3D Point Set Registration based on Hierarchical Descriptors

Somnath Dutta
Technische Universität Dresden, Germany
somnath.dutta@tu-dresden.de

Benjamin Russig
Technische Universität Dresden, Germany
benjamin.russig@tu-dresden.de

Stefan Gumhold
Technische Universität Dresden, Germany
stefan.gumhold@tu-dresden.de

ABSTRACT
Registering partial point clouds is crucial in numerous applications in the field of robotics, vision, and graphics. For arbitrary configurations, the registration problem requires an initial global alignment, which is computationally expensive and often still requires refinement. In this paper, we propose a pair-wise global registration method that combines the fast convergence made possible by global hierarchical surface descriptors with the arbitrarily fine sampling enabled by continuous surface representations. Registration is performed by matching descriptors of increasing resolution – which the continuous surfaces allow us to choose arbitrarily high – while restricting the search space according to the hierarchy. We evaluated our method on a large set of pair-wise registration problems, demonstrating very competitive registration accuracy that often makes subsequent refinement with a local method unnecessary.

Keywords
3D shape registration, surface descriptors, similarity measure.

1 INTRODUCTION
Acquiring 3D geometry of physical objects and environments has become an integral part of numerous disciplines, ranging from applications in mechanical engineering and computer vision to asset generation for movies and games. Rapid development of 3D sensing technology over the last decades led to affordable high precision devices being widely available. Most techniques follow a sequential scanning paradigm where an object or scene is scanned from different viewpoints. The result of this acquisition process is a set of scans, each in a local coordinate system defined by the pose of the scanning device, with different degrees of overlap. These scans have to be aligned with respect to a unique reference to obtain a final reconstructed model of an object. The accuracy of this registration step is of significant importance to the quality of the final model but is often hampered due to low-overlap between scans and inherent noise. Along with that, the registration errors may lead to artificial creases and reconstruction artifacts (oscillations or even holes).

The goal of pairwise registration is to find a transformation that aligns a source scan as close as possible to the surface represented by a target scan. The possible approach that could assist in precise registration accuracy is combination of coarse alignment followed by an iterative local refinement that results in higher convergence. The local refinement is performed effectively using variants of Iterative closest point (ICP) [Bes92] which highly depends on a good initial guess. Coarse registration approaches usually depend on a descriptor [MPD06, Zah12, Tom10, Pet15] associated with a set of feature points encoding useful geometric information from the data based on a local neighborhood. Correspondences between descriptors are established based on similarity, resulting in a rigid transformation that best aligns the set of feature points and using some robust estimator.

The object surface is generally not sampled at the same locations between partial scans (an effect which we refer to as sampling discrepancy), and every sample is subject to measurement noise. Methods that directly rely on descriptor computation are inherently limited by these measurement issues. In addition to these external factors, methods relying on localized feature descriptors often suffer from a certain percentage of false matches, typically caused by insufficient discriminating capabilities of the descriptor and/or an underperforming similarity measure. These factors, to varying degrees, prevent them from achieving higher levels of accuracy, and the resulting rigid transformation negatively affects the convergence of local iterative algorithms employed for refinement. We propose a method that mitigates these drawbacks. Our contributions are as follows:

• novel adaption of an existing hierarchical surface descriptor to continuous surface representations
• based on that, a new pair-wise global registration pipeline for partial scans, featuring:

This work has received funding from the two DFG grants 389792660 (TRR 248, CPEC) and 390696704 (EXC 2050/1, CeTI) as well as from BMBF grant 01/S18026A-F (ScaDS.AI Dresden/Leipzig)
Additionally, the paper contributes a comprehensive evaluation of the proposed approach and compares it to the state-of-the-art global registration method by [Zho16], demonstrating measurable improvements in terms of accuracy and robustness.

2 RELATED WORK

Registration is a widely researched problem with vast applicability in various domains like 3D scanning, shape analysis, or motion capturing. For an overview of existing registration algorithms, we refer the reader to the respective surveys [Rus01], Salvi et al. [Sal07], Tam et al. [Tam13], Bellekens et al. [Bel14], Huang et al. [Hua21]. According to the aforementioned surveys, registration approaches can be classified with respect to the following criteria:

Pairwise / Multi-Way. Pairwise registration methods [Bel14] aim at registering two partially overlapping scans. Overall registration of a data set can be achieved by successively adding scans and registering them to their predecessors or to all preceding scans. On the other hand, multi-way registration [Goj09] optimizes all transforms at the same time, usually by iterating pairwise registrations or by modeling the problem with a single registration objective.

Rigid / Non-Rigid. Rigid registration methods [Bel14] only allow rigid body transforms for the individual scans, whereas non-rigid methods allow arbitrary ones. However, non-rigid approaches [Hua22] usually regularize the objective in order to avoid degenerate solutions (e.g. by demanding the transform to be as rigid as possible).

Local / Global. Local methods [Bes92, Hua11] use an initial transform estimation (e.g. provided by the user) and refine this transformation to optimize an objective. Global methods [Zho16, Pet15] do not require any input other than the geometry and find the registration transforms with arbitrary initial alignment of scans. Such global methods usually produce coarse registrations that can be refined with local methods. Furthermore these methods are not directly based on spatial proximity (but on, e.g., geometric feature descriptors).

Learning-based Registration. In recent years, data-driven approaches are emerging for registering point clouds primarily owing to exemplary results in both 2D and 3D applications. Researchers have used neural networks (NNs) in different ways, either to extract feature from individual point cloud followed by correspondence search or even focused on end-to-end pipeline for transformation estimation. Feature learning methods [Zen17, Hao18, Goj19] use the deep neural network to learn a robust feature correspondence search. Then, the transformation matrix is obtained by one step estimation (SVD, RANSAC) without iteration. Those NNs could provide robust and accurate correspondence searching but are fairly limited by availability of large training data and lesser generalization capabilities, resulting in a large-scale failure for unknown scenes with different distribution compared to the training data. End-to-end learning-based method [Yan19, Wan19] solves the problem of registration with complete end-to-end network, i.e, along with the correspondence search, transformation estimation is also embedded into the framework and differs from feature learning method whose focus in entirely on point feature learning. DeepGMR [Yua20] relies on a neural network to initially learn pose-invariant point-to-distribution parameter correspondences. Then, these correspondences are fed into the GMM optimization module to estimate the transformation matrix.

Descriptors

Many 3D descriptors based on point clouds have been introduced and importantly applied to the task of point cloud registration. A comprehensive reviews for point cloud descriptors in terms of computational efficiency and accuracy are detailed in [Ale12, Guo16, Han18, Ran20, Han18]. Based on the process of descriptor extraction, the descriptors can be approximately classified as local, global, and hybrid. Local descriptors, for every point, embeds spatial distribution or geometric attributes extracted in the neighborhood and is produced by a histogram representing the descriptors for each point. It is less discriminating exactly, but suitable for cluttered scenes. Spin Image [Joh99] is a popular descriptor often used in applications of shape matching, object recognition. It creates a cylindrical support around a keypoint, divides it into radial and vertical volumes, and then counts the number of points that falls within each volume. On the other hand, a global descriptor, for the entire point cloud, is produced with a single context vector by integrating a set of features or incorporating spatial distribution of the complete data. Global-based features are generally used in 3D object recognition and object categorization, additionally some global descriptors can be used for pose estimation of the objects as they also contain local descriptors information, such as Viewpoint Feature Histogram (VFH) and Local-to-Global Signature (LGS) descriptor. Moreover, global-based methods demands less computation compared with the local features and better to describe the whole point cloud but are majorly affected by occlusions and clutter [Had14]. Hybrid-based descriptors incorporate local and global descriptors together with most of the advantages from them.
Sampling Discrepancy

As briefly mentioned when motivating our work in section 1, sampling discrepancy is a major factor limiting the accuracy of point cloud registration, as it means that in general, no exact correspondences can be found. Thus, many strategies have been proposed to reduce its impact. Chen and Medioni [Chen91] project points from the source shape onto their closest position in the target depth image with an iterative algorithm now known as normal shooting [Rus01]. The resulting point will yield an interpolated closest point on the target.

The Adaptive moving least squares (AMLS) surface based approach [Hua11] for ICP registration emphasize on the efficiency of the technique to overcome issues of sampling discrepancy. The essential concept of the AMLS method is to reconstruct a smooth and accurate representation of a surface for ICP based registration by adaptively selecting the width of a Gaussian kernel based on the principal curvature of the MLS surface through local integral invariant analysis [Yan06]. Instead of trying to minimize the distance between corresponding entities, [Mit04, Pot06] estimate the distance field of the underlying surface with local quadratic approximates and directly optimize the distance of the scan and the target surface. Kubacki et al. [Kub12] integrated multiple depth images into a model represented by implicit moving least squares (IMLS) on a grid based on signed distance function updates. Similar ideas appeared in [Par03, Mun07]. However, any such method inherently suffers from discretization artifacts at reasonable resolutions.

3 METHODOLOGY

The proposed method relies on the core concept of descriptor matching in order to estimate the rigid transformation. Our approach is based on the Circon descriptor proposed by Ferrero et al. [Car12] that represents an ordered set of radial contours around a point of interest within a point cloud. The descriptor associated with a particular point-of-interest is used to express the point cloud in a local frame. The environment around a point-of-interest is divided into sectors each representing an angle $\rho \sigma$, and are further divided radially into cells with length $\rho \tau$.

The points are mapped into cells of the descriptor in a way that the cell-value represents the height ($z$-value) with respect to the local reference frame. Figure 1 depicts a descriptor associated with a point-of-interest (cf. section 3.1) encoding the structural information from a point cloud.

We built upon the registration method Ferrero et al. [Car12] tailored for this descriptor. While the core hierarchical approach is the same, the way they use the point cloud directly to build the descriptors ties the

Figure 1: Descriptor spawned by a point-of-interest defining a local reference frame with cell division into sectors represented by green and red shows contour formed by maximum $z$ within each cell.

achievable accuracy to the point cloud resolution. We propose to utilize a continuous surface representation instead, enabling us to construct descriptors of arbitrarily high resolution.

The exact choice of surface representation poses a trade-off of surface expressiveness vs. computational complexity. We opted for NURBS surfaces, as their desirable properties include the ability to exactly represent relevant algebraic shapes (like ellipsoids), and many manufactured objects are likely to have been CAD-designed using NURBS originally. They also render computation of the Circon descriptor non-trivial, which required additional measures to tackle the added computational cost. Our proposed algorithm can be summarized as follows:

AL.1 (pre-processing) Fit NURBS surfaces to the source and target point clouds. We first perform Euclidean clustering [Rad09] (see Appendix A for details) on both to account for highly fragmented scans, and each cluster is fitted with its own surface. We refer to the union of all surfaces fitted to a point cloud as the surface of the point cloud.

AL.2 (pre-processing) Sample each fitted surface at a resolution of $256 \times 256$ in parameter space to get a collection of $(u_0, v_0)$ footpoints needed later for descriptor construction (see section 3.2).

AL.3 (pre-processing) Point-of-interest (poi) selection for both surfaces (section 3.1).
AL.4 (pre-processing) Initial sorting of poi (section 3.4) pairs at a coarse resolution according to their similarity score (section 3.3), so step AL.5 can start with the most promising pairs.

AL.5 Descriptor computation across resolutions (section 3.4) for a pair of poi from source and target surfaces.

AL.6 Estimate transformation from point-pairs with highest similarity score at maximum resolution (user-defined) and evaluate stopping criterion.

AL.7 Repeat from AL.5 until the stopping is criterion satisfied.

We elaborate on the key details of the registration methods in the following sections.

3.1 Selecting Points-of-Interest

Points-of-interest (poi) for the source and target shapes are essential for the proposed approach. They define where descriptors are initially constructed and should be chosen in a way such that the resulting descriptors cover every location on the shape at least once. The original registration algorithm by Ferrero et al. [Car12] adapts a strategy to select non-edge points by performing edge detection and thresholding based on the Laplacian of the normal vectors, essentially restricting their selection to relatively flat areas. We did not find this to satisfy. We use the simple Nelder-Mead algorithm [Nel65] to solve this, since the problem is low-dimensional and the evaluation of the objective function is cheap. This strategy requires an initial guess though, which we obtain in the following way:

During pre-processing, we perform a regular sampling of the NURBS surfaces in parameter space, giving us a database of footpoints on the surface and their associated parameters \((u_0, v_0)\). When transforming a surface into the local frame of the current descriptor, we can transform these footpoints along with it. Using a simple regular grid on the descriptor reference plane, descriptor cells can collect those footpoints that project close to them. We then select the most promising initial guess by applying a multi-criterion sort according to (a) footpoint height \(z\) above the descriptor reference plane and (b) closeness to the ray.

We accomplish (a) by binning the footpoints in the cell according to their height \(z\), and (b) results from sorting them inside each bin. For selecting the initial guess, we begin with the highest \(z\)-bin and choose the first (and thus, closest) sample from it. If no intersection could be found, we continue with the next sample in the bin, or the next lower bin if the current bin does not contain more samples, and so on.

The NURBS surface is then evaluated at the optimized parameter \((u', v')\) to determine corresponding 3D point along with its normal. The cell value \(c_{i,j}\) is obtained from the \(z\)-coordinate of the point quantized with respect to a height resolution \(\rho_z\) as in equation 2, i.e.

\[
c_{i,j} = \lceil \frac{z}{\rho_z} \rceil
\]

The \(x, y\) coordinate of each cell is given by

\[
x = j * \rho_x * \cos((-i - 1) * \rho_y)
\]

\[
y = j * \rho_x * \sin((-i - 1) * \rho_y)
\]
As the descriptor embeds an environment of the poi, it represents a closed sequence such that the first and last row is considered adjacent and also the elements with the same column index corresponds to adjacent cell. Hence, the descriptor exhibits a cyclical property that is crucial for matching and determining the rotation parameter of the alignment transformation it represents (see section 3.4).

### 3.3 Similarity Measure

There exist multiple measure, for instance, correlation coefficient, mutual information, join entropy to obtain a quantitative value and are largely dependent on the proportion of overlap area to arrive at a score. To compare the source and target descriptors for each resolution, we rely on a specific similarity measure [Car09]. In point clouds with low overlap region, it is likely that the environment around the corresponding points of source and target scan appear similar resulting in selection of false pairs. To circumvent false matching that could be detrimental to registration accuracy, the similarity measure incorporates information from the non-overlapping region enlarging its descriptive capabilities. The objective of the algorithm is to maximize similarity $S(D_s, D_t)$ in equation 5 while comparing descriptors from two surfaces. $D(s, p)$ and $D(t, q)$ refers to a descriptor with respect to point $p, q$ from two surfaces $s, t$ respectively.

$$\text{argmax}_{p,q} S(D(s, p \in s), D(t, q \in t)) = (4)$$

The similarity measure $S(D_s, D_t)$ in equation 5 relies on the percentage of overlap ($d_{OL}$) between two surfaces and their proximity ($d_{OL}$), representing the average distance in overlapped area. The affect of overlap in the similarity measure can be modified using $\beta$ whereas $\alpha$ directly influences the average distance $d_{OL}$. The detailed derivation and additional concepts are provided in [Car12] and [Car09].

### 3.4 Multi-resolution Descriptor and Matching

The descriptor computation in section 3.2 and similarity measure in section 3.3 are core blocks of our hierarchical registration pipeline. The algorithm initially performs a descriptor matching at a coarse resolution of $16 \times 16$ with similarity measure described in section 3.3 between two sets of point-of-interest $s_{poi}$ and $t_{poi}$. The pairs $(s_{poi}^i, t_{poi}^j)$ are sorted in decreasing order of their similarity score at this coarse resolution to obtain a feasible set of starting points for the hierarchical matching of descriptors. We try the highest-ranked pairs first, avoiding initial match ups that are unlikely to yield good results.

The actual process of descriptor comparison then works on some $(s_{poi}^i, t_{poi}^j)$ pair in isolation. We name the initial pair of descriptors at some resolution the primary descriptors for source and target, respectively. A search is performed over each cell and its associated point on the surface of the primary source descriptor to find a secondary source descriptor that gives the greatest similarity to the target at some row shift $k$. This row shift represents rotation around the normal of the keypoint that spawned the secondary descriptor, determining the remaining free parameter needed to form a rigid transformation. When the rows of the secondary source descriptor had to be shifted $k$ times for highest similarity, then the rotation angle is $k \rho_k$, where $\rho_k$ is the angular step at the given resolution.

The 3D point and rotation associated with the secondary descriptor that gave the best match at the current resolution becomes the starting point for the primary descriptor at the next higher resolution. Figure 2 depicts a series of primary descriptors from coarse to fine resolution. The process terminates for a specific $(s_{poi}^i, t_{poi}^j)$ pair once the specified maximum resolution is reached.

The final alignment transformation for a points-of-interest pair results from the best-matching secondary source descriptor at the highest resolution. Since we now assume both source and target descriptors to describe the same point on the surface, we can find this alignment by going from world space $W$ into the local reference frame $s$ of the final secondary source descriptor, applying the rotation resulting from the final row shift $k$, and finally concatenating the transformation back to $W$ from the reference frame $t$ of the target descriptor:

$$W T_{align} = W T_s^{-1} \cdot R(k) \cdot W T_t$$

The transformation $W T_{f}$ for some descriptor reference frame $l$ is given directly from the keypoint $p$ that spawned the descriptor and its associated normal $n$ (see section 3.2). The rotational part of the transformation from $W$ to $l$ is obtained as follows:

$$\begin{bmatrix} W R_l \end{bmatrix} = \begin{bmatrix} x_l \quad n \times x_l \quad n \end{bmatrix}^T$$

The full transformation $W T_{f}$ then follows as

$$W T_{f} = \begin{bmatrix} W R_l & 0 \\ 0_{1 \times 3} & 1 \end{bmatrix}$$

Finally, a rotation matrix parametrized by some row shift $k$ is obtained as follows:

$$R(k) = \begin{bmatrix} \cos(\rho_k \cdot k) & \sin(\rho_k \cdot k) & 0 & 0 \\ -\sin(\rho_k \cdot k) & \cos(\rho_k \cdot k) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
Our key contribution of adapting the construction of Circon descriptors to continuous surfaces enables us to extend the hierarchy with much higher resolution descriptors, without being constrained by the original resolution of the point cloud. Like [Car12], we also reduce the search space further down in the hierarchy by halving the number of columns (i.e., radial steps away from the contour center) considered at each level. This is especially important for us since descriptor computation is significantly more expensive, and there is no need to consider points far away from the center of the radial contour at later stages. To mitigate the performance impact of our construction process even further, we also limit descriptor construction to 64 columns (see figure 2d). Every cell further away than 64 radial steps will thus be invalid, which the similarity measure is able to handle naturally as part of its down-weighting of non-overlapping regions (see section 3.3).

3.5 Stopping Criterion

The stopping criterion is essential for the iterative search to be convergent and importantly must embed characteristics of the descriptor. To this end, we choose three non-collinear points from the final source and target descriptors that each form a triangle around the respective Circon descriptor center. Such points can be easily obtained by selecting the same three equidistantly spaced cells from the first column of each descriptor and computing their 3D point equivalents. These triangles form a fictitious correspondence pair in accordance with Ferrero et al. [Car12]. The centroid of each triangle defines a local reference frame with fictitious transformation $T_{fict}$ between them as represented by equation 6. $T_{fict}$ is compared with rigid transformation $T_{align}$ in terms of a delta transformation:

$$\Delta T = T_{align} \cdot T_{fict}^{-1}$$  (10)

This delta transformation is evaluated for rotational and translational difference separately (see Appendix A for details). If the rotational difference $\Delta r$ and translation distance $\Delta t$ extracted from $\Delta T$ are smaller than their respective thresholds, the algorithm terminates and $T_{align}$ is returned as the final alignment transformation.

4 EVALUATION

To demonstrate the effectiveness of the proposed hierarchical registration method (HER), we perform our analysis on a publicly available dataset and compare quantitative measure with state-of-the-art registration method and also the hierarchical method proposed by Ferrero et al. [Car12] and for simplicity name it as ORI–HER. We ran all our experiments on an Intel Core i5-6500 clocked at 3.20GHz. The following section presents how we performed our evaluation including dataset description and the tested methods.

4.1 Dataset

We performed an evaluation on the comprehensive benchmark provided by Petrelli et al. [Pet15], which we refer to as the LRF dataset. This dataset contains multiple objects captured by sensors of different quality, ranging from consumer-level depth cameras to professional laser scanners resulting in point clouds with varying point density. Moreover, the benchmark includes challenging low-overlap registration pairs.


Figure 2: Descriptor cell values mapped to an image at the top and corresponding point cloud represented by the descriptor at the bottom of Bunny model. The restriction of higher resolution descriptors to 64 columns is visible in image (d).
4.2 Methods

We base our comparisons with ORI-HER [Car12] that uses hierarchical descriptors with similarity measure for coarse point cloud registration and state-of-the-art global registration method FGR [Zho16]. This methods has been shown to rival the accuracy of local methods and therefore ideal for comparison to understand the difference of accuracy with our proposed approach. The parameters for each methods and their respective values adapted during the evaluation process are described in separate Appendix-A.

4.3 Metrics

After solving the various registration problems with different methods based on the dataset discussed in section 4.1, we evaluate how well the results match the ground truth transforms with a data-dependent metric. We use root mean square error (RMSE) as a metric to quantify registration accuracy. The RMSE measures the mean error over all source cloud point locations for some computed transformation $T_{out}$ with respect to ground truth transformation $T_{g}$:

$$\text{RMSE}(T_{out}) = \frac{1}{d} \sqrt{\frac{1}{n} \sum_{i=1}^{n} ||T_{out}p_i - T_{g}p_i||^2} \quad (11)$$

The normalization factor $d$ represents the sampling distance. For the LRF dataset, where scans are represented as polygon meshes, we use the average mesh edge length.

5 RESULTS

The section elucidates on the results of our experiments that verify the proposed approach produces more accurate registration results than state-of-the-art approach. We will also examine various characteristics of our algorithm.

5.1 Accuracy

The accuracy comparison is based on datasets introduced in section 4.1 with multiple registration problems and RMSE defined in section 4.3 on the registration output after convergence. We categorize RMSE results into fine ($< 20d$), coarse ($[20d, 50d]$), failed (test) and measure the number of view pairs (Table 1) that fall under these categories with respect to the total number of registration problems. To refine the comparison, we concentrate on the fine registration view-pairs and Figure 5 shows the accuracy results for the LRF datasets falling under this category. Figure 5 asserts that the highest accuracy are obtained by our method compared to FGR and ORI-HER. Furthermore, The statistical measures in Table 2 based on fine RMSE delineates lowest values for our method among others. The continuous representation of scans followed by descriptor computation with finer resolution attributes to the significant improvement in the accuracy. Figure 3 and Figure 4 depict particularly difficult registration problems from the LRF dataset that FGR and ORI-HER failed to align, while our method achieved fine alignment.
Table 1: Categorization of registration view pairs from LRF datasets based on threshold as discussed in section 5.1

<table>
<thead>
<tr>
<th></th>
<th>No. Pairs</th>
<th>Fine</th>
<th>Coarse</th>
<th>Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORI-HER</td>
<td>1840 57</td>
<td>427</td>
<td>1356</td>
<td></td>
</tr>
<tr>
<td>FGR</td>
<td>1840 279</td>
<td>642</td>
<td>919</td>
<td></td>
</tr>
<tr>
<td>HER</td>
<td>1840 355</td>
<td>711</td>
<td>774</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Statistical Comparison of RMSE based on Fine threshold as discussed in section 5.1

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Avg</th>
<th>Std.dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORI-HER</td>
<td>1.162</td>
<td>23.66</td>
<td>5.99</td>
<td>4.07</td>
</tr>
<tr>
<td>FGR</td>
<td>0.48</td>
<td>21.67</td>
<td>3.49</td>
<td>2.77</td>
</tr>
<tr>
<td>HER</td>
<td>0.34</td>
<td>20.24</td>
<td>2.58</td>
<td>2.25</td>
</tr>
</tbody>
</table>

Figure 5: Accuracy comparison of pairwise registration methods for the LRF datasets. The box height represents the 25th and 75th percentile. The whiskers represent the 5th and 95th percentiles. The middle line represents the median. Better registration results are characterized by lower RMSE.

Figure 6: Runtime comparison of the proposed approach with ORI-HER, FGR including the preprocessing step for each methods. Our method exhibits higher runtime in the core algorithm part owing to its multiple descriptor computation compared to FGR but considerably lesser to ORI-HER more expensive to compute for our method, we typically find an optimal solution within the first few keypoint pairs, as we can use higher resolution descriptors independent of the sampling situation, greatly speeding up convergence.

6 CONCLUSION AND OUTLOOK

In this paper, we have presented a hierarchical (coarse-to-fine) global registration method. The 2D descriptors built from a continuous representation of each point cloud enables us to embed higher accuracy into them. Our evaluation based on the diverse LRF dataset shows considerable improvement in the accuracy compared to the state-of-the-art FGR algorithm.

Nonetheless, a handful of improvements could be made, most obviously concerning speed. We chose NURBS surfaces for their versatility and ready availability of implementations, but the fitting process is expensive. Furthermore, as described in section 3.2, intersecting rays with a NURBS surface requires, in general, a numeric solution. A more specialized surface representation could accelerate both fitting and descriptor building, in turn making the use of higher resolution descriptors more feasible and thus improving accuracy even further.

Alternatively, leveraging GPU hardware to compute the descriptors promises immense speedups. This could be achieved using the programmable rasterization pipeline to generate orthographic depth maps of the continuous surfaces as viewed in the descriptor reference frame. The resolution of these depth maps can be chosen as high as needed to enable artifact-free sampling at the Circion cell centers. We expect this would bring down the computation time spend on building descriptors to a fraction of what it is in our current implementation.

7 REFERENCES


