Proceedings of the Third International Conference on 3D Digital Imaging and Modeling, 3DIM 2001.

A Laser Range Scanner Designed for Minimum Calibration Complexity

James Davis, Xing Chen
Computer Graphics Lab, Stanford University
{jedavis, xcchen}@graphics.stanford.edu

Abstract

Laser range scanners are a popular method for acquiring three-dimensional geometry due to their accuracy and robustness. Maximizing scanner accuracy while minimizing engineering costs is a key challenge to future scanner designs. Engineering costs arise from both expensive components and difficult calibration requirements.

We propose a two camera range scanner design, specifically chosen to minimize calibration complexity and cost. This design eliminates all actuated components from the calibrated geometry. Since it is difficult to ensure absolute repeatability of moving parts, a design with only statically arranged components can dramatically reduce the costs associated with calibration.

1 Introduction

Laser triangulation scanners are an increasingly common means for acquiring three-dimensional geometry of objects. The popularity of this method is derived largely from the relative robustness and precision that is attainable. Conceptually, scanner design is very straightforward, employing simple geometry and using commonly available parts.

The accuracy that a laser scanner can achieve is largely affected by the cost of its components. Improved accuracy can be obtained by using higher priced components. Given a specific scanner design, one can build either a low cost, low accuracy device or a high cost, high accuracy device, depending on the application requirements. Because of this relationship, scanner design can be characterized by the ratio of accuracy to cost. A desirable scanner design will optimize this ratio, achieving a higher accuracy given a fixed cost; or equivalently, reducing cost given a fixed accuracy.

One traditional design for laser range scanners is shown in Figure 1. A calibrated geometry exists between the plane of laser light and the camera. By triangulating between the observed laser image and the known laser plane, a single stripe of object depth can be recovered. In order to recover an entire mesh of depth values, the object is placed on a precisely calibrated motion control platform.

Despite the simple geometry and components, laser scanners must be engineered and calibrated with extremely high precision. This calibration involves static geometry and optics, as well as actuated components. The imager lens system, the relationship between laser sheet and imager, the shape of the laser sheet, and the repeatability and motion of any motorized component must be completely characterized in order to produce an accurate depth image. The complexity of the calibration process is correlated with cost. High precision, high cost components are often needed to achieve a given target accuracy.

Using high precision components can easily result in scanners costing tens of thousands of dollars [7] [8] [14]. However, even high quality components do not make the calibration process easy. A broad literature exists specifically regarding calibration of laser scanning gantries [15] [13] [19] [5]. As an example, the Digital Michelangelo project custom designed a scanning gantry with careful attention to precision and repeatability. However after a year of use, Levoy et. al. report calibration as the most prominent difficulty and ongoing problem [14]. In particular, actuated components have produced problems with repeatability and error modeling. Calibration difficulties have caused errors of several millimeters, delaying accurate alignment of captured geometry.

Although it is theoretically possible to improve the precision and calibration of existing scanner designs through improved manufacturing and measurement, an important alternative is to explicitly design laser scanners with ease of engineering and calibration in mind. In particular, proper design can both improve the repeatability and maintainability of a laser scanner, as well as simplify the calibration process itself. Accuracy per unit cost can be increased, resulting in designs which are both more accurate and less expensive.

The overall cost of a scanner design is dependant on component costs. Existing scanners typically employ

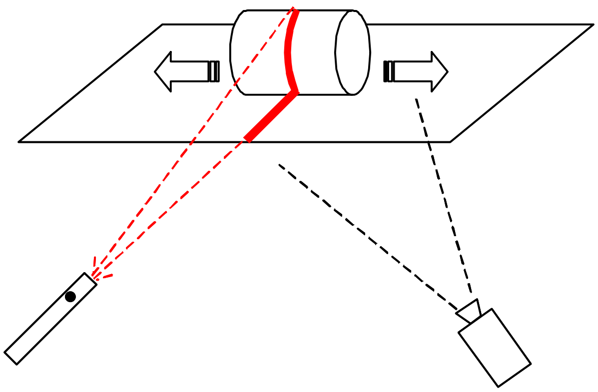

Figure 1. A traditional laser scanner design. Calibrated laser plane, camera and motion platform are used to recover object geometry.

some components that are difficult to calibrate. A better design can be found by eliminating these parts, and choosing components that are inexpensive and easy to calibrate.

Actuated components are notoriously hard to calibrate. They must be characterized in terms of geometry, velocity, and acceleration. In addition, moving parts suffer wear that results in decreased repeatability. Since parameters may change over time, a single calibration in the laboratory is insufficient, and all dynamic subsystems must be continuously monitored.

We can obtain a rough measure of overall calibration difficulty by dividing scanner geometry into statically arranged components and dynamic subsystems. By eliminating all dynamic components, a scanner with simpler calibration requirements can be attained.

This paper proposes a new scanner design that contains no actuated mechanisms requiring calibration. As shown in figure 2a, two static imagers are used to locate a laser stripe as it sweeps over an object. Although the laser stripe is moving, it is not necessary that this motion be calibrated. Depth can be recovered through triangulation between the two known viewpoints.

This design is capable of capturing a depth image using only statically calibrated components. In addition, the static calibration uses only well studied, existing error models and calibration methods, rather than complex new geometries. This in turn results in substantially reduced calibration and engineering costs.

2 Related Work

Most commercial laser triangulation scanners employ some actuated components in order to obtain a range image. Similar to Figure 1, Cyberware's 3030 and Model 15 range scanners employ a linear translation stage to induce relative motion between the scan head and a stationary object. In some cases an additional rotational stage can be used to orient the target [7]. Digibotic's scanners have a pair of translation stages that provide two dimensional motion to the scan head, while the object rests on a rotational platform [8]. As mentioned previously, we seek a design capable of acquiring a two dimensional mesh of depth values using no calibrated moving parts.

One solution was proposed by Bouguet and Perona [3]. Their scanner employs a fixed light source, imager, and ground plane. The user waves an uncalibrated wand between the target object and light source. The shadow of the wand and the light source defines a plane, similar to the plane obtained by projecting a laser in traditional scanners. By observing the shadow on the ground plane, it is possible to derive the shadow plane, and therefore the object depth. The disadvantage to this arrangement is that a flat ground plane is required. Thus objects that can not

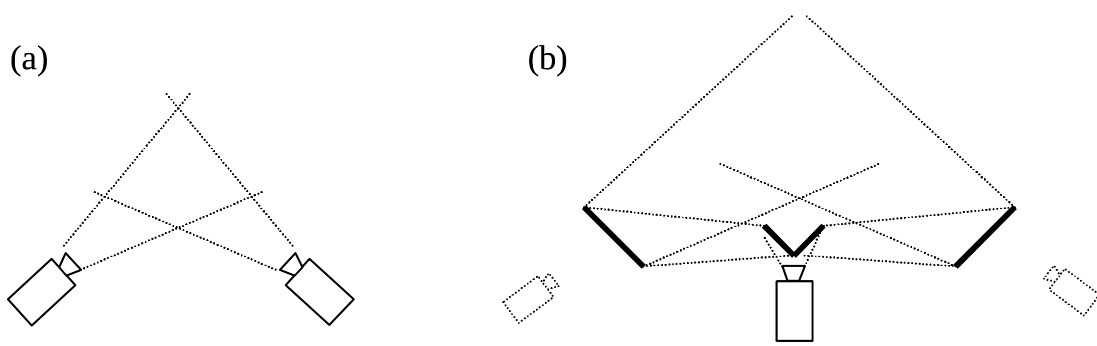

Figure 2. (a)Two static cameras can be used to locate a laser stripe, and recover depth as the uncalibrated stripe sweeps over the object. (b) A system of mirrors can be arranged to split a single camera's imager into two virtual viewpoints.

be moved onto a planar surface or that are too large for a single range image, cannot be scanned.

In order to remove the ground plane requirement, Fisher et. al. proposed tracking the wand itself rather than the ground plane shadow [9]. However this places the wand in the field of view of the imager, potentially occluding the target object. In addition, small errors in estimating wand position are magnified as the shadow plane is extrapolated to the object surface.

Takatsuka et. al. follow a similar course. They use a single imager to view a laser pointer and its projected laser spot on the target object [18]. Locating the laser pointer allows the determination of a single depth value at the object surface. Their system suffers from the same geometry extrapolation errors as mentioned above.

Borghese et. al. use two cameras and specialized hardware to locate a single laser spot [2]. Their use of a simple laser pointer implies a very slow, sparse sampling of the surface, since only a single depth estimate is obtained at each time instant.

3 Scanner design

Laser triangulation scanners rely on a known and calibrated geometry in order to reconstruct depth. The triangulation process requires at least two viewpoints.

In traditional scanners, a fixed laser plane provides one point of view, while an observing camera provides the second. This arrangement, seen in Figure 1, allows triangulation between the laser plane and a set of camera rays, thus allowing the reconstruction of a single stripe of depth values. In order to recover additional depth stripes on the object, the laser plane is swept over the object surface. For this to occur, either the laser or the object must be mounted on an accurately controlled motion platform.

Our design is based on the observation that the need for precisely calibrated motion can be alleviated by replacing the known laser plane with a second calibrated observing camera. In this configuration, the required geometry exists between two cameras and the object, as shown in Figure 2. The cameras are held fixed relative to the object, and the laser plane is now free to sweep over the object without the need for precision. As the laser plane traverses the object, observations are made in both views. For each observed position of the laser plane, a depth stripe can be reconstructed just as in the traditional design.

Since the object and both viewpoints are static in this scanner design, the laser plane is free to move without being calibrated. Irregular motion of the laser plane merely results in irregular sample spacing, rather than the incorrect noisy geometry that is produced by traditional designs. In order to illustrate that precise calibration of the laser plane is not necessary, we sweep the laser plane with a hand held laser pointer in our experiments. Of course, complete automation is possible by placing the laser on an inexpensive motorized platform. As desired, the new configuration creates significantly relaxed calibration requirements.

3.1 Cylindrical lens

As with traditional scanners, the required sheet of laser light can be created by placing a cylindrical lens in the path of a simple laser. If it is necessary that the laser sheet be perfectly planar, the laser optics must be carefully chosen and precisely mounted with respect to the laser. Since traditional systems depend on a planar, or at least calibrated laser sheet, their accuracy depends on careful placement of the optics. In contrast, because we have replaced the laser sheet with a second imager in our known configuration, uncalibrated laser geometry will not adversely affect scanner accuracy. In practice, an approximately planer laser can be obtained by simply gluing a lens onto the end of a laser pointer as shown in Figure 3. The laser stripe produced by our inexpensive lens is approximately 3mm wide.

3.2 Catadioptric layout

Simultaneous observation of the laser stripe from both views is required for accurate geometry. In the two-camera configuration shown in Figure 2a, simultaneous observations can be obtained by synchronizing the two cameras. Two physical cameras provide the highest resolution, however it is well known that an alternate two viewpoint system can be constructed from a single camera, as shown in Figure 2b. In this configuration, mirrors are used to split the camera's optical path,

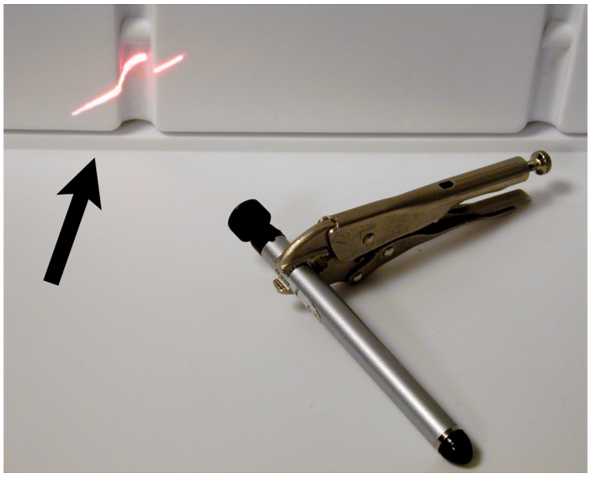

Figure 3. A laser stripe can be produced from an inexpensive laser pointer and cylindrical lens.

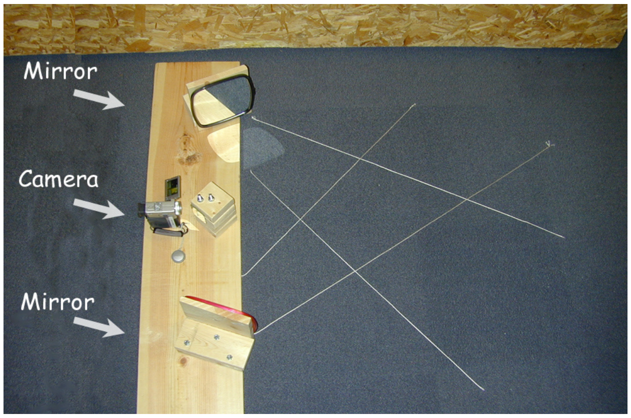

Figure 4. In this photograph of our scanner, the floor has been marked to show the field of view of each virtual camera. The overlapping region is the working volume for the scanner.

providing two virtual viewpoints on a single image plane [12] [10]. There is an reduction by half of resolution from each viewpoint. However, since a single imager is used, the observations are synchronized by construction.

We use a camcorder to record scanning sessions on tape. This allows for a portable system. Rigidly attaching the camera and mirrors to a base allows the scanner to be moved without affecting calibration. A photograph of our scanner prototype can be seen in Figure 4.

4 Depth reconstruction

Since this scanner has been physically constructed to provide two simultaneous observations of the laser stripe, we can extract object depth using stereo triangulation between viewpoints. Unlike traditional scanners that find depth via triangulation between a camera and the known geometry of the laser, in our configuration there is no requirement that the location of the moving laser sheet be known.

4.1 Calibration

In order to reconstruct depth from an observed laser stripe the relative geometry of the two observations must be known. The physical position of the viewpoints, and distortions caused by intervening optics must be accounted for. In order to reduce overall engineering costs, it is desirable to find a configuration in which precise placement of parts is not required.

While the construction shown in Figure 2b has a view frustum folded across a pair of mirrors, it is not required to calibrate the position of these mirrors. Views through a planar mirror produce a view identical to that obtained by placing the camera in a position reflected across the plane of the mirror. Rather than calibrating the physical camera and mirror positions, we can calibrate the complete optical path. The two virtual cameras can be calibrated using standard camera calibration procedures.

After constructing our scanner using only approximate placement of all parts, we calibrate using the method of Tsai [20] or Heikkila[11]. By placing a calibration target so that it lies in the field of view of both virtual cameras, we can calibrate the lens distortion and extrinsic relationship between views. Note that this calibration is identical to that used in stereo imaging systems. Calibration of stereo systems is efficient, well studied, and standard. Our scanner requires no other calibration.

4.2 Stripe processing

Just as in traditional stereo reconstruction, depth estimation requires that correspondence be established between particular locations in each image. However, the projected laser stripe, a form of structured light, nearly eliminates the ambiguities that plague standard stereo techniques.

To find corresponding points in each image, first consider a single scan line in the left camera image. By searching this scan line for the laser stripe, we can find the image plane projection of a single point on the object. Together with the left camera parameters, this point defines a ray in space. The projection of this ray onto the right image plane is an epipolar line on which the object must lie. By finding where the laser stripe in the right image crosses this epipolar line we can determine a pair of corresponding image plane points as shown in Figure 5. These corresponding points define a depth value in 3D.

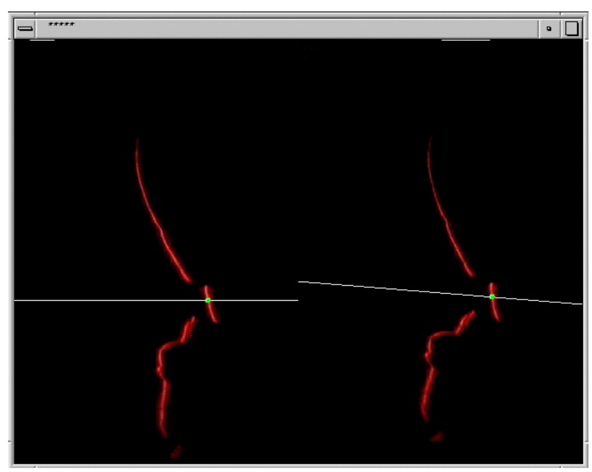

Figure 5. Laser stripe as seen from each viewpoint. The detected laser stripe location in the left viewpoint defines an epipolar line in the right image, along which the corresponding laser peak must lie.

Figure 6. When the laser sheet is swept too quickly over an object, the laser stripe is not observed by all possible pixels.

In order to *ensure* that the laser stripe crosses each line only once, careful positioning of the laser plane would be required. Although the stripe produced by the handheld laser in our system is kept approximately vertical, these precise geometry requirements may not always be satisfied. Despite this, multiple stripe crossings are only occasionally observed. Since these cases are rare and easy to detect, rather than attempting to disambiguate the correspondence, we simply discard the data.

Finding correspondences on every scan line in an image results in a single stripe of depth values. By repeating the procedure as the laser sheet sweeps the object over time, an entire 2D mesh of depth values can be recovered. Note that this mesh is obtained from a single statically arranged viewing geometry, without the need for calibrated moving parts as in traditional scanners.

4.3 Peak finding

Locating the pixels intersected by the laser stripe in each image is straightforward. Illumination from the laser can be isolated from ambient background lighting with a 670nm frequency bandpass filter. Along an image search line, the laser sheet produces an approximately Gaussian intensity distribution. Naidu and Fisher have analyzed the theoretical implications of several sub-pixel accurate peak finding algorithms [16]. Following the results of their evaluation, we locate the stripe peak by fitting a Gaussian function to the three samples surrounding the highest intensity value.

4.4 Filtering

The depth values reconstructed are not perfect. Noise generated in the imager and laser speckle both contribute to errors in the depth estimate. Noise of this sort can be filtered using standard signal processing. Although it is possible to filter the recovered mesh depth values, we have obtained better results by filtering directly in the image domain. A Gaussian filter with a width of a few pixels is applied to captured image data before the peak finding process.

Multiple stripe crossings and specular reflections can cause incorrect correspondence. Although we discard known ambiguities, in some cases it may be impossible to detect these difficult situations due to viewpoint specific occlusion. In this case incorrect depth values will be calculated. However, these depth values are often far from the true surface and thus are best filtered by an outlier rejection policy. We discard any depth estimate that differs by more than some threshold distance from its neighbors.

4.5 Gap Filling

Small gaps can arise in the depth grid due to irregular laser motion. If the laser stripe is swept across the object surface too quickly, there is no guarantee that the stripe will be observed at every possible pixel location. Figure 6 visualizes coverage from the left camera viewpoint, of a laser plane that was swept too quickly across the object. As the laser stripe is detected at each observing pixel, it is marked. Even when the laser plane is swept slowly, pixels are occasionally missed due to imager sampling.

In the case of small holes, it is often desirable to polygonize across the missing data, rather than leave a hole in the final mesh. For holes with only one or two missing samples, we linearly interpolate surrounding depth values to complete the grid.

4.6 Supporting tools

While this work focuses on the acquisition of individual rigid meshes, a great deal of further technology is needed to integrate the raw data into a single geometric object. We manually rotate meshes into a rough alignment, and then use a modified iterative closest point algorithm [4] [1] [14] to find optimal pair-wise alignment. A global optimization is then performed to minimize error over the entire object [17]. After alignment, the individual

Figure 7. This statue was scanned using the scanner shown in figure 4.

meshes are merged into a single polygonal object using a volumetric technique [6]. We have experimented with both surface and volumetric methods for filling holes in the final mesh. Color data can be included by projecting photographs onto the object, and assigning per vertex colors based on various metrics [14].

5 Results

The scanner pictured in Figure 4 as well as an earlier prototype have been used to capture geometry from a number of objects. For example, the statue shown in Figure 7 was scanned from twenty-two directions. Each scan direction resulted in a rigid mesh. Figure 9 shows several rigid scans separately, and then aligned with each other in a single coordinate system. Figure 10a shows a model created by using a volumetric method to merge all twenty-two scans into a single mesh. The volumetric grid was at a lower resolution than the original scans, and low confidence depth data was been discarded. It is clear that many holes still remain in the model. A complete model normally requires many more individual scans in order to ensure that every surface on the object has been adequately observed by the scanner. Algorithmic hole filling techniques can be used to generate a water tight mesh, and the results of one such technique can be seen in Figure 10b.

The speed at which the laser stripe is swept across the object surface determines sample spacing. We have found that sufficient sampling of the surface can be obtained by sweeping the camera field of view over the course of five to ten seconds.

The resolution obtainable by this scanner design is directly correlated to both the size of the working volume and imager resolution. The single camera in our scanner has an effective resolution of approximately 480x240 pixels in each captured field of video. Each imaged field is split between two viewpoints, each covered with a resolution of 240x240 pixels. The working volume of our scanner is approximately 440mm wide and 550mm deep. Based on these measurements, and assuming pixel accuracy for laser stripe location, the expected depth resolution of our scanner can be calculated as approximately 440mm/240 = 1.8mm. If greater resolution is desired, better results can be obtained with a smaller working volume, and higher resolution cameras.

In order to evaluate the actual resolution of our scanner we used a known target with steps of depth ranging from 1mm to 4mm. Figure 8 shows a plot of the recovered depth of a line across the object surface. Even the smallest depth discontinuity, only a 1mm step, is perceivable. This result is better than predicted by the calculation above, indicating that the laser stripe is being located with sub-pixel accuracy. We expect that by using the higher quality imagers found in current commercial laser scanners, competitive high quality accuracy could be obtained with this scanner design.

6 Discussion

The theoretical accuracy with which a laser range scanner can reconstruct object depth. is directly correlated with errors in locating the laser plane. In our two camera design, depth is established by finding the laser peak on each image plane. Camera pixel resolution determines the angular accuracy with which this peak can be found.

One traditional scanner design places the laser on a precise rotational platform, in place of the second camera. In this case the laser plane location is determined by the current rotation of this platform, as well as the pixel resolution of the remaining imager.

These two configurations can be compared by noting that errors due to the first camera are identical in either case. Errors due to the second geometric element can be evaluated in terms of angular error in locating the laser plane. In the case of a camera this error is due to limited pixel resolution. In the case of a rotational platform angular error is caused by poor repeatability introduced from a number of sources.

In our particular scanner, each half of the imager has approximately 240 pixels covering a field of view spanning 21 degrees. The stripe peak can be located with sub-pixel accuracy. Let us call this accuracy ¼ pixel. So the expected angular resolution is on the order of 0.02 degrees. Rotational platforms can of course be obtained with either better or worse specifications. A quick price check in several scientific catalogs suggested that a platform with this accuracy currently costs several hundred US dollars. The mirrors and mounting supplies used in our scanner cost about US$10.

While we have not attempted to characterize the entire space of possible requirements, this informal analysis suggests that our system provides a better accuracy/cost ratio than traditional scanner designs. In addition, the resolution and cost of imagers are tied to the semi-conductor industry, so can be expected to get better and cheaper at a rate far greater than that of mechanical actuated components.

7 Conclusion

Traditional laser triangulation scanners have required precisely engineered components, and complex calibration models. We have proposed a new scanner design which addresses one of the chief difficulties with previous designs, eliminating the need for calibrated moving components. While previous scanners often cost tens of thousands of dollars, we have demonstrated a working configuration that can be built from components costing only a few hundred dollars In addition, and importantly, the components can be assembled with minimal attention to precision, and calibrated using well-studied, robust calibration techniques. Thus, higher reconstruction quality can be achieved given the same cost/complexity constraints.

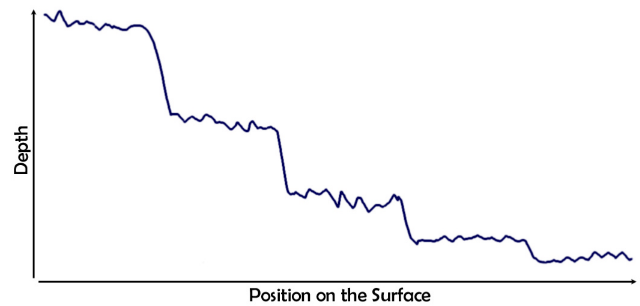

Figure 8. A cross-section of recovered geometry on a target with known steps ranging in depth from 4mm to 1mm.

References

[1] P. J. Besl and H. D. McKay, *A method for registration of 3D shapes*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 14 (1992), pp. 239-56.

[2] N. A. Borghese, G. Ferrigno, G. Baroni, A. Pedotti, S. Ferrari and R. Savare, *Autoscan: a flexible and portable 3D scanner*, IEEE Computer Graphics and Applications, 18 (1998), pp. 38-41.

[3] J. Y. Bouguet and P. Perona, *3D photography on your desk*, *Sixth IEEE International Conference on Computer Vision*, Narosa Publishing House, New Delhi, India, 1998.

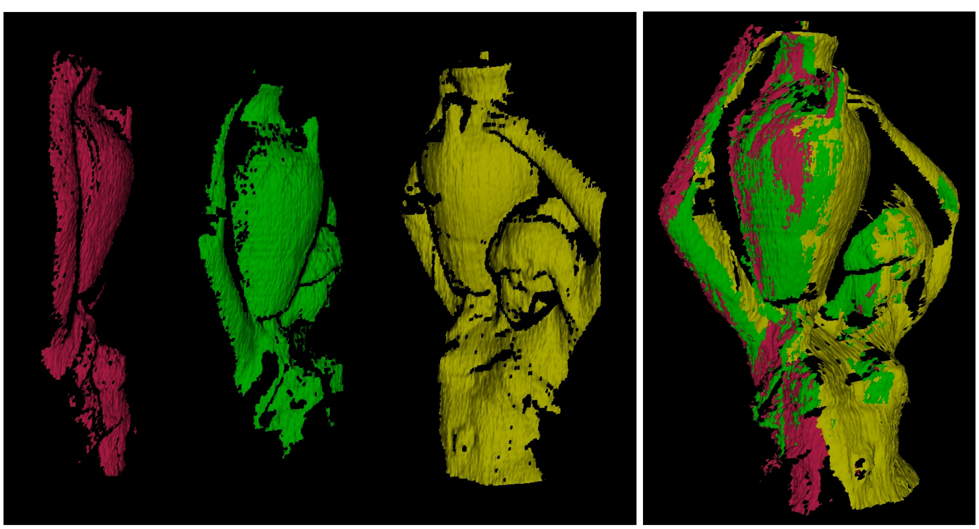

Figure 9. Left: Three meshes scanned from different viewpoints using our scanner. Right: The meshes have been aligned into a single coordinate system.

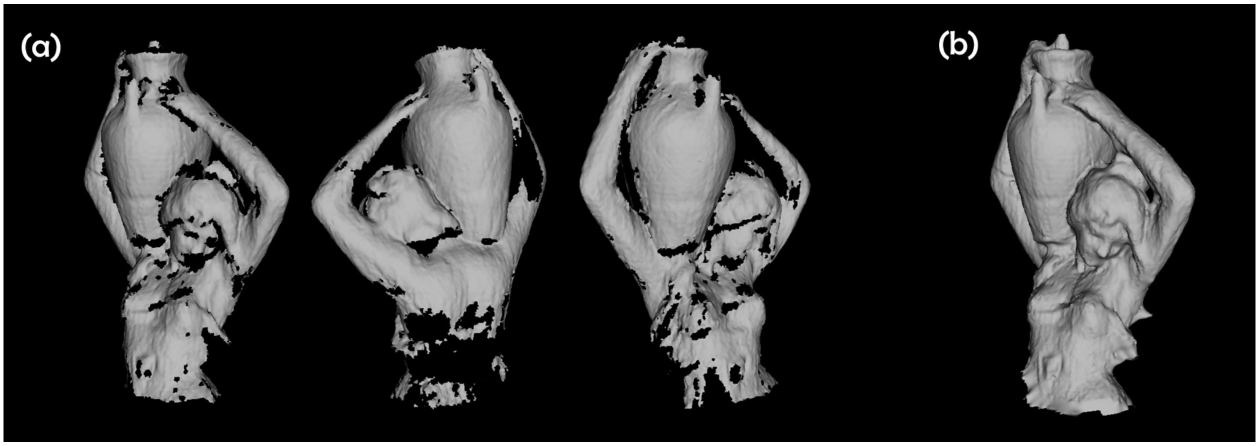

Figure 10. (a) A volumetric method was used to merge twenty-two scans into a single mesh. The single mesh is shown from three viewpoints. (b) A hole filling technique was used to fill holes remaining in areas not observed in the scanning process.

[4] Y. Chen and G. Medioni, *Object modelling by registration of multiple range images*, Image and Vision Computing, 10 (1992), pp. 145-55.

[5] M. Chiarella and K. A. Pietrzak, *An accurate calibration technique for 3-D laser stripe sensors*, *Proceedings of the SPIE Vol. 1194 - The International Society for Optical Engineering*, 1989.

[6] B. Curless and M. Levoy, *A volumetric method for building complex models from range images*, , *Computer Graphics Proceedings. SIGGRAPH '96*, Acm, New York, NY, USA, 1996.

[7] Cyberware, http://www.cyberware.com.

[8] Digibotics, http://www.digibotics.com.

[9] R. B. Fisher, A. P. Ashbrook, C. Robertson and N. Werghi, *A low-cost range finder using a visually located, structured light source*, *Second International Conference on 3-D Digital Imaging and Modeling*, IEEE Comput. Soc, Los Alamitos, CA, USA, 1999.

[10] A. Goshtasby and W. A. Gruver, *Design of a single-lens stereo camera system*, Pattern Recognition, 26 (1993), pp. 923-7.

[11] J. Heikkila and O. Silven, *A Four-Step Camera Calibration Procedure with Implicit Image Correction*, IEEE Computer Vision and Pattern Recognition (1997), pp. 1106-1112.

[12] M. Inaba, T. Hara and H. Inoue, *A stereo viewer based on a single camera with view-control mechanisms*, *IROS '93. Proceedings of the 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems. Intelligent Robots for Flexibility*, Ieee, New York, NY, USA, 1993.

[13] O. Jokinen, *Self-calibration of a light striping system by matching multiple 3-D profile maps*, , *Second International Conference on 3-D Digital Imaging and Modeling*, IEEE Comput. Soc, Los Alamitos, CA, USA, 1999.

[14] M. Levoy, K. Pulli, B. Curless, S. 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*, Computer Graphics. SIGGRAPH '00 (2000).

[15] A. M. McIvor, *Calibration of a laser stripe profiler*, *Second International Conference on 3-D Digital Imaging and Modeling*, IEEE Comput. Soc, Los Alamitos, CA, USA, 1999.

[16] D. K. Naidu and R. B. Fisher, *A Comparative Analysis of Algorithms for Determining the Peak Position of a Stripe to Subpixel Accuracy*, Proc. British Machine Vision Conf. (1991), pp. 217-225.

[17] K. Pulli, *Multiview registration for large data sets*, *Second International Conference on 3-D Digital Imaging and Modeling*, IEEE Comput. Soc, Los Alamitos, CA, USA, 1999.

[18] M. Takatsuka, G. A. W. West, S. Venkatesh and T. M. Caelli, *Low-cost interactive active monocular range finder*, *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Comput. Soc, Los Alamitos, CA, USA, 1999.

[19] E. Trucco and R. B. Fisher, *Acquisition of consistent range data using local calibration*, *Proceedings 1994 IEEE International Conference on Robotics and Automation*, IEEE Comput. Soc. Press, Los Alamitos, CA, USA, 1994.

[20] R. Y. Tsai, *A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses*, IEEE Trans. Robotics and Automation, 3 (1987), pp. 323-344.