Scale-Dependent/Invariant Local 3D Shape Descriptors for Fully Automatic Registration of Multiple Sets of Range Images

John Novatnack and Ko Nishino

Department of Computer Science, Drexel University
{jmn27,kon}@drexel.edu

Abstract. Despite the ubiquitous use of range images in various computer vision applications, little has been investigated about the size variation of the local geometric structures captured in the range images. In this paper, we show that, through canonical geometric scale-space analysis, this geometric scale-variability embedded in a range image can be exploited as a rich source of discriminative information regarding the captured geometry. We extend previous work on geometric scale-space analysis of 3D models to analyze the scale-variability of a range image and to detect scale-dependent 3D features – geometric features with their inherent scales. We derive novel local 3D shape descriptors that encode the local shape information within the inherent support region of each feature. We show that the resulting set of scale-dependent local shape descriptors can be used in an efficient hierarchical registration algorithm for aligning range images with the same global scale. We also show that local 3D shape descriptors invariant to the scale variation can be derived and used to align range images with significantly different global scales. Finally, we demonstrate that the scale-dependent/invariant local 3D shape descriptors can even be used to fully automatically register multiple sets of range images with varying global scales corresponding to multiple objects.

1 Introduction

Range images play central roles in an increasing number of important computer vision applications ranging from 3D face recognition to autonomous vehicle navigation and digital archiving. Yet, the scale variation of geometric structures captured in range images are largely ignored or simply viewed as perturbations of the underlying geometry that need to be accounted for in subsequent processing. Although several methods, mostly, for extracting scale-invariant or multi-resolution features or descriptors from range images based on smoothing 3D coordinates or curvature values of the vertices have been proposed in the past [1,2,3,4,5,6], they are prone to topological errors induced by the lack of canonical scale analysis as discussed in [7]. Most important, they do not fully exploit the rich discriminative information encoded in the scale-variability of local geometric structures that can in turn lead to novel computational methods for processing range images.

D. Forsyth, P. Torr, and A. Zisserman (Eds.): ECCV 2008, Part III, LNCS 5304, pp. 440–453, 2008.

In this paper, we introduce a comprehensive framework for analyzing and exploiting the scale-variability of geometric structures captured in range images based on canonical analysis of its geometric scale-space. We derive novel local 3D shape descriptors that naturally encode the inherent scale of local geometric structures. The geometric scale-space analysis of range images can be viewed as an extension of previous work on geometric scale-space analysis of 3D mesh models by Novatnack and Nishino [7]. The key idea underlying the newly derived geometric scale-space construction and analysis is that a range image is readily a 2D projection of the surface geometry of a 3D shape. We show that we may directly compute a geometric scale-space of a range image in this 2D projection with unique 2D operators defined on the surface geodesics. Based on the geometric scale-space analysis we detect scale-dependent geometric features, more specifically corners, and their inherent spatial extents. We show that we can encode the geometric information within the spatial extent of each feature in a scale-dependent local 3D shape descriptor that collectively form a sparse hierarchical representation of the surface geometry captured in the range images. We demonstrate how this representation can be exploited to robustly register a set of range images with a consistent global scale. Furthermore, we show how we may define a local 3D shape descriptor that is invariant to the variation of the inherent local scale of the geometry, which can be used to register a set of range images with unknown or inconsistent global scales.

We demonstrate the effectiveness of the novel scale-dependent/invariant local 3D shape descriptors by automatically registering a number of models of varying geometric complexity. These registration results can be used as approximate alignments which can then be refined using a global registration algorithm that together realize fully automatic registration of multiple range images without any human intervention. We further demonstrate the effectiveness of our framework by fully automatically registering a set of range images corresponding to multiple 3D models, simultaneously. Note that previous work on fully automatic range image registration assume that the range images capture a single object or scene [8,9,10,11,12,13]. We, on the other hand, show that the novel scale-dependent/invariant descriptors contain rich discriminative information that enables automatic extraction of individual objects from an unordered mixed set of range images capturing multiple objects. To our knowledge, this work is the first to report such capability.

2 Geometric Scale-Space of a Range Image

We first construct and analyze the geometric scale-space of a range image. This part of our framework can be viewed as an extension of the work by Novatnack and Nishino [7] to range images. Readers are referred to [7] for details. The key insight underlying the extension of this approach to range images is that each range image is already a dense and regular projection, mostly a perspective projection, of a single view of the surface of the target 3D shape. Furthermore, the distortion map used for accounting for the distortions induced by the embedding in [7] is unnecessary, as the geodesics can be approximated directly from the range image itself.

We build the geometric scale-space of a range image $\mathbf{R} : D \rightarrow \mathbb{R}^3$, where D is a 2D domain in $\mathbb{R}^2$, by first constructing a normal map $\mathbf{N}$ in the same domain by triangulating the range image and computing a surface normal for each vertex. As noted in [7], it is important to use the surface normals as the base representation since directly smoothing 3D coordinates can result in topological changes and higher-order derivative geometric entities such as the curvatures can be sensitive to noise.

In order to construct a geometric scale-space of a range image that accurately encodes the scale-variability of the underlying surface geometry, we define all operators in terms of the geodesic distance rather than the Euclidean 3D or 2D distances. To efficiently compute the geodesic distance between two points in a range image, we approximate it with the sum of Euclidean 3D distances between vertices along the path joining the two points in the range image; given two points $\mathbf{u}, \mathbf{v} \in D$ we approximate the geodesic distance $d(\mathbf{u}, \mathbf{v})$ as

$$d(\mathbf{u}, \mathbf{v}) \approx \sum_{\mathbf{u}_i \in \mathcal{P}(\mathbf{u}, \mathbf{v}), \neq \mathbf{v}} \| \mathbf{R}(\mathbf{u}_i) - \mathbf{R}(\mathbf{u}_{i+1}) \| , \tag{1}$$

where $\mathcal{P}$ is a list of vertex points in the range image on the path between $\mathbf{u}$ and $\mathbf{v}$. If the path between $\mathbf{u}$ and $\mathbf{v}$ crosses an unsampled point in the range image then we define the geodesic distance as infinity. We also parse the range image and detect depth discontinuities by marking vertex points whose adjacent points lie further than a predetermined 3D distance and define the geodesic distance as infinity if the path crosses such points. When approximating geodesics from a point of interest outwards, the sum can be computed efficiently by storing the geodesic distances of all points along the current frontier and reusing these when considering a set of points further away.

We construct the geometric scale-space of a base normal map $\mathbf{N}$ by filtering the normal map with Gaussian kernels of increasing standard deviation σ, where the kernel is defined in terms of the geodesic distance. The resulting geometric scale-space directly represents the inherent scale-variability of local geometric structures captured in range images and serves as a rich basis for further scale-variability analysis of range image data.

3 Scale-Dependent Features in a Range Image

We may detect geometric features, in our case corner points, and their associated inherent scale, in other words their (relative) natural support sizes, in the geometric scale-space of a range image. The 3D geometric corners are first detected at each discrete scale by applying a corner detector proposed for the geometric scale-space of a 3D model by Novatnack and Nishino [7]. For a point $\mathbf{u}$ in the normal map $\mathbf{N}^\sigma$ at scale σ, the corner response is computed using the Gram matrix $\mathcal{M}(\mathbf{u}; \sigma, \tau)$. The Gram matrix is defined in terms of scale-normalized first-order derivatives in the horizontal $\widetilde{\mathbf{N}_s}^{\sigma}$ and vertical $\widetilde{\mathbf{N}_t}^{\sigma}$ directions, which themselves are the normal curvatures in these directions (see [7] for details):

$$\mathcal{M}(\mathbf{u};\sigma,\tau) = \sum_{\mathbf{v}\in\mathcal{W}} \begin{bmatrix} \widetilde{\mathbf{N}_s}^{\sigma}(\mathbf{v})^2 & \widetilde{\mathbf{N}_s}^{\sigma}(\mathbf{v})\widetilde{\mathbf{N}_t}^{\sigma}(\mathbf{v}) \\ \widetilde{\mathbf{N}_s}^{\sigma}(\mathbf{v})\widetilde{\mathbf{N}_t}^{\sigma}(\mathbf{v}) & \widetilde{\mathbf{N}_t}^{\sigma}(\mathbf{v})^2 \end{bmatrix} g(\mathbf{v};\mathbf{u},\tau)\,, \tag{2}$$

where $\mathcal{W}$ is the local window window to be considered, σ is the particular scale in the geometric scale-space representation, and τ is the weighting of the points in the Gram matrix.

A set of corner points at each scale is detected by searching for spatial local maxima of the corner detector responses. Corners lying along edge points are pruned by thresholding the variance of the second-order partial derivatives. This results in a set of corner points at a number of discrete scales in the geometric scale-space. In order to determine the intrinsic scale of each corner, we then search for local maxima of the corner detector responses across the scales as was originally proposed for 2D scale-space [14]. The result is a comprehensive set of scale-dependent corners, where the support size of each corner follows naturally from the scale in which it was detected. Figure 1 shows the set of scale-dependent corners detected on two range images of a Buddha model. Note that the corners are well dispersed across scales, and that there are a large number of corresponding corner points at the correct corresponding scales.

4 Local 3D Shape Descriptors

Once we detect scale-dependent features via geometric scale-space analysis we may define novel 3D shape descriptors that naturally encode relevant local geometric structure. There are a wide variety of 3D shape descriptors that have been previously proposed [15,16,17,18,19,20]. Many of these suffer from the limitation that they are sensitive to the sampling density of the underlying geometry and the size of their support region cannot be canonically determined. In our case, the associated inherent scale of each scale-dependent corner directly tells us the natural spatial extent (the support size) of the underlying local geometric structure. This information can then in turn be used to identify the size of the neighborhood of each corner that should be encoded in a local shape descriptor. At the same time, we construct dense and regular 2D descriptors that are insensitive to the resolution of the input range images.

4.1 Exponential Map

We construct both our scale-dependent and scale-invariant local 3D shape descriptors by mapping and encoding the local neighborhood of a scale-dependent corner to a 2D domain using the *exponential map*. The exponential map is a mapping from the tangent space of a surface point to the surface itself [21]. Given a unit vector $\mathbf{w}$ lying on the tangent plane of a point $\mathbf{u}$, there is a unique geodesic Γ on the surface such that $\Gamma(0) = \mathbf{u}$ and $\Gamma'(0) = \mathbf{w}$. The exponential map takes

Fig. 1. Scale-dependent corners and scale-dependent local 3D shape descriptors computed based on geometric scale-space analysis of two range images. The scale-dependent corners are colored according to their inherent scales, with red and blue corresponding to the coarsest and finest scales, respectively. The scale-dependent local 3D shape descriptors capture local geometric information in the natural support regions of the scale-dependent features.

a vector $\mathbf{w}$ on the tangent plane and maps it to the point on the geodesic curve at a distance of 1 from $\mathbf{u}$, or $\text{Exp}(\mathbf{w}) = \Gamma(1)$. Following this, any point $\mathbf{v}$ on the surface in the local neighborhood of $\mathbf{u}$ can be mapped to $\mathbf{u}$'s tangent plane, often referred to as the Log map, by determining the unique geodesic between $\mathbf{u}$ and $\mathbf{v}$ and computing the geodesic distance and polar angle of the tangent to the geodesic at $\mathbf{u}$ in a predetermined coordinate frame $\{\mathbf{e}_1, \mathbf{e}_2\}$ on the tangent plane. This ordered pair is referred to as the *geodesic polar coordinates* of $\mathbf{v}$.

The exponential map has a number of properties that are attractive for constructing a 3D shape descriptor, most important, that it is a local operator. Although fold-overs may occur if this neighborhood is too large, the local nature of the scale-dependent and scale-invariant descriptors implies this will rarely happen. In practice we have observed fold-overs on an extremely small number of features, mostly near points of depth discontinuities. Although the exponential map is not, in general, isometric, the geodesic distance of radial lines from the feature point are preserved. This ensures that corresponding scale-dependent corners will have mostly consistent shape descriptors among different views, i.e. different range images. In addition, because the exponential map is defined at the feature point, it does not rely on the boundary of the encoded neighborhood like harmonic images does [22].

4.2 Scale-Dependent Local 3D Shape Descriptor

We construct a scale-dependent local 3D shape descriptor for a scale-dependent corner at $\mathbf{u}$ whose scale is σ by mapping each point $\mathbf{v}$ in the neighborhood of $\mathbf{u}$ to a 2D domain using the geodesic polar coordinates $\mathcal{G}$ defined as

$$\mathcal{G}(\mathbf{u}, \mathbf{v}) = (d(\mathbf{u}, \mathbf{v}), \theta_{\mathcal{T}}(\mathbf{u}, \mathbf{v})) , \tag{3}$$

where again $d(\mathbf{u}, \mathbf{v})$ is the geodesic distance between $\mathbf{u}$ and $\mathbf{v}$ and $\theta_{\mathcal{T}}(\mathbf{u}, \mathbf{v})$ is the polar angle of the tangent of the geodesic between $\mathbf{u}$ and $\mathbf{v}$ defined relative to a fixed bases $\{\mathbf{e}_1, \mathbf{e}_2\}$. In practice we approximate this angle by orthographically projecting $\mathbf{v}$ onto the tangent plane of $\mathbf{u}$ and measuring the polar angle of the intersection point. The radius of the descriptor is set proportional to the inherent scale of the scale-dependent corner σ to encode geometric information in the natural support region of each scale-dependent corner.

After mapping each point in the local neighborhood of $\mathbf{u}$ to its tangent plane we are left with a sparse 2D representation of the local geometry around $\mathbf{u}$. We interpolate a geometric entity encoded at each vertex to construct a dense and regular representation of the neighborhood of $\mathbf{u}$ at scale σ. Note that this makes the descriptor insensitive to resolution changes of the range images. We choose to encode the surface normals from the original range image, rotated such that the normal at the center point $\mathbf{u}$ points in the positive z direction. The resulting dense 2D descriptor is invariant up to a single rotation (the in-plane rotation on the tangent plane). We resolve this ambiguity by aligning the maximum principal curvature direction at $\mathbf{u}$ to the horizontal axis $\mathbf{e}_1$ in the geodesic polar coordinates, resulting in a rotation-invariant shape descriptor. Once this local basis has been fixed we re-express each point in terms of the normal coordinates, with the scale-dependent corner point $\mathbf{u}$ at the center of the descriptor.

We refer to this dense 2D scale-dependent descriptor of the local 3D shape as $\mathbf{G}_{\mathbf{u}}^{\sigma}$ for a scale-dependent corner at $\mathbf{u}$ and with scale σ. Figure 1 shows subsets of scale-dependent local 3D shape descriptors computed at scale-dependent corners in two range images of a Buddha model.

4.3 Scale-Invariant Local 3D Shape Descriptor

The scale-dependent local 3D shape descriptors provides a faithful sparse representation of the surface geometry in different range images when their global scales are the same or are known, e.g. when we know that the range images are captured with the same range finder. In order to enable comparison between range images that do not have the same global scale, we also derive a scale-invariant local 3D shape descriptor $\widehat{\mathbf{G}}_{\mathbf{u}}^{\sigma}$.

We may safely assume that the scales of local geometric structures relative to the global scale of a range image remains constant as the global scale of a range image is altered. Note that this assumption holds as long as the geometry captured in the range image is rigid and does not go under any deformation, for instance, as it is captured with possibly different range sensors. We may then construct a set of scale-invariant local 3D shape descriptors by first building a set of scale-dependent local 3D shape descriptors and then normalizing each descriptor's size to a constant radius. Such a scale-invariant representation of the underlying geometric structures enables us to establish correspondences between a pair of range images even when the global scale is different and unknown.

5 Pairwise Registration

The novel scale-dependent and scale-invariant local 3D shape descriptors contain rich discriminative information regarding the local geometric structures. As a practical example, we show the effectiveness of these descriptors in range image registration, one of the most fundamental steps in geometry processing. In particular, we show how the scale-dependent local 3D shape descriptors form a hierarchical representation of the geometric structures that can be leveraged in a coarse-to-fine registration algorithm. We also show how the scale-invariant local shape descriptors can be used to establish correspondences and compute the transformation between a pair of range images with completely different global scales.

5.1 Similarity Measure

Since each descriptor is a dense 2D image of the surface normals in the local neighborhood we may define the similarity of the local 3D shape descriptors as the normalized cross-correlation of surface normal fields using the angle differences,

$$\mathcal{S}(\mathbf{G}^{\sigma}_{\mathbf{u}_1}, \mathbf{G}^{\sigma}_{\mathbf{u}_2}) = \frac{\pi}{2} - \frac{1}{|A \cap B|} \sum_{\mathbf{v} \in A \cap B} \arccos(\mathbf{G}^{\sigma}_{\mathbf{u}_1}(\mathbf{v}) \cdot \mathbf{G}^{\sigma}_{\mathbf{u}_2}(\mathbf{v}))\,, \tag{4}$$

where A and B are the set of points in the domain of $\mathbf{G}^{\sigma}_{\mathbf{u}_1}$ and $\mathbf{G}^{\sigma}_{\mathbf{u}_2}$, respectively. Here, the similarity measure is defined in terms of the scale-dependent descriptors, but the definition for the scale-invariant descriptors is the same with $\widehat{\mathbf{G}}$ substituted for $\mathbf{G}$.

5.2 Pairwise Registration with Scale-Dependent Descriptors

The hierarchical structure of the set of scale-dependent local 3D shape descriptors can be exploited when aligning a pair of range images $\{\mathbf{R}_1, \mathbf{R}_2\}$ with the same global scale. Note that if we know that the range images are captured with the same range scanner, or if we know the metrics of the 3D coordinates, e.g. centimeters or meters, we can safely assume that they have, or we can covert them to, the same global scale.

Once we have a set of scale-dependent local 3D shape descriptors for each range image, we construct a set of possible correspondences by matching each descriptor to the n most similar[1]. The consistency of the global scale allows us to consider only those correspondences at the same scale in the geometric scale-space, which greatly decreases the number of correspondences that must be later sampled. We find the best pairwise rigid transformation between the two range images by randomly sampling this set of potential correspondences and determining the one that maximizes the area of overlap between the two

[1] In our our experiments n is set in the range of $5 \sim 10$.

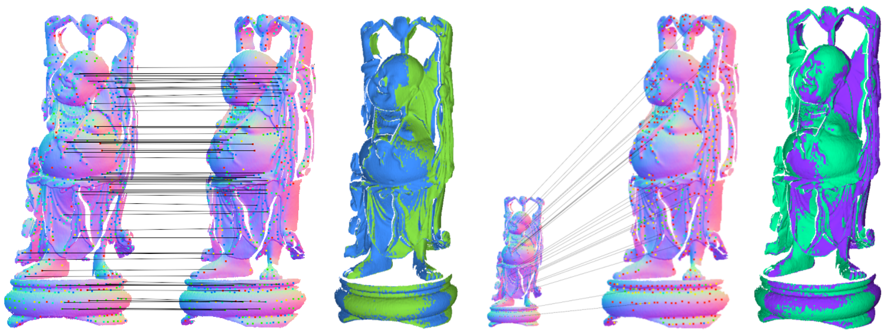

Fig. 2. (a) Aligning two range images with the same global scale using a set of scale-dependent local 3D shape descriptors. On the left we show the 67 point correspondences found with our matching algorithm and on the right the result of applying the rigid transformation estimated from the correspondences. (b) Aligning two range images with inconsistent global scales using a set of scale-invariant local 3D shape descriptors. On the left we show the 24 point correspondences found with our matching algorithm and on the right the results of applying the estimated 3D similarity transformation. Both the scale-dependent and -invariant descriptors realize very accurate and efficient automatic pairwise registration of range images.

range images, similar to RANSAC [23]. However, rather then sampling the correspondences at all scales simultaneously, we instead sample in a coarse-to-fine fashion, beginning with the descriptors with the coarsest scale and ending with descriptors with the finest scale. This enables us to quickly determine a rough alignment between two range images, as there are, in general, fewer features at coarser scales.

For each scale σ_i we randomly construct $N\sigma_i$ sets of 3 correspondences, where each correspondence has a scale between σ_1 and σ_i. For each correspondence set $\mathcal{C}$ we estimate a rigid transformation $\mathcal{T}$, using the method proposed by Umeyama [24], and then add to $\mathcal{C}$ all those correspondences $(\mathbf{u}_j, \mathbf{v}_j, \sigma_j)$ where $\| \mathcal{T} \cdot \mathbf{R}_1(\mathbf{u}_j) - \mathbf{R}_2(\mathbf{v}_j) \| \leq \alpha$ and $\sigma_j \leq \sigma_i$. Throughout the sampling process we keep track of the transformation and correspondence set that yield the maximum area of overlap. Once we begin sampling the next finer scale σ_{i+1} we initially test whether the correspondences at that scale improve the area of overlap induced by the current rigid transformation. This allows us to quickly add a large number of correspondences at finer scales efficiently without drawing an excessive number of samples.

Figure 2(a) shows the results of applying our pairwise registration algorithm to two views of the Buddha model. The number of correspondences is quite large and the correspondences are distributed across all scales. Although the result is an approximate alignment, since for instance slight perturbations in the scale-dependent feature locations may amount to slight shifts in the resulting registration, the large correspondence set established with the rich shape descriptors leads to very accurate estimation of the actual transformation.

5.3 Pairwise Registration with Scale-Invariant Descriptors

We may align a pair of range images $\{\mathbf{R}_1, \mathbf{R}_2\}$ with different global scales using the scale-invariant local 3D shape descriptors, which amounts to estimating the 3D similarity transformation between the range images. Since we no longer know the relative global scales of the range images, we must consider the possibility that a feature in one range image may correspond to a feature detected at a different scale in the second range image. Our algorithm proceeds by first constructing a potential correspondence set that contains, for each scale-invariant local 3D shape descriptor in the first range image $\mathbf{R}_1$, the n most similar in the second range image $\mathbf{R}_2$. We find the best pairwise similarity transformation by applying RANSAC to this potential correspondence set. For each iteration the algorithm estimates the 3D similarity transformation [24] and computes the area of overlap. The transformation which results in the maximum area of overlap is considered the best.

Figure 2(b) shows the result of applying our algorithm to two views of the Buddha model with a relative global scale difference of approximately 2.4. Despite the considerable difference in the relative global scales, we can recover the similarity transformation accurately without any initial alignments or assumptions about the models and their global scales.

6 Multiview Registration

Armed with the pairwise registration using scale-dependent/invariant descriptors, we may derive a fully automatic range image registration framework that exploits the geometric scale-variability. We show that the scale-dependent and scale-invariant descriptors can be used to register a set of range images both with and without global scale variations without any human intervention. Most important, we show that we can register a mixed set of range images corresponding to **multiple 3D models** simultaneously and fully automatically[2].

6.1 Fully Automatic Registration

Given a set of range images $\{\mathbf{R}_1, ..., \mathbf{R}_n\}$, our fully automatic range image registration algorithm first constructs the geometric scale-space of each range image. Scale-dependent features are detected at discrete scales and then combined into a single comprehensive scale-dependent feature set, where the support size of each feature follows naturally from the scale in which it was detected. Each feature is encoded in either a scale-dependent or scale-invariant local shape descriptor, depending on whether the input range images have a consistent global scale or

[2] In all our experiments, we randomized the order of the range images to ensure that no a priori information is given to the algorithm.

not. We then apply the appropriate pairwise registration algorithm, presented in the previous sections, to all pairs of range images in the input set to recover the pairwise transformations. We augment each transformation with the area of overlap resulting from the transformation. Next we construct a graph similar to the model graph [9], where each range image is represented with a vertex and each pairwise transformation and area of overlap is encoded in a weighted edge. We prune edges with an area of overlap less then ϵ. In order to construct the final set of meshes $\{\mathcal{M}_1, ..., \mathcal{M}_m\}$ we compute the maximum spanning tree of the model graph and register range images in each connected component using their estimated corresponding transformations. The alignment obtained by our algorithm is approximate yet accurate enough to be directly refined by any ICP-based registration algorithm without any human intervention, resulting in a fully automatic range image registration algorithm.

6.2 Range Images with Consistent Global Scale

Figure 3 illustrates the results of applying our framework independently to 15 views of the Buddha model and 12 views of the armadillo model, with consistent global scales. Scale-dependent local shape descriptors were detected at 5 discrete

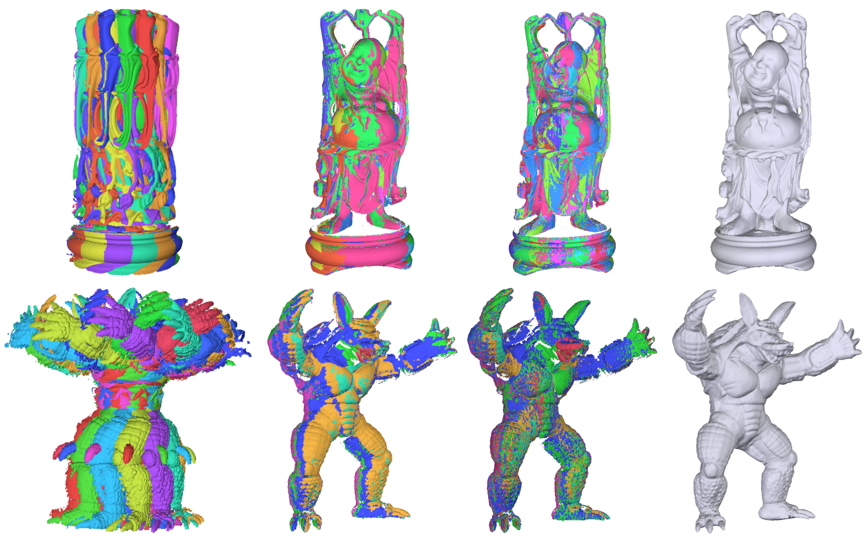

Fig. 3. Fully automatic registration of 15 views of the Buddha model and 12 views of the armadillo model using scale-dependent local descriptors. First column shows the initial set of range images. Note that no initial alignment is given and they are situated as is. Second column shows the approximate registration obtained with our framework, which is further refined with multi-view ICP [25] in the third column. Finally a water tight model is built using a surface reconstruction algorithm [26]. The approximate registration obtained with our framework is very accurate and enables direct refinement with ICP-based methods which otherwise require cumbersome manual initial alignment.

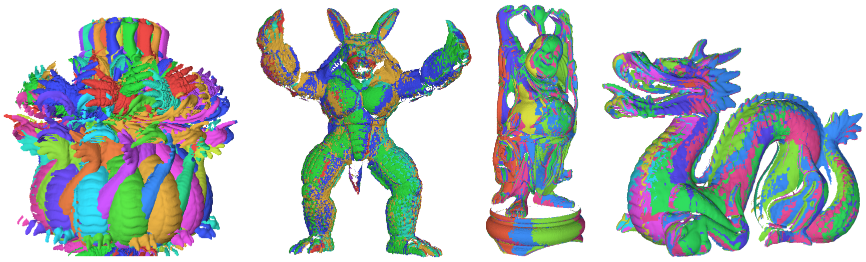

Fig. 4. Automatic registration of a set of range images of multiple objects: total 42 range images, 15 views of the Buddha model, 12 views of the armadillo, and 15 views of the dragon model. The scale-dependent local 3D shape descriptors contain rich discriminative information that enables automatic discovery of the three disjoint models from the mixed range image set. Note that the results shown here have not been post-processed with a global registration algorithm.

scales, $\sigma = \{0.5, 1, 1.5, 2, 2.5\}$, in the geometric scale-space. The approximate registration results after applying our matching method using scale-dependent local 3D shape descriptors are refined using multi-view ICP [25] and a watertight model is computed using a surface reconstruction method for oriented points[26]. We may quantitatively evaluate the accuracy of our approximate registration using the local 3D shape descriptors by measuring the displacement of each vertex in each range image from the final watertight model. The average distances for all the vertices in all range images for the armadillo and Buddha models, relative to the diameter of the models, were 0.17% and 0.29% percent, respectively. The results show that the scale-dependent local 3D shape descriptors provide rich information leading to accurate approximate registration that enables fully automatic registration without any need of initial estimates.

Next, we demonstrate the ability of our framework to simultaneously register range images corresponding to multiple 3D models. In order to automatically discover and register the individual models from a mixed set of range images, we prune the edges on the model graph that correspond to transformations with an area of overlap less then some threshold. In practice, we found this threshold easy to set as our framework results in approximate alignments that are very accurate. Figure 4 summarizes the results. Note that no refinement using global registration algorithms has been applied to these results to clarify the accuracy of our method, but can easily be applied without any human intervention.

6.3 Range Images with Inconsistent Global Scale

Next we demonstrate the effectiveness of our framework for fully automatically registering a number of range images with unknown global scales. Figure 5 illustrates the results of applying our framework to 15 views of the Buddha and dragon models. Each range image was globally scaled by a random factor between 1 and 4 – on average 2.21 and 2.52 for the Buddha and dragon, respectively. For each pair of adjacent range images the average errors in the estimated scales

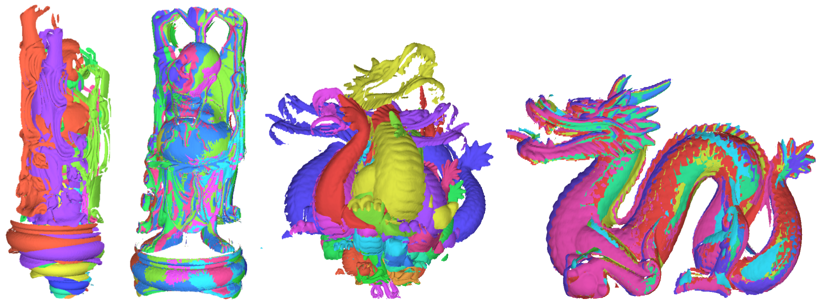

Fig. 5. Automatic registration of 15 views of the Buddha and dragon models each with a random global scaling from 1 to 4. For each model we visualize the initial set of range images and the approximate alignment obtained by our framework. Even with the substantial variations in the global scale, the scale-invariant local 3D shape descriptors enables us to obtain accurate (approximate) registrations without any assumptions about the initial poses.

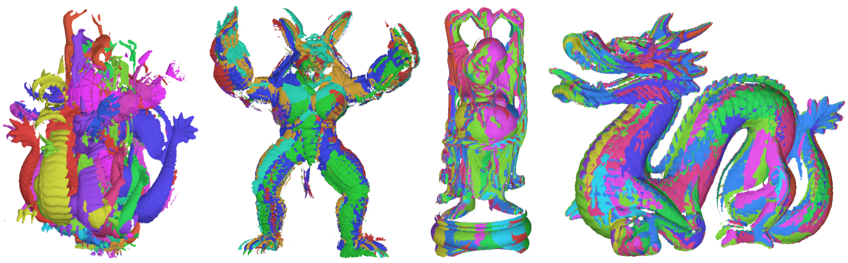

Fig. 6. Automatic approximate registration of 42 randomly scaled range images consisting of 15 views of the Buddha model, 12 views of the armadillo and 15 views of the dragon model. Each range image was randomly scaled by a factor between 1 and 4. The scale-invariant local 3D shape descriptors enables automatic (approximate) registration of the 3 models from this mixed set of range images without any a priori information, which can be directly refined with any ICP-based registration algorithm to arrive at a set of watertight models.

after our approximate registration using scale-invariant local 3D shape descriptors were 1.6% for the dragon and 0.4% for the Buddha model. These results show that even with substantial variations in the global scale, our method successfully aligns the range images with high accuracy, which is good enough for subsequent refinement with ICP-based methods as in the examples shown in Figure 3 without any manual intervention.

Finally, Figure 6 illustrates the results of applying our framework to 42 range images corresponding to three different models that have been randomly scaled by a factor between 1 and 4. Again, despite the significant scale variations, our scale-invariant representation of the underlying local geometric structures enables us to fully automatically discover and register all three models simultaneously without any human intervention.

7 Conclusion

In this paper, we introduced a comprehensive framework for analyzing and exploiting the geometric scale-variability of geometric structures captured in range images. Based on the geometric scale-space analysis of range images, we derived novel scale-dependent and scale-invariant local 3D shape descriptors. We demonstrated the effectiveness of exploiting scale-variability in these descriptors by using them in fully automatic registration of range images. Most important, we showed that the discriminative power encoded in these descriptors are extremely strong, so much so that they enable fully automatic registration of multiple objects from a mixed set of unordered range images. To our knowledge, this work is the first to report such capability. We strongly believe that the results indicate that our framework as well as the descriptors themselves can lead to novel robust and efficient range image processing methods in a variety of important applications beyond registration.

Acknowledgment

The range images used in the experiments are courtesy of Stanford Computer Graphics Laboratory[3]. This work was in part supported by National Science Foundation under CAREER award IIS-0746717.

References

1. Brady, M., Ponce, J., Yuille, A., Asada, H.: Describing Surfaces. Technical Report AIM-822, MIT Artificial Intelligence Laboratory Memo (1985)
2. Ponce, J., Brady, M.: Toward A Surface Primal Sketch. In: IEEE Int'l Conf. on Robotics and Automation, vol. 2, pp. 420–425 (1985)
3. Morita, S., Kawashima, T., Aoki, Y.: Hierarchical Shape Recognition Based on 3D Multiresolution Analysis. In: European Conf. on Computer Vision, pp. 843–851 (1992)
4. Akagunduz, E., Ulusoy, I.: Extraction of 3D Transform and Scale Invariant Patches from Range Scans. In: IEEE Int'l Conf. on Computer Vision and Pattern Recognition, pp. 1–8 (2007)
5. Li, X., Guskov, I.: 3D Object Recognition From Range Images Using Pyramid Matching. In: ICCV Workshop on 3D Representation for Recognition, pp. 1–6 (2007)
6. Dinh, H.Q., Kropac, S.: Multi-Resolution Spin-Images. In: IEEE Int'l Conf. on Computer Vision and Pattern Recognition, pp. 863–870 (2006)
7. Novatnack, J., Nishino, K.: Scale-Dependent 3D Geometric Features. In: IEEE Int'l Conf. on Computer Vision (2007)
8. Chen, C.S., Hung, Y.P., Cheng, J.B.: RANSAC-based DARCES: A New Approach to Fast Automatic Registration of Partially Overlapping Range Images. IEEE Trans. on Pattern Analysis and Machine Intelligence 21(11), 1229–1234 (1999)

[3] `http://www-graphics.stanford.edu/data/3Dscanrep/`

9. Huber, D., Hebert, M.: Fully Automatic Registration of Multiple 3D Data Sets. Image and Vision Computing 21(7), 637–650 (2003)
10. Mian, A., Bennamoun, M., Owens, R.: From Unordered Range Images to 3D Models: A Fully Automatic Multiview Correspondence Algorithm. In: Theory and Practice of Computer Graphics, pp. 162–166 (2004)
11. Gelfand, N., Mitra, N., Guibas, L., Pottmann, H.: Robust Global Registration. In: Symposium on Geometry Processing, pp. 197–206 (2005)
12. Makadia, A., Patterson, A., Daniilidis, K.: Fully Automatic Registration of 3D Point Clouds. In: IEEE Int'l Conf. on Computer Vision and Pattern Recognition, vol. 1, pp. 1297–1304
13. ter Haar, F., Veltkamp, R.: Automatic Multiview Quadruple Alignment of Unordered Range Scans. In: IEEE Shape Modeling International, pp. 137–146 (2007)
14. Lindeberg, T.: Feature Detection with Automatic Scale Selection. Int'l Journal of Computer Vision 30, 77–116 (1998)
15. Stein, F., Medioni, G.: Structural Indexing: Efficient 3-D Object Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 14, 125–145 (1992)
16. Chua, C., Jarvis, R.: Point Signatures: A New Representation for 3D Object Recognition. Int'l Journal of Computer Vision 25(1), 63–85 (1997)
17. Johnson, A., Hebert, M.: Using Spin-Images for Efficient Multiple Model Recognition in Cluttered 3-D Scenes. IEEE Trans. on Pattern Analysis and Machine Intelligence 21(5), 433–449 (1999)
18. Sun, Y., Abidi, M.: Surface Matching by 3D Point's Fingerprint. In: IEEE Int'l Conf. on Computer Vision, vol. 2, pp. 263–269 (2001)
19. Frome, A., Huber, D., Kolluri, R., Bulow, T., Malik, J.: Recognizing Objects in Range Data Using Regional Point Descriptors. In: European Conf. on Computer Vision (May 2004)
20. Skelly, L.J., Sclaroff, S.: Improved Feature Descriptors for 3-D Surface Matching. In: Proc. SPIE Conf. on Two- and Three-Dimensional Methods for Inspection and Metrology, vol. 6762, pp. 63–85 (2007)
21. Do Carmo, M.: Differential Geometry of Curves and Surfaces. Prentice Hall, Englewood Cliffs (1976)
22. Zhang, D., Hebert, M.: Harmonic Maps and Their Applications in Surface Matching. In: IEEE Int'l Conf. on Computer Vision and Pattern Recognition (1999)
23. Fischler, M., Bolles, R.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Commun. ACM 24(6), 381–395 (1981)
24. Umeyama, S.: Least-Squares Estimation of Transformation Parameters Between Two Point Patterns. IEEE Trans. on Pattern Analysis and Machine Intelligence 13(4), 376–380 (1991)
25. Nishino, K., Ikeuchi, K.: Robust Simultaneous Registration of Multiple Range Images. In: Asian Conf. on Computer Vision, pp. 454–461 (2002)
26. Kazhdan, M., Bolitho, M., Hoppe, H.: Poisson Surface Reconstruction. In: Eurographics Symp. on Geometry Processing, pp. 61–70 (2006)