Sampling Relevant Points for Surface Registration

Andrea Torsello, Emanuele Rodolà, and Andrea Albarelli
Dipartimento di Scienze Ambientali, Informatica e Statistica
Università Ca' Foscari Venezia
Venice, Italy
Email: torsello@dsi.unive.it rodola@dsi.unive.it albarelli@unive.it

Abstract—Surface registration is a fundamental step in the reconstruction of three-dimensional objects. This is typically a two-step process where an initial coarse motion estimation is followed by a refinement step that almost invariably is some variant of Iterative Closest Point (ICP), which iteratively minimizes a distance function measured between pairs of selected neighboring points. The selection of relevant points on one surface to match against points on the other surface is an important issue in any efficient implementation of ICP, with strong implications both on the convergence speed and on the quality of the final alignment. This is due to the fact that typically on a surface there are a lot of low-curvature points that scarcely constrain the rigid transformation and an order of magnitude less descriptive points that are more relevant for finding the correct alignment. This results in a tendency of surfaces to "overfit" noise on low-curvature areas sliding away from the correct alignment. In this paper we propose a novel relevant-point sampling approach for ICP based on the idea that points in an area of great change constrain the transformation more and thus should be sampled with higher frequency. Experimental evaluations confront the alignment accuracy obtained with the proposed approach with those obtained with the commonly adopted uniform subsampling and normal-space sampling strategies.

Keywords-Point Sampling; Surface Registration;

I. INTRODUCTION

Surface registration is a fundamental step in the reconstruction of three-dimensional objects using range scanners since, due to occlusions and the limited field of view of the scanners, more range images are necessary to fully cover the object. Registration is typically a two-step process where an initial coarse motion estimation is followed by a refinement step on pairs of range images. Pairwise refinements are almost invariably performed using a variant of the Iterative Closest Point (ICP) algorithm [1], [2] that iteratively minimizes a distance function measured between pairs of selected neighboring points. The performance of ICP depends on several parameters such as the choice of the distance metric, the selection of points on one surface and the selection of mating points in the other. These parameters affect the amount of local minima in the error landscape, the speed of convergence and in general the precision of the resulting alignment in the presence of "complex" geometry. See [3] for a review of some of the variants of ICP.

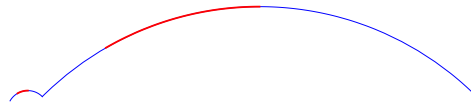In this paper we will concentrate on the sampling process

Figure 1. The two regions compete for the samples, resulting in a larger number of samples on the large low curvature area than on the smaller part. This results in a bias in the distance measure.

used to select points on one surface, with a view of correctly aligning "hard" surfaces that offer very few points that constrain the motion and large areas where the surface can slide. For this reason in all the experiments we will use the point-to-plane distance and the closest point will be used as mate, since these approaches, while not offering the fastest convergence, have been shown to be the most robust combination for "difficult" geometries .

The selection of relevant points on one surface to match against points on the other surface is an important issue in any efficient implementation of ICP with strong implications both on the convergence speed and on the quality of the final alignment. This is due to the fact that typically on a surface there are a lot of low-curvature points that scarcely constrain the rigid transformation and an order of magnitude less descriptive points that are more relevant for finding the correct alignment. See Figure 1 for an illustration of the problem. There we can see a section of a surface composed of two parts with circular profile which in isolation allow the surfaces to slide. The transformation is fully constrained only if enough points are selected from both parts. However, since there is a large size difference between the two parts, uniform sampling will take proportionally more samples from the larger area than from the smaller one. The difference in sampling biases the error term towards fitting the larger region better than the smaller one, but since the region does not constrain the transformation fully, any sliding that would better fit local noise would be preferred to the correct alignment, resulting in a tendency to"overfit" noise on low-curvature areas. In order to better constrain the set of transformations Rusinkiewicz and Levoy [3] propose a normal space sampling approach that attempts to sample uniformly on the sphere of normal directions rather than on the surface. This, however, only partially solves the problem;

to show why we refer again to Figure 1. There the thicker arcs on the surface section refer to points that fall in the same normal bin. Since points in the same bin are sampled uniformly, the points on the smaller arc must compete with the points on the larger arc resulting in the same, albeit a bit reduced, tendency of overfitting noise on the larger region that plagues uniform sampling. In effect, we would like to sample points from the two arcs with the same probability. Further, normal space sampling fully constrains only translational error, but in general cannot limit rotational sliding, and the binning interacts poorly with noise in the normal estimation.

An interesting approach to better constrain the transformation is to select points that best equalize the error covariance matrix. To this effect Guehring [4] proposes to weigh the samples based on their contribution to the covariance matrix, but since the analysis is performed after the sampling, the approach cannot constrain the transformation if too few samples were chosen in a relevant region. On the other extreme, Gelfand *et al.* [5] propose an approach that selects the points that constrain the transformation the most. However, the approach is deterministic and has a tendency of sampling very regularly on isolated regions. Since in general the range surfaces overlap only partially, the optimal constraining property only holds if the sampling is performed only on the overlapping part, which means that the analysis and the sampling cannot be performed only once for each surface, but has to be redone for each pair of surfaces. Further, high levels of noise can make some areas artificially strongly constraining, and for that reason the approach needs the surfaces to be smoothed before the points can be selected.

We propose a different approach to ensure that the rigid transformation is fully constrained, based on the relevance, or local distinctiveness, of points. The idea of point distinctiveness has been extensively used in image processing to develop interest point detectors such as the Harris Operator [6] and Difference of Gaussians [7]. While these approaches work well with 2D intensity images, they cannot be easily extended to handle 3D surfaces since no intensity information is directly available. Several efforts have been made to use other local measures, such as curvature or normals to find relevant points on a surface, but mostly with the end of finding repeatable associations for coarse registration or 3D object recognition. One of the first descriptors to capture the structural neighborhood of a surface point was described by Chua and Jarvis, who with their Point Signatures [8] suggest both a rotation and translation invariant descriptor and a matching technique. Later, Johnson and Hebert introduced Spin Images [9], a rich characterization obtained by a binning of the radial and planar distances of the surface samples respectively from the feature point and from the plane fitting its neighborhood. Given their ability to perform well with both surface registration and
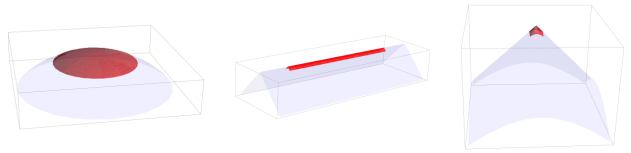

Figure 2. The region A_p grows in all directions in a "flat" part of the surface, in only one direction along edges and boundaries and does not grow much at all on vertices.

object recognition, Spin Images have become one of the most used 3D descriptors. More recently, Pottmann *et al.* proposed the use of Integral Invariants [10], stable multi-scale geometric measures related to the curvature of the surface and the properties of its intersection with spheres centered on the feature point. Finally, Zaharescu *et al.* [11] presented a comprehensive approach for interest point detection (MeshDOG) and description (MeshHOG), based on the value of any scalar function defined over the surface (i.e., curvature or texture, if available). MeshDOG localizes feature points by searching for scale-space extrema over progressive Gaussian convolutions of the scalar function and thus by applying proper thresholding and corner detection. MeshHOG calculates a histogram descriptor by binning gradient vectors with respect to a rotational invariant local coordinate system.

In this paper we propose a local distinctiveness measure that is associated with the average local radius of curvature, and a sampling strategy that samples points according with their distinctiveness. The distinctiveness is computed through an integral measure, and thus is robust with respect to noise.

II. RELEVANCE-BASED SAMPLING

The relevance of a point p is related to how similar points around p are to it. The larger the number of similar points, the less distinctive, and thus the less relevant, p is. For this reason we formalize the idea of distinctiveness of point p in terms of the area of a surface patch around p where points are similar. More specifically, let p be a point of the surface S, we associate to it a connected region A_p such that

$$A_p = \{q \in S | N_p^T N_q > T \text{ and } p \sim q\} \quad (1)$$

where N_p and N_q are the normals of the surface S at points p and q, while $p \sim q$ means that there is a path in A_p connecting p to q, and the dot threshold T is a parameter of the approach. For small values of T the area of A_p is related to the average absolute radius of curvature

$$||A_p|| \approx \bar{r} = \frac{|r_1| + |r_2|}{2} = \frac{|1/k_1| + |1/k_2|}{2}, \quad (2)$$

where $||A_p||$ denotes the area of region A_p, k_1 and k_2 are the principal curvatures of S in p and $r_1 = 1/k_1$ and

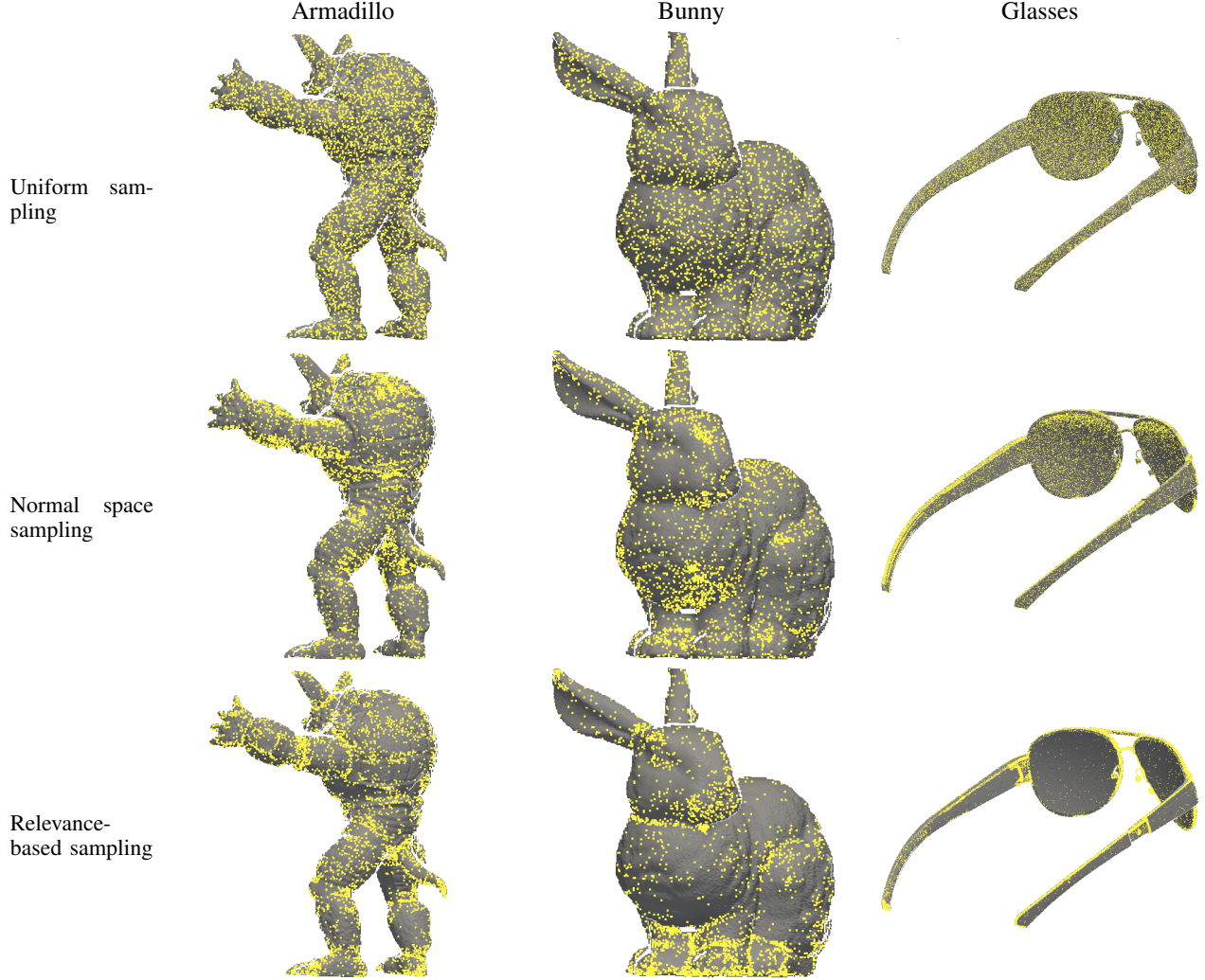|	Armadillo	Bunny	Glasses
Uniform sampling			
Normal space sampling			
Relevance-based sampling			

Figure 3.   Examples of the different sampling approaches.

$r_2 = 1/k_2$ are the radii linked with the principal curvatures. Points within A_p have all the orientations similar to that of p and if the surface orientation varies quickly in one direction the growth of the region in that direction will be limited, thus the size of A_p is linked with the distinctiveness of p. The area will be inversely proportional to the curvature, along edges will extend only in one dimension attaining a size one order of magnitude smaller, and will be almost point-like on vertices, where the transformation is locally completely constrained with the exception of rotations along the point normal (see Figure 2). Hence, the area is inversely proportional to how much the surface is constraining the transformation locally.

With the patches A_p to hand, we can assign to each point p the measure of distinctiveness

$$f(p) = ||A_p||^{-k} \qquad (3)$$

where k is an equalization parameter, changing the relative weight of "common" and "distinctive" point. In particular, the larger the value of k, the more the distinctiveness of points forming a small patch A_p is emphasized.

Moreover, since the region A_p is defined in terms of an angular threshold, $||A_p||$ is invariant with respect to resampling, up to the precision imposed by the new sampling resolution. Further, any scale change varies the areas proportionally, so the ratio between patch areas is scale-invariant.

Finally, the area of A_p is an integral measure, thus being less sensitive to noise, and varies continuously along the surface, with T being a smoothing factor.

When the surface is discretized into points and edges, A_p can be easily computed with a region growing approach starting from each point p. If the regions are big, one could use the continuity and locality of A_p to update the region from neighboring points, but in practice, we add a size threshold D limiting the growth of A_p to points whose

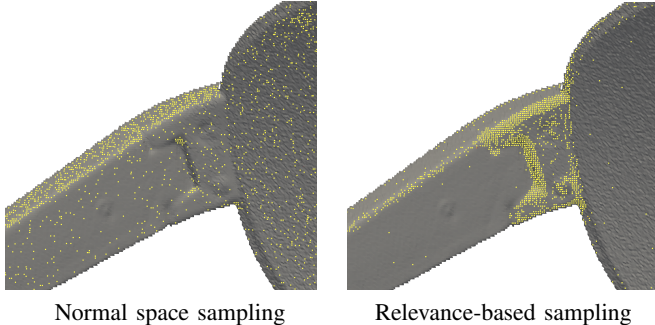Normal space sampling   Relevance-based sampling

Figure 4. Closeup of the samples. Relevance-based sampling concentrates samples along the surfaces' fine structures.

distance from p is less than D. This way we limit the complexity of the region growing process to $O(D^2)$ for each point and we avoid the uncontrolled expansion of A_p on flat surfaces.

Once the areas A_p have been computed, we can assign to each point p the measure of distinctiveness

$$\hat{f}(p) = |A_p|^{-k} \qquad (4)$$

where $|A_p|$ is the number of points in A_p. This approximation works under the assumption that the edge length is uniform through the discretization of surface S.

Once we have computed the distinctiveness of all points in the surface S, we can proceed to sample points from it with a density proportional to $\hat{f}$. To do this we select any order $p_1, \ldots, p_n$ of the points in S and compute the cumulative distribution

$$\hat{F}(p_i) = \sum_{j=1}^{i} \hat{f}(j), \qquad (5)$$

then we sample a number x uniformly in $[0, \hat{F}(p_n)]$ and find the smallest index i such that $\hat{F}(p_i) > x$. To perform the search efficiently, we use interpolation search [12], a variant of binary search that instead of splitting the interval $[i, j]$ in half at each iteration, it splits it at point $i + \frac{x - \hat{F}(p_i)}{\hat{F}(p_j) - \hat{F}(p_i)}$. It is a well known result that interpolation search finds an element in a sorted array in $O(\log \log n)$ on average for near-uniformly distributed data, compared to the $O(\log n)$ complexity of binary search. Further, the search is faster the higher the entropy of f is. This results in an expected $O(m \log \log n)$ complexity when sampling m points from a surface containing n points.

III. EXPERIMENTAL EVALUATION

To evaluate the performance of the sampling approach we created several range images with known ground-truth transformations. To this end we took the 3D models of the Bunny, the Armadillo, the Dragon, and the Buddha from the Stanford 3D scanning repository and range scans

extracted from six sets of glasses scanned using a homebrew scanner built in our lab. The glasses were selected because they are a particularly hard real-world object since it is dominated by large perfectly spherical lenses, while the sliding is constrained only by a very thin rim around them. In the experiments we used 18 scans for each model. For the glasses we used directly the range images provided by the scanner, while for the models taken from the Stanford repository the range images were created by projecting the models onto virtual orthographic cameras placed on a ring around them. Once the range images were to hand, additive Gaussian noise was added along the z dimension to simulate measurement error. In order to avoid having perfect point correspondences, the virtual shots, and thus the points in the various range images, were obtained by projecting equally spaced points on the view-plane of the virtual cameras, and the depths were computed by finding the first intersection of the rays with the model.

All the measures of quality of the alignments are based on the ground-truth alignment, and not the usual Root Mean Square Error (RMSE) because the value of the RMSE depends heavily on the sampling strategy and it is completely blind with respect to the noise overfitting problem.

The proposed Relevance-based sampling (RBS) approach was compared against uniform sampling and normal space sampling (NSS).

Figure 3 shows an example of sampling range scans from the three models using the three sampling strategies. Uniform sampling does exactly what we expect, with the problems we have discussed. Normal space sampling is more selective of the points, but it still wastes quite a few samples on large low curvature areas like the back of the armadillo, the back and the chest of the bunny and the lenses of the spectacles. Relevance-based sampling, on the other hand, concentrates the samples along edges and feature discontinuities, which do a better job at locking the alignment. This difference in behavior can be seen clearly on the closeups in Figure 4. Here normal space sampling samples equally the three main faces of the spectacles, while Relevance-based sampling concentrates the samples along the fine structure details that limit the sliding along the surfaces.

Figure 5 Shows examples of the alignments obtained using the three sampling approaches. To better see the difference in alignment we only show a slice of the aligned surfaces cut approximately orthogonally to the surfaces. From the examples we can clearly see that the use of uniform sampling results on the surfaces sliding along large low-curvature areas. This is particularly evident on the surfaces taken from the spectacles, but it is also evident on the armadillo model. The bunny model is relatively simple and results in good alignments with all the methods, even though uniform sampling has a slightly worse performance here as well. Normal space sampling fares much better, but there is

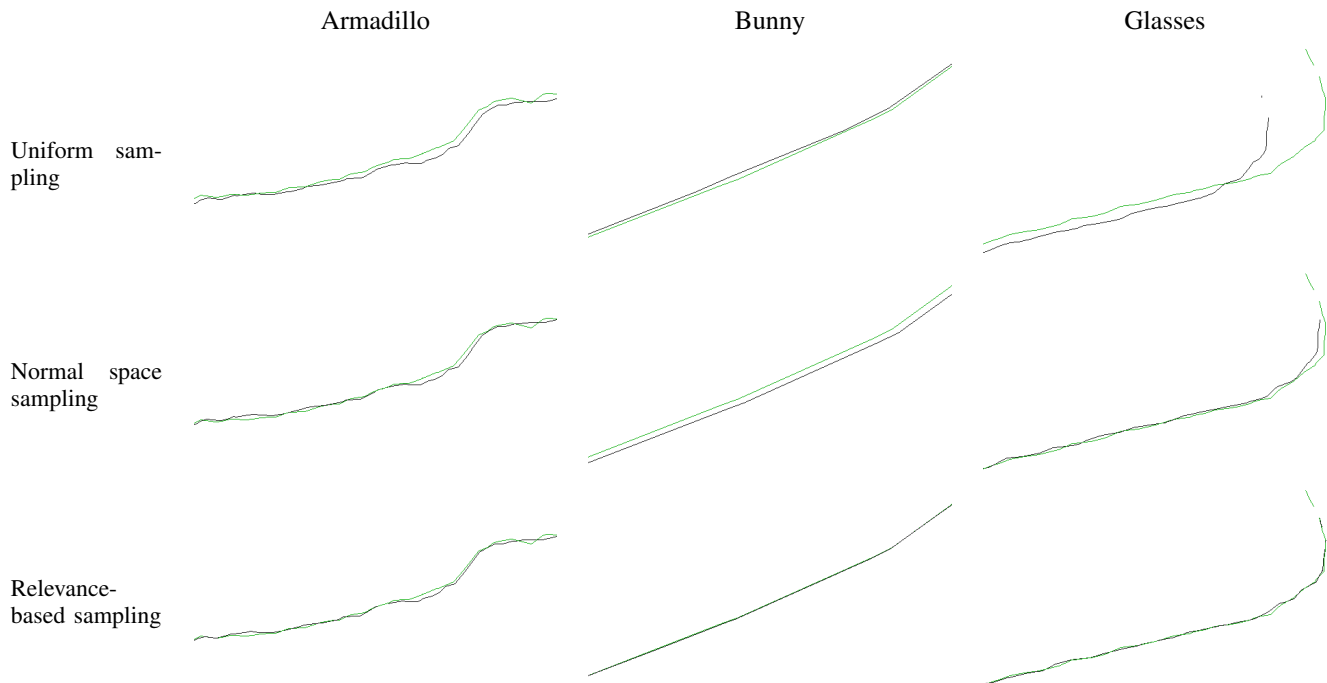| | Armadillo | Bunny | Glasses |
Uniform sampling			
Normal space sampling			
Relevance-based sampling			

Figure 5. Slices of the surfaces aligned with the three sampling strategies.

still some residual misalignment, especially on the spectacles model. Relevance-based sampling, on the other hand, results in an optimal alignment in all cases.

The first set of quantitative experiments was a sensitivity analysis trying to assess the role of the dot threshold T and of the equalization parameter k. Figure 6 shows the angular error (in degrees) and the translation error (in centimeters) as a function of the two parameters. In all the cases the pairs of range images were selected randomly at a distance along the view-circle of at most 3 positions (90 degrees distance in view direction) and were perturbed with additive Gaussian noise along the z dimension with standard deviation equal to 0.4 times the average edge length. We can clearly see that there is an optimal value for the dot threshold at around 10 degrees, and it appears that the optimal value for the equalization parameter k is just slightly below 1. This is due to the fact that noise affects the size of small regions more than larger ones, keeping them smaller, resulting in over-inflated relevance values. A value of k smaller than 1 balances this effect by reducing the relative weight of the smaller regions with respect to larger ones.

Finally, Figure 7 plots the resulting rotation and translation error of the alignments obtained with the three sampling strategies as a function of the level of noise added to the range images. Here the range images were selected using the same strategy adopted for the previous set of experiments, allowing variations in the viewing directions of up to 90 degrees. We can see that the relevance-based sampling consistently outperforms uniform sampling by

a large margin in both rotational and translational error. Normal space sampling, on the other hand, has the same performance as uniform sampling for rotational errors, while it exhibits the same low translational error obtained by the relevance-based sampling for noise levels smaller than 0.4 times the average edge length. This is consistent with the fact that normal space sampling constrains only the translational sliding. Note however, that with larger noise levels the translational error of normal space sampling breaks down to uniform sampling levels. This is probably due to the interaction between bin-size and noise, with high noise spreading neighboring points onto several bins.

IV. CONCLUSIONS

In this paper we have presented a novel sampling strategy for ICP, based on the local distinctiveness of each point. The distinctiveness is gauged through an integral measure that is robust with respect to noise, and the points are then sampled with a density proportional to their distinctiveness. The sampling approach concentrates samples along the surfaces' fine structures, allowing to limit any sliding away from the ideal alignment. Experiments on range images with known ground-truth alignment show that the approach clearly outperforms the most commonly used sampling strategies.

REFERENCES

[1] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," in *Proc. 1991 IEEE Intl. Conf. on Robotics and Automat.*, 1991, pp. 2724–2729 vol.3.
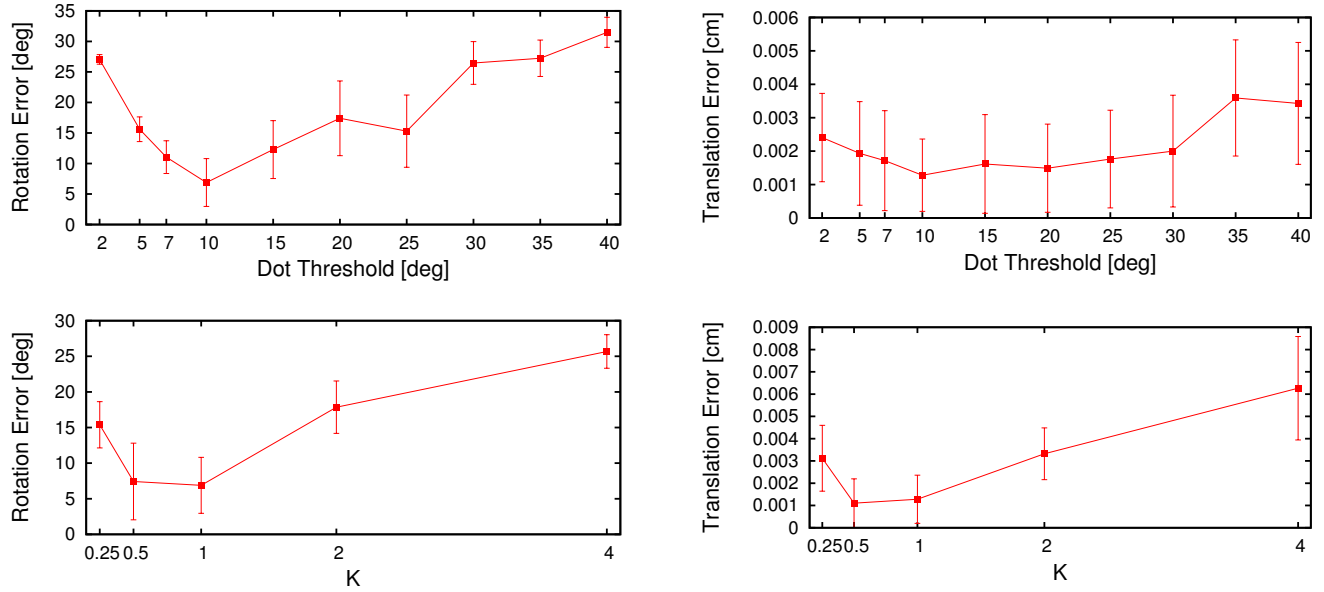
Figure 6. Effects of parameters on rotational and translational error.
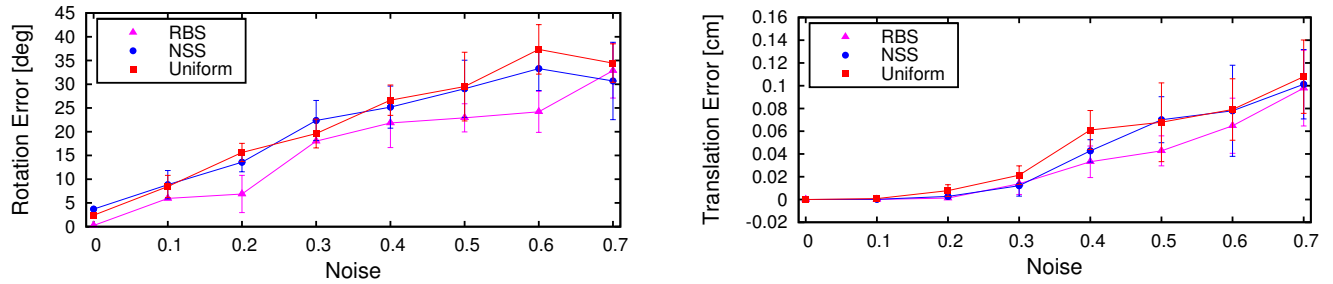

Figure 7. Comparison of rotational and translational error obtained with the three sampling strategies.

[2] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, 1992.

[3] S. Rusinkiewicz and M. Levoy, "Efficient variants of the icp algorithm," in *Proc. of the Third Intl. Conf. on 3D Digital Imaging and Modeling*, 2001, pp. 145–152.

[4] J. Guehring, "Reliable 3d surface acquisition, registration and validation using statistical error models," in *Int. Conf. 3-D Digital Imaging and Modeling*, 2001.

[5] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy, "Geometrically stable sampling for the ICP algorithm," in *Int. Conf. 3-D Digital Imaging and Modeling*, 2003.

[6] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Fourth Alvey Vision Conference*, 1988, pp. 147–151.

[7] D. Marr and E. Hildreth, "Theory of edge detection," *Royal Soc. of London Proc. Series B*, vol. 207, pp. 187–217, 1980.

[8] C. S. Chua and R. Jarvis, "Point signatures: A new representation for 3d object recognition," *Intl. J. of Comput. Vis.*, vol. 25, no. 1, pp. 63–85, October 1997.

[9] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3d scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 5, pp. 433–449, 1999.

[10] H. Pottmann, J. Wallner, Q. Huang, and Y. Yang, "Integral invariants for robust geometry processing," *Comput. Aided Geom. Des.*, vol. 26, pp. 37–60, 2009.

[11] A. Zaharescu, E. Boyer, K. Varanasi, and R. P. Horaud, "Surface feature detection and description with applications to mesh matching," in *Proc. of the IEEE Conf. on Comput. Vis. and Pattern Recognit.*, 2009.

[12] Y. Perl, A. Itai, and H. Avni, "Interpolation search – a log log N search," *CACM*, vol. 21, no. 7, pp. 550–553, 1978.