# Chapter 9

# Conclusion

In this thesis, we have worked from the most basic issues of a prevalent range scanning technology and progressed to a method that reconstructs complex models from vast amounts of range data. We review these contributions in Sections 9.1 and 9.2. In Section 9.3, we describe several areas for future work.

## 9.1 Improved triangulation

Optical triangulation has been used for decades, yet the traditional methods are fundamentally limited in the amount of accuracy they can deliver. Variations in reflectance and shape induce errors on the order of the width of the illuminant, regardless of sensor resolution. In addition, when using coherent illumination, laser speckle corrupts the imaged light reflections and contribute range error, again independent of sensor resolution. Except for very simple surfaces, the errors due to reflectance and shape variations typically dominate; at a minimum, these errors are objectionable because they are coherent.

Many optical triangulation systems sweep the illuminant across an object, leading to consecutive sensor images that record light reflecting from individual points on the object. We have shown that by analyzing the time evolution of these imaged reflections, distortions induced by shape and reflectance changes can be corrected. In theory, for an ideal orthographic sensor with infinite resolution, these errors can be eliminated completely. In practice, we have demonstrated that we can significantly reduce range distortions with existing hardware that uses a perspective sensor of finite resolution. Although our implementation of the spacetime method does not completely eliminate range artifacts, it has proven to reduce the artifacts in all shape and reflectance experiments we have conducted. Further, increases in sensor resolution and reduction of filtering artifacts will undoubtedly improve the accuracy for spacetime analysis, while the same cannot be said for traditional optical triangulation methods. The influence of laser speckle, however, continues to limit triangulation accuracy.

## 9.2 Volumetrically combining range images

Optical triangulation scanning is one of a variety of rangefinder technologies that acquire sets of dense samplings known as range images. Numerous researchers have attacked the problem of reconstructing a single surface from these range images. In this thesis, we have developed a new algorithm for volumetric integration of range images, leading to a surface reconstruction without holes. The algorithm has a number of desirable properties, including the representation of directional sensor uncertainty, incremental and order independent updating, robustness in the presence of sensor errors, and the ability to fill gaps in the reconstruction by carving space. Our use of a run-length encoded representation of the voxel grid and a shear-warp factorization for the scan-conversion of range images makes the algorithm efficient. This in turn allows us to acquire and integrate a large number of range images. In particular, we demonstrate the ability to integrate up to 70 scans into a high resolution voxel grid to generate million polygon models in a few hours. These models are free of holes, making them suitable for surface fitting, rapid prototyping, and rendering.

## 9.3 Future work

In this section, we describe some directions for improving both optical through spacetime analysis and the volumetric method for surface reconstruction. We conclude with a discussion of some open problems in the acquisition of the shape and appearance of objects.

### 9.3.1 Optimal triangulation

The accuracy of the spacetime analysis hinges on reliably capturing the spacetime images. Accordingly, all efforts to enhance the quality of these images are of benefit. In terms of the sensor, higher pixel densities mean more samples in $z$ in the spacetime images (as well as more $y$ samples for the spacetime volume). These improved pixel densities should be accompanied by either slower scanning or higher frame rates for increased resolution in $x$. In addition, greater dynamic range at the sensor will allow for acquisition of surfaces with widely varying reflectances even when oriented at grazing angles to the illumination. Video digitizers with more bits per pixel will also lead to more precise representations of the spacetime images. Experimenting with different illuminants can also lead to greater accuracy. In Chapter 3, we argue that widening the laser sheet may improve results, because the spacetime impulse response acts as a bandlimiting filter. Further, the limitations due to laser speckle can be reduced with partially coherent, perhaps even incoherent illumination. In the presence of spatiotemporal aliasing, methods for registering and deblurring multiple spacetime images should also lead to improved resolution.

When a surface is very bright or shiny, surface interreflections will corrupt the spacetime image. Under these circumstances, recovering accurate range is likely to be very challenging. Research in shape from shading which accounts for surface interreflections may offer some promising avenues toward a solution [Nayar et al. 1990] [Wada et al. 1995]. In the case of shiny surfaces, the errors will depend heavily on the orientation of the surface with respect to the illumination. Acquiring multiple range images can help identify some of these errors as outliers, because they will not be corroborated by the other range images. In addition, the interreflections will tend to distort the shape of the spacetime Gaussians so that they will not behave ideally; this deviation from ideal will be detected, and the samples will be discarded or downweighted in favor of data taken from a different orientation that yields better spacetime Gaussians.

The method of acquiring and analyzing the spacetime images in Chapter 4 requires all frames to be captured and then post-processed. Ideally, we would like to perform the spacetime analysis in hardware, in realtime. Such a system could be implemented as a ring of frame buffers that store the most recent $N$ frames, where $N$ is determined by the time it takes for the illuminant to traverse a point on the surface. To reduce storage costs and improve performance, the frames could be run-length encoded, and the spacetime rotation could be implemented as a shear. In the case of a scanning laser beam with a linear array sensor, the frame buffer would be small – a single array of pixels – and a hardware implementation using a high speed digital signal processing chip could be quite practical.

### 9.3.2 Improvements for volumetric surface reconstruction

The volumetric method described in Chapters 5-8 leads to a variety of directions for future work. Some of the issues that deserve attention are discussed in Sections 6.3 and 6.4. These include choosing better conditioning functions to minimize isosurface artifacts, modifying the algorithm to cope with thin surfaces and sharp corners, and exploring strategies for more aggressive space carving.

Space carving can also be improved to help remove outliers. Currently, the space carving algorithm is conservative in that only unseen regions may be reclassified as empty; any voxels containing a signed distance remain untouched. However, if a voxel is observed to be empty from several points of view, then it is highly unlikely to contain a surface. Recognizing such a circumstance could lead to an algorithm that is more capable of removing false range data, i.e., outliers.

We can also hope to improve the space carving method through the more traditional methods of carving from image silhouettes [Szeliski 1993]. Using geometric backdrops achieves the effect of using image silhouettes, but it has drawbacks. A range scanner can usually digitize large objects that do not fit into the field of view by moving the scanner or the object as needed. Requiring that the scanner always see a backdrop at the back of the field of view, however, makes it more difficult to scan these larger objects. In the case of optical triangulation, the backdrop must be visible to both the laser and the sensor. This means that the backdrop typically has to be positioned far enough away so that the triangulation scanner can see *around* the object. But requiring that the backdrop be far away, yet still visible to the sensor, means wasting valuable resolution, as the pixels must map to a larger field of view. The use of a separate color camera with colored or light-emitting backdrops solves these problems. As long as the backdrop is clearly distinguishable from the object being scanned, then the distance to the backdrop is immaterial, and it can be placed conveniently outside of the working volume of the range scanner.

The hole filling algorithm can be improved by making the surface properties of filler polygons more consistent with the remainder of the surface. When scanning a surface that has a discernable texture (such as the scales of a lizard), there will be a distinct difference between the smoothed hole fill regions and the remainder of the surface. Ideally, we would like these hole fill regions to adapt to the surrounding texture, either automatically or with some user guidance. Recent work on analyzing and synthesizing image textures suggests one approach for propagating surface properties [Heeger & Bergen 1995]. If color information were also available in regions around the hole fillers, then color texture could be grafted as well.

As range scanners become faster, some approaching realtime, a reconstruction algorithm that is fast enough to keep pace also becomes very attractive. One can imagine a motorized acquisition gantry or a hand-held scanner with orientation and position sensors delivering registered range images to a system that could create a detailed model in a very short period of time. Though we have developed a number of optimizations for the volumetric algorithm to ensure good performance, we have yet to explore an increasingly important direction: parallelization. [Lacroute 1996] has parallelized the shear warp volume rendering algorithm and obtained excellent results on a shared memory multiprocessor architecture. We would expect to attain similar improvements with the fast volumetric method described in Chapter 7. One factor that should make implementation easier is the lack of front to back ordering that volume rendering requires. In the case of volumetric surface reconstruction, the voxels are updated in a manner that is completely independent of order.

### 9.3.3 Open problems in surface digitization

A number of important problems remain on the way to automated and hi-fidelity acquisition of the shape and appearance of the surface of an object. Among the most challenging are solving for the next best view and capturing surface reflectance properties.

The "next best view" problem may be stated as follows: given a set of range views of an object, determine the next position and orientation of the object or sensor that is likely to maximize the knowledge about the object and the space around it. Of course, without knowing the shape of the object a priori, it is impossible to determine the next best view in terms of maximizing the visible surface area of the object. Alternative measures are necessary, and one promising measure is the surface area of the boundary surrounding the unknown regions of the space in and around the object. This area is in fact defined by the hole fill polygons computed in the volumetric algorithm described in this thesis. Initial results in solving for next best view are promising [Pito 1996], but more research is warranted.

For faithful reproduction of the appearance of objects, the surface reflectance properties must be acquired. A simple strategy would be to reconstruct the shape of an object, take a few color images, and then "paste" the color onto the object by following the lines of sight from the color camera. For overlapping views, the colors could simply be averaged together. This strategy is flawed, however, because the results are dependent on the lighting and the viewpoint. What we really want are the underlying reflectance properties over the surface of the object so that we can render the computer model under arbitrary lighting and viewing conditions. These reflectance properties can be very complex. For general anisotropic surfaces, the bi-directional reflectance distribution function (BRDF) at a point is five dimensional: 2 dimensions for each of the incoming and outgoing light directions, and one dimension for wavelength (two dimensions for wavelength if fluorescence is possible). Accounting for variations over the surface of the object, the space is seven (possibly eight) dimensional. With controlled lighting and many color images of the object, we can begin to fill this large space with data. The problem is further complicated, however, by the fact that light reflecting from the surface toward the sensor arrives at the surface not only from the light source, but from other points on the surface. Accounting for the effects of these interreflections is crucial to accurately measuring surface BRDF, which in turn requires a very accurate description of the shape of the object. Some results have been obtained for simplified BRDF's and moderate variations across the surface [Nayar et al. 1990] [Sato & Ikeuchi 1996], but the general problem remains open.