# From Range Scans to 3D Models

**Brian Curless**
**University of Washington**

Each year, we see a growing number of 3D range scanning products on the SIGGRAPH exhibition floor. You may find yourself asking "how do these technologies work?" and "how can I make use of the shape data they produce?" In this article, I will describe a few of the more common range scanning technologies, then I will step through a pipeline that takes the range data into a single geometric model, and I will conclude with a discussion of the future of range scanning.

## Range scanners

Computer vision researchers have long studied the problem of determining the shape of a scene from a set of photographs. In a sense, they attempt to take advantage of a wealth of visual cues present in the human visual system: stereo and motion parallax, occlusion, perspective, shading, focus, and so on. These methods generally assume that the sensor simply records light that already exists in the scene; i.e., the sensor is *passive*. In this issue, Debevec [6] discusses a few of these passive techniques. The accuracy of such methods is generally limited by a number of factors that can frequently be resolved by carefully controlling how the scene is illuminated. If we permit ourselves to project special light patterns on our scene, we enter the realm of *active* or *structured light* sensing.

One of the most common forms of active range sensing is *optical triangulation*. The fundamental principle is illustrated in Figure 1a. A focused beam of light illuminates a tiny spot on the surface of an object. For a fairly matte surface, this light is scattered in many directions, and a camera records an image of the spot. We can compute the center pixel of this spot and trace a line of sight through that pixel until it intersects the illumination beam at a point on the surface of the object.

This technique gives us a single range point, but how do we modify the design to scan the surface of an object? Many approaches have been developed to answer this question. One method is to scan the light spot over the surface of the object using mirrors. Another approach is to fan the beam into a plane of laser light, as shown in Figure 1b. This light will cast a stripe onto the surface of the object which is then imaged by a conventional video camera. We can treat each camera scanline separately, find the center of the imaged light, and intersect the line of sight with the laser plane. Thus, each image gives us a range profile (one point per scanline), and by sweeping the light over the surface of the object, we can capture its shape. Figures 1c and 1d show a light stripe cast onto a real object and the reflection observed by the camera.
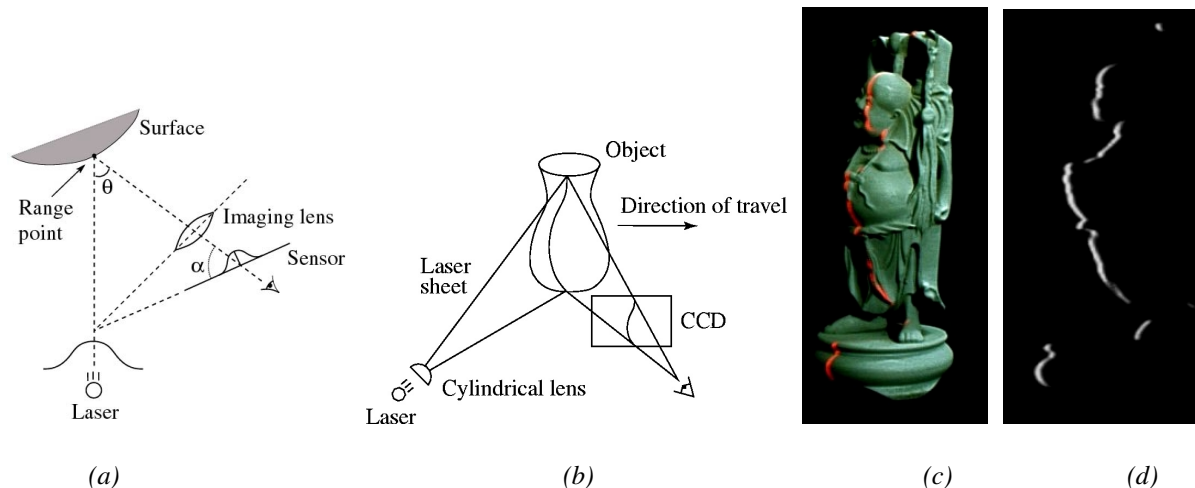
*Figure 1. Optical triangulation. (a) 2D triangulation using a laser beam for illumination. (b) Extension to 3D. (c) Red laser line projected onto small (20cm) statuette. (d) Reflected light seen by CCD camera.*

One of the limitations of the previous two triangulation approaches is the need to capture many frames while sweeping the light over the object. The temptation is to project many points or stripes of light at once to capture as much shape as possible in one shot. One caveat about such approaches is that occlusions may make it difficult to determine which imaged "center" corresponds to which illuminating stripe. With this in mind, researchers have developed multi-stripe systems that work under some assumptions about the smoothness of the surface of an object. A compromise is offered by hierarchical striping methods, i.e., illuminating the surface with very thick stripes, followed by stripes half as thick, and so on. Such methods typically require taking about 10 images, but do not suffer from ambiguities.

Another class of range scanning methods that is becoming increasingly popular is imaging radar. In *time of flight* radar systems, the scanner emits a focused pulse of laser light and waits for it to return to a sensor. The time it takes for the light to return, multiplied by the speed of light in air, tells us how far the pulse traveled. This distance is actually roundtrip, so we have to divide it by two to get the right answer. Another approach is called *amplitude modulation* or *AM* imaging radar. In this case, the laser is operating continuously, but the power of the beam is being modulated sinusoidally over time. After bouncing off of an object and returning to the sensor, the reflected light still has sinusoidal power variation with time, but it is now out of phase with the emitted light. By computing the phase difference between the emitted and reflected power signals, we can again compute roundtrip distance to the surface. Both the time of flight and AM imaging radar systems can scan a scene by using mirrors to sweep the laser beam. More recently, researchers are demonstrating systems that send not a beam or a stripe, but a *plane* of light toward a scene and then use a bank of sensors to estimate the distance for every pixel [13].

Note that many range scanners use laser illumination instead of incandescent or fluoresecent light, because (1) lasers can be focused tightly over very long distances and (2) lasers have an extremely narrow radiation spectrum. The first property allows us to make very tight beams or narrow stripes to improve the resolution of our systems. The second proprerty means that we can achieve relatively high immunity to ambient illumination in the environment. Covered by an optical filter

tuned to pass light only at the laser's wavelength, the sensor effectively sees nothing but the reflected laser light, as demonstrated in Figure 1d.

To be sure, this brief range scanning survey is hardly comprehensive. A plethora of other techniques have been demonstrated to be quite effective as well, including active multi-baseline stereo, moire, shadow striping, and active depth from defocus. The interested reader should turn to the range scanning literature and the Web for more details [2][3][9][12][13].

## Range images and range surfaces

Many of the range scanners described above are designed to return *range images*. Loosely speaking, a range image is like a conventional camera image, except that each pixel stores a depth rather than a color. We can also view a range image as a set of points in 3D (Figure 2a). From a single range image, we can create a *range surface* by connecting nearest neighbors with triangular facets as in Figures 2b-2d. In many places, we will find depth discontinuities that can only be resolved by viewing the scene from a different angle and filling in the data between the points. To avoid making bad assumptions about the shape, we can apply an edge length criterion and omit long skinny triangles that would bridge these discontinuities.



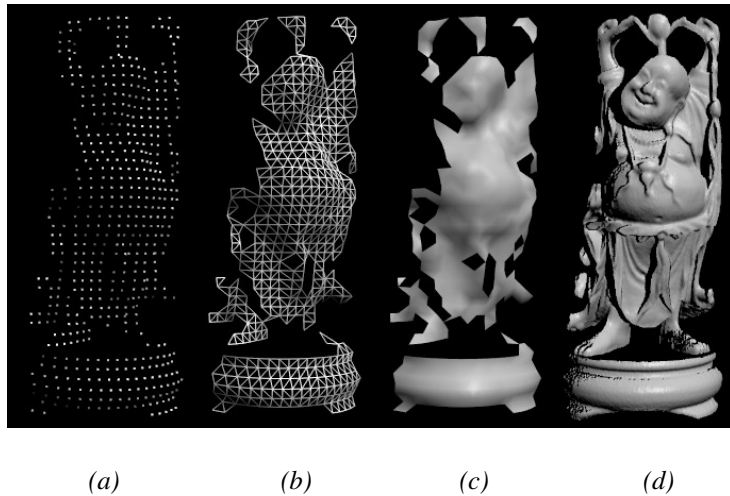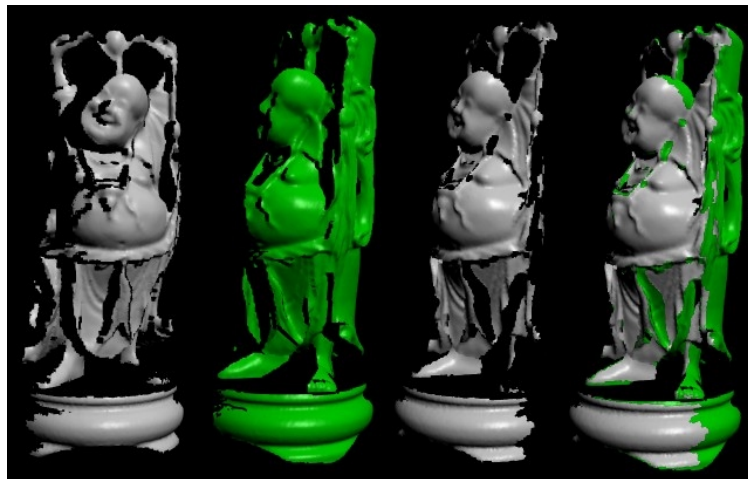*(a)              (b)              (c)              (d)*

*Figure 2. From range image to range surface.  (a) Range image (sub-sampled).
(b) After creating triangles.  (c) Shaded rendering.  (d) High resolution
range surface.*

## Range image registration

Range imaging scanners generally do not capture the shape of an object with one snapshot. To acquire the shape from all sides, indeed to see into every nook and cranny, many scans may be necessary. Each scan captures a portion of the shape of the object, and in order to merge all of the scans into a single shape we must place them in the same coordinate system. The problem of finding each of the rigid transformations to this common coordinate system is called *range registration* or *alignment* [4][11]. If the scanner is under precise computer control or is carefully

tracked, then the problem is solved trivially. However, in many cases, we must compute these transformations using the range data itself.

Consider scanning in a photograph using a traditional flatbed color scanner, and then scanning again after moving the photograph. To align the two digital images of the photograph, we must hold one still while we translate and rotate the other (a similar problem arises when composing images into a single perspective panorama as mentioned in [8] in this issue). When the colors match up, we are finished and can record the translation and rotation. Note that since the translated and rotated image will usually not line up with the pixel grid, we need to compare pixel values of one image with *interpolated* pixel values of the other image. When aligning two range images, we solve a similar problem: find the 3D translation and 3D rotation that bring the points as close together as possible. Much as with the 2D image case, we can compare a point from one range image to *interpolated* points on the other range image; i.e., we use the interpolating range surface when measuring "closeness." Figure 3 shows two ranges images before and after registration.



*(a)*         *(b)*         *(c)*         *(d)*

*Figure 3. Range image registration. (a) First scan. (b) Second scan.*
*(c) First scan after transforming to the coordinate system of the second*
*scan. (d) Both scans overlayed in same coordinate system.*

**Reconstruction**

Once all of the range data is precisely registered into a common coordinate system, we can fuse the data into a single shape, e.g., a dense triangle mesh. This problem is called *surface reconstruction*. Following the analogy of the flatbed scanner, after we successfully align and overlay a set of scanned images, we wish to extract a single image from the result. For the flatbed scanner, we could just average the interpolated values at each pixel location and be done with it. For range data, the problem is more complicated, because the points are not all seen from one direction; thus, we cannot simply average them together along a single line of sight. Indeed, some points may correspond to opposite sides of a thin surface, and we must take care not to average these together.

Researchers have developed numerous solutions to the surface reconstruction problem. Some of the solutions disregard the structure of the range data (e.g., the regular sampling associated with range images) and compute a surface from the "cloud" of range points [1][7]. As a thought experiment, suppose our shape is a closed curve in the plane, and we sample this shape with a range scanner that lives in this plane. We will get a set of points on this curve that correspond to scans from various viewpoints. In general, scanning systems will not capture points that are exactly on the curve, due to noise and systematic errors. Given this set of points, we could play the game of connect the dots by inserting line segments between neighboring points, and we would arrive at an interpolating reconstruction of the curve. Unfortunately, such a reconstruction would probably be very bumpy due to the noise in the samples; further, it may very well have many more line segments than are necessary. A better solution would be to estimate a curve that follows the shape and is very close to the points, but need not interpolate them; i.e., the curve should *approximate* the range data. By freehand, you could "sketch" the curve that is suggested by the points. Numerically, you could compute a small number of line segments that passes near the original data and call this your best guess as to the shape of the curve. These ideas generalize (with some effort!) to 3D, except that we use little triangles instead of line segments to interpolate or approximate the data.

By contrast, a number of solutions to the surface reconstruction problem actually take advantage of the structure of the data. Generally, these algorithms convert the range images to range surfaces (usually triangle meshes) and then merge the surfaces. One approach to merging the surfaces is to stitch or zipper the triangle meshes together [10][11]. Unfortunately, this 3D stitching process can be quite messy and does not always produce high quality meshes.

An alternative approach blends the range surfaces in a sampled volumetric space [5]. The basic idea is to classify each volume element (voxel) as being near the surface, empty, or unseen. Consider a single range scan. The voxels near the captured range surface are regarded as near the reconstructed surface. Along the line of sight from a range point toward the sensor, the voxels must be empty, otherwise we would not have been able to observe the range point. Voxels that are behind the observed range surface or are outside of the field of view of the scanner are left unknown. We can now incrementally build a voxelized version of our object by combining multiple scans in this sampled space. To get a triangle mesh, we can extract a surface from the volume along the observed surface contours as well as at the borders between empty and unseen. One of the advantages of this approach is the fact that the triangle mesh will have no gaps in it, i.e., no missing triangles. In effect, the triangles at the empty/unseen border fill the gaps where our scanner could not reach the object. Figures 4a and 4b illustrate the volume classification and the resulting surface.
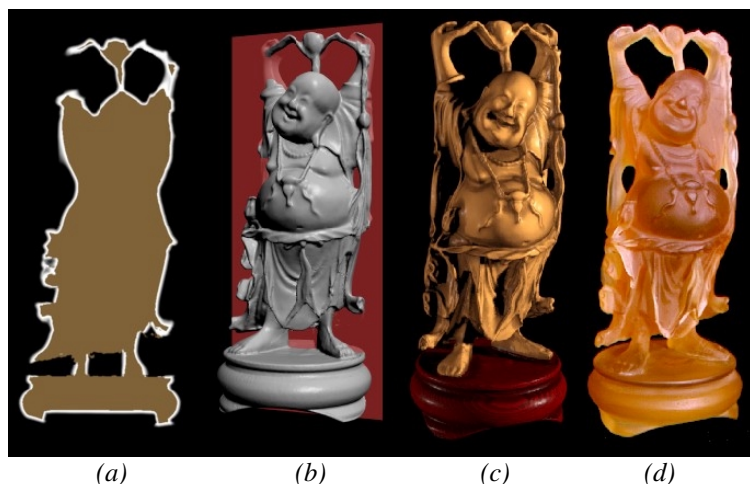
| (a) | (b) | (c) | (d) |

*Figure 4. Surface reconstruction and hardcopy. (a) A slice through the merged volume. Black is empty, brown is unknown, and white are regions near the surface. (b) Reconstructed triangle mesh with red plane indicating the position of the slice in (a). (c) Renderman rendering of the reconstruction. (d) Stereolithography hardcopy.*

## What next?

Once a full geometric model is constructed from the range data, it may be used for a number of purposes. The volumetric method described in the previous section is particularly well-suited for manufacturing high resolution hardcopies using layered manufacturing technologies such as stereolithography, thus yielding a "3D fax" as shown in Figure 4d. Are we far from building a true 3D fax machine: a push-button device the size of a microwave oven that copies and transmits an object for replication at the other end? To some extent, we are close, perhaps already there. For fairly simple objects, one can imagine building such a device -- a box that can move a scanner over the surface of an object or perhaps move the object around so that we can see it from many different angles. However, the ideal 3D fax will also measure the appearance of the object -- the way it interacts with light. The technology for measuring appearance is still a work in progress, but may not be far off (see [6] in this issue). The technology that allows us to manufacture a physical replica with the same surface properties is likely further into the future.

Beyond 3D faxing, range scanning is having a significant impact on how professionals create models for the entertainment industry. Despite advances in geometric modeling software, nothing is quite as intuitive to a sculptor as real clay, the ideal 3D modeling interface. Range scanning allows sculptors to work in their preferred medium but still bring their sculptures into the computer. Post-processing software is now commercially available for taking the resulting scanned models into forms suitable for conventional modeling and animation packages. Will modeling tools and interfaces replace real clay? Time will tell.

Finally, this article has focused largely on the process of geometric model reconstruction, but a number of emerging applications for range scanning do not require models. Real-time scanners in particular offer some exciting alternatives; indeed, RGBZ cameras are on the way [13]. This

realtime Z channel will assist in the compositing process by separating image layers based on their relative depths. Some amount of re-lighting of moving objects is also possible, though shadows may be incorrect, because only some of the scene geometry is observed. Realtime Z will also allow us to study the time evolution of deforming shapes, such as a facial expression. Finally, in combination with image-based rendering techniques (see [8] in this issue), the Z component could permit more freedom of motion for a viewer when playing back an RGBZ video sequence.

In summary, 3D range scanning is having a significant impact on the manufacturing and entertainment industries and may soon be coming to a "television" near you.

**Bibilography**

[1] Amenta, N., Bern, M., and Kamvysselis, M., A new voronoi-based surface reconstruction algorithm. In *Proceedings of SIGGRAPH '98*, p. 415-421, July 1998.

[2] Besl, P. Chapter 1: Active optical range imaging sensors. *Advances in Machine Vision*. pp. 1-63, Springer-Verlag, 1989.

[3] Bouguet, J.-Y. and Perona, P. 3D Photography on your desk. In *Proceedings of ICCV'98*, pp. 43-50, January 1998.

[4] Chen, Y. and Medioni, G., Object modeling by registration of multiple range images. *Image and Vision Computing*, 10(3), pp. 145-155, April, 1992.

[5] Curless, B. and Levoy, M., A volumetric method for building complex models from range images. In *Proceedings of SIGGRAPH '96*, pp. 303-312, August 1996.

[6] Debevec, P., Image-Based Modeling, Rendering, and Lighting. *Computer Graphics*, 33(4), in this issue, November, 1999.

[7] Hoppe, H., DeRose, T.,Duchamp, T., McDonald, J., and Stuetzle, W. Surface reconstruction from unorganized points. In *Proceedings of SIGGRAPH '92*, pp. 71-78, July 1992.

[8] McMillan, L. and Gortler, S., Image Based-Rendering: A New Interface Between Computer Vision and Computer. *Computer Graphics*, 33(4), in this issue, November, 1999.

[9] Nayar, S.K., Watanabe, M., and Noguchi, M. Real-time focus range sensor. In *Proceedings of ICCV'95*, pp. 995-1001.

[10] Soucy, M. and Laurendeau, D., Multi-resolution surface modeling from multiple range views. In *Proceedings of CVPR'92*, pp. 348-353, June 1992.

[11] Turk, G. and Levoy, M., Zippered polygon meshes from range images. In *Proceedings of SIGGRAPH'94*, pp. 311-318, July 1994.

[12] Virtuoso camera, Visual Interface, www.visint.com.

[13] ZCAM, 3DV Systems, www.3dvsystems.com.