 ±0.0187 | 3.8764 ±0.5655 | 0.7047 ±0.0337 | **0.8433 ±0.0188** | **1.0534 ±0.0559** | 1.3419 ±0.0063 |
| JRMPC | $e_R$ | 0.0062 ±0.0003 | 0.0118 ±0.0004 | 0.0165 ±0.0004 | 0.0154 ±0.0002 | 0.0157 ±0.0003 | 0.0048 ±0.0002 |
| | $e_t$ | 1.9628 ±0.0526 | 3.0543 ±0.0542 | 1.7404 ±0.0320 | 0.9694 ±0.0240 | 1.7923 ±0.0272 | 0.7178 ±0.0067 |
| TMM | $e_R$ | 0.0062±0.0004 | 0.0062±0.0006 | 0.0086±0.0006 | 0.0104±0.0004 | 0.0155±0.0006 | 0.0073±0.0003 |
| | $e_t$ | 2.4865±0.0845 | 1.9913±0.0454 | 0.7626±0.0665 | 1.1950±0.0478 | 1.0686±0.0213 | 0.7802±0.0076 |
| LRS | $e_R$ | 0.0123±0.0001 | 0.0398±0.0008 | 0.0127±0.0008 | 0.0181±0.0006 | 0.0153±0.0012 | 0.0265±0.0002 |
| | $e_t$ | 2.0559±0.0287 | 9.1627±0.1068 | 0.9905±0.0371 | 1.0551±0.0248 | 1.5105±0.0699 | 1.3817±0.0062 |
| K-means | $e_R$ | 0.0057 ±0.0004 | 0.0067 ±0.0006 | 0.0107 ±0.0005 | 0.0142 ±0.0008 | 0.0151 ±0.0008 | 0.0056 ±0.0004 |
| | $e_t$ | 1.0750 ±0.0827 | 2.4698 ±0.1109 | 1.7043 ±0.0716 | 1.1212 ±0.0866 | 1.5158±0.1099 | 0.4690 ±0.0152 |
| EMPMR | $e_R$ | **0.0012±0.0001** | **0.0055±0.0003** | **0.0037±0.0002** | **0.0069±0.0002** | **0.0090±0.0004** | **0.0020±0.0002** |
| | $e_t$ | **0.2253±0.0273** | **0.7202±0.0215** | **0.3444±0.0198** | 1.1001±0.0251 | 1.1443±0.0458 | **0.1426±0.0052** |

in multi-view registration. But $\lg K$ is always smaller than $M'\lg N_j$. When the number of iterations is relatively similar in all compared methods, the K-means is the most efficient method. EMPMR has almost the same computation complexity with LRS, they are more efficient than the other three compared methods. Besides, the efficiency of all these compared method are gradually reduced with the increase of the scan number involved in multi-view registration. Experimental results shown in Fig. 6 also verify these conclusions. It should be noted that K-means is more efficient yet less accurate than EMPMR, even less accurate than some other methods due to the information loss in clustering.

*5) Robustness to Noises:* To illustrate its robustness, EMPMR is tested on six data sets, which are added with random Gaussian noises. It also compared with other five methods. They are tested on each data set, which is added with random Gaussian noises. To eliminate randomness, each group of experiment is carried out by 30 independent tests. Experimental results are reported in the form of the average registration errors as well as the standard deviations. Tables IX and X display all statistics results of these compared methods under two levels of Gaussian noises. As shown in Tables IX and X, the registration accuracy of all compared methods tends to be decreased with the increase of noise. Under the same level of noises, EMPMR is able to obtain almost the most accurate registration results, except for translation errors of Stanford Buddha and Dragon data set. As shown in Tables IX and X, both MATrICP and LRS seem robust to noises for some data sets. As mentioned before, these two methods recover global motions from a set of relative motions, which are estimated by the pair-wise registration algorithm. Although the added noises may reduce the accuracy of pair-wise registration results, their influence on multi-view registration may be eliminated by the motion averaging or LRS matrix decomposition algorithm. However, the added noise may result in unreliable and insufficient relative motions, which will lead to the undesired registration results in MATrICP and LRS, respectively.

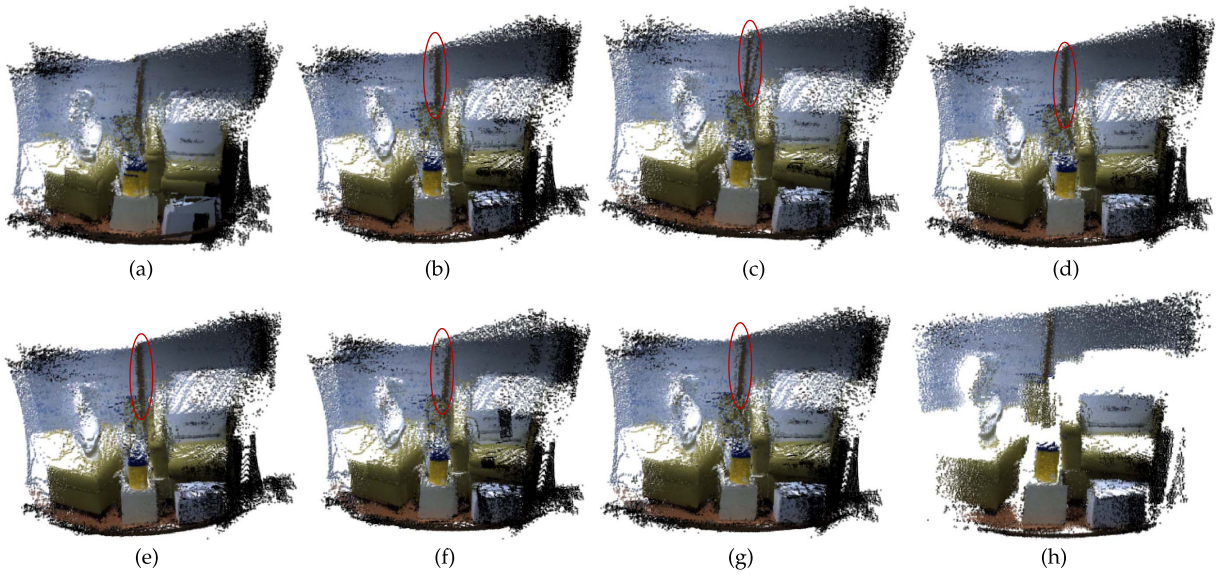


Fig. 7. Results of different resgistration methods tested on EXBI data set, where inconsistency may appear in the red rectangle area. (a) Initial results. (b) MATrICP result. (c) JRMPC result. (d) TMM result. (e) LRS result. (f) K-means result. (g) EMPMR result. (h) Scene reconstruction obtained by denoising EMPMR result.

TABLE XI

COMPARISON OF MAPPING RESULTS FOR DIFFERENT METHODS TESTED ON GAZEBO DATA SET, WHERE BOLD NUMBERS DENOTE THE BEST PERFORMANCE

| | Initial | K-means | JRMPC | LRS | MATrICP | TMM | EMPRM |
|---|---|---|---|---|---|---|---|
| $e_\theta$ (rad.) | 0.0241 | 0.0097 | 0.0216 | 0.0092 | 0.0089 | 0.0145 | **0.0087** |
| $e_t$ (m) | 0.1069 | 0.0318 | 0.1005 | 0.0557 | 0.0511 | 0.0630 | **0.0185** |
| Time (min.) | / | 0.8817 | 7.8113 | 1.6324 | 1.9533 | 19.2284 | **0.4258** |

As probabilistic methods, both JRMPC and TMM take noises into account for multi-view registration, they are robust to noises. However, both of them require to estimate many components as well as rigid transformations, which makes them tend to be trapped into local optimum. Therefore, they are unable to obtain promising registration results. Besides, K-means does not consider noise, so it is susceptible to noises. Accordingly, this method is also difficult to obtain promising registration results.

While, EMPMR utilizes the NN search to specify the centroids for all GMMs, it only requires to estimate $M$ rigid transformations as well as one GMM's covariance. Accordingly, it is more likely to obtain promising registration results. What's more, this method considers noises in data points to estimate rigid transformations, so it is robust to noises. Tables IX and X further verify these conclusions. Therefore, EMPMR is the most robust one among all these compared methods.

### *B. Environmental Data Sets*

In practice, it is required to deal with different kinds of data sets. To illustrate its ability for real applications, EMPMR is tested on EXBI data set [11] and Gazebo data set [46] for scene reconstruction and robot SLAM (Simultaneous Localization And Mapping), respectively. Besides, we also present the results of other compared methods for comparison.

The EXBI data set contains 10 RGB-D scans recorded by the Kinect sensor, which acquires point sets with associated color information. During data acquirement, the Kinect sensor was manually moving around an indoor scene. In this data set, all scan pairs contain high overlapping percentages. For the scene reconstruction, only range information is processed by EMPMR and color information is utilized to assist for the final assessment. Given initial rigid transformations, all registration methods are utilized to achieve multi-view registration, which is the prerequisite for the scene reconstruction. As the ground truth of registration is unavailable in this data set, we only present visual registration results for comparison. Fig. 7 displays the registration results of all compared methods. It is not surprising that both MATrICP and LRS can obtain the most accurate results. This is because the inputs of these two methods are pair-wise registration results, which are all available and reliable in this data set due to high overlapping percentages of each scan pair. Besides, EMPMR can also obtain promising registration result, which is more accurate than other compared methods. Since the raw RGB-D data contains noises, registration result should be filtered for good scene reconstruction. Specifically, we delete these data points from aligned point sets, where its distance to the 5th NN is larger than the given threshold. Here, the threshold is set to be $3d_l$, where $d_l$ denotes the mean distance of all NN pairs in the aligned point sets. As shown in Fig. 7, EMPMR has the potential for the multi-view registration of RGB-D data and can be applied to the scene reconstruction.

The Gazebo data set contains 32 range scans, which were acquired from an outdoor environment by a robot equipped

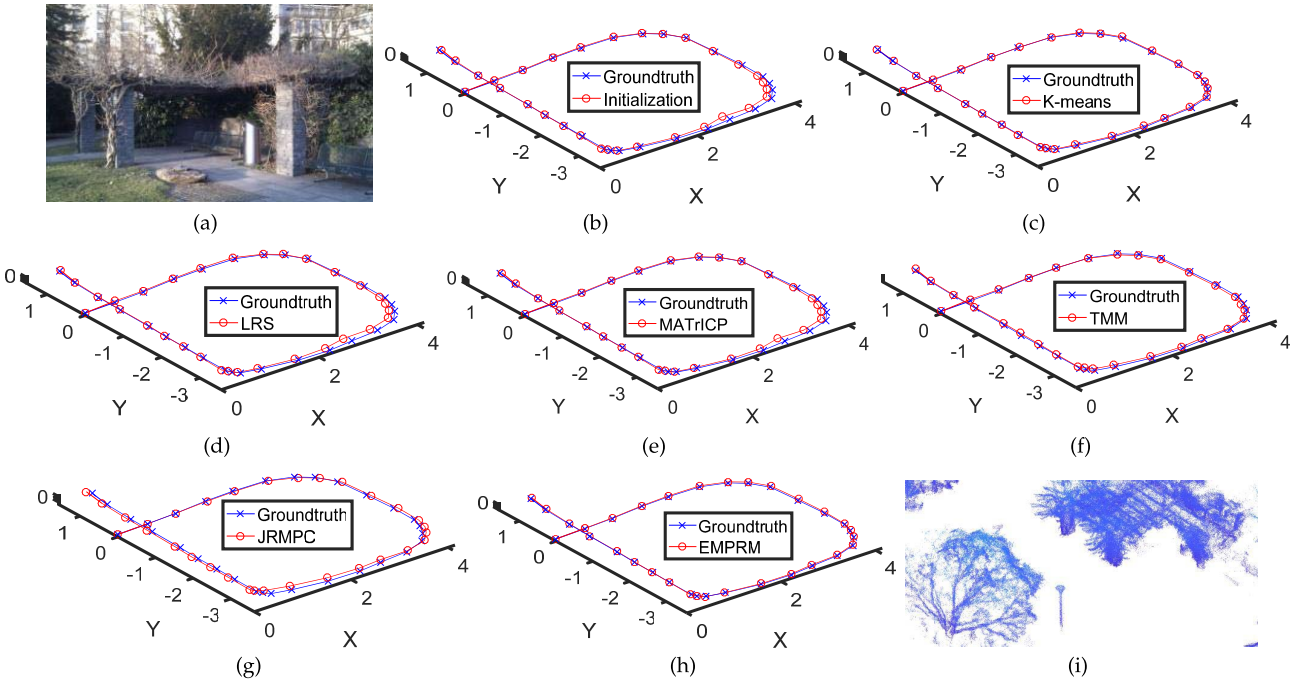


Fig. 8. Comparison of localization results for different methods tested on Gazebo data set. (a) Scene image. (b) Initialization. (c) K-means result. (d) LRS result. (e) MATrICP result. (f) TMM result. (g) JRMPC result. (h) EMPRM result. (i) Corresponding mapping result of EMPMR.

with one laser range finder and location equipment (IMU and GPS). To acquire this data set, the robot mainly moved on a 2D plane and followed by the path to form a closed loop ($4 \times 5$x0.09 m). For accuracy evaluation, ground truth of location is also provided in this data set. Different from the EXBI data set, only some scan pairs contain high overlapping percentages. As the raw range scan contains data indicating ground surface, they should be filtered by the height value for accurate SLAM. To eliminate accumulation error, all multi-view registration methods utilized to refine initial rigid transformations and obtain SLAM result. Table XI displays the localization results of EMPRM with other compared methods. To view comparison results in a more intuitive way, Fig. 8 illustrates localization results of all compared methods tested on Gazebo data set. As shown in Table XI and Fig. 8, EMPRM is a little more accurate than k-means, they are more accurate than other compared methods, especially for the translation vector. It should be noted both MATrICP and LRS only tend to reduce the localization error near the closed loop. While, EMPRM as well as some compared methods are able to reduce localization error along the whole robot path. Therefore, EMPMR has the potential for the SLAM application.

## VI. Conclusion

Under the EM perspective, this paper proposes a novel method for multi-view registration. The main idea of this method is that each data point is generated from one unique GMM, where its NNs are regarded as Gaussian centroids with equal covariance and membership probabilities. Subsequently, all point sets are treated on an equal footing and the multi-view registration problem is cast into the framework of maximum likelihood estimation, which is optimized by the EM algorithm. Since the GMM component number is automatically determined by the point set number, there is no trade-off between registration accuracy and efficiency. What's more, it establishes the relationship between one data point and several NNs to define Gaussian components, which seriously reduces the number of model parameters required to be estimated. Because this method takes data noise into account by the Gaussian distribution, it is more likely to obtain promising registration results. Experimental results illustrate its superior performance over state-of-the-art methods on accuracy, efficiency, and robustness. What's more, it can be applied to SLAM and scene reconstruction.

## Acknowledgment

The authors would like to thank anonymous reviewers and AE for their valuable comments and suggestions.

## References

[1] J. Yang, H. Li, D. Campbell, and Y. Jia, "Go-ICP: A globally optimal solution to 3D ICP point-set registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 11, pp. 2241–2254, Nov. 2016.

[2] H. Lei, G. Jiang, and L. Quan, "Fast descriptors and correspondence propagation for robust global point cloud registration," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 3614–3623, Aug. 2017.

[3] Z. Jiang, J. Zhu, Y. Li, J. Wang, Z. Li, and H. Lu, "Simultaneous merging multiple grid maps using the robust motion averaging," *J. Intell. Robotic Syst.*, vol. 94, nos. 3–4, pp. 655–668, Jun. 2019.

[4] F. Yu, J. Xiao, and T. Funkhouser, "Semantic alignment of LiDAR data at city scale," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1722–1731.

[5] D. Aiger, N. J. Mitra, and D. Cohen-Or, "4-points congruent sets for robust pairwise surface registration," *ACM Trans. Graph.*, vol. 27, no. 3, p. 85, Aug. 2008.

[6] A. Dai, M. Nießner, M. Zollhöfer, S. Izadi, and C. Theobalt, "Bundlefusion: Real-time globally consistent 3d reconstruction using on-the-fly surface reintegration," *ACM Trans. Graph. (ToG)*, vol. 36, no. 3, p. 24, 2017.

[7] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," *Proc. SPIE*, vol. 1611, pp. 586–606, Apr. 1992.

[8] Y. Chen and G. Medioni, "Object modelling by registration of multiple range images," *Image Vis. Comput.*, vol. 10, no. 3, pp. 145–155, Apr. 1992.
[9] A. Rasoulian, R. Rohling, and P. Abolmaesumi, "Group-wise registration of point sets for statistical shape models," *IEEE Trans. Med. Imag.*, vol. 31, no. 11, pp. 2025–2034, Nov. 2012.
[10] X. Mateo, X. Orriols, and X. Binefa, "Bayesian perspective for the registration of multiple 3D views," *Comput. Vis. Image Understand.*, vol. 118, pp. 84–96, Jan. 2014.
[11] G. D. Evangelidis and R. Horaud, "Joint alignment of multiple point sets with batch and incremental expectation-maximization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 6, pp. 1397–1410, Jun. 2018.
[12] N. Ravikumar, A. Gooya, S. Çimen, A. F. Frangi, and Z. A. Taylor, "Group-wise similarity registration of point sets using Student's t-mixture model for statistical shape models," *Med. Image Anal.*, vol. 44, pp. 156–176, Feb. 2018.
[13] Z. Min, J. Wang, and M. Q.-H. Meng, "Joint rigid registration of multiple generalized point sets with hybrid mixture models," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 1, pp. 334–347, Jan. 2020.
[14] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *Proc. 3DIM*, vol. 1, May/Jun. 2001, pp. 145–152.
[15] D. Chetverikov, D. Stepanov, and P. Krsek, "Robust Euclidean alignment of 3D point sets: The trimmed iterative closest point algorithm," *Image Vis. Comput.*, vol. 23, no. 3, pp. 299–309, Mar. 2005.
[16] E. Lomonosov, D. Chetverikov, and A. Ekárt, "Pre-registration of arbitrarily oriented 3D surfaces using a genetic algorithm," *Pattern Recognit. Lett.*, vol. 27, no. 11, pp. 1201–1208, Aug. 2006.
[17] R. Sandhu, S. Dambreville, and A. Tannenbaum, "Point set registration via particle filtering and stochastic dynamics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1459–1473, Aug. 2010.
[18] R. B. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (FPFH) for 3D registration," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2009, pp. 3212–3217.
[19] Y. Aoki, H. Goforth, R. A. Srivatsan, and S. Lucey, "PointNetLK: Robust & efficient point cloud registration using PointNet," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 7163–7172.
[20] A. Kurobe, Y. Sekikawa, K. Ishikawa, and H. Saito, "CorsNet: 3D point cloud registration by deep neural network," *IEEE Robot. Autom. Lett.*, vol. 5, no. 3, pp. 3960–3966, Jul. 2020.
[21] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–2275, Dec. 2010.
[22] B. Jian and B. C. Vemuri, "Robust point set registration using Gaussian mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1633–1645, Aug. 2011.
[23] W. Gao and R. Tedrake, "FilterReg: Robust and efficient probabilistic point-set registration using Gaussian filter and twist parameterization," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 11095–11104.
[24] R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau, "Towards a general multi-view registration technique," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 5, pp. 540–547, May 1996.
[25] J. Williams and M. Bennamoun, "Simultaneous registration of multiple corresponding point sets," *Comput. Vis. Image Understand.*, vol. 81, no. 1, pp. 117–142, Jan. 2001.
[26] J. Zhu, Z. Li, S. Du, L. Ma, and T. Zhang, "Surface reconstruction via efficient and accurate registration of multiview range scans," *Opt. Eng.*, vol. 53, no. 10, Apr. 2014, Art. no. 102104.
[27] S. Krishnan *et al.*, "Global registration of multiple 3D point sets via optimization-on-a-manifold," in *Proc. Symp. Geometry Process.*, 2005, pp. 187–196.
[28] Y. Tang and J. Feng, "Hierarchical multiview rigid registration," in *Computer Graphics Forum*, vol. 34, no. 5. Hoboken, NJ, USA: Wiley, 2015, pp. 77–87.
[29] S.-W. Shih, Y.-T. Chuang, and T.-Y. Yu, "An efficient and accurate method for the relaxation of multiview registration error," *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 968–981, Jun. 2008.
[30] A. Torsello, E. Rodola, and A. Albarelli, "Multiview registration via graph diffusion of dual quaternions," in *Proc. CVPR*, Jun. 2011, pp. 2441–2448.
[31] F. Wang, B. C. Vemuri, and A. Rangarajan, "Groupwise point pattern registration using a novel CDF-based Jensen-Shannon divergence," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2006, pp. 1283–1288.
[32] T. Chen, B. C. Vemuri, A. Rangarajan, and S. J. Eisenschenk, "Group-wise point-set registration using a novel CDF-based Havrda-Charvát divergence," *Int. J. Comput. Vis.*, vol. 86, no. 1, p. 111, 2010.
[33] L. G. Sanchez Giraldo, E. Hasanbelliu, M. Rao, and J. C. Principe, "Group-wise point-set registration based on Rényi's second order entropy," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2454–2462.
[34] V. M. Govindu and A. Pooja, "On averaging multiview relations for 3D scan registration," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1289–1302, Mar. 2014.
[35] F. Arrigoni, B. Rossi, and A. Fusiello, "Global registration of 3D point sets via LRS decomposition," in *Proc. Eur. Conf. Comput. Vis.* Amsterdam, The Netherlands: Springer, 2016, pp. 489–504.
[36] F. Arrigoni, B. Rossi, P. Fragneto, and A. Fusiello, "Robust synchronization in SO(3) and SE(3) via low-rank and sparse matrix decomposition," *Comput. Vis. Image Understand.*, vol. 174, pp. 95–113, Sep. 2018.
[37] R. Guo, J. Zhu, Y. Li, D. Chen, Z. Li, and Y. Zhang, "Weighted motion averaging for the registration of multi-view range scans," *Multimedia Tools Appl.*, vol. 77, no. 9, pp. 10651–10668, May 2018.
[38] C. Jin, J. Zhu, Y. Li, S. Pang, L. Chen, and J. Wang, "Multi-view registration based on weighted LRS matrix decomposition of motions," *IET Comput. Vis.*, vol. 13, no. 4, pp. 376–384, Jun. 2019.
[39] Z. Min and M. Q.-H. Meng, "Joint alignment of multiple generalized point sets with anisotropic positional uncertainty based on expectation maximization," in *Proc. Int. Conf. 3D Vis. (3DV)*, Sep. 2018, pp. 170–179.
[40] Z. Min, J. Wang, and M. Q.-H. Meng, "Joint registration of multiple generalized point sets," in *Proc. Int. Workshop Shape Med. Imag.* Granada, Spain: Springer, 2018, pp. 169–177.
[41] J. Zhu, Z. Jiang, G. D. Evangelidis, C. Zhang, S. Pang, and Z. Li, "Efficient registration of multi-view point sets by K-means clustering," *Inf. Sci.*, vol. 488, pp. 205–218, Jul. 2019.
[42] A. Nuchter, K. Lingemann, and J. Hertzberg, "Cached k-d tree search for ICP algorithms," in *Proc. 6th Int. Conf. 3-D Digit. Imag. Modeling (3DIM)*, Aug. 2007, pp. 419–426.
[43] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 4, pp. 376–380, Apr. 1991.
[44] D. Q. Huynh, "Metrics for 3D rotations: Comparison and analysis," *J. Math. Imag. Vis.*, vol. 35, no. 2, pp. 155–164, Oct. 2009.
[45] B. C. M. Levoy, J. Gerth, and K. Pull. (2005). *The Stanford 3D Scanning Repository*. [Online]. Available: http://www-graphics.stanford.edu/data/3dscanrep
[46] F. Pomerleau, M. Liu, F. Colas, and R. Siegwart, "Challenging data sets for point cloud registration algorithms," *Int. J. Robot. Res.*, vol. 31, no. 14, pp. 1705–1711, Dec. 2012.

**Jihua Zhu** (Member, IEEE) received the B.E. degree in automation from Central South University, China, in 2004, respectively, and the Ph.D. degree in pattern recognition and intelligence systems from Xi'an Jiaotong University, China, in 2011. He is currently an Associate Professor with the School of Software Engineering, Xi'an Jiaotong University. His research interests include computer vision and machine learning.

**Rui Guo** received the M.S. degree from Lanzhou Jiaotong University, Lanzhou, China, in 2016. He is currently pursuing the Ph.D. degree in software engineering with Xi'an Jiaotong University, China. His research interests include computer vision and image processing.

**Zhongyu Li** photograph and biography not available at the time of publication.

**Jing Zhang** photograph and biography not available at the time of publication.

**Shanmin Pang** (Member, IEEE) photograph and biography not available at the time of publication.