Density Adaptive Point Set Registration

Felix Järemo Lawin, Martin Danelljan, Fahad Shahbaz Khan, Per-Erik Forssén, Michael Felsberg
Computer Vision Laboratory, Department of Electrical Engineering, Linköping University, Sweden
{felix.jaremo-lawin, martin.danelljan, fahad.khan, per-erik.forssen, michael.felsberg}@liu.se

This document is a supplement to the CVPR 2018 paper "Density Adaptive Point Set Registration". Here we provide derivations of the proposed EM procedure and the observation weight function, based on both the empirical estimates and the sensor model. We also provide an evaluation of re-sampling methods and an analysis of the parameters introduced by our proposed model. Further, we present additional results and examples.

1. Derivation of EM procedure

We will here describe how the proposed objective in equation (7) in the paper, can be maximized using Expectation Maximization. To simplify the derivation, we first study the maximization of a single term i in the objective (7) in the paper and drop the index i to avoid clutter. We denote the likelihood in the local coordinate system of the point set as $p_X(x|\Theta) = p_X(x|\omega, \theta) = p_V(\phi(x;\omega)|\theta)$. The objective is then to maximize,

$$\mathcal{E}(\Theta) = \int_{\mathbb{R}^3} \log\left(p_V(\phi(x;\omega)|\theta)\right) \frac{q_V(\phi(x;\omega))}{q_X(x)} q_X(x)\,\mathrm{d}x = \int_{\mathbb{R}^3} \log\left(p_X(x|\Theta)\right) \frac{q_V(\phi(x;\omega))}{q_X(x)} q_X(x)\,\mathrm{d}x\,. \tag{1}$$

Here, $f(x) = \frac{q_V(\phi(x;\omega))}{q_X(x)}$ is the observation weight function. We now introduce a latent random variable $Z \in \{1,\dots,K\}$ that assigns a 3D-point V to a particular mixture component k. Using the relation $\log p_X(x|\Theta) = \log(p_{Z,X}(k,x|\Theta)) - \log(p_{Z|X}(k|x,\Theta))$, equation (1) can be written as

$$\mathcal{E}(\Theta) = \int_{\mathbb{R}^3} \log\left(p_{Z,X}(k,x|\Theta)\right) f(x) q_X(x)\,\mathrm{d}x + \int_{\mathbb{R}^3} -\log\left(p_{Z|X}(k|x,\Theta)\right) f(x) q_X(x)\,\mathrm{d}x \tag{2}$$

Following the derivation in [2] we first take the expectation of both sides in equation (2) with respect to a distribution $Z \sim \tilde{p}(k)$. We then add and subtract $\int_{\mathbb{R}^3} \sum_k \tilde{p}(k) \log\left(\tilde{p}(k)\right) f(x) q_X(x)\,\mathrm{d}x$ on the right hand side of (2). Since the left hand side does not depend on Z, and $\tilde{p}(k)$ sums to one we get:

$$\begin{aligned}\mathcal{E}(\Theta) &= \int_{\mathbb{R}^3} \underbrace{\sum_{k=1}^{K} \tilde{p}(k) \log\left(\frac{p_{Z,X}(k,x|\Theta)}{\tilde{p}(k)}\right)}_{=:\mathcal{L}(\tilde{p},\Theta)} f(x) q_X(x)\,\mathrm{d}x + \int_{\mathbb{R}^3} \underbrace{-\sum_{k=1}^{K} \tilde{p}(k) \log\left(\frac{p_{Z|X}(k|x,\Theta)}{\tilde{p}(k)}\right)}_{=\mathrm{KL}(\tilde{p}||p)} f(x) q_X(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^3} \mathcal{L}(\tilde{p},\Theta) f(x) q_X(x)\,\mathrm{d}x + \int_{\mathbb{R}^3} \mathrm{KL}(\tilde{p}||p) f(x) q_X(x)\,\mathrm{d}x\,,\end{aligned} \tag{3}$$

where $\mathrm{KL}(\tilde{p}||p)$ is the KL divergence from $\tilde{p}$ to the posterior distribution $p_{Z|X}(k|x,\Theta)$. We know that $\mathrm{KL}(\tilde{p}||p) \geq 0$ with equality if and only if $\tilde{p}(k) = p_{Z|X}(k|x,\Theta)$. Hence, $\int_{\mathbb{R}^3} \mathcal{L}(\tilde{p},\Theta) f(x) q_X(x)\,\mathrm{d}x$ is a lower bound of $\mathcal{E}(\Theta)$. In the E-step we maximize the lower bound by setting $\tilde{p} = p$ given the current parameters Θ^n. This leads to equality between the objective $\mathcal{E}(\Theta)$ and the lower bound $\mathcal{Q}(\Theta,\Theta^n)$ at the current parameter estimate $\mathcal{E}(\Theta^n) = \mathcal{Q}(\Theta^n,\Theta^n)$. Here, $\mathcal{Q}(\Theta,\Theta^n)$ is the lower bound obtained with $\tilde{p}(k) = p_{Z|X}(k|x,\Theta^n)$,

$$\begin{aligned}\mathcal{Q}(\Theta,\Theta^n) &= \int_{\mathbb{R}^3} \sum_{k=1}^{K} p_{Z|X}(k|x,\Theta^n) \log\left(p_{Z,X}(k,x|\Theta)\right) f(x) q_X(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^3} E_{Z|x,\Theta^n}\left[\log\left(p_{X,Z}(x,Z|\Theta)\right)\right] f(x) q_X(x)\,\mathrm{d}x\end{aligned} \tag{4}$$

In the M-step we maximize the lower bound (4) with respect to Θ to update the parameters $\Theta^{n+1} = \arg\max_\Theta \mathcal{Q}(\Theta, \Theta^n)$.

Since the exact value of the integral in (4) is intractable, we treat the observations x_j as a Monte Carlo sampling. This results in an approximation of the lower bound,

$$\mathcal{Q}(\Theta, \Theta^n) \approx Q(\Theta, \Theta^n) = \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{K} \alpha_{jk}^n f(x_j) \log\left(p_{X,Z}(x_j, k|\Theta)\right) , \tag{5}$$

where $\alpha_{jk}^n = p_{Z|X}(k|x_j, \Theta^n)$. As described in the paper, with independent point samples we obtain the latent posteriors as

$$p_{Z|X}(k|x, \Theta) = \frac{p_{X,Z}(x, k|\Theta)}{\sum_{k=1}^{K} p_{X,Z}(x, k|\Theta)} = \frac{\pi_k \mathcal{N}(\phi(x; \omega_k); \mu_k, \Sigma_k)}{\sum_{l=1}^{K} \pi_l \mathcal{N}(\phi(x; \omega_l); \mu_l, \Sigma_l)} . \tag{6}$$

Maximizing the lower bound will cause $\mathcal{E}(\Theta)$ to increase unless it is at a maximum. Note that, during the whole procedure described above, we only evaluated the observation weight function $f(x)$ at the Monte Carlo sampling of $\mathcal{Q}(\Theta, \Theta^n)$. Although, $f(x)$ affects the maximization of $\mathcal{Q}(\Theta, \Theta^n)$, we see that it does not influence the derivation of the EM algorithm.

The derivation can trivially be generalized to multiple point sets. In this case, the M-step extends to,

$$\mathcal{Q}(\Theta, \Theta^n) \approx Q(\Theta, \Theta^n) = \sum_{i=1}^{M} \frac{1}{N_i} \sum_{j=1}^{N_i} \sum_{k=1}^{K} \alpha_{ijk}^n f_i(x_{ij}) \log\left(p_{V,Z}(\phi(x_{ij}; \omega_i), k|\theta)\right) . \tag{7}$$

We then apply the optimization procedure proposed in [6] to maximize (7).

2. Derivation of sensor model

Here, we derive the expression for the sensor model described in section 3.4.1 in the paper. We denote the measures $q_V(A) = \mathbb{P}(V \in A)$ and $q_{X_i}(A) = \mathbb{P}(V_i \in A)$ for the latent scene distribution q_V and the sampling density q_{X_i} respectively. By changing the variables in the integral to the reference frame $v = \phi_i(x)$, the measure $q_{X_i}(A)$ can be written as,

$$q_{X_i}(A) = \int_{S_i \cap A} \frac{g_i}{|S|} \,\mathrm{d}S_i = \int_{\phi_i(S_i \cap A)} g_i \circ \phi_i^{-1} \frac{\mathrm{d}S}{|S|} = \int_{S \cap \phi_i(A)} g_i \circ \phi_i^{-1} \frac{\mathrm{d}S}{|S|} . \tag{8}$$

Here, we have used the fact that $\phi_i(x) = R_i x + t_i$ is an isometric bijection.

From the definition $q_V(A) = \frac{1}{|S|} \int_{S \cap A} \mathrm{d}S$, we see that $\frac{\mathrm{d}q_V}{\mathrm{d}S} = \frac{1}{|S|}$ on S and zero elsewhere. From (8) we thus obtain,

$$q_{X_i}(A) = \int_{\phi_i(A)} g_i \circ \phi_i^{-1} \frac{\mathrm{d}q_V}{\mathrm{d}S} \mathrm{d}S = \int_{\phi_i(A)} g_i \circ \phi_i^{-1} \mathrm{d}q_V = \int_A g_i \,\mathrm{d}(q_V \circ \phi_i) . \tag{9}$$

In the last equality in (9), we have performed another change of variables back to the sensor-based coordinate frame. Here, $q_V \circ \phi_i(A) = q_V(\phi_i(A))$ denotes the composed measure.

Since, q_V is derived from the Lebesgue measure on the surface dS, it is σ-finite [4]. Furthermore, we also see that $q_V \circ \phi_i$ is absolutely continuous with respect to q_{X_i}, since the definition of g_i (eq. (12) in the paper) guarantees that $q_{X_i}(A) = 0 \implies q_V \circ \phi_i(A) = 0$. We also see that $q_{X_i}(A)$ is σ-finite since g_i is bounded everywhere except for the singularity in the origin of the sensor reference frame (the sensor position). The singularity of g_i in the origin is not a problem if we assume that S_i do not intersect the sensor center. From (9) we obtain the Radon-Nikodym derivative $\frac{\mathrm{d}q_{X_i}}{\mathrm{d}(q_V \circ \phi_i)} = g_i$. We can thus conclude, from the properties of the Radon-Nikodym derivative [4], that $f_i = \frac{\mathrm{d}(q_V \circ \phi_i)}{\mathrm{d}q_{X_i}} = \frac{1}{g_i}$.

3. Derivation of empirical sample model

Here, we add more details to the description of the empirical sensor model in section 3.4.2. As it is sufficient to study a single point set, we drop the index i. The observation weight function that we estimate is $f(x) = \frac{q_V(\phi(x))}{q_X(x)}$. In the local neighborhood of a point on a surface $\bar{v} \in S$ we approximate the latent scene distribution as a Gaussian distribution in the normal direction

$$q_V(v) \approx \frac{1}{|S|} \mathcal{N}(\hat{n}_{\bar{v}}^{\mathrm{T}}(v - \bar{v}); 0, \sigma_{\hat{n}}^2(\bar{v})) . \tag{10}$$

Assuming a rigid transformation $v = \phi(x) = Rx + t$ we can write (10) as

$$q_V(\phi(x)) \approx \frac{1}{|S|}\mathcal{N}(\hat{n}^{\mathrm{T}}_{\phi(\bar{x})}(\phi(x) - \phi(\bar{x})); 0, \sigma^2_{\hat{n}}(\phi(\bar{x}))) = \frac{1}{|S|}\mathcal{N}(\underbrace{(R^{\mathrm{T}}\hat{n}_{\phi(\bar{x})})}_{\hat{n}_x}{}^{\mathrm{T}}(x - \bar{x}); 0, \sigma^2_{\hat{n}}(\phi(\bar{x}))) = \tag{11}$$

$$= \frac{1}{|S|\sqrt{2\pi\sigma^2_{\hat{n}_{\bar{x}}}(\bar{x})}} e^{-\frac{1}{2\sigma^2_{\hat{n}_{\bar{x}}}(\bar{x})}(x-\bar{x})^{\mathrm{T}}\hat{n}_x\hat{n}_x^{\mathrm{T}}(x-\bar{x})}, \tag{12}$$

where $\hat{n}_{\bar{x}}$ is the surface normal vector at $\bar{x}$ in the reference frame of the point set $\mathcal{X}$.

For each 3D point x in the point cloud we approximate the sampling density as a Gaussian distribution from the L nearest neighbors

$$q_X(x) \approx \frac{L}{N}\mathcal{N}(x; \bar{x}, C)\,. \tag{13}$$

Here, the covariance matrix is calculated as $C = \frac{1}{L-1}\sum_j (x_j - \bar{x})(x_j - \bar{x})^{\mathrm{T}}$ and $\bar{x}$ is the mean of the L nearest neighbors. We now perform an eigenvalue decomposition of C to obtain

$$C = BDB^{\mathrm{T}} = (\hat{b}_1, \hat{b}_2, \hat{b}_3)\begin{pmatrix}\sigma_1^2 & 0 & 0\\ 0 & \sigma_2^2 & 0\\ 0 & 0 & \sigma_3^2\end{pmatrix}(\hat{b}_1, \hat{b}_2, \hat{b}_3)^{\mathrm{T}} = \sum_{i=1}^{3}\sigma_i^2\hat{b}_i\hat{b}_i^{\mathrm{T}} \quad \text{and} \tag{14}$$

$$C^{-1} = (\hat{b}_1, \hat{b}_2, \hat{b}_3)\begin{pmatrix}\sigma_1^{-2} & 0 & 0\\ 0 & \sigma_2^{-2} & 0\\ 0 & 0 & \sigma_3^{-2}\end{pmatrix}(\hat{b}_1, \hat{b}_2, \hat{b}_3)^{\mathrm{T}} = \sum_{i=1}^{3}\sigma_i^{-2}\hat{b}_i\hat{b}_i^{\mathrm{T}}\,, \tag{15}$$

where the eigenvalues are sorted in descending order. The sample distribution in (13) can be expanded to

$$q_X(x) = \frac{L}{N\sqrt{2\pi\det C}}e^{-\frac{1}{2}(x-\bar{x})^{\mathrm{T}}C^{-1}(x-\bar{x})} = \frac{L}{N\sqrt{2\pi\sigma_1^2\sigma_2^2\sigma_3^2}}e^{-\frac{1}{2}(x-\bar{x})^{\mathrm{T}}\left(\sum_{i=1}^{3}\sigma_i^{-2}\hat{b}_i\hat{b}_i^{\mathrm{T}}\right)(x-\bar{x})}\,. \tag{16}$$

The eigenvector $\hat{b}_3$ corresponding to the smallest eigenvalue σ_3 approximates the surface normal, and the squared eigenvalue corresponds to the variance in this direction. We set $\hat{n}_{\bar{x}} = \hat{b}_3$ and $\sigma^2_{\hat{n}}(\bar{v}) = \sigma_3^2$. By inserting this into (11) we get

$$q_V(\phi(x)) \approx \frac{1}{|S|\sqrt{2\pi\sigma_3^2}}e^{-\frac{1}{2}(x-\bar{x})^{\mathrm{T}}\sigma_3^{-2}\hat{b}_3\hat{b}_3^{\mathrm{T}}(x-\bar{x})}\,, \tag{17}$$

The observation weight function can now be calculated from (16) and (17)

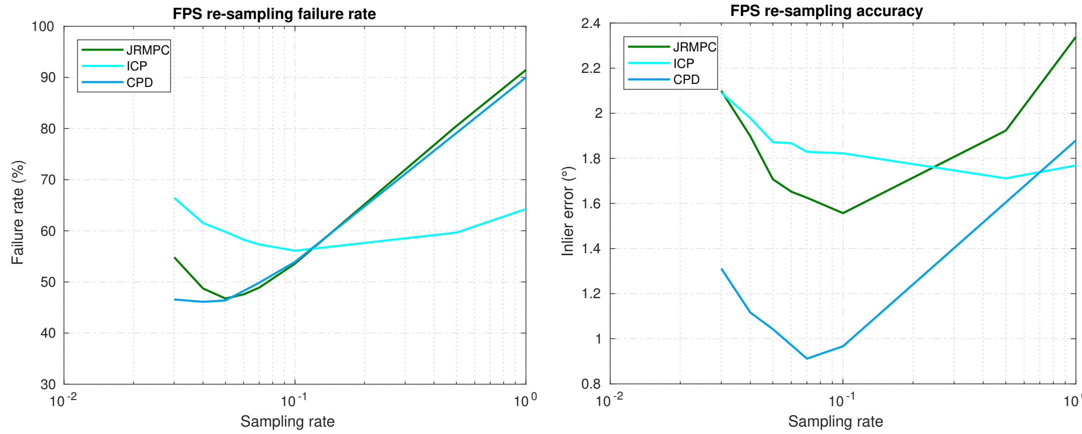
$$\begin{aligned} f(x) = \frac{q_V(\phi(x))}{q_X(x)} &\approx \frac{N\sqrt{2\pi\sigma_1^2\sigma_2^2\sigma_3^2}}{L|S|\sqrt{2\pi\sigma_3^2}}e^{\frac{1}{2}(x-\bar{x})^{\mathrm{T}}\left(\sum_{i=1}^{3}\sigma_i^{-2}\hat{b}_i\hat{b}_i^{\mathrm{T}}\right)(x-\bar{x})-\frac{1}{2}(x-\bar{x})^{\mathrm{T}}\sigma_3^{-2}\hat{b}_3\hat{b}_3^{\mathrm{T}}(x-\bar{x})} = \\ &= \frac{N\sigma_1\sigma_2}{L|S|}e^{\frac{1}{2}(x-\bar{x})^{\mathrm{T}}\left(\sum_{i=1}^{2}\sigma_i^{-2}\hat{b}_i\hat{b}_i^{\mathrm{T}}\right)(x-\bar{x})} = \frac{N\sigma_1\sigma_2}{L|S|}e^{\frac{1}{2}(x-\bar{x})^{\mathrm{T}}\left(\sum_{i=1}^{2}\sigma_i^{-2}\hat{b}_i\hat{b}_i^{\mathrm{T}}\right)(x-\bar{x})} \\ &\propto \sigma_1\sigma_2 e^{\frac{1}{2}(x-\bar{x})^{\mathrm{T}}\left(\sum_{i=1}^{2}\sigma_i^{-2}\hat{b}_i\hat{b}_i^{\mathrm{T}}\right)(x-\bar{x})}\,. \end{aligned} \tag{18}$$

This is equivalent to equation (13) in the paper.

4. Experiments

In this section we present more detailed results and analysis, complementary to the experiments provided in the main paper. We provide an evaluation of different re-sampling methods at varying sampling rate. Further, we analyze the impact of the parameters introduced by our proposed model. We present additional results for both pairwise and multi-view registration.

For quantitative comparison between different methods we provide recall curves for both pairwise and multi-view registration. In addition to rotation error recall curves (see section 4.1 in the main paper for description), we also provide recall curves for the translation error. The translation error is calculated as the Euclidean norm of the difference between the ground truth and the found translation.



(a) Failure rate (b) Inlier error

Figure 1. Failure rates (a) and average inlier error (b) as a function of the sample rate using the FPS re-sampling method.

4.1. Re-sampling evaluation

In the main paper we compare our approach to existing methods with re-sampling preprocessing. However, the selection of re-sampling technique is cumbersome and depends both on the dataset and the registration method. We evaluate JRMPC [6], ICP [1] and CPD [8], using FPS [5], GSS [7] and the voxel grid re-sampling at different sampling rates. The evaluation is performed for pairwise registration on the same dataset as in the main paper, with a reduced number of pairs compared to the experiment in the main paper. This includes the facade and office datasets from TLS ETH and the indoor and outdoor dataset from VPS. In figure 1-3, we present the performance both in terms of failure rate and mean inlier error as a function over the sampling rate.

Figure 1 shows the performance of the registration methods using FPS re-sampling for different sampling rates. For JRMPC the robustness increases as the sampling rate decreases, with the lowest failure rate at sampling rate 0.05 (see figure 1a). In 1b we see that the average inlier error increases as the sampling rate is significantly reduced. The best performance gains are observed for CPD, both in terms of robustness and accuracy, while ICP is only marginally affected.

Further, figure 2 shows the robustness and accuracy for the registration methods using voxel grid re-sampling. We varied the voxel side length between 0 to 1.5 meters. For JRMPC we observe similar performance as in the FPS case, with the lowest failure rate at a voxel side length of 1.0 meters. As in the FPS case, the accuracy is degrading for low sampling rate (e.g large voxel side length). CPD strongly benefited from the voxel grid re-sampling for large voxel sizes, with the best performance recorded at 1.25 meters.

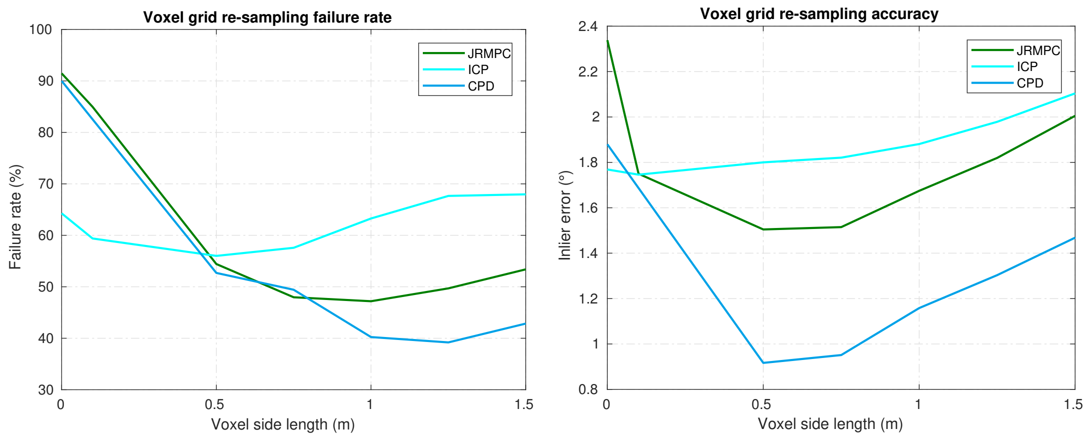
Finally, the GSS re-sampling method did not improve the robustness for JRMPC and CPD as we see in figure 3. With a sampling rate at 0.5, we observe a small improvement in robustness for ICP.

From the results of the re-sampling evaluation we deduce that both the choice of sampling rate and the re-sampling approach have a significant impact on the performance of point set registration methods. The best performing re-sampling setting for JRMPC is FPS with a sampling rate of 0.05. For CPD and ICP the best performance is achieved using voxel grid re-sampling, with voxel side length 1.25 and 0.5 meters respectively. In the following experiments, we denote the methods using these re-sampling settings as the empirical upper bounds.

4.2. Parameter analysis

Our density adaptive approach introduces two additional parameters to the JRMPC framework. First, we introduce a threshold T, which is used for clipping all observation weights larger than T times the mean of the weights. This way we reduce the impact of potential outliers. Second, we introduce L, which is the number of neighbors used for calculating the empirical observation weights.

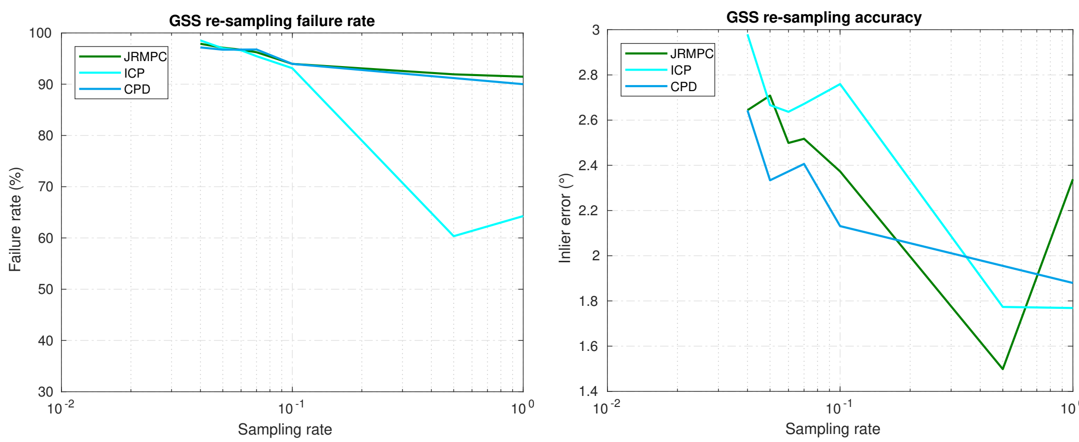
Table 1 shows the performance of our method in terms of average inlier error and failure rate for different values on T and L. The experiment was performed for pairwise registration on the facade and office datasets from TLS ETH, the VPS indoor dataset, and the VPS outdoor dataset, with a reduced number of pairs compared to the experiment in the main paper. We see that by increasing T the robustness is slightly improved, to the cost of lower accuracy. By increasing the number of neighbors L the accuracy is improved, to the cost of reduced robustness. We set $T = 8$ and $L = 10$ for the experiments in



(a) Failure rate

(b) Inlier error

Figure 2. Failure rates (a) and average inlier error (b) as a function of grid side length using the voxel grid re-sampling method.



(a) Falure rate

(b) Inlier error

Figure 3. Failure rates (a) and average inlier error (b) as a function of the sample rate using the GSS re-sampling method.

T	L	avg inlier err	failure rate
6	10	1.49	43.8
8	10	1.49	42.3
10	10	1.56	41.3
12	10	1.59	42.0
8	5	1.53	42.0
8	20	1.46	42.8
8	40	1.45	43.5

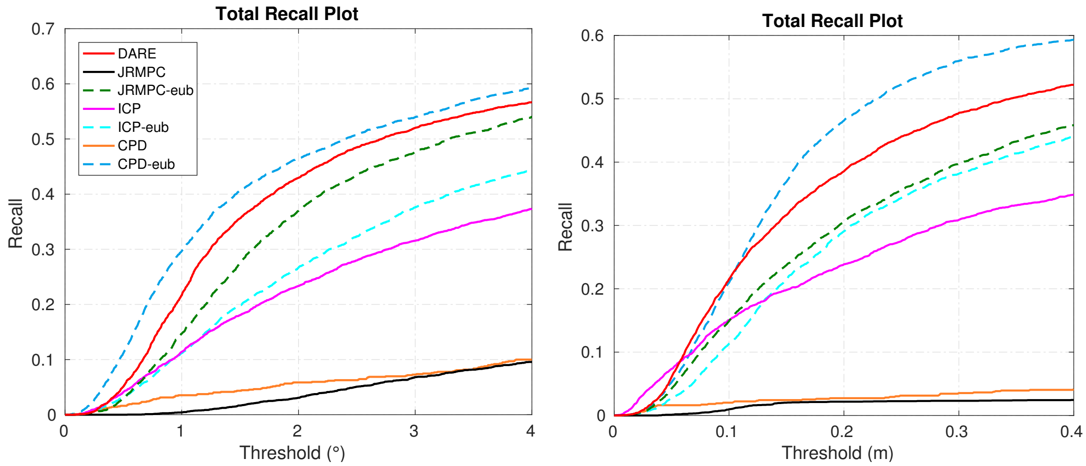
Table 1. Analysis of the parameters T and L for our DARE method. By increasing T the robustness is slightly improved, to the cost of lower accuracy. By increasing the number of neighbors L the accuracy can be improved, at the cost of reduced robustness. Overall, our method is not sensitive to the values of these parameters.

the main paper.

In the experiments we have observed that our DARE approach occasionally benefits from re-sampling when T is set to a lower value. This is expected, since a lower T-value leads to a behavior more similar to JRMPC.

4.3. Pairwise Registration

In table 1 in the main paper we present numerical values for the methods in terms of average inlier angular error and the failure rate on the combined VPS and TLS ETH dataset. In figure 4 we provide the corresponding recall plot. For all



(a) Rotation recall

(b) Translation recall

Figure 4. Recall curves with respect to (a) the angular error and (b) translation error on the joint facade and office dataset in TLS ETH and the indoor and outdoor dataset in VPS. Our DARE method shows improved results with respect to the baseline JRMPC, with and without re-sampling both in terms of rotation and the translation error.

methods, the results are also presented using empirically optimized re-sampling for each method from the experiment in section 4.1. Methods using empirically optimized re-sampling denoted by adding *-eub* in the method name. We see that our density adaptive method consistently improves over the baseline JRMPC with and without re-sampling, both in with respect to the rotation and translation error.

Complementary to the results on the combined VPS and TLS ETH dataset, we present recall plots for each dataset separately. We provide results for the facade dataset in TLS ETH, office dataset in TLS ETH, VPS indoor dataset, and VPS outdoor dataset. Our approach (DARE) is compared with the following methods: JRMPC [6], ICP [1] and CPD [8]. The recall curves for the separated datasets are summoned in figure 4.3.

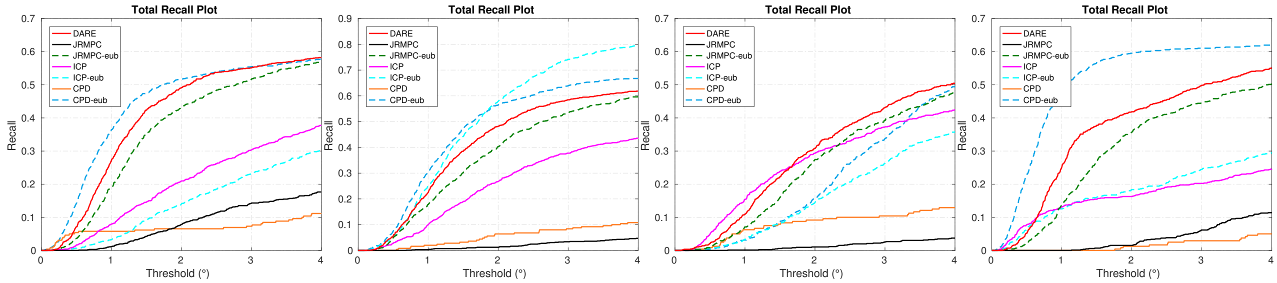
Figure 6 shows a qualitative comparison of our approach, when performing pairwise registration, on the facade dataset in TLS ETH. The baseline JRMPC method fails to register the point clouds. Our approach successfully performs the registration task on this dataset. We also provide a situation where our approach fails to align the point sets. Figure 7 shows a pairwise registration example on the VPS outdoor dataset. In this example, the point sets had a very limited overlap due to the placement of the Lidar sensor during acquisition. As we can see in figure 7, both the baseline JRMPC and our approach struggle, since none of the methods is designed to handle these extreme cases.

Another failure case for our DARE method is shown in figure 8. In this example the registration procedure has converged in a local minimum, where two of the aligned walls in the office room are shifted. This partly explains the reduced recall with respect to the translation error in comparison to the recall with respect to the rotation error, seen in figure 4.

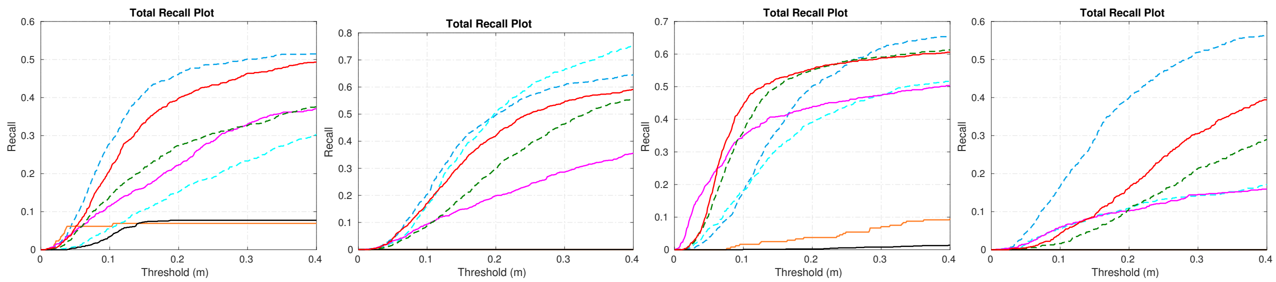
4.4. Multi-view registration

Finally, we present results, in case of multi-view registration, on the VPS indoor dataset. We compare our DARE approach with JRMPC and our color based approach DARE-color to CPPSR [3]. We also compare our approaches to JRMPC and CPPSR with the optimal re-sampling settings from 4.1.

The evaluation is performed over 500 registrations. The ground-truth is generated by first selecting a random rotation axis for each point set. We then rotate the point sets with rotation angles within 0-45 degrees around the rotation axes and apply random translations drawn from a multivariate normal distribution with standard deviation 1.0 in all directions. Figures 9 and 10 shows recall curves on the VPS indoor and outdoor datasets respectively. Our DARE method provides improved registration results compared to both CPPSR and JRMPC.

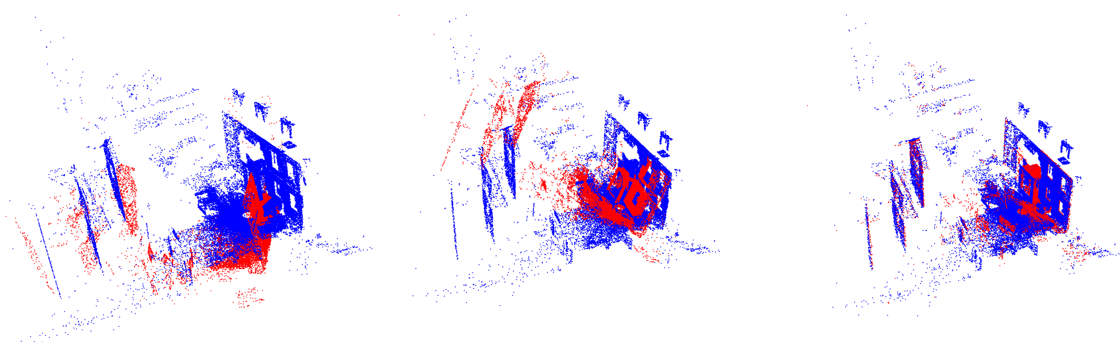


(a) **Facade**: Rotation recall (b) **Office**: Rotation recall (c) **VPS indoor**: Rotation recall (d) **VPS outdoor**: Rotation recall



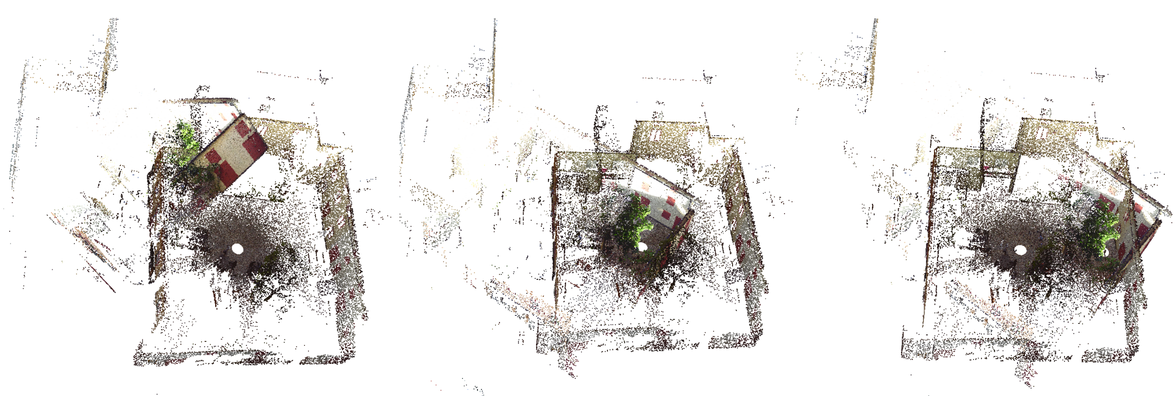
(e) **Facade**: Translation recall (f) **Office**: Translation recall (g) **VPS indoor**: Translation recall (h) **VPS outdoor**: Translation recall

Figure 5. Recall curves for the individual dataset with respect to the rotation error (top row) and the translation error (bottom row). Our DARE method shows improved results with respect to rotation and translation error compared to the baseline JRMPC, even when using empirically optimal re-sampling settings.



(a) Initital point sets (b) Failed JRMPC registration (c) Successful DARE registration

Figure 6. Pairwise registration example on the facade dataset in TLS ETH. To distinguish between the different point sets they are colored in red and blue. Despite the large initial transformation error, seen in (a), our DARE method (c) successfully registers the two point sets.



(a) Initiat point sets (b) Failed JRMPC registration (c) Failed DARE registration

Figure 7. Pairwise registration example on the indoor dataset in VPS with very limited overlap and large initial transformation error. Both the baseline JRMPC and our approach fails to align the point sets.

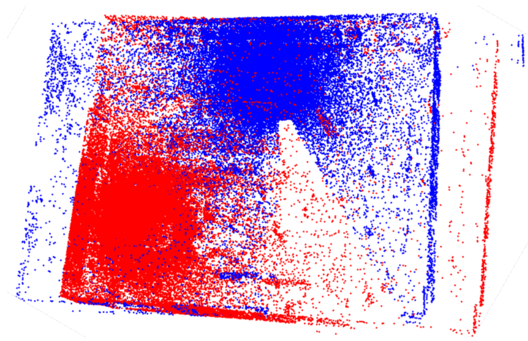
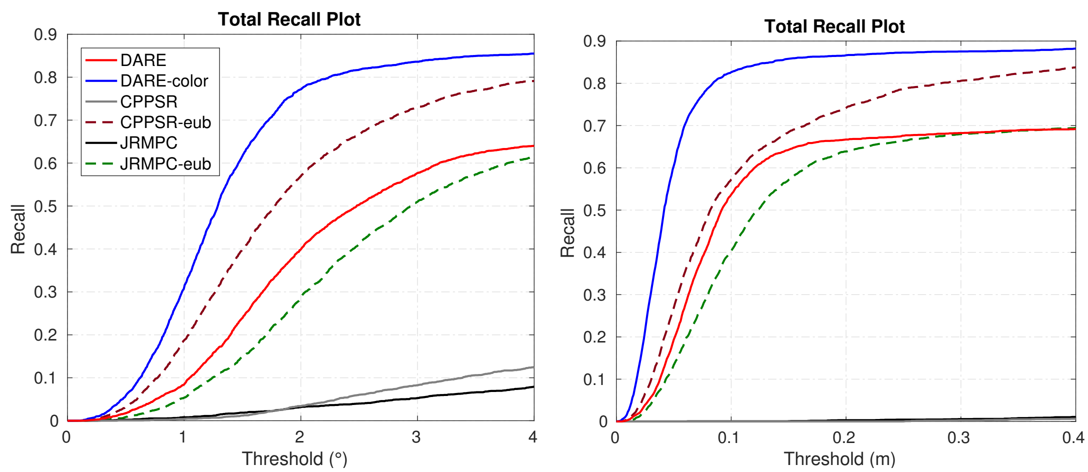
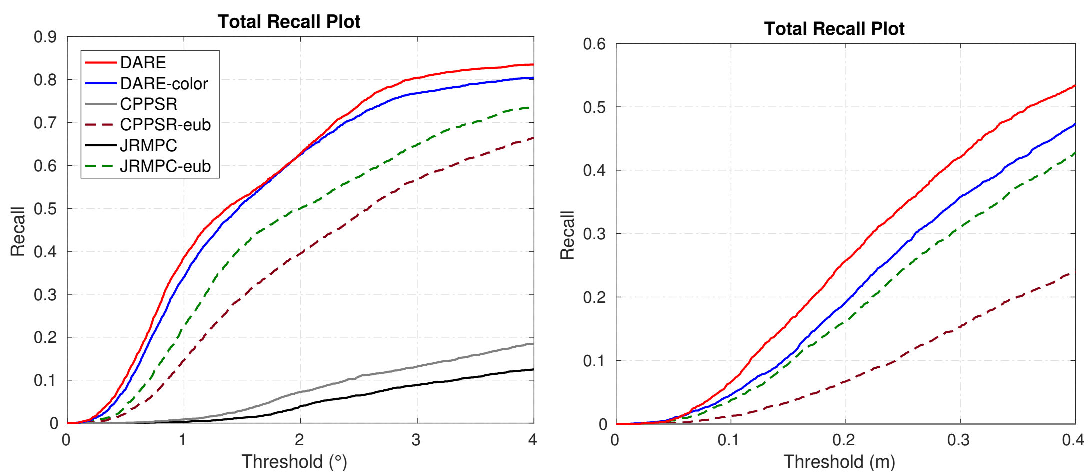


Figure 8. Failure example for our DARE method on the office dataset. The registration procedure has converged in a local minimum, where two of the aligned walls in the office room are are shifted, leading to a large translation error.



(a) Rotation recall (b) Translation recall

Figure 9. Recall curves with respect to (a) the angular error and (b) translation error on the indoor dataset in VPS. Our DARE and DARE-color (DARE with color) approaches show significant improvement compare to both JRMPC and CPPSR. Our approach also shows improved results compared to optimal re-sampling.



(a) Rotation recall (b) Translation recall

Figure 10. Recall curves with respect to (a) the angular error and (b) translation error on the indoor dataset in VPS. Our DARE and DARE-color (DARE with color) approachess show significant improvement compare to both JRMPC and CPPSR. Our approach also shows improved results compared to optimal re-sampling.

References

[1] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *PAMI*, 14(2):239–256, 1992. 4, 6

[2] C. M. Bishop. *Pattern recognition and machine learning*. springer, 2006. 1

[3] M. Danelljan, G. Meneghetti, F. Shahbaz Khan, and M. Felsberg. A probabilistic framework for color-based point set registration. In *CVPR*, 2016. 6

[4] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010. 2

[5] Y. Eldar, M. Lindenbaum, M. Porat, and Y. Y. Zeevi. The farthest point strategy for progressive image sampling. *TIP*, 6(9):1305–1315, 1997. 4

[6] G. D. Evangelidis, D. Kounades-Bastian, R. Horaud, and E. Z. Psarakis. A generative model for the joint registration of multiple point sets. In *European Conference on Computer Vision*, pages 109–122. Springer, 2014. 2, 4, 6

[7] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy. Geometrically stable sampling for the icp algorithm. In *3DIM 2003*, 2003. 4

[8] A. Myronenko and X. Song. Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence*, 32(12):2262–2275, 2010. 4, 6