

Robust 3D Object Registration Without Explicit Correspondence Using Geometric Integration

Dirk Breitenreicher · Christoph Schnörr

Received: date / Accepted: date

Abstract 3D vision guided manipulation of components is a key problem of industrial machine vision. In this paper, we focus on the localization and pose estimation of known industrial objects from 3D measurements delivered by a scanning sensor. Because local information extracted from these measurements is unreliable due to noise, spatially unstructured measurements and missing detections, we present a novel objective function for robust registration without using correspondence information, based on the likelihood of model points. Furthermore, by extending Runge-Kutta type integration directly to the group of Euclidean transformation, we infer object pose by computing the gradient flow directly on the related manifold. Comparison of our approach to existing state of the art methods shows that our method is more robust against poor initializations while having comparable run-time performance.

Keywords Registration · Iterative Closest Point (ICP) · Kernel-Based Similarity Measures · Geometric Integration

1 Introduction

1.1 Overview and Motivation

3D object recognition, pose estimation, quality inspection and bin picking are typical tasks that frequently occur in industrial machine vision applications. In each case, reliability and accuracy are important, besides a sufficiently short processing time. While one often knows *a priori* the object as

well as a rough estimate of the position, accurately determining the 3D object pose without prior information is the major difficulty of bin picking.

This difficulty results from the fact that a vision-based bin picking system has to cope with multiple instances of arbitrary, but known objects, that are randomly assembled in a bin. Figure 1 shows a typical set-up. Although prior knowledge of the objects shape simplifies the problem, a considerable amount of noise, symmetries of the object's shape, and mutual occlusion renders the tasks far from being trivial. A close-up view depicted in the right panel of Fig. 1 illustrates these issues.

A common way to address such a problem is to determine significant local landmarks in both the scene and the model, and to match them. Landmark detectors in 3D include spin images [1], shape context [2], point signatures [3], and volume integral descriptors [4]. Having established the correspondence of the landmarks, determining the Euclidean transformation that aligns measurements and object model is straightforward [5].

However, in the scenario considered in this paper, landmark based approaches are likely to fail due to a high degree of occlusion, object symmetry, and sparse noisy measurements (Fig. 1). An alternative is to work directly with the observed point clouds in order to register objects and to determine their pose. Still, this involves the problem to establish the correspondence between object model and observations.

An alternative class of methods to determine the rigid body transformation involves problems that do not require explicit point-to-point correspondences. Based on the work of Jian and Vemuri [6], in this paper we investigate a corresponding novel probabilistic objective function for the robust registration of point clouds. According to this criterion, optimal configurations for the pose are achieved by directly minimizing a distance measure without the need to

University of Heidelberg
Image & Pattern Analysis Group (IPA)
Heidelberg Collaboratory for Image Processing (HCI)
Speyerer Strasse 6, 69115 Heidelberg, Germany
Tel.: +49-6221-54-8875
Fax: +49-6221-54-5276
E-mail: {breitenreicher, schnoerr}@math.uni-heidelberg.de

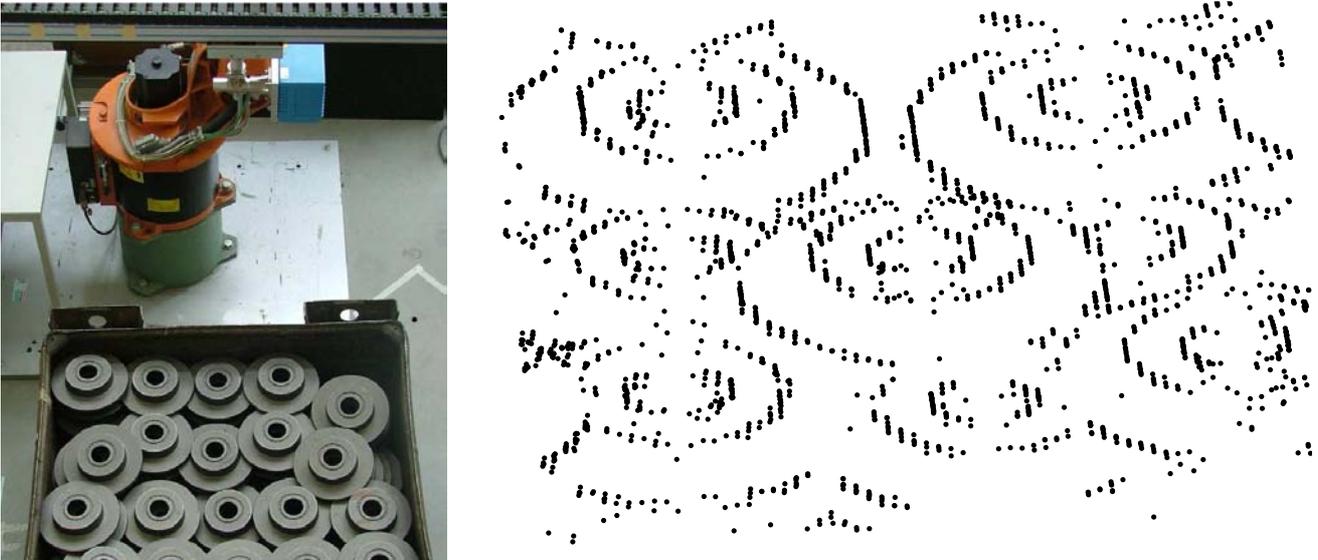


Fig. 1 Industrial Bin-Picking System (left) and the corresponding filtered edge data obtained with a SICK LMS 400 scanning device (right). The data are noisy and components occlude each other. This paper studies the problem to compute likely hypotheses of 3D pose of individual objects.

establish point correspondences. This minimization is carried out by numerically computing the gradient flow on the special Euclidean group of transformation, without resorting to particular parametrizations of rotations like Euler angles or quaternions, that are common in graphics and vision but not necessarily convenient for numerical purposes.

1.2 Related Work

Aligning two point sets amounts to the '*chicken and egg*' problem of determining simultaneously point-to-point correspondences and a rigid transformation. Having solved either problem, the other one becomes trivial.

As a consequence, most approaches proceed in an iterative fashion where given an estimate of the transformation, correspondence is estimated by some heuristic, followed by updating the transformation estimate, and so forth. The prototypical representant of this class of approaches is the Iterative Closest Point (ICP) algorithm developed in parallel by Besl and McKay [7] and Chen and Medioni [8]. Due to its simplicity and fast convergence, ICP is still a state of the art algorithm [9]. Further variants for speeding up the ICP algorithm have been suggested in the literature [10]. Because this work does not address our main concern, robustness with respect to initialization and local minima, we do not consider it in this paper, however.

As is well known, this basic two-step iteration is susceptible to noise and poor initialization. To overcome these issues Rusinkiewicz and Levoy [9], Rangarajan et al. [11] and Fitzgibbon [12] proposed robust variations that are more tolerant against imprecise initializations as well as against

noise and outlying structures. These improvements are based on smoothing the cost function of ICP but require determining accurate annealing schedules or noise thresholds.

In order to model uncertainties more precisely, Granger and Pennec [13] considered a robust ICP approach in an expectation maximization framework (EM-ICP). Using a probabilistic framework facilitates the interpretation of involved thresholds. Yet, despite increased robustness, there appears to be still room for improvement, especially concerning the requirement of accurate initializations of correspondences.

To alleviate these problems, Tsin and Kanade [14] and Jian and Vemuri [6] suggested to use an approach that does not require explicit point-to-point correspondences. The respective point clouds of both, the scene and the model are represented by mixture distributions, and the registration is achieved by optimizing the correlation of both distributions.

1.3 Contribution

The contribution of this paper is twofold. Firstly, we investigate a novel distance measure for point set alignment that does not involve explicit point-to-point correspondences.

Secondly, we show how to optimize this criterion numerically on the manifold of rigid transformations, using numerically convenient local charts rather than special parametrizations of the rotational part like, e.g., Euler angles.

Our distance measure is closely related to the work of Jian and Vemuri [6], who proposed a whole family of probabilistic distance measures for the registration of point sets that are continuously represented by kernel density estimates

or mixture distributions, and therefore do not need explicit point-to-point correspondences. Specifically, the L^2 -distance was studied as a particular case in [6], because it admits closed form computations of the distance between Gaussian mixture distributions and, therefore, is computationally efficient. In the present paper, we investigate a more general objective function for the registration of point sets in terms of the Kullback-Leibler distance between mixture distributions and show that it is more robust with respect to noise and poor initializations while having run-time performance similar to state of the art algorithms like EM-ICP [13], Softassign [11] and Kernel Correlation [6]. We also show that our distance measure includes the one proposed in [15] as another special case.

The second major contribution of this paper concerns the method to infer the rigid transformation optimizing our cost function. We apply a geometric Runge-Kutta type integration method [16, 17] directly on the underlying manifold of special Euclidean transformations, rather than using particular parametrizations of the involved orthogonal group. This leads to accurate and numerically stable gradient minimization at reasonably large step sizes.

Finally we apply our approach to artificial and real world point sets to evaluate its performance and compare it with related state of the art methods including ICP [7], EM-ICP [13], Softassign [11], and Kernel Correlation [6].

1.4 Organization

Our paper is organized as follows. In Sec. 2 we define our novel alignment measure. We show the relations to ICP and Kernel Correlation and examine the computational complexity of evaluating this measure. An algorithm for determination the optimal configuration is developed in Sec. 3. We exploit the manifold structure of the group of rigid transformations in order to optimize numerically the objective function. To evaluate the performance of our approach in Sec. 4, we compare our approach to existing point set registration algorithms on both, artificial and real world point sets. We conclude and point out further work in Sec. 5.

2 Rigid Point Set Registration

Let $\{\mathbf{u}_i, i = 1, \dots, N\} \subset \mathbb{R}^3$ denote the set of scene measurements obtained by a scanning device, and let $\{\mathbf{v}_j, j = 1, \dots, M\} \subset \mathbb{R}^3$ be the set of samples specified by a given model description, i.e. a CAD file or a sample scan (Fig. 2). In practice, explicit geometric mesh models are typically not available (cf. Fig. 2, left). Furthermore, generating meshes from thousands of 3D point measurements may not be error-free and causes a time-consuming preprocessing step. We



Fig. 2 Examples of 3D point cloud models obtained by a sample scan (left) and by discretizing a CAD model (right).

therefore only consider models in terms of 3D point measurements in order to uniformly handle all practically relevant situations.

The objective of the registration step is to determine a rigid body transformation $\theta \in SE(3)$ such that model and scene samples are aligned best with respect to a certain distance measure. Here, $SE(3)$ denotes the special Euclidean group of rigid transformations $\theta = \mathbf{T}_t \mathbf{R}$ whose elements are uniquely given by a proper rotation $\mathbf{R} \in SO(3)$ of 3-space, followed by a translation $\mathbf{T}_t(\mathbf{x}) = \mathbf{t} + \mathbf{x}$.

2.1 Correspondence Based Registration

Typically, point set alignment is achieved through correspondence based registration. A common distance measure involves sums of point-to-point distances $d(\cdot, \cdot)$ given by

$$\sum_{i=1}^N \sum_{j=1}^M \delta_{ij} d(\mathbf{u}_i, \theta(\mathbf{v}_j)), \quad (1)$$

where $\theta(\mathbf{v}_j)$ rigidly transforms \mathbf{v}_j , and $\delta_{ij} \in \{0, 1\}$ denotes that \mathbf{v}_j corresponds to \mathbf{u}_i .

To compute θ by minimizing (1), we have to determine the point-to-point correspondences $\{\delta_{ij}\}$. Determining these correspondences, on the other hand, requires to know the transformation θ . In general, this problem is solved in an iterative manner, see e.g. [7, 8, 13, 11].

2.2 Registration Without Explicit Correspondence

To alleviate the correspondence estimation problem, recent approaches adopted continuous representations of point sets using mixture distributions [14, 6]. These representations for the model and the scene, respectively, read

$$m(\mathbf{x}; \theta, \sigma_m) = \sum_{j=1}^M w_j^m k(\mathbf{x}, \theta(\mathbf{v}_j); \sigma_m), \quad (2a)$$

$$s(\mathbf{x}; \sigma_s) = \sum_{i=1}^N w_i^s k(\mathbf{x}, \mathbf{u}_i; \sigma_s), \quad (2b)$$

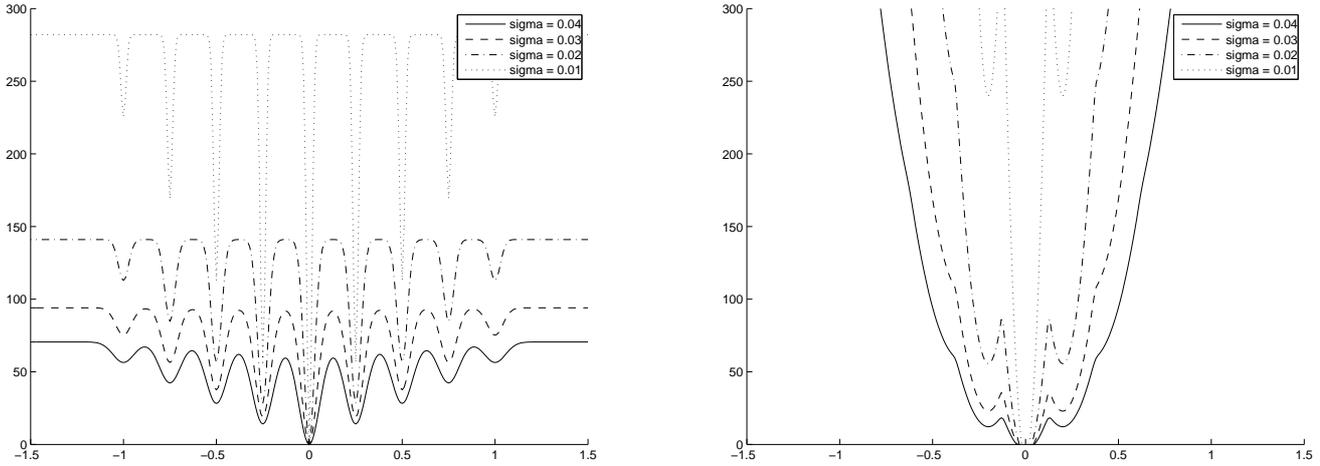


Fig. 3 L^2 distance (left) and KL distance (right) for one-dimensional mixture distributions with respect to translation. We can see that for increasing σ_m , the KL distance effectively convexifies the L^2 distance, leading to increased robustness of registration.

with weights $w_i^s, w_j^m \geq 0$ and kernel functions $k(\cdot, \cdot; \sigma_s)$, $k(\cdot, \cdot; \sigma_m)$ whose scale is controlled by parameters σ_s, σ_m .

In [6], Gaussian mixture distributions with covariance matrices $\sigma^2 \mathbf{I}$ were chosen in (2), and a parametrized family of distances d_α was proposed that includes the KL-distance ($\alpha = 0$)

$$D(s||m) = \int s(\mathbf{x}; \sigma_s) \log \frac{s(\mathbf{x}; \sigma_s)}{m(\mathbf{x}; \theta, \sigma_m)} d\mathbf{x} \quad (3)$$

and the L^2 distance ($\alpha = 1$). The latter distance was preferred because it can be evaluated in closed form. In this paper, we study the natural distance (3) between distributions [18] and show the relation to the standard distance measure (1). This enables to interpret the ICP algorithm as an special instance of the KL-distance.

The continuous point set representations (2) contain a scale parameters σ_s, σ_m that take into account noise and uncertainty of the point correspondences. It is reasonable to confine this degree of freedom to either the scene or the model. Without loss of generality, taking into account uncertainty and noise only in the representation of the model points, and choosing uniform weights, minimizing the distance (3) leads to the problem

$$\max_{\theta} F(\theta) , \quad (4a)$$

$$F(\theta) := \sum_{i=1}^N \log \sum_{j=1}^M \frac{1}{M} k(\theta(\mathbf{v}_j), \mathbf{u}_i; \sigma_m) . \quad (4b)$$

This can be seen as follows. Since the first term of (3),

$$\int s(\mathbf{x}; \sigma_s) \log s(\mathbf{x}; \sigma_s) d\mathbf{x} - \int s(\mathbf{x}; \sigma_s) \log m(\mathbf{x}; \theta, \sigma_m) d\mathbf{x} , \quad (5)$$

does not depend on θ , minimizing (3) amounts to maximize the second term in (5). Confining the modeling of uncertainty and noise to the representation of the model points, as discussed above, and choosing uniform weights, reduces the representation of the scene points to

$$s(\mathbf{x}; \sigma_s) = \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x} - \mathbf{u}_i) , \quad (6)$$

where $\delta(\cdot)$ denotes the Dirac delta. Insertion into the second term of (5) and dropping the constant, leads to the problem

$$\theta^* = \max_{\theta} \sum_{i=1}^N \log m(\mathbf{u}_i; \theta, \sigma_m) , \quad (7)$$

which together with (2) gives (4).

In order to highlight the role of the scale parameter σ_m , we drop the constant M^{-1} and write out the term of the first sum,

$$\begin{aligned} & \log \sum_{j=1}^M k(\theta(\mathbf{v}_j), \mathbf{u}_i; \sigma_m) \\ &= const. + \log \sum_{j=1}^M \exp \left(-\frac{1}{2\sigma_m^2} \|\theta(\mathbf{v}_j) - \mathbf{u}_i\|^2 \right) . \end{aligned} \quad (8)$$

Up to scaling, the latter term on the right hand side corresponds to the log-exponential function having well-known properties [19]. As a result, $\forall \sigma_m > 0$, scaling this term im-

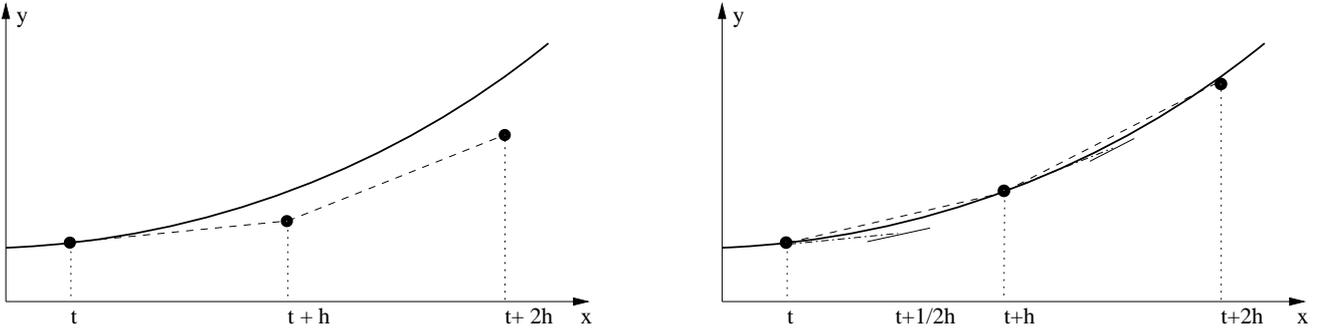


Fig. 4 Sketch for obtaining the gradient flow using Euler's (left) and the Improved Euler (right) algorithm for a 1D curve. We can see that the obtained values (circle) of the flow are closer to the true function (solid curve) for the improved Euler Method than for the ordinary algorithm.

mediately yields the estimate

$$\sigma_m^2 \log \sum_{j=1}^M \exp \left(-\frac{1}{2\sigma_m^2} \|\theta(\mathbf{v}_j) - \mathbf{u}_i\|^2 \right) - \sigma_m^2 \log M \quad (9a)$$

$$\leq \max_{j=1, \dots, M} \left\{ -\frac{1}{2} \|\theta(\mathbf{v}_j) - \mathbf{u}_i\|^2 \right\} \quad (9b)$$

$$\leq \sigma_m^2 \log \sum_{j=1}^M \exp \left(-\frac{1}{2\sigma_m^2} \|\theta(\mathbf{v}_j) - \mathbf{u}_i\|^2 \right). \quad (9c)$$

These lower and upper bounds reveal that each term in (4) provides a smoothed version and (including scaling) uniformly approximates for $\sigma_m \rightarrow 0$ the minimum distance of transformed model points $\theta(\mathbf{v}_j)$ to the measurement \mathbf{u}_i , which corresponds to the basic ICP objective function. Figure 3 illustrates this robustness property of our objective function (4). This result also shows that our approach (4) includes as special case the objective function suggested in [15].

Concerning computational complexity, similar to original version of ICP [7], EM-ICP [13], Softassign [11] and Kernel Correlation [6], our distance measure can be evaluated in $O(NM)$ operations, leading to an acceptable runtime performance.

3 Optimization by Geometric Integration

In this section, we derive an algorithm for optimizing the objective function (4). Due to the kernel function involved, computing the optimum in closed form is not possible, and we have to resort to numerical techniques.

Typically, objective functions of pose parameters are iteratively minimized using standard algorithms like Nelder-Mead [6] or Levenberg-Marquardt [12], based on a particular parametrization of the group $SO(3)$ of rotations. As standard algorithms do not take into account the curved structure of the underlying space, they require substantial modifications. We refer to [20] for typical examples.

Our plan in this paper, therefore, is to apply proper modifications of ordinary integration schemes [17, 16] in order to perform gradient ascent with respect to the objective function (4) directly on the group of rigid body transformations, viewed as a smooth manifold in the space of 4×4 matrices.

Specifically, we adopt an improved version [21] of Euler's algorithm for numerical integration to compute the gradient flow on $SE(3)$. To elucidate the exposition, we first sketch the common Euclidean case, followed by explaining the required modifications for transferring the scheme to the manifold.

3.1 The Euclidean Case

Let $\mathbf{X} \in \mathbb{R}^{n \times m}$ be an arbitrary real valued matrix. A basic approach for maximizing a smooth objective function $F(\cdot) : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$ is to perform gradient ascent

$$\dot{\mathbf{X}}(t) = \nabla F(\mathbf{X}(t)), \quad (10)$$

until a stationary point $\dot{\mathbf{X}}(t) = \mathbf{0}$ is reached and the necessary optimality condition $\nabla F(\mathbf{X}) = \mathbf{0}$ is satisfied.

Many schemes for numerically integrating equations of the form (10) exist [22, 21]. In view of our objective to apply such schemes to the curved space of pose parameters $SE(3)$, we explain next the most basic scheme, Euler's method, and its improvement through the mid-point rule.

Let $\mathbf{X}(t)$ denote a known point on the unknown path solving (10). First-order expansion around this point yields

$$\mathbf{X}(t+h) = \mathbf{X}(t) + h\dot{\mathbf{X}}(t) + O(h^2). \quad (11)$$

Omitting higher order terms, we obtain with (10)

$$\mathbf{X}(t+h) \approx \mathbf{X}(t) + h\nabla F(\mathbf{X}(t)), \quad (12)$$

defining Euler's scheme to compute the next point $\mathbf{X}(t+h)$ along the solution path based on the current iterate $\mathbf{X}(t)$. As is well-known, however, very small step sizes h have to

Algorithm 1 Improved Euler Algorithm on $\mathbb{R}^{n \times m}$ **Require:** $F(\cdot) : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$ **Require:** $h > 0$.

```

1: set  $\mathbf{X}_0 \in \mathbb{R}^{n \times m}$ 
2:  $t \leftarrow 0$ 
3: repeat
4:    $\mathbf{A} \leftarrow \nabla F|_{\mathbf{X}_t}$ 
5:
6:    $\mathbf{A} \leftarrow \nabla F|_{\mathbf{X}_t + \frac{1}{2}h\mathbf{A}}$ 
7:
8:    $\mathbf{X}_{t+h} \leftarrow \mathbf{X}_t + h\mathbf{A}$ 
9:    $t \leftarrow t + h$ 
10: until convergence

```

be chosen to avoid accumulation of numerical integration errors. This holds, in particular, when integrating paths on curved spaces.

Improving the Euler algorithm can be done as follows. At each point $\mathbf{X}(t)$ the search direction $\nabla F(\mathbf{X}(t))$ is computed. However, instead of moving with step size h along the gradient direction $\nabla F(\mathbf{X}(t))$, we move half the distance $\frac{1}{2}h$ to obtain a new direction $\nabla F\left(\mathbf{X}(t) + \frac{1}{2}h\nabla F(\mathbf{X}(t))\right)$ at this point. The corresponding update is given by

$$\mathbf{X}(t+h) = \mathbf{X}(t) + h\nabla F\left(\mathbf{X}(t) + \frac{1}{2}h\nabla F(\mathbf{X}(t))\right). \quad (13)$$

Figure 4 illustrates that the numerical scheme (13) is significantly more accurate than (12), at moderate additional costs.

Algorithm 1 details each step of the numerical scheme (13). Here $\nabla F|_{\mathbf{X}_t}$ denotes the computation of ∇F at $\mathbf{X}(t)$. We describe next how to modify and complement these steps when $\mathbf{X}(t)$ is defined on a manifold.

3.2 Extension to Rigid Body Transformations

In this section, we modify the scheme (13) such that it numerically integrates (10) on a manifold. For the general mathematical background, we refer to [17, 16].

We present the modified scheme first, followed by discussing the differences to the Euclidean case (13), viz. algorithm 1.

3.2.1 Improved Euler Algorithm on SE(3)

A key property of the Euclidean space \mathbb{R}^n , viewed as a flat manifold, is that the tangent space at any point can be identified with \mathbb{R}^n , i.e. the underlying manifold, itself. This is no longer the case for non-flat manifolds, like SE(3).

Besides clarifying the structure of tangent vectors, the basic operation additionally needed is the so-called exponential mapping $\exp(\mathbf{X})$, taking tangents to the manifold M at point \mathbf{X} to points $\exp(\mathbf{X}) \in M$ in the neighborhood of \mathbf{X} . For the mathematical details, we refer to [23].

Algorithm 2 Improved Euler Algorithm on SE(3)**Require:** $F(\cdot) : \text{SE}(3) \rightarrow \mathbb{R}$ **Require:** $h > 0$.

```

1: set  $\mathbf{Q}_0 \in \text{SE}(3)$ 
2:  $t \leftarrow 0$ 
3: repeat
4:    $\mathbf{G} \leftarrow \frac{\partial}{\partial \mathbf{Q}} F|_{\mathbf{Q}_t}$ 
5:    $\mathbf{A} \leftarrow \pi_{\mathbf{Q}_t}(\mathbf{G})$ 
6:    $\mathbf{G} \leftarrow \frac{\partial}{\partial \mathbf{Q}} F|_{\exp(\frac{1}{2}h\mathbf{A})\mathbf{Q}_t}$ 
7:    $\mathbf{A} \leftarrow \pi_{\exp(\frac{1}{2}h\mathbf{A})\mathbf{Q}_t}(\mathbf{G})$ 
8:    $\mathbf{Q}_{t+h} \leftarrow \exp(h\mathbf{A})\mathbf{Q}_t$ 
9:    $t \leftarrow t + h$ 
10: until convergence

```

We will represent as usual elements $\theta \in \text{SE}(3)$ by matrices

$$\mathbf{Q} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{pmatrix} \quad (14)$$

composed of rotation matrices $\mathbf{R} \in \text{SO}(3)$ and translation vectors $\mathbf{t} \in \mathbb{R}^3$. In the sequel, it will be convenient to write \mathbf{Q}_t instead of $\mathbf{Q}(t)$ to improve readability of formulas.

For the readers convenience, we derive the structure of tangents to SE(3) in appendix A. The exponential mapping for matrix \mathbf{X} is in general given by

$$\exp(\mathbf{X}) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{X}^k, \quad (15)$$

which for the specific case of a matrix \mathbf{A} tangent to $\mathbf{Q}_t \in \text{SE}(3)$ takes the form

$$\exp(\mathbf{A}) := \begin{pmatrix} \exp(\Delta_{\mathbf{R}}) & \mathbf{A}\Delta_{\mathbf{t}} \\ \mathbf{0}^\top & 1 \end{pmatrix}, \quad (16)$$

where, for $\gamma = \|\Delta_{\mathbf{R}}\|$

$$\exp(\Delta_{\mathbf{R}}) = \mathbf{I} + \frac{\sin(\gamma)}{\gamma} \Delta_{\mathbf{R}} + \frac{\sin^2(\frac{\gamma}{2})}{\frac{\gamma}{2}} \Delta_{\mathbf{R}}^2, \quad (17a)$$

$$\mathbf{A} = \mathbf{I} + \frac{1 - \cos(\gamma)}{\gamma^2} \Delta_{\mathbf{R}} + \frac{\gamma - \sin(\gamma)}{\gamma^3} \Delta_{\mathbf{R}}^2. \quad (17b)$$

All tangents \mathbf{A} and, in turn, matrices $\Delta_{\mathbf{R}}, \Delta_{\mathbf{t}}$ result from projections as detailed in appendix B.

Let $\frac{\partial}{\partial \mathbf{Q}} F|_{\mathbf{Q}_t}$ denote the usual matrix derivative of the objective function F at \mathbf{Q}_t , and $\pi_{\mathbf{Q}_t}(\mathbf{G})$ denote the orthogonal projection of a matrix \mathbf{G} onto the tangent space of SE(3) at \mathbf{Q}_t , as derived in appendix B. Then, together with the basic operation (16), algorithm 2 is well-defined and corresponds to the scheme (13) for integrating (10) on SE(3).

3.2.2 Differences to the Euclidean Case

The step-by-step comparison of algorithms 1 and 2 reveal the modifications necessary to transfer the numerical integration scheme (13) to the manifold SE(3).

- Usual gradient computations of the objective functional have to be interleaved with projections onto tangent spaces, because for objective functions defined on manifolds, the gradient actually is a vector field to make sense of (10), cf. [23].
- Addition is replaced by the group operation, given by matrix multiplication. This modification reflects that shortest paths along a prescribed direction are straight lines in the Euclidean space, but smooth curves when defined on manifolds.

4 Evaluation and Application

In most industrial applications, estimating the pose of arbitrary but known objects is difficult due to measurement errors and salient outlying structures. In order to align a model to the scene under these conditions, sufficiently good initializations are crucial. In this section, we therefore study the behavior of the objective function (4) with respect to initializations of varying accuracy. Moreover, we compare our approach to state of the art algorithms including ICP [7] based on a k-D-tree implementation, EM-ICP [13], Softassign [11] and Kernel Correlation [6].

Finally, we consider real world applications where model points are taken from CAD files or a reference scan (Fig. 2) and the corresponding scene is obtained by a SICK-LMS 400 laser scanner (Fig. 1).

4.1 Point Set Registration with Ground Truth

In this subsection we compare our approach to established state of the art algorithms, including ICP [7], EM-ICP [13], Softassign [11] and Kernel Correlation [6]. To this end, we used models freely available [24] in order to analyze the performance in a fully controlled environment. The models we used for registration are visualized in Fig. 5.

Each model consists of up to 50.000 samples. To speed up the whole process, we sampled each model down to the size of 200 points. Scenes corresponding to the model were placed arbitrarily in space. The primary concern in industrial bin picking applications are outliers. To simulate this, we randomly added up to 50 % outliers to each scene.

In order to analyze all algorithms with respect to sensitivity to initialization, we repeated the alignment process 150 times, where for each iteration the scene was transformed to a different location. Random rotations were obtained by sampling the complete space of rotations, while the translation varied randomly between 2 times the diameter of the model.

Two sets were considered to be properly aligned if the length of the curve connecting both transformations was less

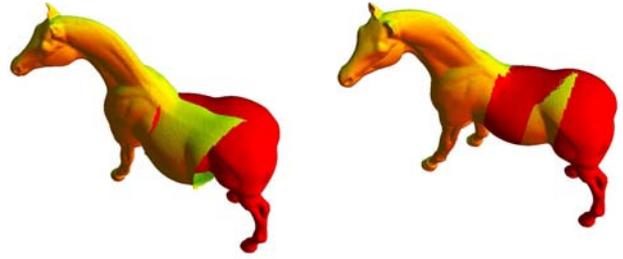


Fig. 7 To demonstrate the ability of our novel cost function to handle highly structured outliers we used a discrete set of points sampled from slightly overlapping back and head of horse as scene and model respectively (left). Instead of fitting the means of the point sets, our method accurately merges both point sets and reconstructs the original model (right).

than 0.2. In Euler angles, this distance corresponds to an error of about 5 degrees in each angle and a total deviation of 5 % of the model’s size in translation. This was chosen empirically based on our experience with the industrial application.

Furthermore, the parameters necessary for most algorithms were all tuned by hand to optimize performance and to guarantee a fair comparison. Accordingly, with respect to the annealing schedules required for the Softassign [11] and the EM-ICP Algorithm [13], we adopted conservative schedules at the cost of slower convergence, in order to better escape from local minima.

The choice of the parameter σ cannot be specified in general but depends on the given data and its scale. As a rule, too small values will impair robustness of the approach, making it sensitive to local minima like the ICP algorithm. Exceedingly large values, on the other hand, limit the accuracy of the registration. Our experiments show the existence of a reasonably large interval of values resulting in good performance.

The results of our experiments are presented in Fig. 6. It shows that our novel objective function (4) show a significantly increased robustness against inaccurate initializations and uniformly outperforms related state-of-the-art approaches like ICP and Kernel Correlation. Moreover, it reveals comparable performance to Softassign and EM-ICP without using an annealing schedule.

Moreover, to demonstrate the robustness of our novel objective function to structured outliers, we applied the registration procedure to the horse model, with a small overlap as visualized in Fig. 7. Using an additional background kernel supplemented to the model mixture, allows to accurately align both point sets and reconstruct the original model.

Finally, the time, necessary to evaluate our objective function is similar to those arising from EM-ICP, Softassign and Kernel Correlation, namely $O(NM)$.

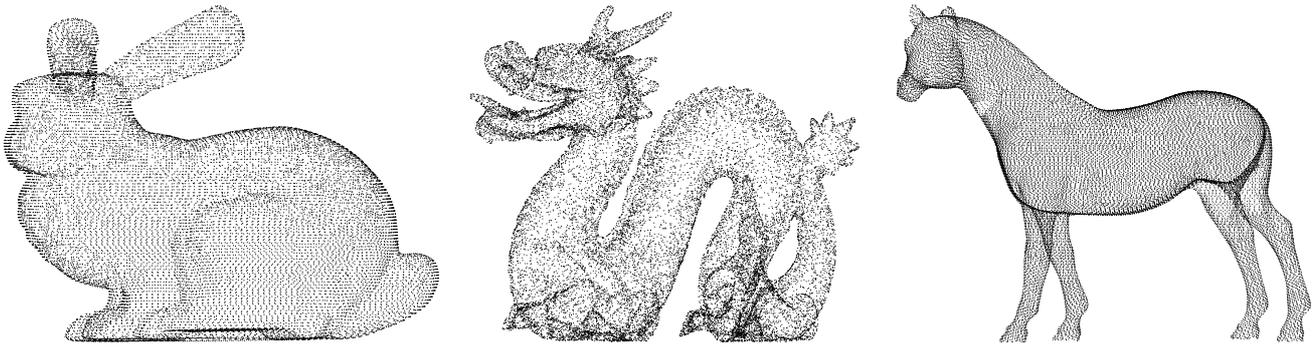


Fig. 5 Synthetic data set obtained from [24]. For our experiments, all data sets were down-sampled to 200 points and supplemented with outlier points placed arbitrarily inside the bounding box of each model.

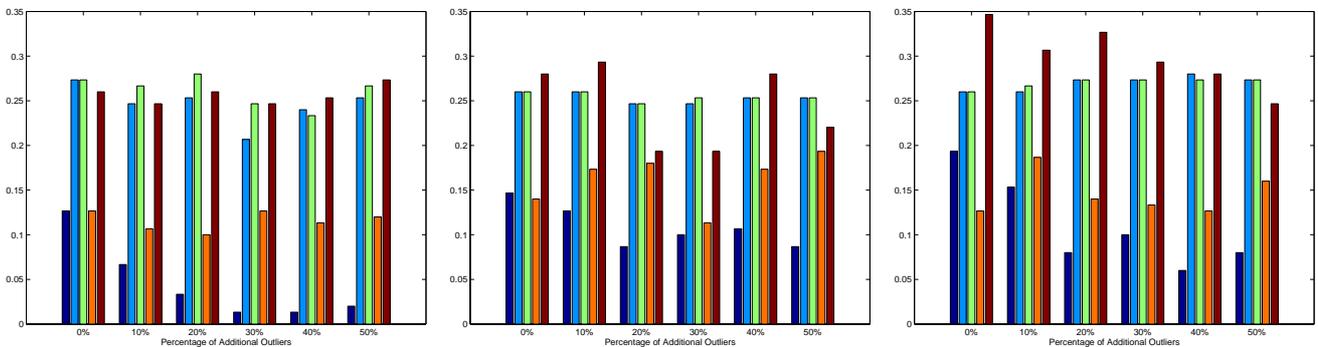


Fig. 6 Percentage of experiments by applying from left to right ICP (dark blue), EM-ICP (light blue), Softassign (green), Kernel Correlation (red) and our approach (brown) to the bunny (left), dragon (middle) and horse set (right) that converged to the true solution. All corresponding scenes were supplemented with noisy points in the range up to 50 % of the model size.

4.2 Real World Applications

In real industrial applications, the rigid alignment of a model point set to scene samples becomes more difficult due to the type of noise. As Fig. 1 shows, the scene recorded by a SICK LMS 400 scanning device contains noise as well as multiple structures similar to model parts. Most structures have to be considered as outliers, except for those corresponding to the model. Quantitatively, nearly 80% of the samples are structures not belonging to the object of interest.

In order to alleviate these issues, a common procedure is to fit the scene to the model instead of fitting the model to the scene, and to introduce a further background kernel as already mentioned above. This improves the capability to properly assign even occluded samples to parts.

In the particular scenario depicted in Fig. 1, experts from industry expect a maximum variability of rotation of ± 6 degree. Consequently, to be on the safe side, we randomly sampled the rotation space within ± 25 degrees around the model reference position and placed the model arbitrarily in the scene. Similar to the synthetic experiments, we tuned the parameters by hand.

The results of applying our approach as well as a robust implementation of ICP [9] are visualized in Fig. 8. The statistics resulting from these experiments, presented in table 4.2, reveal that according to our objective function 36 of 50 experiments converged to a visually correct solution. In contrast, the robust implementation of ICP only converged in 12 of 50 experiments.

This also leads to the fact that our approach locates 7 of 8 model instances in the scene with at least one experiment. The robust ICP approach in contrast, is only able to extract 4 different model instances accurately.

Even more complicated models can be handled using our novel objective function, see Fig. 9 and Fig. 10. While we used the discretized model visualized on the right-hand side of Fig. 2 to fit the scene measurements of Fig. 10 best, in Fig. 9 we demonstrate the applicability to automatic processes. The model, depicted on the left of Fig. 9 is used to specify the transformation to subsequent sample scans. As initialization we used the recorded position of the model.

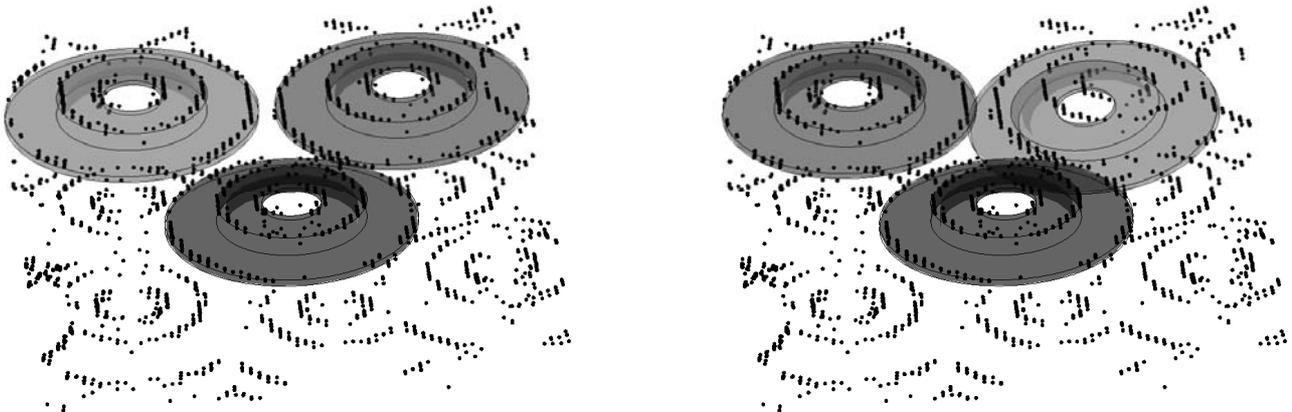


Fig. 8 Typical results of aligning a brake disc model to sample scans obtained by a SICK LMS 400 scanning device using our novel cost function (left) and the robust ICP algorithm [9] (right). While our approach locates the disk in the upper right corner accurately, robust ICP failed to detect it for the same initialization.

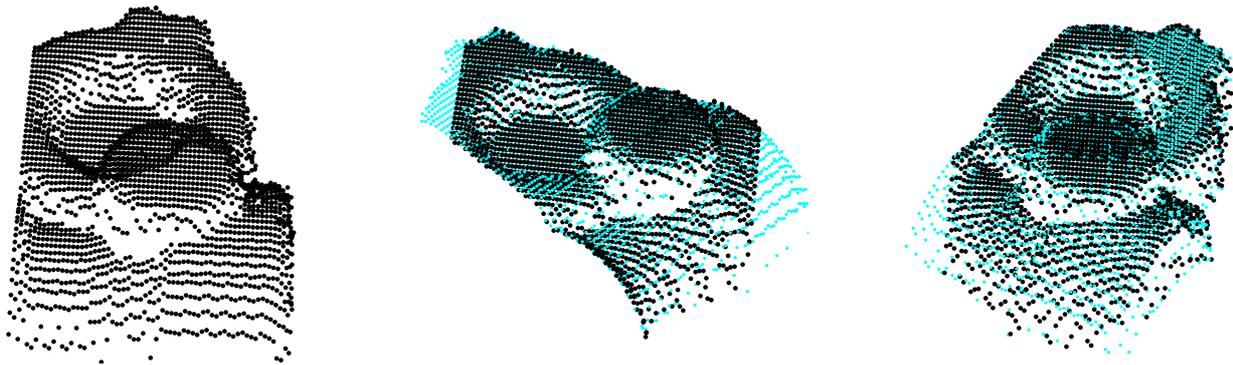


Fig. 9 Typical results of aligning a model, recorded by a SICK LMS 400 scanning device (left), to subsequent scans of the same object (middle and right) using our novel objective function. The model (light blue) accurately fits to the recorded scene (black), even if parts of the scene are missing.

	Our Approach	Robust ICP [9]
# located objects	7	4
# positive detections	36	12
# negative detections	14	38

Table 1 Statistical results of aligning a brake disc model to real world data for our approach and a robust implementation of ICP using outlier rejection [9].

4.3 Limitations of our Approach

Despite increased robustness in comparison to ICP, EM-ICP, Softassign and Kernel Correlation, our approach still may fail for too inaccurate initializations. Typical results where our approach failed to converge to true solutions are shown in Fig. 11.

In these cases, objects are placed such that parts of the model accurately fit various salient structures in the scene that, however, do not belong to a single object instance. Due to the local influence of the kernels involved in our objective function, we cannot currently escape from such pronounced local optima.

5 Conclusion and Outlook

We briefly summarize the results obtained and give a short outlook to our future work.

In this paper, we proposed a novel objective function for rigid point set registration based on Jian and Vemuri’s [6] work on Kernel Correlation. The major benefit of our criterion is its ability to align point sets robustly even when a

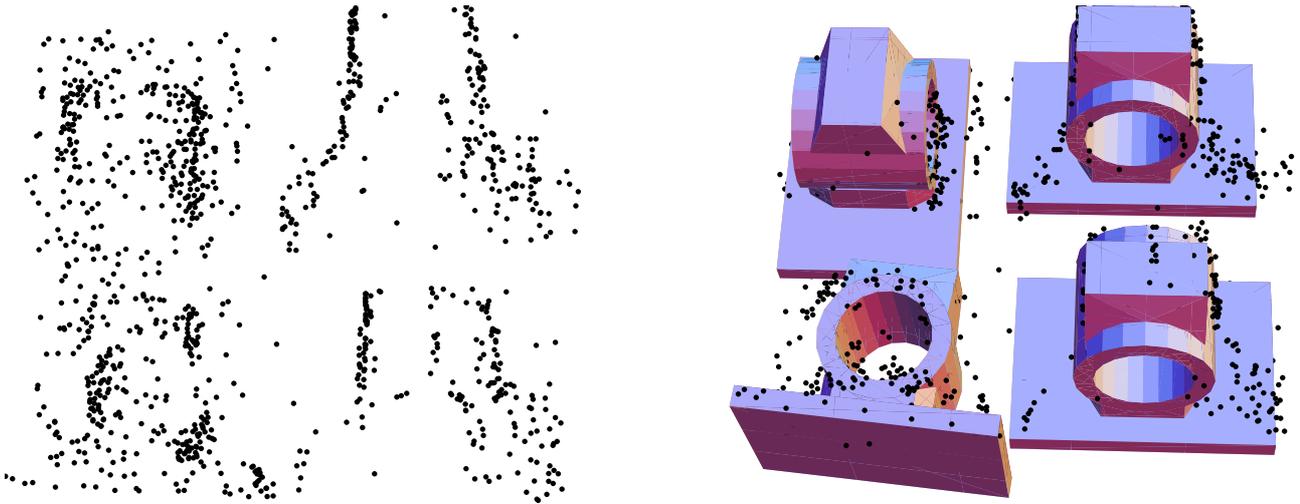


Fig. 10 Typical results of aligning a more complicated models to sample scans (left) obtained by a SICK LMS 400 scanning device using our novel cost function. Due to high amount of occlusions, more than half of the model is invisible. Nevertheless, our method is able to accurately align the model to the scene (right) given proper initializations.

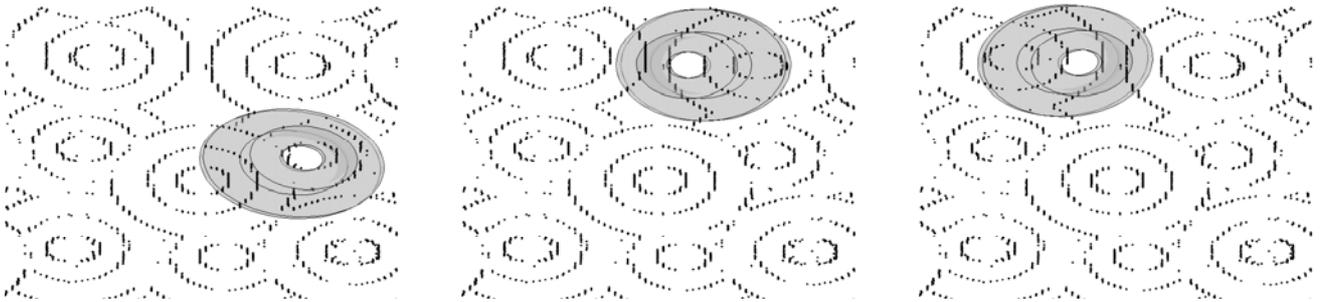


Fig. 11 Typical negative detections occurring during experiments with too inaccurate initializations. Salient structures of different objects are placed, such that they fit few parts of the model. Similar to state-of-the-art approaches, this causes our algorithm to get stuck in local optima.

considerable amount of spatial outlying structure is present in the scene. This was demonstrated numerically in comparison to a range of established methods, and implemented by a numerical algorithm that takes into account the geometry of the underlying manifold of transformations.

In our future work, we will address the persistent local minima problem sketched in Sec. 4.3. A natural idea is to apply an annealing procedure with respect to the kernel size, similar to [11]. Parameter σ_m in (4) appears to be a suitable candidate for this purpose.

Another point to be considered concerns the convergence rate of our geometrical integration method. In this paper, we used a stabilized version (13) of the basic Euler method (12), adapted so as to integrate (10) on $SE(3)$. This leaves room for improvements by considering higher-order methods in order to become competitive with the linear convergence rate of basic ICP as well as the quadratic convergence of the improved scheme presented in [25].

Acknowledgements This research work was supported by the VMT Vision Machine Technic Bildverarbeitungssysteme GmbH, a company of the Pepperl+Fuchs Group.

A Tangent Space of $SE(3)$

We represent rigid transformations $\theta \in SE(3)$ by matrices \mathbf{Q} composed of rotations \mathbf{R} and translations \mathbf{t} ,

$$\mathbf{Q} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{pmatrix}. \quad (18)$$

The inverse is given by

$$\mathbf{Q}^{-1} = \begin{pmatrix} \mathbf{R}^\top & -\mathbf{R}^\top \mathbf{t} \\ \mathbf{0}^\top & 1 \end{pmatrix}. \quad (19)$$

Let $\mathbf{Q}_t := \mathbf{Q}(t)$ denote any smooth path parametrized by t . Computing the derivative of $\mathbf{Q}\mathbf{Q}^{-1} = \mathbf{I}$ yields

$$\dot{\mathbf{Q}}\mathbf{Q}^{-1} + \mathbf{Q}\dot{\mathbf{Q}}^{-1} = \mathbf{0}. \quad (20)$$

Together with (14) and (19), (20) can be simplified to

$$\Delta = \begin{pmatrix} \Delta_{\mathbf{R}} & \Delta_{\mathbf{t}} \\ \mathbf{0}^\top & 0 \end{pmatrix} = \dot{\mathbf{Q}}\mathbf{Q}^{-1}, \quad (21)$$

where Δ is rotational skew symmetric, i.e. $\Delta_{\mathbf{R}}^\top = -\Delta_{\mathbf{R}}$. Multiplying both sides by \mathbf{Q} reveals the general structure of tangent vectors,

$$\dot{\mathbf{Q}} = \Delta\mathbf{Q}. \quad (22)$$

As a result, the space of tangents to $\text{SE}(3)$ at \mathbf{Q} is given by

$$T_{\mathbf{Q}} = \{ \dot{\mathbf{Q}} \mid \dot{\mathbf{Q}} = \Delta\mathbf{Q}, \Delta_{\mathbf{R}}^\top = -\Delta_{\mathbf{R}} \}, \quad (23)$$

where Δ has the form (21).

B Projections to $T_{\mathbf{Q}}$

The gradient ∇F of a function $F: \text{SE}(3) \rightarrow \mathbb{R}$ is defined by the relation [23]

$$\langle \partial F, \mathbf{Y} \rangle = \langle \nabla F, \mathbf{Y} \rangle, \forall \mathbf{Y} \in T_{\mathbf{Q}}, \quad (24)$$

where $\langle \cdot, \cdot \rangle$ is the Riemannian metric of the underlying manifold, in our case simply the ordinary matrix inner product

$$\langle \mathbf{A}, \mathbf{A} \rangle = \text{tr}(\mathbf{A}^\top \mathbf{A}). \quad (25)$$

$T_{\mathbf{Q}}$ denotes the tangent space at \mathbf{Q} defined in (23), and ∂F is the usual matrix derivative of F given by

$$(\partial F)_{ij} = \frac{\partial}{\partial Q_{ij}} F(\mathbf{Q}). \quad (26)$$

To obtain ∇F in (24), we consider

$$\nabla F = \arg \min_{\mathbf{Z} \in T_{\mathbf{Q}}} \frac{1}{2} \|\mathbf{Z} - \partial F\|^2. \quad (27)$$

Analogously to the decomposition in (14), ∂F can be factorized as

$$\partial F = \begin{pmatrix} \partial F_{11} & \partial F_{12} \\ \partial F_{21} & \partial F_{22} \end{pmatrix}, \quad (28)$$

where $\partial F_{11} \in \mathbb{R}^{3 \times 3}$, $\partial F_{12} \in \mathbb{R}^{3 \times 1}$, $\partial F_{21} \in \mathbb{R}^{1 \times 3}$, $\partial F_{22} \in \mathbb{R}$.

Using the general form of (23) and (21), together with (28), the solution to (27) is given by

$$\Delta_{\mathbf{t}} = \partial F_{12} - \Delta_{\mathbf{R}} \mathbf{t}, \quad (29)$$

$$\Delta_{\mathbf{R}} = \frac{1}{2} \left(\partial F_{11} \mathbf{R}^\top - \mathbf{R} \partial F_{11}^\top \right). \quad (30)$$

As a result, the projection of ∂F to $\nabla F \in T_{\mathbf{Q}}$ reads

$$\pi_{\mathbf{Q}}(\partial F) = \begin{pmatrix} \Delta_{\mathbf{R}}(\partial F_{11}) & \partial F_{1,2} - \Delta_{\mathbf{R}}(\partial F_{11})\mathbf{t} \\ 0 & 0 \end{pmatrix} \mathbf{Q}, \quad (31)$$

where $\Delta_{\mathbf{R}}(\partial F_{11}) = \frac{1}{2} (\partial F_{11} \mathbf{R}^\top - \mathbf{R} \partial F_{11}^\top)$.

By inserting (31) into (24), direct computation shows that (24) holds for all $\mathbf{Y} \in T_{\mathbf{Q}}$. Moreover, for each $\mathbf{Z} \in T_{\mathbf{Q}}$,

$$\mathbf{Z} = \pi_{\mathbf{Q}}(\mathbf{Z}). \quad (32)$$

References

- Johnson, A., Hebert, M.: Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**, 433–449 (1999)
- Frome, A., Huber, D., Kolluri, R., Bulow, T., Malik, J.: Recognizing Objects in Range Data using Regional Point Descriptors. In: *Proc. Europ. Conf. on Comp. Vision* (2004)
- Chua, C.S., Jarvis, R.: 3D Free-Form Surface Registration and Object Recognition. *Intl. J. of Comput. Vision* **17**, 77–99 (1996)
- Gelfand, N., Mitra, N.J., Guibas, L.J., Pottmann, H.: Robust Global Registration. In: *Proc. 3rd Eurogr. Symp. on Geometry Processing* (2005)
- Eggert, D.W., Lorusso, A., Fisher, R.B.: Estimating 3-D Rigid Body Transformations: A Comparison of Four Major Algorithms. *Mach. Vision Appl.* **9**, 272–290 (1997)
- Jian, B., Vemuri, B.C.: A Robust Algorithm for Point Set Registration Using Mixture of Gaussians. In: *Proc. 10th IEEE Intl. Conf. on Comput. Vision* (2005)
- Besl, P.J., McKay, N.D.: A Method for Registration of 3-D Shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**, 239–256 (1992)
- Chen, Y., Medioni, G.: Object Modelling by Registration of Multiple Range Images. *Image Vision Comput.* **10**, 145–155 (1992)
- Rusinkiewicz, S., Levoy, M.: Efficient Variants of the ICP Algorithm. In: *Proc. 3rd Intl. Conf. on 3D Digital Imaging and Modeling* (2001)
- Jost, T., Hügli, H.: A Multi-Resolution Scheme ICP Algorithm for Fast Shape Registration. In: *Intl. Symp. 3D Data Processing Visualization and Transmission* (2002)
- Rangarajan, A., Chui, H., Bookstein, F.L.: The Softassign Procrustes Matching Algorithm. In: *Proc. 15th Intl. Conf. on Inf. Processing in Med. Imaging* (1997)
- Fitzgibbon, A.W.: Robust Registration of 2D and 3D Point Sets. *Image Vision Comput.* **21**, 1145–1153 (2003)
- Granger, S., Pennec, X.: Multi-scale EM-ICP: A Fast and Robust Approach for Surface Registration. In: *Proc. 7th Europ. Conf. on Comput. Vision-Part IV*, pp. 418–432 (2002)
- Tsin, Y., Kanade, T.: A Correlation-Based Approach to Robust Point Set Registration. In: *Proc. Europ. Conf. on Comput. Vision* (2004)
- Goldberger, J., Gordon, S., Greenspan, H.: An Efficient Image Similarity Measure Based on Approximations of KL-Divergence Between Two Gaussian Mixtures. In: *Proc. 9th IEEE Intl. Conf. on Comput. Vision* (2003)
- Munthe-Kaas, H.: Runge-Kutta Methods on Lie Groups. *BIT* **38**, 92–111 (1998)
- Crouch, P.E., Grossman, R.: Numerical Integration of Ordinary Differential Equations on Manifolds. *J. of NonLinear Science* **3**, 1–33 (1993)
- Cover, T., Thomas, J.: *Elements of Information Theory*. John Wiley & Sons, Inc. (1991)
- Rockafellar, R., Wets, R.B.: *Variational Analysis*. Springer (1998)
- Edelman, A., Arias, T.A., Smith, S.T.: The Geometry of Algorithms with Orthogonality Constraints. *SIAM J. Matrix Anal. Appl.* **20**(2), 303–353 (1999)
- Schwarz, H.R.: *Numerische Mathematik*. B. G. Teubner Stuttgart (1986)
- Hairer, E., Nørsett, S.P., Wanner, G.: *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer Verlag (2008)
- Matsushima, Y.: *Differentiable Manifolds*. Marcel Dekker, INC. New York (1972)
- Turk, G., Mullins, B.: *Large Geometric Models Archive*. http://www-static.cc.gatech.edu/projects/large_models/
- Pottmann, H., Huang, Q.X., Yang, Y.L., Hu, S.M.: Geometry and Convergence Analysis of Algorithms for Registration of 3D Shapes. *Intl. J. Computer Vision* **67**, 277–296 (2006)