

# Dual-Beam Structured-Light Scanning for 3-D Object Modeling

Johnny Park, Guilherme N. DeSouza, and Avinash C. Kak

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

## Abstract

*In this paper, we present our Dual-Beam Structured-Light Scanner (DSLS), a scanning system that generates range maps much richer than those obtained from a conventional structured-light scanning system. Range maps produced by DSLS require fewer registrations for 3-D modeling. We show that the DSLS system more easily satisfies what are often difficult-to-satisfy conditions for determining the 3-D coordinates of an arbitrary object point. Two specific advantages of DSLS over conventional structured-light scanning are: 1) A single scan by the DSLS system is capable of generating range data on more surfaces than possible with the conventional approach using the same number of camera images. And 2) since the data collected by DSLS is more free of self-occlusions, the object needs be examined from a smaller number of viewpoints.*

## 1 Introduction

The existing technology for 3-D modeling and bin-picking has improved significantly in the last few years. The electronics developed to date for structured-light scanners, range scanners, etc., has allowed for acquisition of range data with resolution as fine as 0.05mm, [10]. All this new technology has made it possible to model objects with sizes varying from the large statue of Buddha [12] and David [10], to small industrial parts to be picked from a conveyor-belt [9]. However, despite the growing number of applications found today and the apparently impressive results reported, there still exist a few challenging problems in 3-D modeling. One of these is multiview registration.

Multiview registration is a problem that has caught the attention of many researchers in recent years [2, 6, 16]. The need for multiview registration stems from the intrinsic inability of sensors to perceive the entire object from one single view angle. Frequently, an object contains details that are occluded by other parts of the object. Sometimes, occluded surfaces are *extrapolated* from those that are visible and labeled as “unimaged surfaces” [13], but eventually the

information regarding such surfaces needs to be replaced by actual data and the problem of aligning the two sets of range data has to be faced again.

One alternative to multiview registration is to construct a scanning system in such a way that the transformation matrices corresponding to the different viewpoints are known in advance. However, this condition is difficult to satisfy in practice, especially when the viewpoints are chosen with special criteria such as the minimization of the occluded areas, as used in the notions of the *Best Position* (BP) of an object and its *Next-best-view* [5], and in the *Next-best-pose* [15] for range data collection.

In order to attack the problem of multiview registration, different methods have been proposed. The early methods devised for combining multiview range data came from Chen and Medioni [7], where views are incrementally merged into larger views (metaviews), and from Besl and McKay’s Iterative Closest Points (ICP) algorithm [2], where features from different views are paired based on their distances and then used to compute a rigid 3-D transformation. Many other researchers improved these methods or proposed yet new ones, such as: Bergevin *et al.* [1] who improved [7] by bringing information from previously registered views in the merging of metaviews; Carmichael *et al.* [3] who proposed an algorithm for view registration based on local 3-D signatures; etc. The method in [3], for example, which was derived from the work by Johnson *et al.* [8], improved the computation of local surface signatures – called spin-images – by efficiently dealing with data sets with large variations in resolution and cluttered scenes.

From the description above, one can immediately point out the two major difficulties in multiview registration. The first difficulty is how to efficiently process the large amount of overlapping range data that is acquired by a scanning system for different poses of an object. The range data for the successive poses must overlap since otherwise it would be impossible to carry out multiview registration. The second difficulty is the accumulation of error during view-to-view registration. Although techniques based on curvature patches [11], multi-resolution [14], etc, have been proposed to solve these problems, satisfactory solutions remain to be

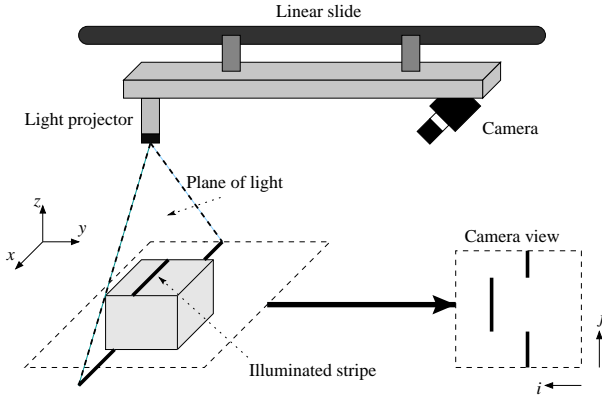


Figure 1: A conventional structured-light scanning system.

found.

In this paper, we propose a scanning system that by generating richer range maps attacks these difficulties at their very root. The proposed system, DSLS, is composed of two light projectors and one camera. The devices are calibrated with respect to each other and a much richer range map can be obtained with a single scan. We believe that this setting significantly reduces the possibility of occlusions, and therefore, the number of necessary view angles and consequent data sets is minimized. Also, by minimizing the number of data sets, the number of views used in the registration process is reduced. This has the effect of reducing the accumulation of the view-to-view error.

## 2 Structured-Light Scanning

structured-light scanners are widely used for various applications in robotics and computer vision. They are especially effective in 3-D object bin picking and 3-D object modeling applications because of the accuracy and reliability of the range data yielded. A typical structured-light scanner system is shown in Figure 1. In this system, a plane of light parallel to the  $xz$ -plane is projected onto the object being scanned. The intersection of the plane of light and the object creates a stripe of illuminated points on the object surface. The plane of light sweeps the object as the linear slide carries the scanning system in the  $y$  direction while a sequence of images is taken by the camera at discrete steps. An index number  $k$  is assigned to each of the images in the order they are taken. Therefore, each  $k$  corresponds to a  $y$  position of the plane of light. For each image  $k$ , a set of image coordinates  $(i, j)$  of the pixels in the illuminated stripe is obtained. The triples  $(i, j, k)$ 's are converted to  $(x, y, z)$  world coordinates using a calibration matrix.

In order to obtain the position of any point on the object surface, the following two conditions must be satisfied:

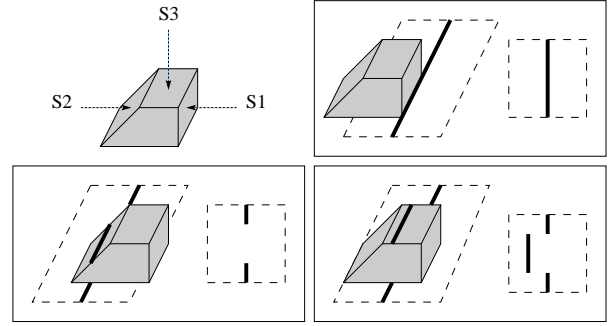


Figure 2: Example of three basic cases of occlusions. **Upper right:** occlusion with respect to light projector, **lower left:** occlusion with respect to camera, **lower right:** no occlusion.

1. The object point must be illuminated by the plane of light.
2. The camera must be able to see the illuminated point.

In other words, the object point cannot be occluded either with respect to the light projector or with respect to the camera. Consider, for example, three basic cases as shown in Figure 2. In the first case, the surface  $S1$  can be seen by the camera, but there is no intersection with the plane of light and the above condition 1 is not satisfied. Thus, no points on the surface  $S1$  are detected. In the second case, the plane of light intersects the surface  $S2$  and it creates a stripe of illuminated points on the surface. However, the stripe cannot be seen by the camera, violating condition 2. Again, points on the surface  $S2$  cannot be detected. Finally, in the third case, the plane of light intersects the surface  $S3$  creating a stripe of illuminated points that can be seen by the camera. In this case, both conditions are satisfied and all points on the surface  $S3$  are detected.

Some researchers try to reduce occlusions with respect to the camera by adding a second camera on the other side of the light projector. The motivation for the second camera is that some of the object surfaces that cannot be seen by the initial camera may be seen by the second camera. However, the second camera generates twice as many images to be processed and it does not reduce occlusions with respect to the light projector (Condition 1).

Our proposed system, Dual-Beam Structured-Light Scanner (DSLS), substantially reduces occlusions with respect to the light projector while using the same number of images. In the next section, we present the DSLS system and its advantages in 3-D object modeling.

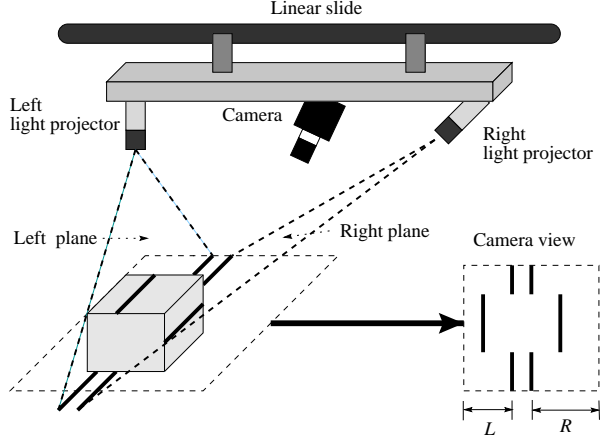


Figure 3: Dual-beam structured-light scanning system

### 3 Dual-Beam Structured-Light Scanner (DSLS)

#### 3.1 System Integration

The DSLS system we have developed is shown in Figure 3. An additional light projector is mounted on the right end of the linear slide. The additional plane of light (right plane) generated by this projector intersects the initial plane of light (left plane) right below the ground plane, thus creating two stripes that are very close to each other on the ground plane. The camera is positioned in the middle of the projectors and it observes the illuminated stripes created by the left and the right planes. Since the two planes do not intersect above the ground, the illuminated stripes generated by the left plane and those by the right plane never overlap in the camera view. In other words, the illuminated stripes by the two planes are closest to each other when they both hit the ground. It must be further noted that the illuminated stripes created by the left plane appear only in region  $L$  (see Figure 3). On the other hand, the illuminated stripes created by the right plane appear only in region  $R$ . No stripes are observed in the region between  $L$  and  $R$  and this region should be minimized in order to maximize the height above the ground plane for which dual data would be available. Although we have worked with the laser beam orientations as shown in Figure 3, one could also design a DSLS like system with other orientations as long as the two beams do not intersect above the ground plane.

#### 3.2 Data Acquisition

The data acquisition process of the DSLS is easily modified from the conventional structured-light scanner. In fact, the only modification comes from realizing that the  $L$  region and the  $R$  region (See Figure 3) provide two different sets

of data. For each image  $k$ , the  $L$  region is searched and a set of triples  $(i, j, k)_L$  is obtained. Similarly, the  $R$  region is searched to obtain the set  $(i, j, k)_R$ . These two sets of triples form two range maps. This is an attractive feature since the processing time for obtaining two different range maps by the DSLS is practically the same as the time for obtaining one range map with a conventional system.

	Average
DSLS	14.03 sec
Conventional System	13.27 sec

The table shows the average processing time for the DSLS and the conventional system over 5 trials. The object scene was the same for all the trials and 200 images were taken. The processing time was recorded from the start of the scan until the range map was generated.

#### 3.3 System Calibration

The calibration of the dual structured-light scanner is done by modifying the method described in [4]. In this method,  $n$  data points are used to solve a  $4 \times 3$  transformation matrix  $T$ . Let the  $i$ th data point in the world coordinate be denoted by  $(x_i, y_i, z_i)$  and the corresponding image coordinate be denoted by  $(r_i, c_i)$ . Also, let variables  $t_{11}$  to  $t_{42}$  be the elements of the matrix  $T$ . Then, we have

$$\begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \\ t_{41} & t_{42} & 1 \end{bmatrix} \begin{bmatrix} r_i \\ c_i \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{x}_i \\ \bar{y}_i \\ \bar{z}_i \\ \rho \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} \bar{x}_i / \rho \\ \bar{y}_i / \rho \\ \bar{z}_i / \rho \end{bmatrix} \text{ or } \begin{bmatrix} \bar{x}_i - \rho x_i \\ \bar{y}_i - \rho y_i \\ \bar{z}_i - \rho z_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (2)$$

We use the free variable  $\rho$  to account for the non-uniqueness of the homogeneous coordinate expressions. Expanding Eq. (1) and rearranging it using Eq. (2), we have

$$\begin{bmatrix} RC & 0 & 0 & RCX \\ 0 & RC & 0 & RCY \\ 0 & 0 & RC & RCZ \end{bmatrix} \begin{bmatrix} t_{11} \\ t_{12} \\ t_{13} \\ t_{21} \\ t_{22} \\ t_{23} \\ t_{31} \\ t_{32} \\ t_{33} \\ t_{41} \\ t_{42} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ y_1 \\ y_2 \\ \vdots \\ y_n \\ z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \quad (3)$$

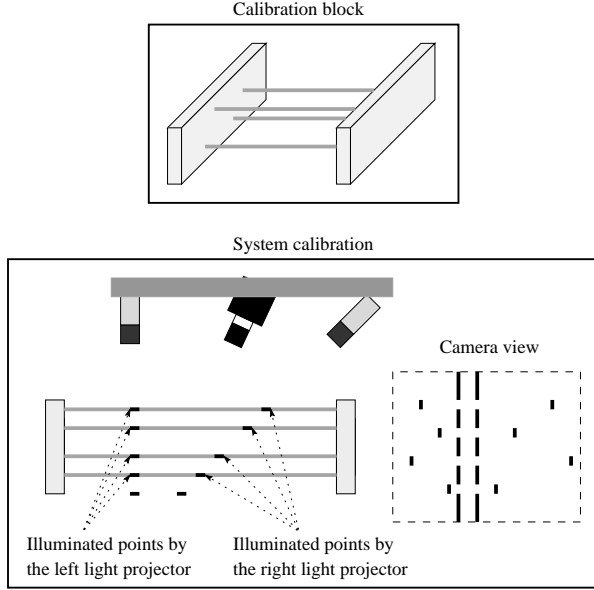


Figure 4: System calibration

where

$$RC = \begin{bmatrix} r_1 & c_1 & 1 \\ r_2 & c_2 & 1 \\ \vdots & \vdots & \vdots \\ r_n & c_n & 1 \end{bmatrix}, RCX = \begin{bmatrix} -r_1 x_1 & -c_1 x_1 \\ -r_2 x_1 & -c_2 x_2 \\ \vdots & \vdots \\ -r_n x_n & -c_n x_n \end{bmatrix}$$

$$RCY = \begin{bmatrix} -r_1 y_1 & -c_1 y_1 \\ -r_2 y_2 & -c_2 y_2 \\ \vdots & \vdots \\ -r_n y_n & -c_n y_n \end{bmatrix}, RCZ = \begin{bmatrix} -r_1 z_1 & -c_1 z_1 \\ -r_2 z_2 & -c_2 z_2 \\ \vdots & \vdots \\ -r_n z_n & -c_n z_n \end{bmatrix}$$

If we rewrite Eq. (3) as  $Ax = b$ , then our problem is to solve for  $x$  in  $Ax = b$ . We can form the normal equations and find the linear least squares solution by solving  $(A^t A)x = A^t b$  where  $A^t$  is the transpose of  $A$ . The resulting solution  $x$  forms the transformation matrix  $T$ . Note that Equation 3 contains  $3n$  equations and 11 unknowns, therefore the minimum number of data points needed to solve this equation is 4.

For the dual-beam structured-light scanner, we need to find two transformation matrices for the left and the right light projectors. It is possible to compute one transformation matrix and solve for the other if we know the exact relative positions of two light projectors. This approach, however is not practical since finding the exact relative positions is very difficult.

The calibration block we have devised is shown in Figure 4. Using this calibration block, we measure the illuminated points on the rods generated by the left light projector and their corresponding points in the camera view. Then, a transformation matrix  $T_{left}$  can be computed using those measured data points. Similarly, we compute a transformation matrix  $T_{right}$  using the data points that were created by the right light projector.

We attached 9 rods in the calibration block such that the camera is able to view all 18 illuminated points. Also, all the rods are assumed to be parallel to the world coordinate  $xy$ -plane.

### 3.4 Advantages of DSLS

In general, with the DSLS system we are able to generate range data on more surfaces than possible with the conventional approach and do so in a single scan using the same number of camera images.

To see the second and more important advantage of DSLS, we need to first describe briefly the shortcomings of the current best practice for combining range images for object modeling: 3-D modeling requires that all of the external surfaces of an object be range mapped. Since the different range maps would in all likelihood be taken from different viewpoints, there is then the problem of registering the range maps into a common coordinate frame, a problem for which no fully automatic procedure has yet been devised. One may compute the registration by selecting the corresponding points manually; however, that can be tedious and difficult since, for complex objects, humans are not always good at visualizing 3D points in 2D projections of the data collected. To avoid this painful process, many researchers are using the calibration between the sensor and the object to compute the registration. A popular approach consists of placing an object on a turntable which rotates in front of a structured-light scanner. To enhance the accuracy of registration achieved in this manner, one can also use the ICP (Iterative Closest Point) algorithm [2]. If the object of the example of Figure 2 was placed on a turntable, the surface S2 would be detected by rotating the turntable by  $180^\circ$  but the surface S1 would not be since the turntable would be rotating perpendicular to the plane of light. The only way to detect the surface S1 using a single structured-light scanner would be to change the object's orientation with respect to the turntable in such a way that the surface S1 would intersect with the plane of light and the illuminated stripe would be seen by the camera. But changing the pose of the object would alter the transformation from the object to the turntable, and, if this new transformation matrix is not available, would require manual registration. The DSLS will reduce the need of changing the object's orientation with respect to the turntable.

## 4 Experimental Results

We first want to show that a single scan by the two beams of DSLS can produce a range map for more surfaces of an object than possible with just one beam. Consider, for example, the object shown in Figure 5(a). A single scan of the

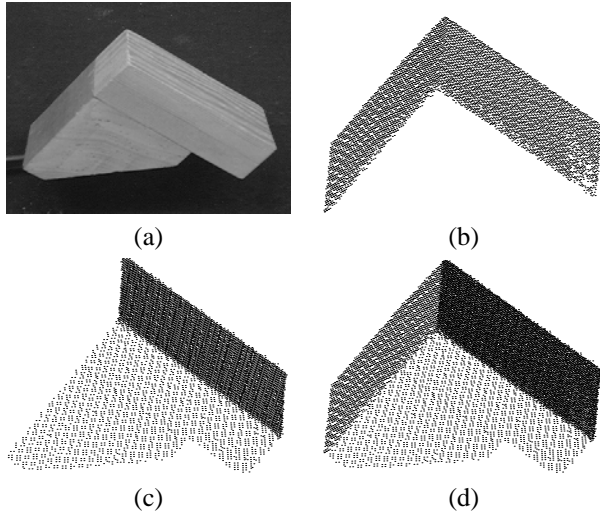


Figure 5: (a): photograph of object, (b): cloud of points detected by left light projector, (c): cloud of points detected by right light projector, (d): superposition of (b) and (c).

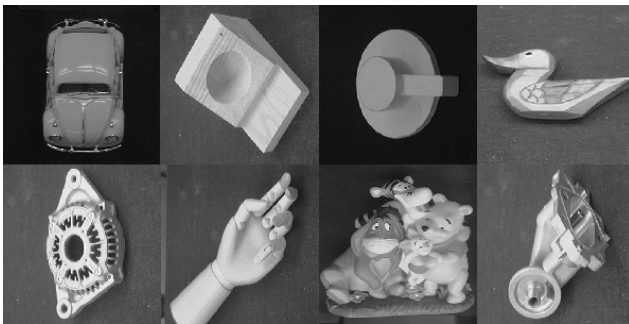


Figure 6: Eight different objects to be tested. **Top left to top right:** Obj.1 - Obj.4, **bottom left to bottom right:** Obj.5 - Obj.8

object with just the left light projector produces the range map shown in Figure 5(b) and a single scan with only the right light projector the map shown in Figure 5(c). However, a single scan with the two light beams together produces the range map of Figure 5(d). The fact that a single DSLS scan is able to capture more surfaces in some of the poses of the object means that one would need fewer scans for modeling the entire object and, consequently, one would have to register fewer range maps.

We now illustrate how DSLS improves upon traditional structural-light scanning with regard to the single-scan efficiency of data collection. This is done by using eight different objects of different shapes, colors, and surface textures (Figure 6). Each of these objects is scanned 5 times and the pose of the object changed randomly for each scan. Two range maps are collected for each pose: from the left light projector and the right light projector. The two range maps are analyzed to find the points that were detected only by

the left light projector, the points detected only by the right light projector, and the points detected by both light projectors. Figure 7 shows the result of this experiment. What the bar graphs depict is explained by the legends at the bottom of the figure. In each pair of bars, the left bar shows the number of object points detected by the left light projector, the right bar the number of points by the right light projector. The gray bottom portion in each bar shows the number points that are detected by both light projectors, and the black top portions the number of points unique to each light projector. The figure shows that the right light projector, which projects a beam at a slant angle with respect to the direction of the scan, consistently detects more object points than the left light projector. This is not surprising since the vertically projected beam by the left light projector will fail to see any vertical surfaces on the objects. In Figure 8, the DSLS range maps of the eight objects are shown.

To show the complementary roles played by the left and the right light projectors, we show in the left column of Figure 9 the range maps for the free-form object (labeled Obj.7 in Figure 6). The top entry in the column is the range map produced by the left light projector, the middle entry the range map produced by the right light projector, and the bottom entry the composite range map by DSLS. As the reader can see, the occluded parts of the left-projector range map are covered by data in the right-projector range map. The fact that the opposite of this statement is also true is made evident by examining the same range maps but from a different perspective, as shown in the right column of the Figure 9.

## 5 Conclusions

In this paper, we described the Dual-Beam Structured-Light Scanning System. Both quantitative and qualitative results were presented to illustrate the advantages of using a second light projector. The results showed that the number of registrations required for 3-D modeling can be significantly reduced. This reduction was possible because of the extra range data that is obtained by using both projectors as opposed to one single projector. The DSLS added in average over 40% more points to the range data than the conventional scanning system.

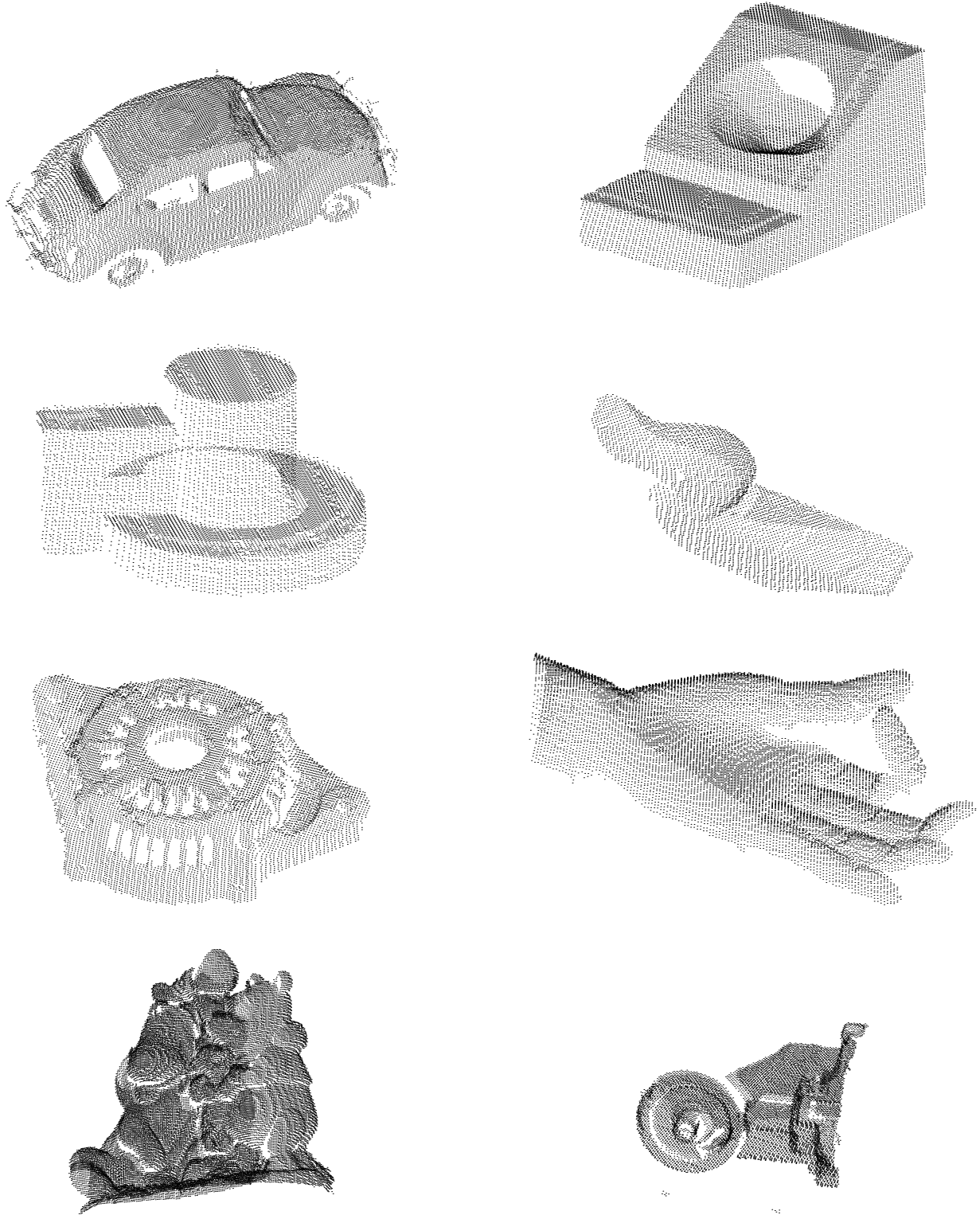


Figure 8: Single scanned DSLS range maps for objects shown in Figure 6

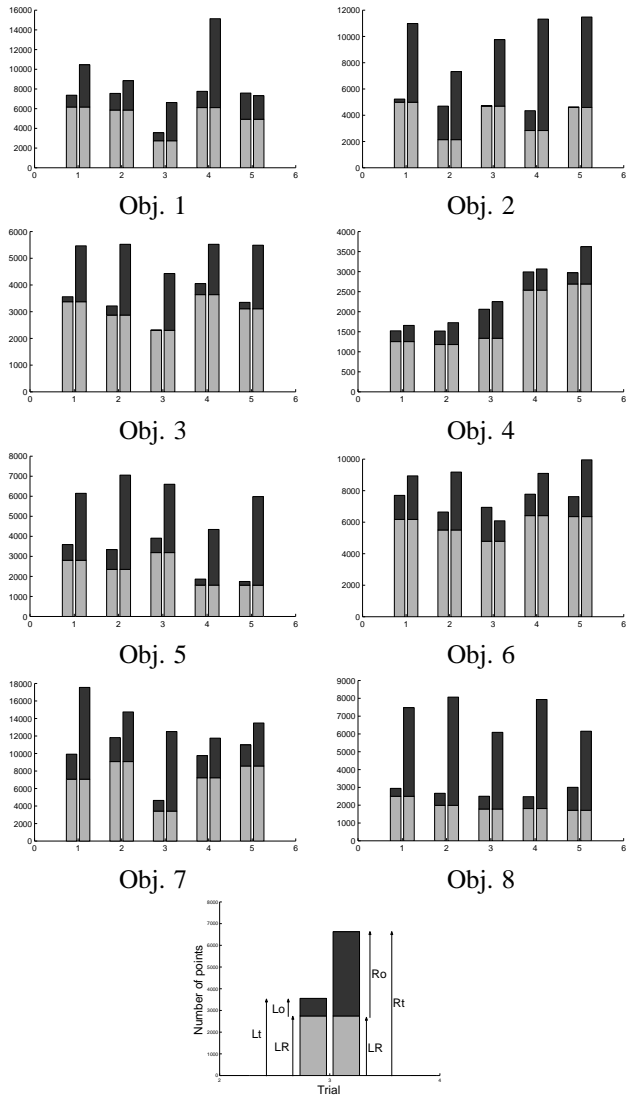


Figure 7: Number of points detected by DSLS. **Lt**: total number of points detected by left light projector, **Lo**: number of points detected only by left light projector, **Rt**: total number of points detected by right light projector, **Ro**: number of points detected only by right light projector, **LR**: number of points detected by both light projectors.

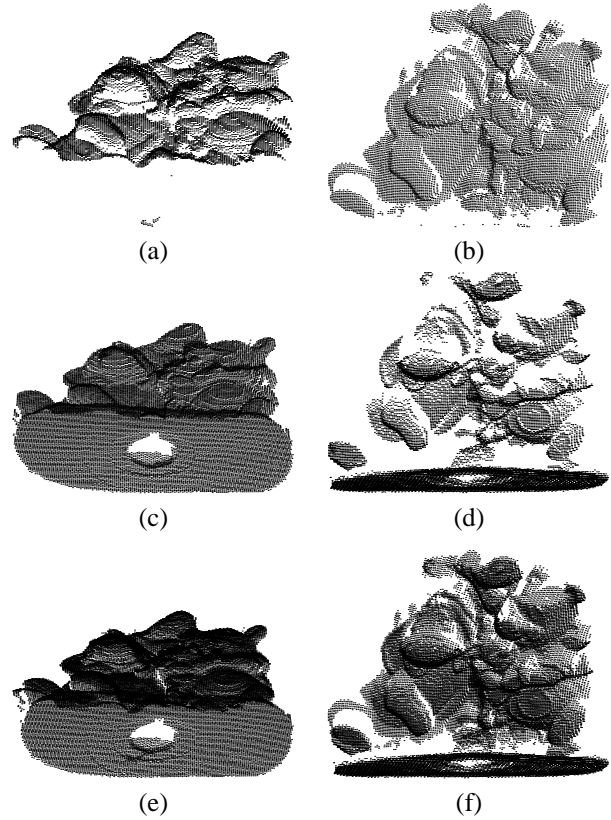


Figure 9: Object points detected by DSLS in a single scan. The left column shows in (a) the cloud of points yielded by the left light projector, in (c) the cloud of points yielded by the right light projector, and in (e), the composite DSLS range map. The right column shows exactly the same data but from a different perspective.

## References

- [1] R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau, "Towards a General Multi-View Registration Technique", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 5, May 1996, pp. 540-47.
- [2] P.J. Besl, and N.D. McKay, "A Method for Registration of 3-D Shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, February 1992, pp. 239-56.
- [3] O. Carmichael, D. Huber, and M. Hebert, "Large Data Sets and Confusing Scenes in 3-D Surface Matching and Recognition," *Proceedings of the Second International Conference on 3-D Imaging and Modeling*, Ottawa, Canada, October 1999, pp.358-67.
- [4] C.H. Chen, and A.C. Kak, "Modeling and Calibration of a Structured Light Scanner for 3-D Robot Vision", *Proceedings of the IEEE International Conference on Robotics and Automation*, Raleigh NC, March 1987, pp. 807-15.
- [5] B.T. Chen, W.S. Lou, C.C. Chen, and H.C. Lin, "A 3D Scanning System Based on Low-Occlusion Approach", *Proceedings of the Second International Conference on 3-D Imaging and Modeling*, Ottawa, Canada, October 1999, pp. 506-15.
- [6] Y. Chen, and G. Medioni, "Object Modeling by Registration of Multiple Range Views", *Proceedings of the IEEE International Conference on Robotics and Automation*, Sacramento, California, April 1991, pp. 2724-9.
- [7] Y. Chen, and G. Medioni, "Object Modeling by Registration of Multiple Range Views," *Image and Vision Computing*, vol. 10, no. 3, April 1992, pp. 145-55, UK.
- [8] A. Johnson, and M. Hebert, "Surface Matching for Object-Recognition in Complex Three-dimensional Scenes," *Image & Vision Computing*, vol. 16, no. 9-10, July. 1998, pp. 635-51.
- [9] A.C. Kak, J.L. Edwards, "Experimental State of the Art in 3D Object Recognition and Localization Using Range Data", *Proceedings of the Workshop on Vision and Robots*, Pittsburgh, PA, 1995
- [10] M. Levoy, K. Pulli, B. Curless, Z. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk, "The Digital Michelangelo Project: 3D Scanning of Large Statues", *Proceedings of SIGGRAPH*, 2000, pp. 131-44.
- [11] V. Nguyen, V. Nzomigni, and C. Stewart, "Fast and robust registration of 3-D surfaces using low curvature patches," *Proceedings of the Second International Conference on 3-D Imaging and Modeling*, Ottawa, Canada, October 1999, pp.
- [12] K. Nishino, Y. Sato, and K. Ikeuchi, "Appearance Compression and Synthesis Based on 3D Model for Mixed Reality", *Proceedings of the International Conference on Computer Vision*, Corfu, Greece, September 1999, pp. 38-45.
- [13] M. Reed, and P. Allen, "3-D Modeling from Range Imagery: An Incremental Method with a Planning Component", *Image and Vision Computing*, February 1999, 17(1): pp. 99-111.
- [14] H. Zha, Y. Makimoto, and T. Hasegawa, "Dynamic Gaze-Controlled Levels of Detail of Polygonal Objects in 3-D Environment Modeling," *Proceedings of the Second International Conference on 3-D Imaging and Modeling*, Ottawa, Canada, October 1999, pp. 321-30.
- [15] H. Zha, K. Morooka, T. Hasegawa, and T. Nagata, "Active modeling of 3-D objects: planning on the next best pose (NBP) for acquiring range images," *Conference on Recent Advances in 3-D Digital Imaging and Modeling*, Ottawa, Canada, May, 1997, pp. 68-75.
- [16] Z. Zhang, "Iterative Point Matching for Registration of Free-Form Curves and Surfaces," *International Journal of Computer Vision*, vol. 13, no. 2, 1994, pp. 119-52.