A Generative Model for the Joint Registration of Multiple Point Sets

Georgios Evangelidis, Dionyssos Kounades-Bastian, Radu Horaud, Emmanouil Psarakis

A Generative Model for the Joint Registration of Multiple Point Sets⋆

Georgios D. Evangelidis[1], Dionyssos Kounades-Bastian[1,2], Radu Horaud[1],
Emmanouil Z. Psarakis[2]

[1]INRIA Grenoble Rhône-Alpes, France
[2]CEID, University of Patras, Greece

Abstract. This paper describes a probabilistic generative model and its associated algorithm to jointly register multiple point sets. The vast majority of state-of-the-art registration techniques select one of the sets as the "model" and perform pairwise alignments between the other sets and this set. The main drawback of this mode of operation is that there is no guarantee that the model-set is free of noise and outliers, which contaminates the estimation of the registration parameters. Unlike previous work, the proposed method treats all the point sets on an equal footing: they are realizations of a Gaussian mixture (GMM) and the registration is cast into a clustering problem. We formally derive an EM algorithm that estimates both the GMM parameters and the rotations and translations that map each individual set onto the "central" model. The mixture means play the role of the registered set of points while the variances provide rich information about the quality of the registration. We thoroughly validate the proposed method with challenging datasets, we compare it with several state-of-the-art methods, and we show its potential for fusing real depth data.

Keywords: point set registration, joint registration, expectation maximization, Gaussian mixture model.

1 Introduction

Registration of point sets is an essential methodology in computer vision, computer graphics, robotics, and medical image analysis. To date, while the vast majority of techniques deal with two sets, e.g., [4, 10, 26, 23, 15, 18], the multiple-set registration problem has comparatively received less attention, e.g., [30, 28]. There are many practical situations when multiple-set registration is needed, nevertheless the problem is generally solved by applying pairwise registration repeatedly, either sequentially [6, 20, 17], or via a *one-versus-all* strategy [3, 7, 16].

Regardless of the particular pairwise registration algorithm that is being used, their use for multiple-set registration has limited performance. On the one
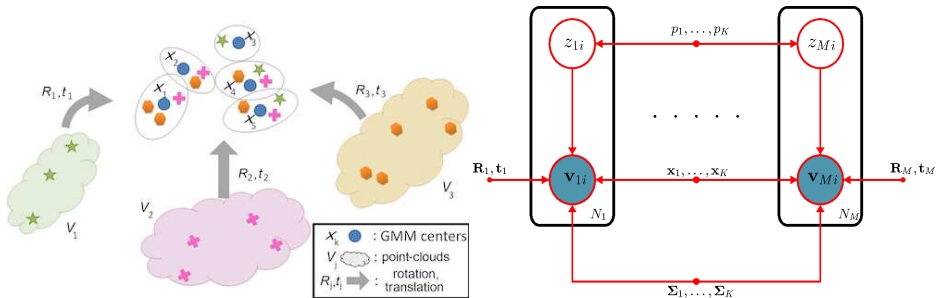

Fig. 1. The proposed generative model for joint registration of multiple point clouds (left) and the associated graphical model (right). Unlike pairwise registration strategies, the proposed model simultaneously registers an arbitrary number of point clouds with partial or total overlap and optimally estimates both the GMM and registration parameters. Hence, the solution is not biased towards a particular cloud.

hand, sequential register-then-integrate strategies suffer from error propagation while they are optimal only locally, i.e., between point-set pairs. On the other hand, one-versus-all registration apparently leads to a biased estimator since one of the sets governs the registration and the solution is optimal only for this reference set. Therefore, an unbiased solution that evenly distributes the errors across all point sets is particularly desirable.

In this paper we propose a generative approach to the problem of joint registration of multiple 3D point sets. An arbitrary number of point sets are assumed to be generated from the *same* Gaussian mixture model (GMM). More precisely, an observed point i from set j, once rotated ($\mathbf{R}_j$) and translated ($\mathbf{t}_j$), is generated from the k-th component of a GMM, e.g., Fig. 1. Therefore, the GMM parameters are conditioned by the registration parameters (rotations and translations). This can be cast into a maximum likelihood formulation that is efficiently solved via an expectation conditional maximization (ECM) algorithm that jointly and optimally estimates all the GMM and registration parameters.

Unlike existing approaches to point registration that constrain the GMM means to coincide with the points of one set, the parameters of the proposed mixture model are not tight to a particular set. Existing approaches have the danger that noise and outliers, inherently present in the point set chosen to be the GMM means, contaminates the solution in an irrevocable way. It is well known that noisy data and outliers can be very robustly handled with GMMs by including a uniform component [1]. This has already been proposed in the framework of pairwise registration [23, 15], in which case one set is supposed to be "bad" while the other one is supposed to be "perfect". In the proposed model all the sets are treated similarly and the GMM means are obtained by averaging over several transformed points belonging to different sets. Therefore, the proposed approach puts all the data on an equal footing and hence it is more robust. This is particularly beneficial when the task is to align a large number of point clouds, e.g., gathered with a depth camera.

The remainder of this paper is organized as follows: Section 2 discusses the related work. Section 3 formulates the problem in a generative probabilistic framework while Section 4 presents the proposed formulation and the associated algorithm. Experiments are presented in Section 5 and Section 6 concludes the paper.

2   Related work

Modern point registration methods adopt soft assignment strategies, thus generalizing ICP [4]. In all these methods one set is the "model" and the other set is the "data" [29, 11, 9, 23, 15], to cite just a few. This non-symmetric treatment leads to biased solutions. Alternatively, [18, 14] consider two GMMs, one for each point set and the rigid transformation is applied to one of these mixtures. This leads to a non-linear optimization problem, hence requiring proper initialization. Moreover, outliers are not explicitly taken into account by these methods.

Multiple point-set registration is often solved using a sequential pairwise registration strategy [6, 20, 17, 24, 8]. Whenever an additional set is available, the model parameters are updated using either ICP or a probabilistic scheme. In addition to the drawbacks associated with pairwise registration, this incremental mode of operation is subject to error propagation. Another possible strategy is to register pairs of views and subsequently to refine the registration parameters by solving an optimization problem that updates the parameters with respect to a reference set [3]. [16] starts with pairwise registrations to build a connected graph of overlapping sets, while a global optimization step over this graph representation eliminates matches that are not globally consistent. Similarly and more efficiently, [7] and [25] globally refine the registration between overlapping sets by working only in the transformation space. Despite the global refinement step, these methods except [25] suffer from the same limitation, namely one of the point sets is chosen as a reference set and hence the final parameters are biased.

Multiple point-set registration was also addressed in [30, 28]. Both these methods estimate a transformation for each point set, such that the transformed sets can be jointly registered. Starting from some *known* point correspondences between the sets, [30] estimates the transformation parameters through the minimization of a criterion that relates *any* two overlapping sets, and optionally integrates confidence of points. Since point correspondences are provided by pairwise ICP, this approach is referred to as multi-set ICP. As with pairwise ICP, one-to-one correspondences lead to the aforementioned limitations. Notice that the same formulation but with a different optimization method is proposed in [19], and was recently extended in [21] to deal with unknown correspondences. As in [30], [2] registers matched shapes by estimating transformations (one per shape) of an unknown reference shape. [28] shares a lot of similarities with [18] in the sense that it represents each point set as a GMM and the transformations are applied to these mixtures rather than to individual points. The model parameters are estimated by minimizing the Jensen-Shannon divergence. A by-product

of the algorithm is a probabilistic atlas defined by a convex combination of the mixtures. To the best of our knowledge, this is the only method that achieves joint multiple-set registration without recourse to a pairwise strategy. However the GMM representation of a point set inherently encapsulates the set's noisy and outlier observations, and hence the registration of point sets with different amounts of noise and of outliers is problematic, as well as sets with large non overlapping regions.

3   Problem formulation

Let $\mathbf{V}_j = [\mathbf{v}_{j1} \dots \mathbf{v}_{ji} \dots \mathbf{v}_{jN_j}]$ be a $\mathbb{R}^{3 \times N_j}$ matrix of N_j points associated with the j-th point set and let M be the number of sets. We denote with $\mathbf{V} = \{\mathbf{V}_j\}_{j=1}^M$ the union of all the data points. It is assumed that there is a rigid transformation $\phi_j : \mathbb{R}^3 \to \mathbb{R}^3$ that maps a point set j onto a scene-centered model. The objective is to estimate the set-to-scene transformations under the constraint that the sets are jointly registered. It is assumed that the point sets are rigidly-transformed realizations of an unknown "central" GMM. Hence, one can write

$$P(\mathbf{v}_{ji}) = \sum_{k=1}^K p_k \mathcal{N}\big(\phi_j(\mathbf{v}_{ji}) | \mathbf{x}_k, \mathbf{\Sigma}_k\big) + p_{K+1} \mathcal{U}(a - b), \tag{1}$$

where $\phi_j(\mathbf{v}_{ji}) = \mathbf{R}_j \mathbf{v}_{ji} + \mathbf{t}_j$ (a 3×3 rotation matrix $\mathbf{R}_j$ and a 3×1 translation vector $\mathbf{t}_j$), p_k are the mixing coefficients $\sum_{k=1}^{K+1} p_k = 1$, $\mathbf{x}_k$ and $\mathbf{\Sigma}_k$ are the means and covariance matrices, and $\mathcal{U}$ is the uniform distribution parameterized by $a - b$. Here we take $a - b = h$, where h is the volume of the 3D convex hull encompassing the data [15]. We now define γ as the ratio between outliers and inliers, that is,

$$p_{K+1} = \gamma \sum_{k=1}^K p_k. \tag{2}$$

This allows to balance the outlier/inlier proportion by choosing γ. To summarize, the model parameters are

$$\Theta = \left(\{p_k, \mathbf{x}_k, \mathbf{\Sigma}_k\}_{k=1}^K, \{\mathbf{R}_j, \mathbf{t}_j\}_{j=1}^M\right). \tag{3}$$

We stress that the deterministic nature of ϕ_j does not affect the statistical properties of the mixture model. Fig. 1 shows a graphical representation of the proposed model.

    This problem can be solved in the framework of expectation-maximization. In particular, we define hidden variables $\mathcal{Z} = \{z_{ji} | j = 1 \dots M, i = 1 \dots N_j\}$ such that $z_{ji} = k$ assigns observation $\phi_j(\mathbf{v}_{ji})$ to the k-th component of the mixture model, and we aim to maximize the expected complete-data log-likelihood

$$\mathcal{E}(\Theta | \mathbf{V}, \mathcal{Z}) = \mathbb{E}_{\mathcal{Z}}[\log P(\mathbf{V}, \mathcal{Z}; \Theta) | \mathbf{V}] = \sum_{\mathcal{Z}} P(\mathcal{Z} | \mathbf{V}, \Theta) \log(P(\mathbf{V}, \mathcal{Z}; \Theta)) \tag{4}$$

in order to estimate the parameters Θ.

4   Joint Multiple-Set Registration

Assuming that the observed data $\mathbf{V}$ are independent and identically distributed, it is straightforward to write (4) as

$$\mathcal{E}(\Theta|\mathbf{V}, \mathcal{Z}) = \sum_{j,i,k} \alpha_{jik} \Big(\log p_k + \log P(\phi_j(\mathbf{v}_{ji})|z_{ji} = k; \Theta) \Big) \tag{5}$$

where $\alpha_{jik} = P(z_{ji} = k|\mathbf{v}_{ji}; \Theta)$ are the posteriors. By replacing the standard expressions of the likelihoods [5] and by ignoring constant terms, (5) can be written as an objective function of the form

$$f(\Theta) = -\frac{1}{2} \sum_{j,i,k} \alpha_{jik} \Big(\|\phi_j(\mathbf{v}_{ji}) - \mathbf{x}_k)\|_{\mathbf{\Sigma}_k}^2 + \log|\mathbf{\Sigma}_k| - 2\log p_k \Big)$$
$$+ \log p_{K+1} \sum_{j,i} \alpha_{ji(K+1)} \tag{6}$$

where $|\cdot|$ denotes the determinant and $\|y\|_A^2 = y^\top A^{-1} y$. Therefore, one has to solve the following constrained optimization problem:

$$\begin{cases} \max_\Theta & f(\Theta) \\ \text{s.t.} & \mathbf{R}_j^\top \mathbf{R}_j = \mathbf{I}, \;\; |\mathbf{R}_j| = 1, \forall j = 1 \ldots M. \end{cases} \tag{7}$$

Direct optimization of $f(\Theta)$ via a closed-form solution is difficult owing to the induced non-linearities. Therefore, we adopt an expectation conditional maximization (ECM) scheme to solve (7). ECM is more broadly applicable than EM, while it is well suited for our problem owing to the extended parameter set. Notice that ECM replaces the M-step of EM with a series of conditional maximization (CM) steps, that is, an M-substep for each parameter. We will refer to this algorithm as *joint registration of multiple point clouds* (JR-MPC); its outline is given in Algorithm 1. JR-MPC maximizes $f(\Theta)$, and hence $\mathcal{E}(\Theta|\mathbf{V}, \mathcal{Z})$, sequentially with respect to each parameter, by clamping the remaining ones to their current values. Commonly, such an iterative process leads to a stepwise maximization of the observed-data likelihood as well [22]. At each iteration, we first estimate the transformation parameters, given the current GMM parameters, and then we estimate the new GMM parameters, given the new transformation parameters. It is of course possible to adopt a reverse order, in particular when a rough alignment of the point sets is provided. However, we consider no prior information on the rigid transformations, so that the pre-estimation of the registration parameters favors the estimation of the GMM means, $\mathbf{x}_k$, that should be well distributed in space.

Now that our objective function is specified, we are going to present in detail each step of JR-MPC. We restrict the model to isotropic covariances, i.e., $\mathbf{\Sigma}_k = \sigma_k^2 \mathbf{I}$, since it leads to a more efficient algorithm, while experiments with non-isotropic covariance [15] showed that there is no significant accuracy gain.

Algorithm 1 Joint Registration of Multiple Point Clouds (JR-MPC)

Require: Initial parameter set Θ^0
1: $q \leftarrow 1$
2: **repeat**
3:    *E-step*: Use Θ^{q-1} to estimate posterior probabilities $\alpha_{jik}^q = P(z_{ji} = k|\mathbf{v}_{ji}; \Theta^{q-1})$
4:    *CM-step-A*: Use α_{jik}^q, $\mathbf{x}_k^{q-1}$ and $\mathbf{\Sigma}_k^{q-1}$ to estimate $\mathbf{R}_j^q$ and $\mathbf{t}_j^q$.
5:    *CM-step-B*: Use α_{jik}^q, $\mathbf{R}_j^q$ and $\mathbf{t}_j^q$ to estimate the means $\mathbf{x}_k^q$.
6:    *CM-step-C*: Use α_{jik}^q, $\mathbf{R}_j^q$, $\mathbf{t}_j^q$ and $\mathbf{x}_k^q$ to estimate the covariances $\mathbf{\Sigma}_k^q$.
7:    *CM-step-D*: Use α_{jik}^q to estimate the priors p_k^q.
8:    $q \leftarrow q + 1$
9: **until** Convergence
10: **return** Θ^q

E-step: By using the definitions for the likelihood and prior terms, and the decomposition of the marginal distribution, $P(\phi_j(\mathbf{v}_{ji})) = \sum_{s=1}^{K+1} p_s P(\phi_j(\mathbf{v}_{ji})|z_{ji} = s)$, the posterior probability α_{jik} of $\mathbf{v}_{ij}$ to be an *inlier* can be computed by

$$\alpha_{jik} = \frac{p_k \sigma_k^{-3} \exp\left(-\frac{1}{2\sigma_k^2}\|\phi_j(\mathbf{v}_{ji}) - \mathbf{x}_k\|^2\right)}{\sum_{s=1}^{K}\left[p_s \sigma_s^{-3}\exp\left(-\frac{1}{2\sigma_s^2}\|\phi_j(\mathbf{v}_{ji}) - \mathbf{x}_s\|^2\right)\right] + \beta}, \quad k = 1, \ldots, K, \quad (8)$$

where $\beta = \gamma/h(\gamma + 1)$ accounts for the outlier term, while the posterior to be an *outlier* is simply given by $\alpha_{ji(K+1)} = 1 - \sum_{k=1}^{K} \alpha_{jik}$. As shown in Alg. 1, the posterior probability at the q-th iteration, α_{jik}^q, is computed from (8) using the parameter set Θ^{q-1}.

CM-step-A: This step estimates the transformations ϕ_j that maximize $f(\Theta)$, given current values for α_{jik}, $\mathbf{x}_k$, $\mathbf{\Sigma}_k$. Notice that this estimation can be carried out independently for each set j, since ϕ_j associates each point set with the common set of GMM means.

By setting the GMM parameters to their current values, we reformulate the problem to estimate the roto-translations. It can be easily shown that the maximizers $\mathbf{R}_j^*, \mathbf{t}_j^*$ of $f(\Theta)$ coincide with the minimizers of the following constrained optimization problems

$$\begin{cases} \min_{\mathbf{R}_j, \mathbf{t}_j} & \|(\mathbf{R}_j \mathbf{W}_j + \mathbf{t}_j \mathbf{e}^\top - \mathbf{X})\mathbf{\Lambda}_j\|_F^2 \\ \text{s.t.} & \mathbf{R}_j^\top \mathbf{R}_j = \mathbf{I}, \quad |\mathbf{R}_j| = 1 \end{cases} \quad (9)$$

where $\mathbf{\Lambda}_j$ is a $K \times K$ diagonal matrix with elements $\lambda_{jk} = \frac{1}{\sigma_k}\sqrt{\sum_{i=1}^{N_j} \alpha_{jik}}$, $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_K]$, $\mathbf{e}$ is a vector of ones, $\|\cdot\|_F$ denotes the Frobenius norm, and $\mathbf{W}_j = [\mathbf{w}_{j1}, \ldots, \mathbf{w}_{jK}]$, with $\mathbf{w}_{jk}$, is a virtual 3D point given by

$$\mathbf{w}_{jk} = \frac{\sum_{i=1}^{N_j} \alpha_{jik} \mathbf{v}_{ji}}{\sum_{i=1}^{N_j} \alpha_{jik}}, \quad (10)$$

or the weighted average of points in point set j, being the weights proportional to the posterior probabilities in terms of the k-th component. The above problem is an extension of the problem solved in [27] as here we end up with a weighted case due to $\mathbf{\Lambda}_j$. The problem still has an analytic solution. In specific, the optimal rotation is given by

$$\mathbf{R}_j^* = \mathbf{U}_j^L \mathbf{S}_j \mathbf{U}_j^{R\top}, \tag{11}$$

where $\mathbf{U}_j^L$ and $\mathbf{U}_j^R$ are the left and right matrices obtained from the singular value decomposition of matrix $\mathbf{X}_j \mathbf{\Lambda}_j \mathbf{P}_j \mathbf{\Lambda}_j \mathbf{W}^\top$, with $\mathbf{P}_j = \mathbf{I} - \frac{\mathbf{\Lambda}_j \mathbf{e}(\mathbf{\Lambda}_j \mathbf{e})^\top}{(\mathbf{\Lambda}_j \mathbf{e})^\top \mathbf{\Lambda}_j \mathbf{e}}$ being a projection matrix, and $\mathbf{S}_j = \mathrm{diag}(1, 1, |\mathbf{U}_j^L||\mathbf{U}_j^R|)$. Once the optimum rotation is known, the optimum translation is computed by

$$\mathbf{t}_j^* = \frac{1}{tr(\mathbf{\Lambda}_j^2)}(\mathbf{X} - \mathbf{R}_j^* \mathbf{W}_j)\mathbf{\Lambda}_j^2 \mathbf{e}. \tag{12}$$

Note that ϕ_j aligns the GMM means $\{\mathbf{x}_k\}_{k=1}^K$ with the virtual points $\{\mathbf{w}_{jk}\}_{k=1}^K$. Hence, our method deals with point sets of different cardinalities and the number of components K in the GMM can be chosen independently of the cardinality of the point sets.

CM-step-B and CM-step-C: These steps estimate the GMM means and variances given the current estimates of the rigid transformations and of the posteriors. By setting $\partial f/\partial \mathbf{x}_k = 0$, $k = 1 \ldots, K$, we easily obtain the optimal means. Then, we replace these values and obtain optimal variances by setting $\partial f/\partial \sigma_k = 0$. This leads to the following formulas for the means and the variances

$$\mathbf{x}_k^* = \frac{\sum\limits_{j=1}^M \sum\limits_{i=1}^{N_j} \alpha_{jik}(\mathbf{R}_j^* \mathbf{v}_{ji} + \mathbf{t}_j^*)}{\sum\limits_{j=1}^M \sum\limits_{i=1}^{N_j} \alpha_{jik}}, \quad \sigma_k^{*2} = \frac{\sum\limits_{j=1}^M \sum\limits_{i=1}^{N_j} \alpha_{jik}\|\mathbf{R}_j^* \mathbf{v}_{ji} + \mathbf{t}_j^* - \mathbf{x}_k^*\|_2^2}{3\sum\limits_{j=1}^M \sum\limits_{i=1}^{N_j} \alpha_{jik}} + \epsilon^2, \tag{13}$$

with ϵ^2 being a very low positive value to efficiently avoid singularities [15].

CM-step-D: This step estimates the priors p_k. From (2) we obtain $\sum\limits_{k=1}^K p_k = 1/(1+\gamma)$. By neglecting the terms in (6) that do not depend on the priors and by using a Lagrange multiplier, the dual objective function becomes

$$f_L(p_1, \ldots, p_K, \mu) = \sum_{k=1}^K \left(\log p_k \sum_{i,j} \alpha_{jik} \right) + \mu \left(\sum_{k=1}^K p_k - \frac{1}{1+\gamma} \right). \tag{14}$$

Setting $\partial f_L/\partial p_k = 0$ yields the following optimal priors

$$p_k^* = \frac{\sum\limits_{j,i} \alpha_{jik}}{\mu}, \quad k = 1 \ldots K \quad \text{and} \quad p_{K+1}^* = 1 - \sum_{k=1}^K p_k^*, \tag{15}$$

with $\mu = (\gamma + 1)(N - \sum_{j,i} \alpha_{ji(K+1)})$ and $N = \sum_j N_j$ being the cardinality of $\mathbf{V}$.
Note that if $\gamma \to 0$, which means that there is no uniform component in the mixture, then $\mu \to N$, which is in agreement with [5]. Based on the pseudocode of Alg. 1, the above steps are iterated until a convergence criterion is met, e.g., a sufficient number of iterations or a bound on the improvement of $f(\Theta)$.

5   Experiments with Synthetic and Real Data

For quantitative evaluation, we experiment with the 3D models "Bunny", "Lucy" and "Armadillo" from the Stanford 3D scanning repository[1]. We use fully viewed models in order to synthesize multiple point sets, as follows. The model point coordinates are shifted at the origin, the points are downsampled and then rotated in the xz-plane; points with negative z coordinates are rejected. This way, only a part of the object is viewed in each set, the point sets do not fully overlap, and the extent of the overlap depends on the rotation angle, as in real scenarios. It is important to note that the downsampling is different for each set, such different points are present in each set and the sets have different cardinalities (between 1000 and 2000 points). We add Gaussian noise to point coordinates based on a predefined signal-to-noise ratio (SNR), and more importantly, we add outliers to each set which are uniformly distributed around five randomly chosen points of the set.

For comparison, we consider the 3D rigid registration algorithms ICP [4], CPD [23], ECMPR [15], GMMReg [18] and the simultaneous registration algorithm of [30] abbreviated here as SimReg. Note that CPD is exactly equivalent to ECMPR when it comes to rigid registration and that SimReg internally uses an ICP framework. Other than SimReg, the rest are pairwise registration schemes that register the first point set with each of the rest. Sequential ICP (seqICP) does the known register–then–integrate cycle. Although GMMReg is the version of [28] for two point sets, the authors provide the code only for the pairwise case. We choose GMMReg for comparison since, as showed in [18], Levenberg-Marquardt ICP [10] performs similarly with GMMReg, while [28] shows that GMMReg is superior to Kernel Correlation [26].

As far as the registration error is concerned, we use the root–mean–square error (RMSE) of rotation parameters since translation estimation is not challenging. For all algorithms, we implicitly initialise the translations by transferring the centroids of the point clouds into the same point, while identity matrices initialize the rotations. The only exception is the SimReg algorithm which fails without a good starting point, thus the transformations are initialized by pairwise ICP. GMMReg is kind of favored in the comparison, since it uses a two-level optimization process and the first level helps the algorithm to initialize itself. Notice that both SimReg and the proposed method provide rigid transformations for *every* point set, while ground rotations are typically expressed in terms

---

[1] https://graphics.stanford.edu/data/3Dscanrep/3Dscanrep.html

of the first set. Hence, the product of estimations $\hat{R}_1^\top \hat{R}_j$ is compared with the ground rotation R_j, and the error is $\|\hat{R}_1^\top \hat{R}_j - R_j\|_F$.

We consider a tractable case of jointly registering four point sets, the angle between the first set and the other sets being 10°, 20° and 30° respectively. Since JR-MPC starts from a completely unknown GMM, the initial means $\mathbf{x}_k$ are distributed on a sphere that spans the convex hull of the sets. The variances σ_k are initialized with the median distance between $\mathbf{x}_k$ and all the points in $\mathbf{V}$. For our experiments, we found that updating priors do not drastically improve the registration, thus we fix the priors equal to $1/(K+1)$ and $\gamma = 1/K$, while h is chosen to be the volume of a sphere whose radius is 0.5; the latter is not an arbitrary choice since the point coordinates are normalized by the maximum distance between points of the convex hull of $\mathbf{V}$. CPD and ECMPR deal with the outliers in the same way. The number of the components, K, is here equal to 60% of the mean cardinality. We use 100 iterations for all algorithms excepting GMMReg, whose implementation performs 10 and 100 function evaluations for the first and second optimization levels respectively.

Fig. 2 shows the final log-RMSE averaged over 100 realisations and all views as a function of outlier percentage for each 3D model. Apparently, ICP and SimReg are more affected by the presence of outliers owing to one-to-one correspondences. CPD and GMMReg are affected in the sense that the former assigns outliers to any of the GMM components, while the latter clusters together outliers. The proposed method is more robust to outliers and the registration is successful even with densely present outliers. The behavior of the proposed algorithm in terms of the outliers is discussed in detail below and showed on Fig. 4. To visualize the convergence rate of the algorithms, we show curves for a typical setting ($SNR = 10dB$ and 20% outliers). Regarding GMMReg, we just plot a line that shows the error in steady state. There is a performance variation as the model changes. "Lucy" is more asymmetric than "Bunny" and "Armadillo", thus a lower floor is achieved. Unlike the competitors, JR-MPC may show a minor perturbation in the first iterations owing to the joint solution and the random initialization of the means $\mathbf{x}_k$. However, the estimation of each transformation benefits from the proposed joint solution, in particular when the point sets contain outliers, and JR-MPC attains the lowest floor.

It is also important to show the estimation error between non overlapping sets. This also shows how biased each algorithm is. Based on the above experiment (SNR=10db, 20% outliers), Table 1 reports the average rotation error for the pairs (V_2, V_3) and (V_3, V_4), as well as the standard deviation of these two errors as a measure of bias. All but seqICP do not estimate these direct mappings. The proposed scheme, not only provides the lowest error, but it also offers the most symmetric solution.

A second experiment evaluates the robustness of the algorithms in terms of rotation angle between two point sets, hence the degree of overlap. This also allows us to show how the proposed algorithm deals with the simple case of two point sets. Recall that JR-MPC does not reduce to CPD/ECMPR in the two-set case, but rather it computes the poses of the two sets with respect to the
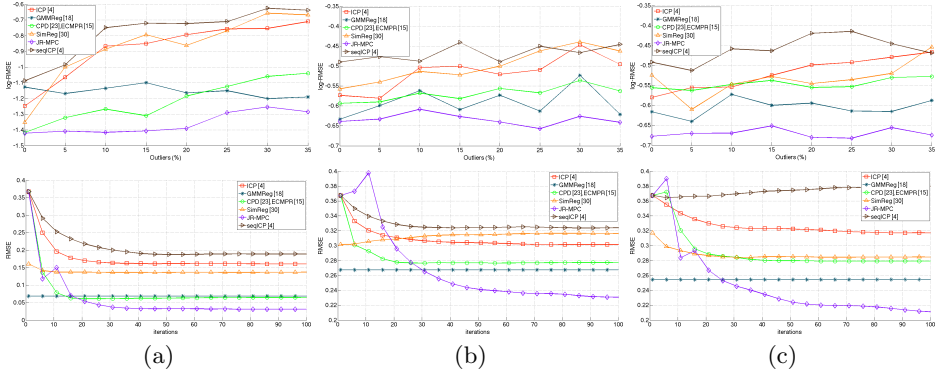
Fig. 2. *Top:* log-RMSE as a function of outlier percentage when SNR=10*dB*. *Bottom:* The learning curve of algorithms for a range of 100 iterations when the models are disturbed by SNR=10*dB* and 20% outliers. (a) "Lucy", (b) "Bunny" (c) "Armadillo".

	Bunny			Lucy			Armadillo		
ICP [4]	0.329	0.423	0.047	0.315	0.297	0.009	0.263	0.373	0.055
GMMReg [18]	0.364	0.303	0.030	0.129	0.110	0.009	0.228	0.167	0.031
CPD [23], ECMPR [15]	0.214	0.242	0.014	0.144	0.109	0.017	0.222	0.204	0.009
SimReg [30]	0.333	0.415	0.041	0.354	0.245	0.055	0.269	0.301	0.016
JR-MPC	0.181	0.165	**0.008**	0.068	0.060	**0.004**	0.147	0.147	**0.000**

Table 1. Registration error of indirect mappings. For each model, the two first columns show the rotation error of $V_2 \to V_3$ and $V_3 \to V_4$ respectively, while the third column shows the standard deviation of these two errors ($SNR = 10db$, 30% outliers).

"central" GMM. Fig. 3 plots the average RMSE over 50 realizations of "Lucy" and "Armadillo", when the relative rotation angle varies from -90° to 90°. As for an acceptable registration error, the proposed scheme achieves the widest and shallowest basin for "Lucy", and competes with GMMReg for "Armadillo". Since "Armadillo" consists of smooth and concave surface parts, the performance of the proposed scheme is better with multiple point sets than the two-set case, hence the difference with GMMReg. The wide basin of GMMReg is also due to its initialization.

As mentioned, a by-product of the proposed method is the reconstruction of an outlier-free model. In addition, we are able to detect the majority of the outlying points based on the variance of the component they most likely belong to. To show this effect, we use the results of one realization of the first experiment with 30% outliers. Fig 4 shows in (a) and (b) two out of four point sets, thereby one verifies the distortion of the point sets, as well as how different the sets may be, e.g., the right hand is missing in the first set. The progress of $\mathbf{x}_k$ estimation is shown in (d-f). Apparently, the algorithm starts by reconstructing the scene model (observe the presence of the right hand). Notice the size increment of the hull of the points $\mathbf{x}_k$, during the progress. This is because the posteriors in
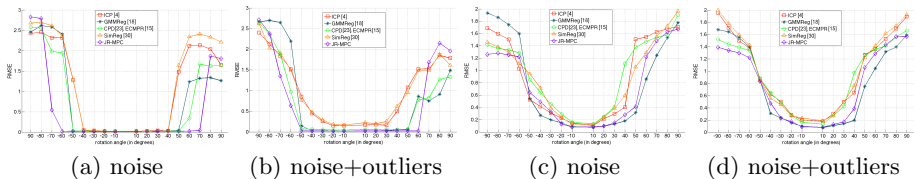
Fig. 3. RMSE as a function of the overlap (rotation angle) when two point sets are registered (SNR=20dB, 30% outliers) (a),(b) "Lucy" (c), (d) "Armadillo"

the first iteration are very low and make the means $\mathbf{x}_k$ shrink into a very small cell. While the two point sets are around the points $(0, 0, 0)$ and $(40, 40, 40)$, we build the scene model around the point $(5, 5, 5)$. The distribution of the final deviations σ_k is shown in (c). We get the same distribution with any model and any outlier percentage, as well as when registering real data. Although one can fit a pdf here, e.g., Rayleigh, it is convenient enough to split the components using the threshold $T_\sigma = 2 \times median(\mathcal{S})$, where $\mathcal{S} = \{\sigma_k | k = 1, \ldots K\}$. Accordingly, we build the scene model and we visualize the binary classification of points $\mathbf{x}_k$. Apparently, whenever components attract outliers, even not far from the object surface, they tend to spread their hull by increasing their scale. Based on the above thresholding, we can detect such components and reject points that are assigned with high probability to them, as shown in (g). Despite the introduction of the uniform component that prevents the algorithm from building clusters away from the object surface, locally dense outliers are likely to create components outside the surface. In this example, most of the point sets contain outliers above the shoulders, and the algorithm builds components with outliers only, that are post-detected by their variance. The integrated surface is shown in (h) and (i) when "bad" points were automatically removed. Of course, the surface can be post-processed, e.g., smoothing, for a more accurate representation, but this is beyond of our goal.

We report here CPU times obtained with unoptimized Matlab implementations of the algorithms. ICP, CPD (ECMPR), SimReg, and JR-MPC require 14.7s, 40.6s, 24.6s, and 20.9s respectively to register four point sets of 1200 points, on an average. The C++ implementation of GMMReg requires 6.7s. JR-MPC runs faster than repeating CPD(ECMPR) since only one GMM is needed and the number of components is less than the number of points. Of course, ICP is the most efficient solution. However, SimReg needs more time as it enables every pair of overlapping point sets.

We also tested our method with real depth data captured from a time-of-flight (TOF) camera that is rigidly attached to two color cameras. Once calibrated [13, 12], this sensor provides 3D point clouds with associated color information. We gathered ten point clouds by manually moving the sensor in front of a scene, e.g., Fig. 5. Multiple-set registration was performed with all the above methods. While only depth information is used for the registration, the use of color information helps the final assessment and also shows the potential for fusing RGB-D data.
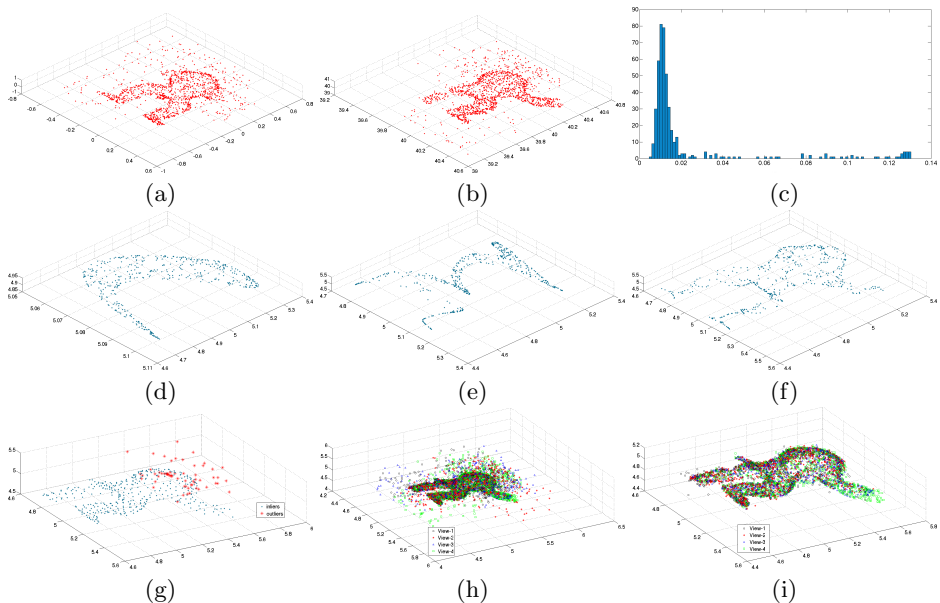
Fig. 4. (a),(b) Two point sets (out of four) with outliers; (c) distribution of estimated variances; instances of GMM means after (d) 5, (e) 15, and (f) 30 iterations; (g) the splitting of model points into inliers and outliers; joint-registration of four point sets (h) before and (i) after removing "bad" points (*best viewed on-screen*).

Fig. 5 shows the results obtained with JR-MPC before (a) and after (b) rejecting outliers, seqICP (c) and SimReg (d). The proposed method successfully register the point clouds, while it automatically removes most of the jump-edge errors contained in range images. SimReg registers the majority of point sets, but it fails to register a few sets that appear flipped in the integrated view. Using the 5-th set as a reference for symmetry reasons, CDP/ECMPR and ICP also fail to register all the clouds while GMMReg yields low performance with too many misalignments. SeqICP causes weak misalignments, since it estimates weak geometric deformations between successive captures. However, the registration is not very accurate and further processing may be necessary, e.g., [17]. We refer the reader to the supplementary material for the integrated set of all the algorithms, viewed by several viewpoitns.

6   Conclusions

We presented a probabilistic generative model and its associated algorithm to jointly register multiple point sets. The vast majority of state-of-the-art techniques select one of the sets as the model and attempt to align the other sets onto this model. However, there is no guarantee that the model set is free of noise and outliers and this contaminates the estimation of the registration parameters.
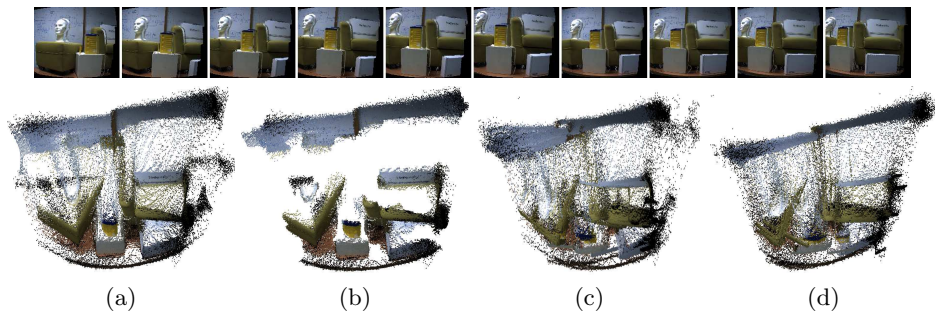
(a)	(b)	(c)	(d)

Fig. 5. The integrated point clouds from the joint registration of 10 TOF images that record a static scene. *Top*: color images that roughly show the scene content of each range image (occlusions due to cameras baseline cause some texture artefacts). *Bottom:* top-view of joint registration obtained from (a) JR-MPC, (b) JR-MPC+outlier rejection, (c) sequential ICP and (d) SimReg.

Unlike previous work, the proposed method treats all the point sets on an equal footing: they are realizations of a GMM and the registration is cast into a clustering problem. We formally derive an expectation-maximization algorithm that estimates the GMM parameters as well as the rotations and translations between each individual set and a "central" model. In this model the GMM means play the role of the registered points and the variances provide rich information about the quality of the registration. We thoroughly validated the proposed method with challenging datasets, we compared it with several state-of-the-art methods, and we showed its potential for fusing real depth data.

Supplementary material. Datasets, code and videos are publicly available at https://team.inria.fr/perception/research/jrmpc/

References

1. Banfield, J.D., Raftery, A.E.: Model-based Gaussian and non-Gaussian clustering. Biometrics 49(3), 803–821 (1993)
2. Bartoli, A., Pizzaro, D., Loog, M.: Stratified generalized procrustes analysis. IJCV 101(2), 227–253 (2013)
3. Bergevin, R., Soucy, M., Gagnon, H., Laurendeau, D.: Towards a general multi-view registration technique. IEEE-TPAMI 18(5), 540–547 (1996)
4. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE-TPAMI 14, 239–256 (1992)
5. Bishop, C.: Pattern Recognition and Machine Learning. Springer (2006)
6. Blais, G., D. Levine, M.: Registering multiview range data to create 3d computer objects. IEEE-TPAMI 17(8), 820–824 (1995)
7. Castellani, U., Fusiello, A., Murino, V.: Registration of multiple acoustic range views for underwater scene reconstruction. CVIU 87(1-3), 78–89 (2002)
8. Chen, Y., Medioni, G.: Object modelling by registration of multiple range images. IVC 10(3), 145–155 (1992)

9. Chui, H., Rangarajan, A.: A new point matching algorithm for non-rigid registration. CVIU 89(2-3), 114–141 (2003)
10. Fitzgibbon, A.W.: Robust registration of 2D and 3D point sets. IVC 21(12), 1145–1153 (2001)
11. Granger, S., Pennec, X.: Multi-scale EM-ICP: A fast and robust approach for surface registration. In: ECCV (2002)
12. Hansard, M., Horaud, R., Amat, M., Evangelidis, G.: Automatic detection of calibration grids in time-of-flight images. CVIU 121, 108–118 (2014)
13. Hansard, M., Horaud, R., Amat, M., Lee, S.: Projective alignment of range and parallax data. In: CVPR (2011)
14. Hermans, J., Smeets, D., Vandermeulen, D., Suetens, P.: Robust point set registration using em-icp with information-theoretically optimal outlier handling. In: CVPR (2011)
15. Horaud, R., Forbes, F., Yguel, M., Dewaele, G., ZhangI, J.: Rigid and articulated point registration with expectation conditional maximization. IEEE-TPAMI 33(3), 587–602 (2011)
16. Huber, D.F., Hebert, M.: Fully automatic registration of multiple 3d data sets. IVC 21(7), 637–650 (2003)
17. Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., Fitzgibbon, A.: Kinectfusion: Real-time 3d reconstruction and interaction using a moving depth camera. In: ACM Symposium on UIST (2011)
18. Jian, B., Vemuri, B.C.: Robust point set registration using gaussian mixture models. IEEE-TPAMI 33(8), 1633–1645 (2011)
19. Krishnan, S., Lee, P.Y., Moore, J.B.: Optimisation-on-a-manifold for global registration of multiple 3d point sets. Int. J. Intelligent Systems Technologies and Applications 3(3/4), 319–340 (2007)
20. Masuda, T., Yokoya, N.: A robust method for registration and segmentation of multiple range images. CVIU 61(3), 295–307 (1995)
21. Mateo, X., Orriols, X., Binefa, X.: Bayesian perspective for the registration of multiple 3d views. CVIU 118, 84 – 96 (2014)
22. Meng, X.L., Rubin, D.B.: Maximum likelihood estimation via the ECM algorithm: a general framework. Biometrika 80, 267–278 (1993)
23. Myronenko, A., Song, X.: Point-set registration: Coherent point drift. IEEE-TPAMI 32(12), 2262–2275 (2010)
24. Newcombe, R.A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A.J., Kohli, P., Shotton, J., Hodges, S., Fitzgibbon, A.: Kinectfusion: Real-time dense surface mapping and tracking. In: IEEE ISMAR (2011)
25. Torsello, A., Rodol, E., A, A.: Multiview registration via graph diffusion of dual quaternions. In: CVPR (2011)
26. Tsin, Y., Kanade, T.: A correlation-based approach to robust point set registration. In: ECCV (2004)
27. Umeyama, S.: Least-squares estimation of transformation parameters between two point patterns. IEEE-TPAMI 13(4), 376–380 (1991)
28. Wang, F., Vemuri, B.C., Rangarajan, A., Eisenschenk, S.J.: Simultaneous nonrigid registration of multiple point sets and atlas construction. IEEE-TPAMI 30(11), 2011–2022 (2008)
29. Wells III, W.: Statistical approaches to feature-based object recognition. IJCV 28(1/2), 63–98 (1997)
30. Williams, J., Bennamoun, M.: Simultaneous registration of multiple corresponding point sets. CVIU 81(1), 117–142 (2001)