

MULTI-PEAK RANGE IMAGING FOR ACCURATE 3D RECONSTRUCTION OF SPECULAR OBJECTS

John Park and Avinash C. Kak

Robot Vision Laboratory, Purdue University
1285 EE Building, West Lafayette, IN. 47907-1285
{jpark,kak}@purdue.edu

ABSTRACT

In this paper, we present new range acquisition and 3D reconstruction methods that are capable of modeling specular objects. Multi-peak range imaging, a new concept in range data collection that accounts for the effects of interreflections, is used to acquire range data of an object. False measurements caused by interreflections are then iteratively eliminated by using constraints based on local smoothness and global consistency. Experimental results show that our method can generate accurate and reliable 3D models of objects of complex shapes with specular surfaces.

1. INTRODUCTION

In recent years, there has been considerable interest in constructing accurate three-dimensional models of real-world objects for applications where the focus is primarily on visualization of the objects by humans. This interest is fueled by the technological advances in range sensors and the rapid increase of computing power. Researchers have shown that the state-of-the-art techniques can now construct detailed 3D models of objects ranging from small figurines to large statues. Although they have established the feasibility of constructing accurate 3D models, there still remain several challenging issues. One of these challenging issues is a result of the fact that many objects have surface materials that are not ideal for range sensors. Various surface properties that cause difficulties in range imaging include specular surfaces, highly absorptive surfaces, translucent surfaces, and transparent surfaces. Some researchers have tried to simply do away with such surface-related problems by painting the object or coating the objects with removable powder to ensure that the surfaces reflect the light source diffusely. Obviously, this approach is not desirable and may not even be feasible for real-world objects outside the laboratory.

Of the various surface-related properties we mentioned above, surface specularity is one of the more problematic material properties. Specularity causes interreflections (i.e., mutual reflections) that give rise to ghosts in the measured structured-light data. Depending on the extent of specularity, the presence of these ghosts can make it difficult to localize a data point that corresponds to the object point that was actually illuminated.

The effect of interreflections with a typical optical triangulation range sensor is illustrated in Fig. 1(a). Laser is projected onto a specular object as shown. Some of the laser light scatters off the object in the form of a diffuse reflection, while the rest scatters off to illuminate another part of the object. This secondary illumination can also get partly reflected in a diffuse manner, with the

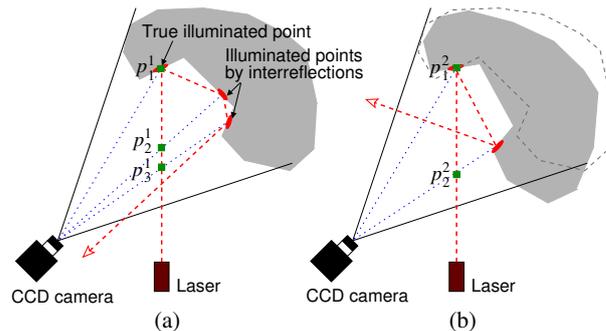


Fig. 1. (a): Due to the interreflection, three illuminated points are observed by the CCD. (b): The same object point as in (a) is being scanned from a different position where the dotted line indicates the position in (a). Only the coordinates of the range measurements from true illuminated points (p_1^1 and p_1^2) are consistent.

rest of it creating a tertiary illumination for yet another part of the object. That illumination can create yet another diffuse reflection visible to the camera. When such a scene is captured by a CCD camera, three illuminated points will be observed. Choosing the point with the highest intensity value – which is the standard approach of peak detection – does not guarantee that this is the true illuminated point. If one of the illuminated points caused by interreflection is chosen, an erroneous coordinate (p_2^1 or p_3^1) will be computed as the object surface point. From now on, we will refer to the illuminated points caused directly by the light source as the true illuminated points, and the illuminated points caused by the interreflections as the false illuminated points.

As an improvement over the conventional methods for structured-light range imaging, in this paper we propose new range acquisition and 3D reconstruction methods that can also handle surface specularities. The contribution of this paper is two-fold: (1) a new range acquisition method that accounts for the effects of interreflections without using any special hardware; and (2) a new 3D reconstruction method that iteratively eliminates measurements produced by interreflections and refines the registration of multiple range images.

2. RELATED WORK

2.1. Range Acquisition of Specular Objects

Determining the shape of specular objects has long been a challenging problem in computer vision. Nayar *et al.* [6] proposed an

This work was supported by the Ford Motor Company.

iterative algorithm that recovers the shape and reflectance properties of surfaces in the presence of interreflections. This algorithm is useful for the shape-from-intensity approach to range acquisition; this approach, however, does not produce dense and accurate range maps compared to the optical triangulation methods. Additionally, the proposed algorithm was tested only on Lambertian surfaces of simple geometry.

Clark *et al.* [4] developed a laser scanning system that uses polarization analysis to disambiguate the true illuminated points from those caused by interreflections. Their system was tested on shiny aluminum objects with concavities, and false illuminated points were successfully discriminated. However, the system requires special equipment such as linear polarizers, and multiple images need to be captured at each position of the laser. In their experiment, three images were acquired at three different angles of the linear polarizer.

Trucco and Fisher [7] proposed a number of consistency tests for acquiring reliable range images of specular objects. Their range sensor consists of two CCD cameras observing a laser stripe from opposite sides. The consistency tests are based on the fact that the range measurements obtained from the two cameras will be consistent only if the measurements correspond to the true illuminated point. Their method was tested on a polished aluminum block with holes. However, their method does not consider the situation where more than one illuminated point is observed, for example, the case illustrated earlier in Fig. 1(a). The consistency tests, therefore, are applied only to the measurement corresponding to a single illuminated point observed per camera scan line. In our experiments, we have noticed that the illuminated points caused by interreflections occur very frequently in the vicinity of the true illuminated points, and thus they are seen together along the same camera scan line. Eliminating all points whenever multiple peaks are observed in the same camera scan line may result in too few range measurements.

2.2. 3D Reconstruction

Much prior work has been done in 3D reconstruction of object models from range images. Due to the space limit, we will therefore refer the reader to a good survey paper [3].

3. RANGE ACQUISITION AND 3D RECONSTRUCTION METHODS

3.1. Overview

Fig. 2 shows the flowchart of our system. First, several scans from different viewpoints are performed so that the entire surface of the object can be captured. For each scan, the multi-peak range imaging as explained in Section 3.2 is carried out to account for the effects of interreflections. The local smoothness test (Section 3.3) is then invoked. This test iteratively eliminates all points that are determined to be locally non-smooth. The determination is made while estimating surface normals.

After performing the steps described above for each scan, the registration (Section 3.4) of all the range data acquired from N different views is carried out first interactively to yield approximate registration, and then via a multiview registration step to fine-tune the registration. The isolated region test (Section 3.5) is then applied to the registered data which eliminates all isolated region points. The global consistency test (Section 3.6) is then followed.

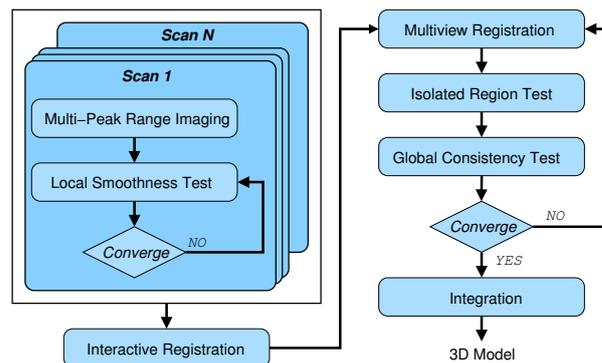


Fig. 2. System Overview

The test eliminates the points that are determined to be highly inconsistent based on two criteria – namely the coordinate and the visibility consistency – using the information of all the range data collected from the N viewpoints. If any points were eliminated during the above two tests, the multiview registration step is carried out again on the new data set followed by another application of the isolated region test and the global consistency test. This iteration is continued until no points are eliminated. Finally, the integration is performed on the resulting output. We adapted the method of Curless and Levoy [5] for the integration.

3.2. Multi-Peak Range Imaging

Multi-peak range imaging is a new concept in range data collection that accounts for the effects of interreflections. The basic idea is to store multiple measurements per rigel (i.e., range image element) if multiple peaks are detected in the same camera scan line. Existing peak detection algorithms usually search for the highest intensity peak in each scan line. If that peak's intensity value is higher than some predetermined threshold, its image coordinates are converted into the world coordinates of the corresponding object point. However, as we already mentioned before, choosing the highest intensity peak does not guarantee that this is the true illuminated point especially when strong specular reflections are present. Multi-peak range imaging, by storing multiple range measurements computed from all the detected peaks, guarantees that one of them is an object point that corresponds to a true illuminated point if it exists. A simple data structure such as a linked-list can be used to store multiple range measurements.

We have experimented with several peak detection methods [8], and found that the center-of-mass method works most reliably when interreflections are present. Our peak detection method finds positions of multiple peaks by computing the center-of-mass (or weighted average) of continuous pixels that satisfy both of the following two conditions: (1) each pixel's intensity value is greater than a preset threshold, and (2) the intensity value is monotonically increasing from each end of the continuous pixels to the highest intensity pixel in the middle.

3.3. Local Smoothness Test

Local smoothness is a legitimate assumption we can make about a true range measurement. That is, an object point must lie on a smooth plane with respect to its neighbors. A common surface normal estimation method computes the best fitting plane for a

small window (e.g., 3×3 or 5×5) centered at the point of interest. Thus we can carry out the local smoothness test while estimating surface normals.

There are two constraints in the local smoothness test. The first constraint states the condition that the number of elements in a fitting window must be greater than τ_{member} ; we shall discuss how to select all the thresholds shortly. Valid data elements in a fitting window are those that have distances to the point of interest less than a threshold $\tau_{neighbor}$. In case multiple valid points exist in a rigel, the one with the minimum distance is chosen.

The second constraint in the local smoothness test requires that the fitting error of the best fitting plane be less than τ_{error} . All range measurements that do not satisfy either of these two constraints are eliminated.

The threshold τ_{member} should be a number between 3 (small number needed to compute best fitting plane) and the maximum number of data elements in a fitting window (e.g., 25 for 5×5 window). In our experiments, τ_{member} is set as 50% of the maximum number of data elements in a fitting window. The parameter $\tau_{neighbor}$ represents the maximum distance allowed between any two neighbors; thus it should be larger than the range sensor resolution. We empirically chose $\tau_{neighbor}$ to be four times the resolution. Finally, τ_{error} should be set to eliminate the points with large fitting errors; so we set τ_{error} to be equal to the resolution. In general, all the thresholds should be set in such a way that only severely non-smooth points are eliminated in each iteration even if this implies that only a small portion of false measurements will be eliminated. In some sense, the main task of the local smoothness test is to reduce the number of false measurements as much as possible without eliminating true measurements so that the global consistency test, which is computationally more expensive, would have less burden.

When the local smoothness test converges, a weight (or confidence value) is assigned to each range measurement. A common approach is to take the dot product of the surface normal and the direction of laser projection [9]. Since the accuracy of the surface normal greatly influences the weight, and since the fitting error of the best fitting plane is available from the previous local smoothness test, we combine the dot product and the fitting error to compute the weight. Let \mathbf{N} be the surface normal, \mathbf{L} the direction of laser projection, and ε the fitting error. A weight of a range measurement is computed as

$$w = (\mathbf{N} \bullet \mathbf{L}) \left(\frac{\tau_{error}}{\tau_{error} + \varepsilon} \right)$$

where τ_{error} is the fitting error threshold described in the local smoothness test.

3.4. Registration

The registration process in our system consists of two steps: the interactive step that yields approximate registration, and the multiview registration step which fine-tunes the registration. The interactive step allows a user to look at a set of range images that need to be registered and to click on the corresponding points between the anchor data set and the moving data set. The approximate registration provided by the human interaction serves as the initial registration for the multiview registration step based on the ICP algorithm. Our multiview registration is similar to the one proposed by [2]. In each i 'th iteration of ICP, the correspondence between two data sets is determined by two criteria: (1) the distance be-

tween two corresponding points must be less than τ_D^i , and (2) the angle between surface normals be less than τ_θ^i . These thresholds are dynamically set in each iteration by

$$\begin{aligned} \tau_D^i &= \mu_D^{i-1} + 3\sigma_D^{i-1} \\ \tau_\theta^i &= \mu_\theta^{i-1} + 3\sigma_\theta^{i-1} \end{aligned} \quad (1)$$

where μ_D^{i-1} and σ_D^{i-1} are the mean and the standard deviation of the distances between corresponding points at the previous iteration $i-1$, and analogously, μ_θ^{i-1} and σ_θ^{i-1} for the angles of surface normals. The two thresholds at the termination of the multiview registration, τ_D^t and τ_θ^t , will be used in the following tests as values indicating approximate registration error.

Adapting from the method proposed by Benjema and Schmitt [1], our multiview registration step randomizes the processing order of range images to reduce any bias toward any particular range image. Also, the weights of the range measurements are taken into account when computing the transformation that minimizes the distances between the corresponding points.

3.5. Isolated Region Test

The main purpose of the isolated region test is to eliminate all points that are far separated from the true object points. The test involves constructing a 3D volumetric grid that contains the entire data with each voxel having a binary value 1 if any point exists in the voxel, and 0 otherwise. A voxel segmentation is then performed to cluster connected voxels. In order to ensure that true measurements belong to a common region, the resolution of volumetric grid is set as $\max(\tau_D^t, \tau_{neighbor})$ where τ_D^t is the distance threshold at the termination of multiview registration, and $\tau_{neighbor}$ is the same threshold introduced in Section 3.3. The isolated region test eliminates all measurements except the ones that belong to the largest sized region. It should be noted that the test assumes the entire surface of the object has been captured by multiple range images, and that these range images are approximately registered.

3.6. Global Consistency Test

The global consistency test is based on two criteria: the coordinate consistency and the visibility consistency.

3.6.1. Coordinate Consistency

The coordinate consistency states that the coordinates of true measurements are always consistent between all registered range images that captured the same object surface. On the other hand, the coordinates of false measurements are inconsistent since the locations where interreflections occur depend on the object surface normal relative to the direction of the light source. For example, consider the scene illustrated in Fig. 1(b) where the same object shown in Fig. 1(a) is now being observed from a different viewpoint. The dotted line represents the previous position of the object in (a). Notice that the same object point is being sampled. As the surface normal relative to the direction of the light source changes, the bouncing light direction also changes, and creates one false illuminated point at a different location from the case shown in (a). Consequently only the coordinate p_1^1 is consistent with the coordinate p_1^2 , assuming that the two range images are registered, and all the other coordinates p_2^1 , p_3^1 , and p_2^2 are inconsistent with one another.

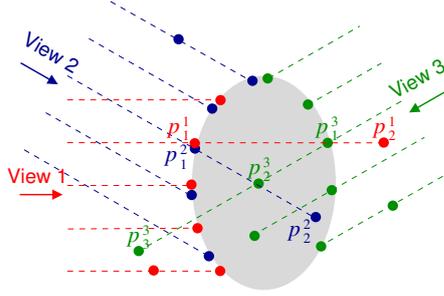


Fig. 3. Visibility consistency

Let \mathcal{C} represent the coordinate consistency, and w the weight. Assuming we have range data from N different views and there are M_i range measurements in i 'th view, the coordinate consistency of j 'th measurement from i 'th range image is defined as

$$\mathcal{C}(p_j^i) = \sum_{k, k \neq i}^N \sum_l^{M_k} \delta_c(p_j^i, p_l^k) w(p_l^k) \quad (2)$$

The test function δ_c is given by

$$\delta_c(p_j^i, p_l^k) = \begin{cases} 1 & \text{if } D(p_j^i, p_l^k) < \tau_D^t \text{ and } \Theta(p_j^i, p_l^k) < \tau_\Theta^t \\ 0 & \text{otherwise} \end{cases}$$

where D is the Euclidean distance function, and Θ the angle between surface normals.

3.6.2. Visibility Consistency

The second criterion of the global consistency test, we call it visibility consistency, is based on the fact that the line space between the sensor and a true measurement is empty, and that the line space beyond a true measurement is invisible to the sensor. Consider the example shown in Fig. 3 where an object has been scanned from three different viewpoints using multi-peak range imaging. The dotted lines represent line-of-sight (e.g., direction of laser projection). Suppose in the first scan, among other detected measurements, p_1^1 and p_2^1 were detected in the same rigel due to interreflection. If p_1^1 is a true measurement (we do not know yet which one is true), p_2^1 should not have been visible to the sensor. On the other hand, if p_2^1 is a true measurement, the space at p_1^1 should have been empty. Thus, p_1^1 and p_2^1 are inconsistent with each other. In fact, all multiple measurements in a same rigel are inconsistent among themselves. Now, let's consider the visibility consistency between the first and the second view. For the reasons mentioned above, there is a visibility inconsistency between p_1^1 and p_2^2 , and between p_2^1 and p_1^2 . Notice that between p_1^1 and p_1^2 , there is a coordinate consistency. For another example, consider between the second and the third view. A visibility inconsistency is observed between p_2^2 and p_2^3 , the visibility consistency is not violated between p_1^2 and p_2^3 .

In order to test the visibility consistency of a range measurement p_j^i , we first need to check for any measurements in the same rigel; there is always a visibility inconsistency between measurements in the same rigel. We then need to check for measurements from other range images that lie on the line-of-sight where p_j^i was sampled. The distance between each of those measurements and the sensor along the line-of-sight need to be computed, and if this

distance is smaller than that between p_j^i and the sensor, there is a visibility inconsistency. Following the same notations as Eq. (2), the visibility consistency \mathcal{V} is defined as

$$\mathcal{V}(p_j^i) = \sum_k^N \sum_l^{M_k} \delta_v(p_j^i, p_l^k) w(p_l^k)$$

and the test function δ_v is given by

$$\delta_v(p_j^i, p_l^k) = \begin{cases} -1 & \text{(1) if } p_j^i \text{ and } p_l^k \text{ are in the same rigel or} \\ & \text{(2) if } p_l^k \text{ is on the line-of-sight of } p_j^i \text{ and} \\ & r^i(p_j^i) > r^i(p_l^k) \\ 0 & \text{otherwise} \end{cases}$$

where r^i is the distance to the sensor of i 'th view. Notice that \mathcal{V} can only have a zero or a negative value since the test function δ_v checks for visibility inconsistency.

3.6.3. Global Consistency Test and its Implementation

The global consistency value, denoted as \mathcal{G} , is simply the sum of the weight, the coordinate consistency and the visibility consistency:

$$\mathcal{G}(p_j^i) = w(p_j^i) + \mathcal{C}(p_j^i) + \mathcal{V}(p_j^i)$$

The global consistency test eliminates all range measurements with low consistency values. Specifically, a range measurement p_j^i is eliminated if $\mathcal{G}(p_j^i) < \tau_G$. The threshold τ_G is set as $\max(\mu_G - c\sigma_G, 0)$ where μ_G and σ_G are the mean and the standard deviation of \mathcal{G} , and c is a small positive constant. We empirically chose $c = 2$ in our experiments. Note that by adding the weight to the global consistency value, and by forcing $\tau_G \leq 0$, we are preserving all measurements that do not have any neighbors from other range images (i.e., $\mathcal{C}(\cdot) = 0$) as long as there is no visibility inconsistency with other measurements of higher weight than itself (i.e., $\mathcal{V}(\cdot) \leq w(\cdot)$). This property lifts a requirement – which may be difficult to satisfy – that every part of the object surface must be sampled at least twice from different viewpoints.

Since the amount of range data in 3D modeling typically reaches to several millions of points, an efficient algorithm for the global consistency test is necessary. Due to the space limit, we will provide only the outline of our implementation: For each range image i , the rest of range images are projected onto the range image coordinate system of i 'th view. For each projected point p_l^k , if its projection coordinate (u', v') is inside the boundary of i 'th range image, the coordinate consistency is tested between p_j^i and the points in the rigel corresponding to (u', v') of i 'th range image. Since projection coordinates will most likely be in floating points, all four rigels surrounding the projection coordinate, each weighted by linear interpolation, are considered. The visibility consistency is tested only if the coordinate consistency test fails (i.e., $\delta_c = 0$). This process is performed for all projected points, and for all range images.

4. EXPERIMENTAL RESULTS

We will now report experimental results on two objects: a polished stainless steel bowl and a glazed ceramic seashell, both shown in Fig. 4. Both objects have surface materials and shapes that are highly likely to generate interreflections.



Fig. 4. Objects used in the experiment

First, in order to show the improvement of our multi-peak range imaging over the conventional range imaging methods, five range images, by rotating 20 degrees each time, were acquired for the bowl. Fig. 5(a) to (d) show one of the five range images produced by our multi-peak range imaging: (a) shows the intensity-coded number of range measurements stored in each rigel, (b) the 3D points viewed from an angle slightly different from the line-of-sight, (c) the 3D points after the local smoothness test convergence, and (d) the 3D points after the global consistency test convergence. Fig. 5(e) to (h) show the same result, but of the range image produced by a conventional method that stores at most one range measurement corresponds to the highest intensity peak on a camera scan line. Notice that the range image produced by multi-peak range imaging, after eliminating false measurements, is much more dense compared to that produced by a conventional method. The regions of missing range data in (h) correspond to the rigels in (a) where multiple measurements were stored.

Fig. 5(i) to (k) show all five range images produced by multi-peak range imaging where different colors represent different range images; (i) shows the 3D points of original data, (j) after the local smoothness test convergence, and (k) the final result after the global consistency test convergence. Fig. 5(l) shows the final result using the conventional method. The figures in the next row, (m) to (p), show the same data as (i) to (l) visualized from a different viewpoint. It is evident that our range imaging method produces much more dense and accurate data compared to the conventional method. Notice in the final result that the false measurements on the left corner were not eliminated. This is mainly due to the fact that we did not sample the entire surface of the bowl, and as a consequence, there was no data that would have provided visibility constraints to these false measurements, for example, the left side of the bowl. This also agrees with our claim that the global consistency preserves all measurements that do not have any neighbors from other range images as long as there is no visibility inconsistency with other measurements.

In the second experiment, twenty seven range images capturing the entire surface of the seashell were acquired using our method. Fig. 5(q) shows the original data, (r) shows the data after the local smoothness test convergence, (s) after the global consistency test convergence, and finally (t) shows the integrated 3D model. The figures in the next row, (u) to (x), again show the same data visualized from a different row viewpoint. The reconstructed 3D model captures the fine details on the real object, even the regions with severe concavities where interreflections are highly likely to occur.

5. CONCLUSION AND FUTURE WORK

In this paper, we presented new range acquisition and 3D reconstruction methods that are capable of modeling specular objects.

Our multi-peak range imaging approach stores all the detected peaks in each camera scan line, thus ensuring that one of the peaks will correspond to a true illuminated point if it exists. A series of tests, then, iteratively eliminates false measurements generated by interreflections. We have shown through our experiments that our approach can generate accurate 3D models of specular objects with complex shapes.

Although the most of the thresholds are set dynamically using the current data statistics, some values in the tests still need to be defined *a priori*; for example, the thresholds in the local smoothness test and the constant c in the global consistency test. The values used in our experiment, in general, should work well, yet a more robust approach for setting these values could improve the performance.

Once we successfully remove all false measurements generated by interreflections and create a 3D model, we can trace back the false measurements that were eliminated, and compute the locations on the object surface where the corresponding interreflections occurred. This information in return can be used to correct or refine the geometry of current 3D model, and the process can iterate until it converges. This can also provide useful information about the surface property.

Another important future work is to extend our methods to other types of optically challenging surfaces such as translucent or highly absorptive surfaces.

6. REFERENCES

- [1] R. Benjemaa and F. Schmitt. Fast global registration of 3D sampled surfaces using a multi-Z-buffer technique. In *Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pages 113–120, 1997.
- [2] R. Bergevin, M. Soucy, H Gagnon, and D. Laurendeau. Towards a general multiview registration technique. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):540–547, 1996.
- [3] F. Bernardini and H. Rushmeier. The 3D model acquisition pipeline. *Computer Graphics Forum*, 21(2):149–172, 2002.
- [4] J. Clark, E. Trucco, and L. B. Wolff. Using light polarization in laser scanning. *Image and Vision Computing*, 15(1):107–117, 1997.
- [5] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In *SIGGRAPH'96*, pages 303–312, 1996.
- [6] S. Nayar, K. Ikeuchi, and T. Kanade. Recovering shape in the presence of interreflections. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 1814–1819, April 1991.
- [7] E. Trucco and R. B. Fisher. Acquisition of consistent range data using local calibration. In *IEEE International Conference on Robotics and Automation*, pages 3410–3415, 1994.
- [8] E. Trucco, R. B. Fisher, A. W. Fitzgibbon, and D. K. Naidu. Calibration, data consistency and model acquisition with laser stripes. *International Journal of Computer Integrated Manufacturing*, 11(4):293–310, 1998.
- [9] G. Turk and M. Levoy. Zippered polygon meshes from range images. In *SIGGRAPH'94*, pages 311–318, 1994.

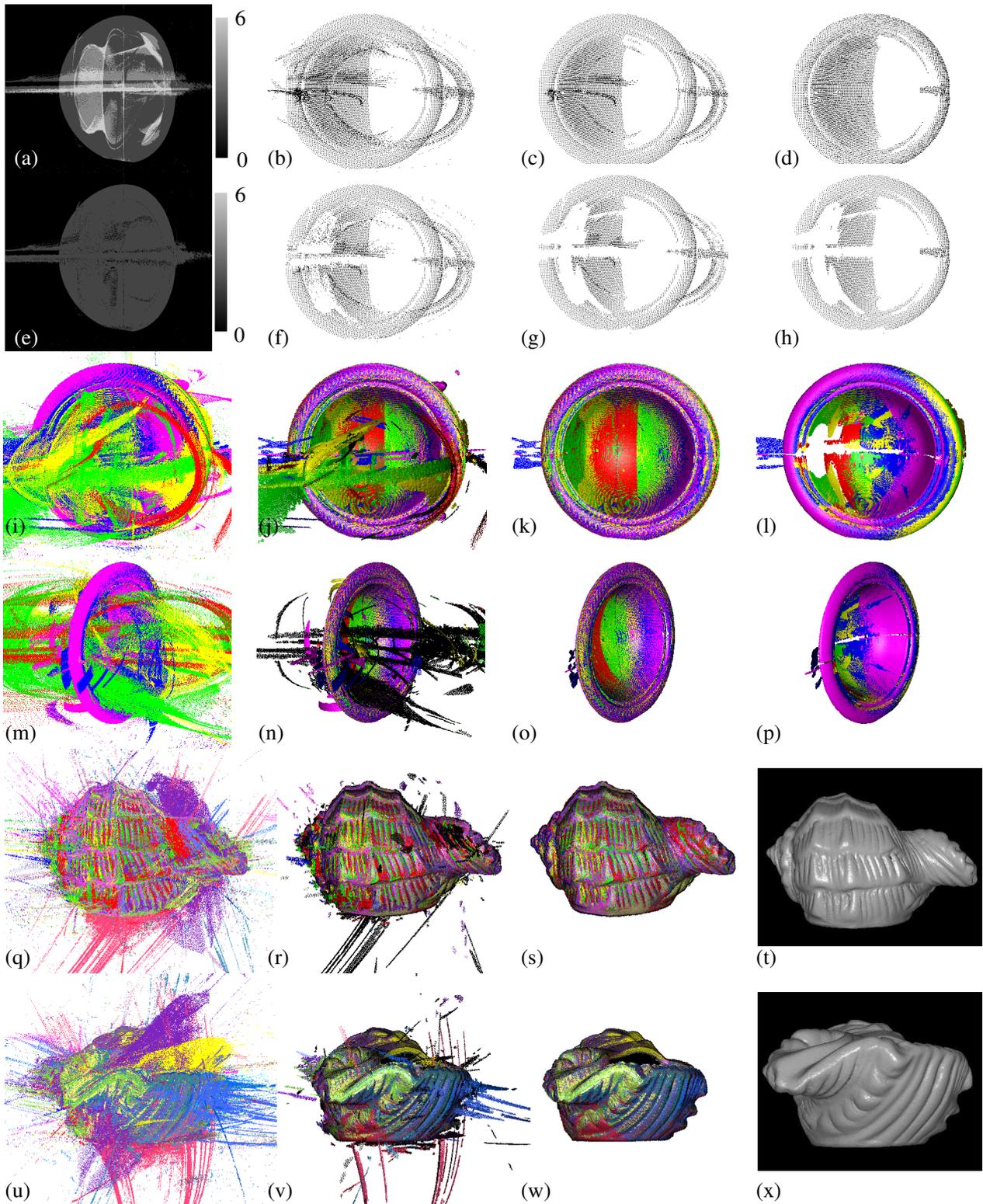


Fig. 5. (a)~(d): One of five range images acquired for the bowl using our multi-peak range imaging. (a) shows the intensity-coded number peaks (ranges), (b) the original 3D points, (c) after the local smoothness test convergence, and (d) after the global consistency test convergence. (e)~(h): The same result as in the first row, but of the range image produced by a conventional method. (i)~(k): All five range images using multi-peak range imaging. (i) the original data, (j) after the local smoothness test convergence, and (k) after the global consistency test convergence. (l): After the global consistency test convergence from the range images of conventional method. (m)~(p): The same data as in the third row visualized from a different viewpoint. (q)~(r): Twenty seven range images acquired for the seashell. (q) the original data, (r) after the local smoothness test convergence, and (s) after the global consistency test convergence. (t): The integrated 3D model. (u)~(x): The same data visualized from a different viewpoint.