The Digital Michelangelo Project: 3D Scanning of Large Statues

Marc Levoy [1 *] *Kari Pulli* [1] *Brian Curless* [2]
Szymon Rusinkiewicz [1] *David Koller* [1] *Lucas Pereira* [1]
Matt Ginzton [1] *Sean Anderson* [1] *James Davis* [1]
Jeremy Ginsberg [1] *Jonathan Shade* [2] *Duane Fulk* [3]

[1]Computer Science Department
Stanford University

[2]Department of Computer Science and Engineering
University of Washington

[3]Cyberware Inc.

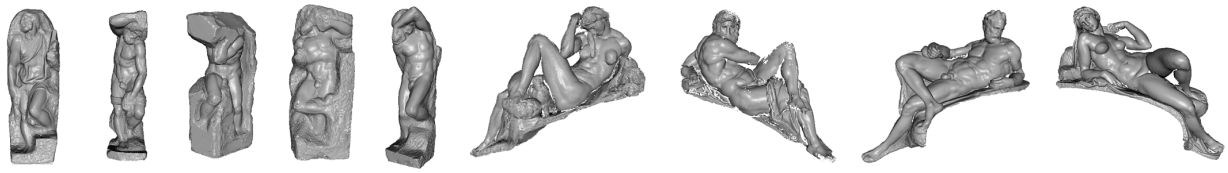


Figure 1: Renderings of the statues we scanned (except the David). Our raw database (including the David) contains 10 billion polygons and 40,000 color images, occupying 250 gigabytes. From left to right: St. Matthew, Bearded Slave, Slave called Atlas, Awakening Slave, Youthful Slave, Night, Day, Dusk, and Dawn.

Abstract

We describe a hardware and software system for digitizing the shape and color of large fragile objects under non-laboratory conditions. Our system employs laser triangulation rangefinders, laser time-of-flight rangefinders, digital still cameras, and a suite of software for acquiring, aligning, merging, and viewing scanned data. As a demonstration of this system, we digitized 10 statues by Michelangelo, including the well-known figure of David, two building interiors, and all 1,163 extant fragments of the Forma Urbis Romae, a giant marble map of ancient Rome. Our largest single dataset is of the David - 2 billion polygons and 7,000 color images. In this paper, we discuss the challenges we faced in building this system, the solutions we employed, and the lessons we learned. We focus in particular on the unusual design of our laser triangulation scanner and on the algorithms and software we developed for handling very large scanned models.

CR Categories: I.2.10 [Artificial Intelligence]: Vision and Scene Understanding — modeling and recovery of physical attributes; I.3.3 [Computer Graphics]: Picture/Image Generation — digitizing and scanning; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism — color, shading, shadowing, and texture; I.4.8 [Image Processing]: Scene Analysis — range data

Additional keywords: 3D scanning, rangefinding, sensor fusion, range images, mesh generation, reflectance and shading models, graphics systems, cultural heritage

[*] Email: levoy@cs.stanford.edu Web: http://graphics.stanford.edu/projects/mich/

1. Introduction

Recent improvements in laser rangefinder technology, together with algorithms for combining multiple range and color images, allow us to accurately digitize the shape and surface characteristics of many physical objects. As an application of this technology, a team of 30 faculty, staff, and students from Stanford University and the University of Washington spent the 1998-99 academic year in Italy digitizing sculptures and architecture by Michelangelo.

The technical goal of this project was to make a 3D archive of as many of his statues as we could scan in a year, and to make that archive as detailed as scanning and computer technology would permit. In particular, we wanted to capture the geometry of his chisel marks, which we found to require a resolution of 1/4 mm, and we wanted to scan the David, which stands 5 meters tall without its pedestal. This implies a dynamic range of 20,000:1. While not large for a computer model, it is very large for a scanned model.

Why did we want to capture Michelangelo's chisel marks? On his finished or nearly finished statues, especially those in the Medici Chapel (first through fourth from the right in figure 1), Michelangelo often left the surface deliberately bumpy. The tiny shadows cast by these bumps deepen the shading of curved surfaces. If we wanted our computer models to look realistic under arbitrary lighting, we had to capture these bumps geometrically. On his unfinished statues, for example St. Matthew and the Slaves (first through fifth from the left), his chisel marks tell us how he worked. Starting from a computer model, it might be possible to segment the statue surface according to the chisels used to carve each region (figure 2).

In addition to capturing shape, we also wanted to capture color. More specifically, we wanted to compute the surface reflectance of each point on the statues we scanned. Although extracting reflectance is more difficult than merely recording color, it permits us to relight the statue when rendering it. It also constitutes a unique and useful channel of scientific information. Old statues like

the David are covered with a complex brew of marble veining, dirt, waxes and other materials used in prior restorations, and, since it sat outside for 400 years, discoloration and other effects of weathering [Dorsey99]. These tell us a story about the history of the statue. To help uncover this story, we scanned the David under white light and, separately, under ultraviolet light (figure 14). Unfinished statues, like St. Matthew (figure 9), have different stories to tell. The bottoms of its chisel marks are whiter than the surrounding marble due to the crushing of marble crystals under the impact of the chisel. The characteristics of these whitened areas might tell us how Michelangelo held his chisel and how hard he struck it.

Although digitization of 2D artwork is a mature field and is widely deployed in the museum and library communities, relatively few groups have tackled the problem of digitizing large 3D artworks. Two notable exceptions are the National Research Council of Canada (NRC) and IBM. The NRC efforts are interesting because they focus on building robust, field-deployable systems, and consequently their papers echo some of the same concerns raised in this paper [Beraldin99]. The IBM efforts are interesting first because they scanned a statue under field conditions, and second because they used a structured-light scanner in conjunction with photometric stereo, producing geometry at 2.0 mm and a normal vector field at sub-millimeter resolution [Rushmeier97]. Although their resulting models are not as detailed as ours, their equipment is lighter-weight and therefore more portable.

In the remaining sections, we describe the scanner we built (section 2), the procedure we followed when scanning a statue (section 3), and our post-processing pipeline (section 4). In section 5, we discuss some of the strategies we developed for dealing with the large datasets produced by our scanning system. In addition to scanning the statues of Michelangelo, we acquired a light field of one statue, we scanned two building interiors using a time-of-flight scanner, and we scanned the fragments of an archeological artifact central to the study of ancient Roman topography. These side projects are described briefly in figures 12, 15, and 16, respectively.

2. Scanner design

The main hardware component of our system was a laser triangulation scanner and motorized gantry customized for digitizing large statues. Our requirements for this scanner were demanding; we wanted to capture chisel marks smaller than a millimeter, we wanted to capture them from a safe distance, and we wanted to reach the top of Michelangelo's David, which is 23 feet tall on its pedestal. In the sections that follow, we describe the range and color acquisition systems of this scanner, its supporting mechanical gantry, and our procedure for calibrating it.

2.1. Range acquisition

To a first approximation, marble statues present an optically cooperative surface: light-colored, diffuse (mostly), and with a consistent minimum feature size imposed by the strength of the material. As such, their shape can be digitized using a variety of non-contact rangefinding technologies including photogrammetry, structured-light triangulation, time-of-flight, and interferometry. Among these, we chose laser-stripe triangulation because it offered the best combination of accuracy, working volume, robustness, and portability. Our design, built to our specifications by Cyberware Inc., employed a 5 mW 660-nanometer laser diode, a 512 x 480 pixel CCD sensor, and a fixed triangulation angle. Although based on Cyberware's commercial systems, it differed in two important respects: we used a triangulation angle of 20° rather than 30°, and our sensor viewed the laser sheet from only one side rather than combining views from both sides using a beam splitter. These changes were made to reduce our baseline, which in turn reduced the size and weight of our scan head.

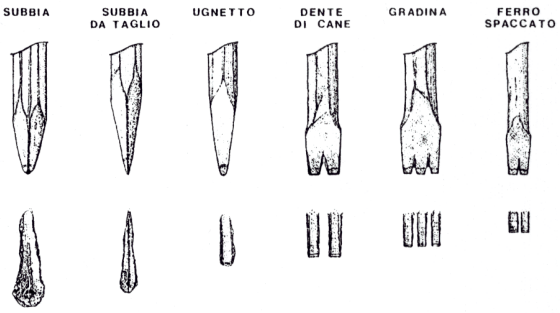
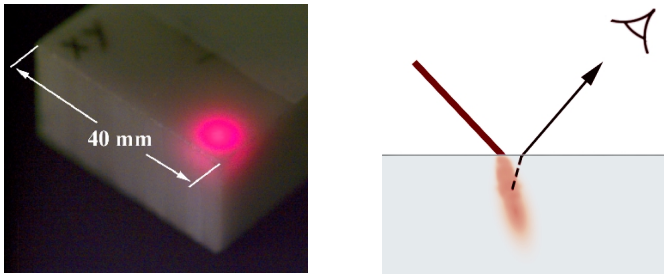


Figure 2: Some of the chisels that Michelangelo may have used when carving St. Matthew (figure 9). At top are the tools themselves, labeled with their Italian names. At bottom are sketches of the characteristic trace left by each tool. The traces are 2-10 mm wide and 1-5 mm deep [Giovannini99].

Resolution and field of view. One of our goals was to capture Michelangelo's chisel marks. It is not known exactly what tools Michelangelo used, but they almost certainly included the single-point and multi-point chisels shown in Figure 2. We wished not only to resolve the traces left by these chisels, but to record their shape as well, since this gives us valuable clues about how Michelangelo held and applied his chisels. After testing several resolutions, we decided on a Y sample spacing (along the laser stripe) of 1/4 mm and a Z (depth) resolution at least twice this fine [1]. This gave us a field of view 14 cm wide (along the laser stripe) by 14 cm deep. In retrospect, we were satisfied with the resolution we chose; anything lower would have significantly blurred Michelangelo's chisel marks, and anything higher would have made our datasets unmanageably large.

Standoff and baseline. The ability of lasers to maintain a narrow beam over long distances gave us great latitude in choosing the distance between the camera and the target surface. A longer standoff permits access to deeper recesses, and it permits the scanner to stay further from the statue. However, a longer standoff also implies a longer baseline, making the scan head more cumbersome, and it magnifies the effects of miscalibration and vibration. Keeping these tradeoffs in mind, we chose a standoff of 112 cm - slightly more than half the width of Michelangelo's David. This made our baseline 41 cm. In retrospect, our standoff was sometimes too long and other times not long enough. For an inward-facing surface near the convex hull of a statue, the only unoccluded and reasonably perpendicular view may be from the other side of the statue, requiring a standoff equal to the diameter of the convex hull. In other cases, the only suitable view may be from near the surface itself. For example, to scan the fingertips of David's upraised and curled left hand, we were forced to place the scan head uncomfortably close to his chest. A scanner with a variable standoff would have helped; unfortunately, such devices are difficult to design and calibrate.

[1] As built, our Y sample spacing was 0.29 mm. Our CCD was interlaced, so samples were acquired in a zigzag pattern and deinterlaced by interpolation. Our Z (depth) resolution was 50 microns.



(a) (b)

Figure 3: Subsurface scattering of laser light in marble. (a) Photograph of a focused 633-nanometer laser beam 120 microns in diameter striking an unpolished sample of Carrara Statuario marble. (Photo courtesy of National Research Council of Canada.) (b) The scattered light forms a volume below the marble surface, leading to noise and a systematic bias in derived depth.

2.2. How optically cooperative is marble?

Although marble is light-colored and usually diffuse, it is composed of densely packed transparent crystals, causing it to exhibit subsurface scattering. The characteristics of this scattering greatly depend on the type of marble. Most of Michelangelo's statues were carved from Carrara Statuario, a highly uniform, non-directional, fine-grain stone. Figure 3(a) shows the interaction of a laser beam with a sample of this marble. We observe that the material is very translucent. Fortunately, the statues we scanned were, with the exception of Night, unpolished, which increased surface scattering and thus reduced subsurface scattering. Moreover, several of them, including the David, were coated with dirt, reducing it more.

In the context of our project, subsurface scattering had three implications: it invalidated our assumption that the surface was ideal Lambertian (see section 4.2), it changed the way we should render our models if we wish to be photorealistic, and it degraded the quality of our range data. Given the goals of our project, the latter effect was important, so working in collaboration with the Visual Information Technology lab of the National Research Council of Canada (NRC), we have been analyzing the effect of subsurface scattering on 3D laser scanning.

When a laser beam enters a marble block, it creates a volume of scattered light whose apparent centroid is below the marble surface, as shown in figure 3(b). This has two effects. First, the reflected spot seen by the range camera is shifted away from the laser source. Since most laser triangulation scanners operate by detecting the center of this spot, the shift causes a systematic bias in derived depth. The magnitude of this bias depends on angle of incidence and angle of view. On this marble sample, we measured a bias of 40 microns at roughly normal incidence and 20° viewing obliquity. Second, the spot varies in shape across surface of the block due to random variations in the crystalline structure of the marble, leading to noise in the depth values. Our scanner exhibited a 1-sigma noise of 50 microns on an optically cooperative surface. However, this noise was 2-3 times higher on Michelangelo's statues, more on polished statues, and even more if the laser struck the surface obliquely. The latter effect made view planning harder.

For a statue of reasonably homogeneous composition, it should be possible to correct for the bias we describe here. However, we know of no way to completely eliminate the noise. These effects are still under investigation.

2.3. Color acquisition

Some Cyberware laser-stripe scanners acquire range and color in a single pass using a broadband luminaire and a 1D color sensor. Simultaneous acquisition makes sense for moving objects such as faces, but avoiding cross-talk between the laser and luminaire is difficult, and consequently the color fidelity is poor. Other scanners employ RGB lasers, acquiring color and shape at once and avoiding cross-talk by sensing color in three narrow bands. However, green and blue lasers, or tunable lasers, are large and complex; at the time we designed our system no portable solution existed. We therefore chose to acquire color using a broadband luminaire, a separate sensor, and a separate pass across the object. The camera we chose was a Sony DKC-5000 - a programmable 3-CCD digital still camera with a nominal resolution of 1520 x 1144 pixels [1].

Standoff. Having decided to acquire color in a separate pass across the object, we were no longer tied to the standoff of our range camera. However, to eliminate the necessity of repositioning the scanner between range and color scans, and to avoid losing color-to-range calibration, we decided to match the two standoffs. This was accomplished by locking off the camera's focus at 112 cm.

Resolution and field of view. Our color processing pipeline (section 4.2) uses the surface normals of our merged mesh to convert color into reflectance. Since the accuracy of this conversion is limited by the accuracy of these normals, we decided to acquire color at the same resolution as range data. To achieve this we employed a 25 mm lens, which at 112 cm gave a 25 cm x 19 cm field of view on the statue surface. The spacing between physical CCD pixels was thus 0.31 mm. By contrast, the IBM group acquired color at a higher resolution than range data, then applied photometric stereo to the color imagery to compute high-resolution normals [Rushmeier97]. Our decision to match the two resolutions also simplified our 3D representation; rather than storing color as a texture over parameterized mesh triangles [Sato97, Pulli97, Rocchini99], we simply stored one color per vertex.

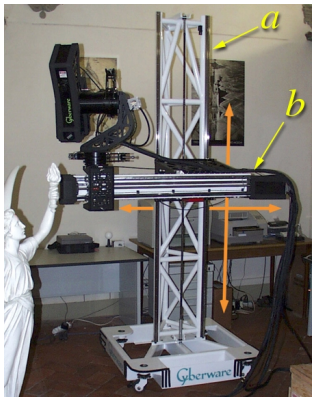
Lighting and depth of field. When acquiring color, it is important to control the spatial and spectral characteristics of the illumination. We employed a 250-watt quartz halogen lamp focused to produce as uniform a disk as possible on the statue surface. Since we planned to acquire color and range data from the same standoff, it would be convenient if the color camera's depth of field matched or exceeded the Z-component of the field of view of our range camera. For our lighting, we achieved this by employing an aperture of f/8. This gave us a circle of confusion 0.3 mm in diameter at 10 cm in front of and behind the focused plane.

2.4. Gantry: geometric design

Although our scan head was customized for scanning large statues, its design did not differ greatly from that of other commercial laser-stripe triangulation systems. Our mechanical gantry, on the other hand, was unusual in size, mobility, and reconfigurability.

Scanning motions. Most laser-stripe scanners sweep the laser sheet across the target surface by either translating or rotating the scan head. Rotational tables are easy to build, but curved working volumes don't work well for scanning flat or convex surfaces, and motion errors are magnified by the lever arm of the standoff

[1] The 3 CCDs actually have a physical resolution of only 795 x 598 pixels; the camera's nominal resolution is achieved by offsetting the CCDs diagonally and interpolating.



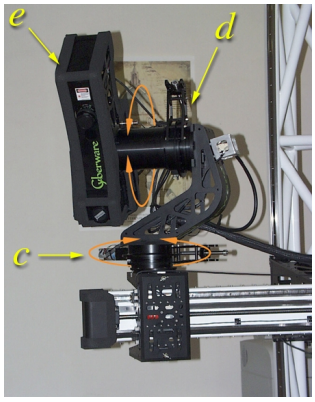


Figure 4: Our laser triangulation scanner and motorized gantry. The scanner, built to our specifications by Cyberware Inc., consisted of an 8-foot vertical truss (*a*), a 3-foot horizontal arm (*b*) that translated vertically on the truss, a pan (*c*) and tilt (*d*) assembly that translated horizontally on the arm, and a scan head (*e*) that mounted on the pan-tilt assembly, The scan head contained a laser, range camera, white spotlight, and digital color camera. The four degrees of freedom are shown with orange arrows.

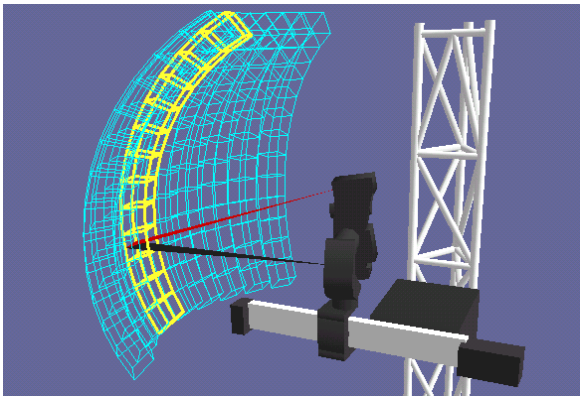


Figure 5: The working volume of our scanner. The volume scannable using our tilt motion was a curved shell 14 cm wide, 14 cm thick, and 195 cm long (yellow). Our pan axis increased the width of this shell to 195 cm (blue). Our horizontal translation table increased its thickness to 97 cm (not shown), assuming the scan head was looking parallel to the table. Including vertical motion, all truss extensions, and all scan head reconfigurations, our working volume was 2 meters x 4 meters x 8.5 meters high.

distance. Translational tables avoid these problems, but they are harder to build and hold steady at great heights. Also, a translating scan head poses a greater risk of collision with the statue than a rotating scan head. Mainly for this reason, we chose a rotating scanner. Our implementation, shown in figure 4, permits 100° of tilting motion, producing the working volume shown in yellow in figure 5. To increase this volume, we mounted the scan head and tilt mechanism on a second rotational table providing 100° of panning motion, producing the working volume shown in blue in the figure. This was in turn mounted on horizontal and vertical translation tables providing 83 cm and 200 cm of linear motion, respectively.

Extensions and bases. To reach the tops of tall statues, the 8-foot truss supporting our vertical translation table could be mounted above a 2-foot or 4-foot non-motorized truss (or both), and the horizontal table could be boosted above the vertical table by an 8-foot non-motorized truss (see figure 6). The entire assembly rested on a 3-foot x 3-foot base supported by pads when scanning or by wheels when rolling. To maintain a 20° tipover angle in its tallest configuration, up to 600 pounds of weights could be fitted into receptacles in the base. To surmount the curbs that surround many statues, the base could be placed atop a second, larger platform with adjustable pads, as shown in the figure. Combining all these pieces placed our range camera 759 cm above the floor, and 45 cm higher than the top of David's head, allowing us to scan it.

Scan head reconfigurations. Statues have surfaces that point in all directions, and laser-stripe scanning works well only if the laser strikes the surface nearly perpendicularly. We therefore designed our pan-tilt assembly to be mountable above or below the horizontal arm, and facing in any of the four cardinal directions. This enabled us to scan in any direction, including straight up and down. To facilitate scanning horizontal crevices, e.g. folds in carved drapery, the scan head could also be rolled 90° relative to the pan-tilt assembly, thereby converting the laser stripe from horizontal to vertical.

Discussion. The flexibility of our gantry permitted us to scan surfaces of any orientation anywhere within a large volume, and it gave us several ways of doing so. We were glad to have this flexibility, because we were often constrained during scanning by various obstructions. On the other hand, taking advantage of this flexibility was arduous due to the weight of the components, dangerous since some reconfigurations had to be performed while standing on a scaffolding or by tilting the gantry down onto the ground, and time-consuming since cables had to be rerouted each time. In retrospect, we should probably have mechanized these reconfigurations using motorized joints and telescoping sections. Alternatively, we might have designed a lighter scan head and mounted it atop a photographic tripod or movie crane. However, both of these solutions sacrifice rigidity, an issue we consider in the next section.

2.5. Gantry: structural design

The target accuracy for our range data was 0.25 mm. Given our choice of a rotating scanner with a standoff of 112 cm, this implied knowing the position and orientation of our scan head within 0.25 mm and 0.013°, respectively. Providing this level of accuracy in a laboratory setting is not hard; providing it atop a mobile, reconfigurable, field-deployable gantry 7.6 meters high is hard.

Deflections. Our scan head and pan-tilt assembly together weighed 15.5 kg. To eliminate deflection of the gantry when panning or tilting, the center of gravity of each rotating part was made coincident with its axis of rotation. To eliminate deflection during horizontal motion, any translation of the scan head / pan-tilt assembly in one direction was counterbalanced by translation in the opposite direction of a lead counterweight that slid inside the horizontal arm. No attempt was made to eliminate deflections during vertical motion, other than by making the gantry stiff.

Vibrations. Our solutions to this problem included using high-grade ball-screw drives for the two scanning motions (pan and tilt), operating these screws at low velocities and accelerations, and

keeping them well greased. One worry that proved unfounded was the stability of the museum floors. We were fortunate to be operating on marble floors supported below by massive masonry vaults.

Repeatability. In order for a mechanical system to be calibratable, it must be repeatable. Toward this end, we employed high-quality drive mechanisms with vernier homing switches, we always scanned in the same direction, and we made the gantry stiff. Ultimately, we succeeded in producing repeatable panning, tilting, and horizontal translation of the scan head, even at maximum height. Repeatability under vertical translation, including the insertion of extension trusses, was never assumed. However, reconfiguring the pan-tilt assembly proved more problematic. In retrospect, this should not have surprised us; 11 microns of play - 1/10 the diameter of a human hair - in a pin and socket joint located 5 cm from the pan axis will cause an error of 0.25 mm at our standoff distance of 112 cm. In general, we greatly underestimated the difficulty of reconfiguring our scanner accurately under field conditions.

2.6. Calibration

The goal of calibrating our gantry was to find a mapping from 2D coordinates in its range and color images to 3D coordinates in a global frame of reference. Ideally, this frame of reference should be the (stationary) statue. However, we did not track the position of the gantry, so it became our frame of reference, not the statue. The final mapping from gantry to statue was performed in our system by aligning new scans with existing scans as described in section 4.1.

Calibration of the range and motion systems. To calibrate any system, one must first choose a mathematical model that approximates the system behavior, then estimate the parameters of that model by measuring the behavior of the system. In our case, the natural mathematical model was a parameterized 3D geometric model of the scan head and gantry. If the components of the system are sufficiently independent, then calibration can be partitioned into stages corresponding to each component. For us, independent meant rigid - yet another reason to build a stiff gantry. Partitioning calibration into stages reduces the degrees of freedom in each stage and therefore the number of measurements that must be made to calibrate that stage. For a mechanical system, it also reduces the physical volume over which these measurements must be taken, a distinct advantage since our gantry was large. Finally, multi-stage calibration is more resistant to the replacement of individual components; if our laser had failed in the field, only one part of our calibration would have been invalidated. We had six calibration stages:

(1) a 2D mapping from pixel coordinates in the range camera image to physical locations on the laser sheet

(2) a 2D -> 3D rigid transformation from the laser sheet coordinate system to steel tooling balls attached to the scan head

(3) a 3D rigid transformation to accommodate rolling the scan head 90° (by remounting it) relative to the pan-tilt assembly

(4) the location of the tilting rotation axis and the nonlinear mapping from motion commands to physical rotation angles

(5) the location of the panning rotation axis and the mapping from its motion commands to physical rotation angles

(6) the location of the translation axis, which also depended how the pan-tilt assembly was mounted on the horizontal arm

We chose not to calibrate our vertical translation axis, since its motion induced deflections in the gantry that exceeded our error budget. The results of our calibration procedure can be visualized as the concatenation of six 4 x 4 transformation matrices:

$$\begin{bmatrix}\text{horizontal}\\ \text{translation}\end{bmatrix}\begin{bmatrix}\text{panning}\\ \text{rotation}\end{bmatrix}\begin{bmatrix}\text{tilting}\\ \text{rotation}\end{bmatrix}\begin{bmatrix}\text{rolling}\\ \text{rotation}\end{bmatrix}\begin{bmatrix}\text{laser to}\\ \text{scan head}\end{bmatrix}\begin{bmatrix}\text{image}\\ \text{to laser}\end{bmatrix}$$

Calibration of the color system.

- To correct for geometric distortion in our color camera, we photographed a planar calibration target, located a number of feature points on it, and used these to calculate the camera's intrinsic parameters. Our model included two radial and two tangential distortion terms, off-center perspective projection, and a possibly non-uniform (in X and Y) scale [Heikkilä97].
- To obtain a mapping from the color camera to the scan head, we scanned the target using our laser and range camera. Since our scanner returned reflected laser intensity as well as depth, we were able to calculate the 3D coordinates of each feature point.
- To correct for spatial radiometric effects, including lens vignetting, angular non-uniformity and inverse-square-law falloff of our spotlight, and spatial non-uniformity in the response of our sensor, we photographed a white card under the spotlight and built a per-pixel intensity correction table.

Discussion. How well did our calibration procedures work? Only moderately well; in fact, this was the weakest part of our system. The fault appears to lie not in our geometric model, but in the repeatability of our system. Comparing scans taken under different conditions (different scan axes, translational positions, etc.), we have observed discrepancies larger than a millimeter, enough to destroy Michelangelo's chisel marks if they cannot be eliminated. Fortunately, we have been able to use our software alignment process to partially compensate for the shortcomings of our calibration process, as discussed in section 4.1. An alternative solution we are now investigating is self-calibration - using scans taken under different conditions to better estimate the parameters of our geometric model [Jokinen99]. We also learned a few rules of thumb about designing for calibration: store data in the rawest format possible (e.g. motion commands instead of derived rotation angles) so that if the calibration is later improved, it can be applied to the old data (we did this), check the calibration regularly in the field (we didn't do this), and be wary of designing a reconfigurable scanner. Finally, we found that partitioning calibration into stages, and our particular choice of stages, forced us to measure scan head motions to very fine tolerances. We are currently exploring alternative partitionings.

3. Scanning procedure

Figure 6 shows our typical working environment in a museum. The basic unit of work was a "scan"; an efficient team could complete 10-15 scans in an 8-hour shift. Here are the steps in a typical scan:

Scan initiation. An operator interactively moved the scan head through a sequence of motions, setting the limits of the volume to be scanned. The volume that could be covered in a single scan was constrained by four factors:

- the field of view and limits of motion of the scanner
- the falloff in scan quality with increasing laser obliquity
- occlusions of either the laser or the line of sight to the camera
- physical obstructions such as walls, the statue, or the gantry

Once a scan was planned, a scanning script ran automatically, taking from a few minutes to an hour or more to complete, depending on how large an area was to be covered.

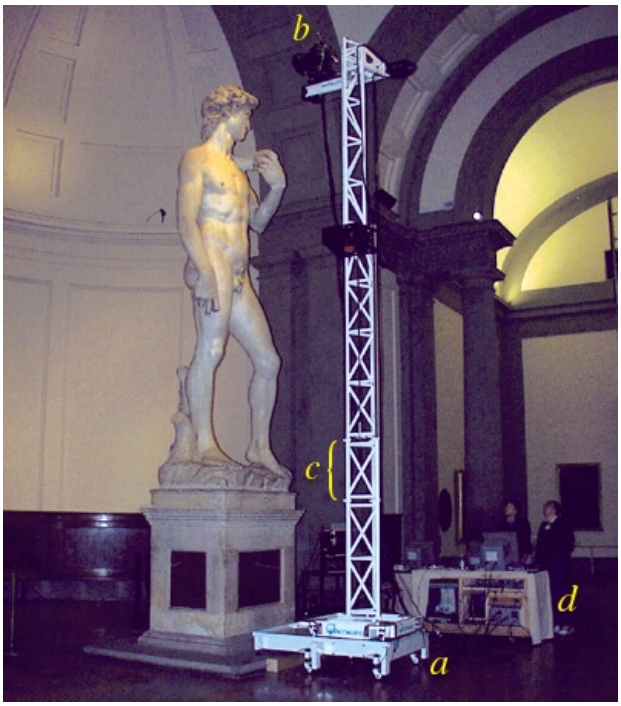


Figure 6: Our Cyberware gantry standing next to Michelangelo's David. In this photograph the gantry is sitting atop its curb-hopping platform (*a*) and is extended to its maximum height, thus placing the scan head (*b*) 25 feet above the floor. The truss section at (*c*) was added at the last minute when we discovered that the David was taller than we thought [1]. The scanner was tethered to a mobile workbench (*d*), which contained the scanner electronics, a Silicon Graphics Octane, and 72 GB of disk storage.

Range scanning. A typical range scan consisted of several concentric curved shells separated by translational motion of the scan head along the horizontal table. Each shell in turn consisted of several horizontally adjacent vertical sweeps of the laser, as shown in figure 5. If the laser line was turned vertically, then the sweeps were horizontal instead. We decided to overlap adjacent sweeps and shells by 40% and 15%, respectively - enough to align them in software in the absence of precisely calibrated motion. Since scanning was slow (1 cm per second), we preceded each sweep with a high-speed (10 cm per second), low-resolution pre-scan that conservatively determined which part of the sweep actually contained data.

Color scanning. To maintain color-to-range calibration, we interspersed color and range scanning. Since the field of view and depth of field of the color camera were greater than the field of view of the range camera, we acquired color images more sparsely than range sweeps. To compensate for ambient lighting (we often scanned during the day), we shot each image twice, once with and once without our spotlight. By subtracting these two images, we obtained an image as if illuminated only by the spotlight.

[1] We designed our gantry according the height given in Charles De Tolnay's 5-volume study of Michelangelo [Tolnay45] and echoed in every other book we checked, including the official guidebook sold at the museum. However, the David is not 434cm without his pedestal, as given by these sources; he is 517cm, an error of nearly 3 feet! We do not know the original source of this error.

```
for horizontal = min to max by 12 cm
  for pan = min to max by 4.3°
    for tilt = min to max continuously
      perform fast pre-scan (5°/sec)
    search pre-scan for range data
    for tilt = all occupied intervals
      perform slow scan (0.5°/sec)
  on every other horizontal position,
  for pan = min to max by 7°
    for tilt = min to max by 7°
      take color image without spotlight
  warm up spotlight
  for pan = min to max by 7°
    for tilt = min to max by 7°
      take color image with spotlight
```

Figure 7: The sequence of events executed by a typical scanning script. For this script, the scan head is assumed to be mounted above the horizontal arm, looking parallel to it, and the laser stripe is assumed to be horizontal.

Discussion. Figure 7 summarizes the sequence of events executed by a typical scanning script. In general, our scanning procedure worked smoothly. Figure 8(a) and 8(b) show our scanner acquiring range and color data on St. Matthew.

Our biggest failure was the lack of an automated method for planning scans. View planning is a well-known computational geometry problem. Recent papers covering the special case of 3D scanning include [Maver93, Pito96]. Most of these methods focus on what might be called the "midgame": given a partially complete 3D model, what are the next best *n* scans to acquire? However, our experience suggests that the "endgame" is more important: given a 3D model that contains a few holes, how can these holes be filled? Since we did not have an automated view planning system, we planned scans by eye - a slow and error-prone process. We often spent hours positioning the gantry in fruitless attempts to fill holes in our model of the David. A view planner might have saved 25% of the man-hours we spent in the museum.

A mixed success was our attempt to use a commercial handheld laser triangulation scanner for hard-to-reach places. (Ours was a 3D Scanners ModelMaker mounted on a Faro Silver Series digitizing arm.) Although its Y-resolution (along the laser stripe) matched that of our Cyberware scanner, its X-resolution depended on how slowly the user swept the stripe across the surface. In practice, hand tremors made it difficult to obtain smooth and monotonic motion. The latter introduced folds into the range data, complicating our post-processing pipeline. Moreover, it was fatiguing to hold the scanner for long periods of time. We used it on St. Matthew and the Slaves, but not on the David or the Medici Chapel statues.

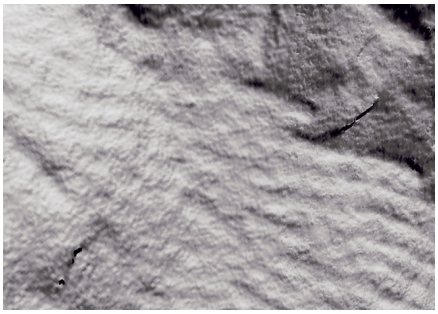
3.1. Safety for the statues

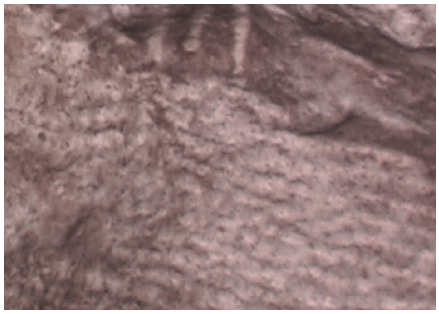
An overriding concern throughout our project was to avoid harming the statues we were digitizing. Laser triangulation is fundamentally a non-contact digitization method; only light touches the artwork. Nevertheless, light and heat can potentially damage art, so their levels must be controlled. Our scanning beam was a 5 mW red semiconductor laser, but its power was spread into a line 20 cm wide at the statue surface, and it moved nearly continuously during scanning. Our white light source was a 250 W incandescent bulb,

(a) Acquiring range data. The laser line sweeps downward at 1 cm per second, acquiring 14,400 points per second. After a sweep, the head pans to the right and performs another sweep.

(b) Acquiring color data. One 1520 x 1144 pixel image is acquired every 3 seconds. Our white spotlight can be seen illuminating the statue. Figures (c) and (d) are taken from the upper neck.

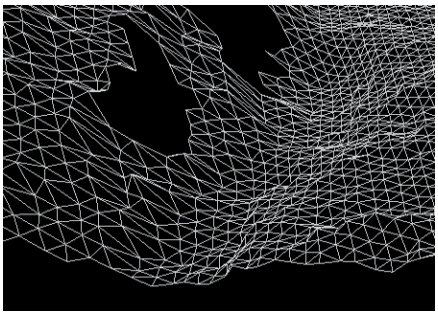
(c) Computer rendering, with artificial lighting and reflectance, of the scan from (a). The spacing between samples is 0.29 mm, and the depth resolution is 50 microns. This area was carved with a gradina (see figure 2).

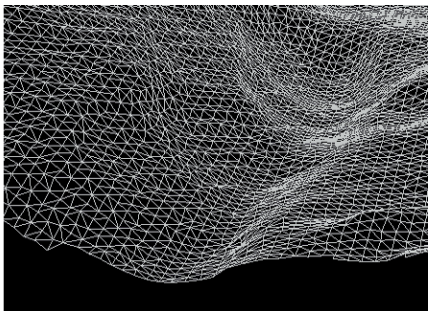




(d) Typical color image acquired in (b). The nominal X,Y pixel spacing is 0.15 mm; the physical spacing is 0.31 mm. The image seems flat because the spotlight is nearly co-axial with the camera.

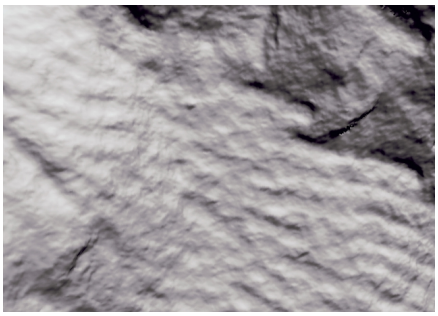
(e) Closeup of (c). This is a single sweep, so it is a regular 2D array (note that the triangles are in rows). The gaps are missing data due to occlusions.





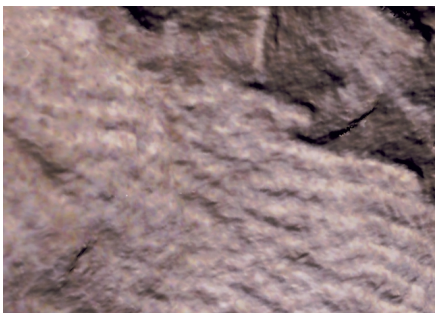
(f) Closeup of a merged mesh combining (c) and other scans. It is irregular since it is the isosurface of a volume. It is made finer than the raw scans to avoid aliasing.

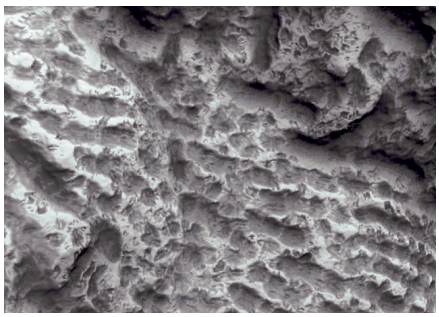
(g) Rendering of merged mesh. It is slightly blurrier than (c), due to miscalibration and misalignment of the scans, but the chisel marks are still clearly visible. Not all of our scans were used here, so some holes remain.



(h) Merged mesh from (g) with per-vertex reflectances blended from (d) and other color images. This view is not lit, so the whitening at the bottom of chisel marks is easy to see, e.g. at the extreme right and top.

(i) Mesh from (h), but lit and rendered like (c) and (g). The whitening, which is easy to see in (h), is masked here by lighting effects. However, it becomes obvious if the light is moved interactively.





(j) A non-photorealistic visualization of the chisel marks. The geometry is the same as (g), but per-vertex colors are computed using accessibility shading [Miller94].

Figure 8: Our scanning procedure and post-processing pipeline. The statue is Michelangelo's unfinished apostle St. Matthew, in the Accademia gallery in Florence. It took us 6 days to scan this statue and another week to post-process it. The full model is shown in figure 9.

but its power was conducted to the scan head through a fiber-optic cable, which effectively blocked heat, and its light was spread into a disk 50cm wide at the statue surface. In both cases, energy deposition on the statue was negligible compared to ambient lighting.

A more serious danger was posed by accidental collisions between the scanner and the statue. Our primary defense against such accidents was our long standoff, but as discussed earlier, we often ended up working closer to the statue than we expected. To reduce the chance of collisions in these situations, we used manual rather than software motor controls to plan scan head motions, we mounted pressure-sensitive motion cutoff switches on the rails of the horizontal and vertical translation tables, and we assigned one member of each scanning team to operate as a spotter. To reduce the chance of damage in case of inadvertent contact, our scan head was encased in foam rubber. Finally, we established rigid operating protocols, we tried (mostly unsuccessfully) to avoid putting ourselves under time pressure, and we tried to get enough sleep.

In the end, we succeeded in scanning nearly around-the-clock for five months without damaging anything. However, we consider ourselves lucky. There is no silver bullet for these safety issues. Scanners employing longer standoffs can reduce the risks during certain phases, but most statues cannot be scanned entirely from outside their convex hull. Eventually, either a scanner or a mirror must be positioned close to the statue.

4. Post-processing

As soon as a scan was acquired, it entered a lengthy post-processing pipeline, whose eventual goal was to produce a polygon mesh with a reflectance value at each mesh vertex.

4.1. Range processing pipeline

Our range processing pipeline consisted of aligning the scans taken from different gantry positions, first interactively, then automatically, merging these scans using a volumetric algorithm, and filling holes using space carving. The output of this pipeline was a watertight irregular triangle mesh.

Aligning scans. The first step was to bring the hundreds of scans for a statue, which were acquired with the gantry in different (untracked) locations, into a common coordinate system. Alignment was done in four steps, as follows:

(1) As each scan completed, an operator interactively aligned it to the existing model by translating and rotating it until it lay in roughly the correct position. Alternatively, the operator could identify three corresponding points on the new scan and any existing scan. This sounds easy, but it's not, especially if the scans are smooth or the overlap with existing scans is slight.

(2) Once the new scan was approximately positioned, the operator chose one existing scan that substantially overlapped the new scan and invoked a modified iterated-closest-points (ICP) algorithm [Besl92, ChenMed92] to refine their alignment. ICP algorithms operate by finding matching points on two meshes, computing the rigid transformation that minimizes the sum of squared distances between these point pairs, and iterating until some convergence criterion is met.

(3) After the statue was completely scanned and the data brought back to the laboratory, a script was run to find every pair of substantially overlapping scans, isolate them from the rest of the data, and run ICP on them to find their mutual alignment. For every such alignment we stored the rigid transformation between the two scans and a list of matching point pairs.

(4) Finally, these lists of matching point pairs were used as constraints in an iterative relaxation algorithm designed to bring the scans into global alignment while evenly spreading out the error among the pairwise alignments [Pulli99].

Previous solutions to the global alignment problem typically generalize pairwise ICP [Bergevin96], that is, each scan is aligned in turn with all other nearby scans, and the process is iterated until it converges. However, this approach requires all scans to be resident in memory, and the repeated use of ICP makes it slow. Our algorithm produces comparable results, and it is more efficient in space and time. However, it is not infinitely scalable; we must at least fit our lists of matching points into memory. Our justification for separating pairwise alignment from the relaxation step is that pairwise alignment of scans finds the best local matching of points that can be found; repeating this step can only make these matches worse.

Merging scans. To reconstruct a surface from a set of aligned range scans, we employed the method described in [Curless96]. Beginning with a dense volumetric grid, we visit each voxel near one or more range surfaces and store a weighted sum of the signed distances from that voxel to each range surface, where distances are taken along lines of sight from the voxel to the range camera. Voxels in front of the surface but on a line of sight are marked as empty, a process known as space carving. All remaining voxels are marked as unseen. We can extract from the volume either an accurate isosurface that corresponds to the observed surface, or we can extend the isosurface to include the boundaries between empty and unseen regions, resulting in a watertight (hole-free) surface. The triangles added in this way can be flagged as "reconstructed" rather than "observed", to differentiate them for scientific purposes. To enable this algorithm to handle large datasets, we broke large volumes into blocks and processed them independently Each block yielded a piece of reconstructed surface, which we stitched together by identifying and merging common vertices between neighboring blocks.

Discussion. How well did our pipeline work? In most cases, it worked well. However, it was time-consuming, occupying about 25% of our man-hours in the museum and several man-days per statue in the laboratory afterwards. Some time could have been saved if we had tracked our gantry; even a rough estimate of the position and orientation of the scan head would have permitted us to bypass step 1 of our alignment procedure. To eliminate step 2 we would have had to track the scan head with 25 micron accuracy from a distance of up to 10 meters[1]. We know of no field-deployable technology with this accuracy. Alternatively, we could have replaced step 1, possibly combined with step 2, with searching globally for an alignment to the existing model using heuristics such as in [Besl92, Hebert95, Johnson97, Zhang99]. The problem with global search methods is that they require more surface overlap than local methods. This is an area for further research.

To our surprise, our software alignment process proved sufficiently robust that we could use it to partially compensate for our calibration problems. Miscalibration manifested itself in our system as misalignments between sweeps and as warpage of individual sweeps. The first effect could be reduced by treating sweeps as

[1] An error of 25 microns in the position of a tooling ball located on the scan head 11 cm from the tilt axis would produce an angular error of 0.013° (see section 2.5) in the tilt of the scan head, leading in turn to an error of 0.25 mm in the position of digitized objects located 112 cm from this axis.

separate scans in steps 3 and 4 of our alignment procedure. This strategy works only if a sweep contains enough fine detail for the ICP algorithm to lock on to. So far, we have used it successfully to align St. Matthew, which contains 104 scans and 2285 sweeps (see figure 9). Aligning this data took 20 hours on one processor of an SGI Onyx2, and volumetric merging took 50 hours. Note that Michelangelo's chisel marks are preserved. For statues without fine geometric detail, like some portions of the David, it might be possible to use color imagery to guide alignment [Pulli97, Bernardini00]. In the future, we envision breaking sweeps into smaller pieces and aligning them separately in order to distribute the warp more uniformly across the statue. We also plan to experiment with non-rigid alignment. Finally, we plan to use theodolite data as constraints, to ensure that the additional degrees of freedom introduced by using these techniques do not cause distortion of the entire model [Beraldin97] [1].

Although our merging process worked well, we were disappointed by the number of holes, some several centimeters in size, that remained after we had done our best job scanning a statue. A sculptor can, using chisels and a drill, carve recesses too deep to scan using any triangulation rangefinder. The David has many such recesses, especially around his hands, hair, and scrotum. We can bridge these gaps using space carving as described earlier, but the bridge surfaces sometimes look objectionable. We believe they can be ameliorated by relaxing their shape, subject to maintaining tangency where they meet observed surfaces, but we have not tried this.

4.2. Color processing pipeline

Our color processing pipeline consisted of correcting our color images geometrically and radiometrically, discarding pixels that were occluded with respect to the camera or light, projecting the remaining pixels onto our merged mesh, and converting color to reflectance. The output of this pipeline was an RGB reflectance triplet for each vertex of the mesh.

For this last step we needed a model of the bidirectional reflectance distribution function (BRDF) at each point on the surface. Clean marble is a dielectric material with varying color (due to veining) and a roughness that depends on the level of polish. For surfaces of this type, Sato and Ikeuchi have used range and color measurements to estimate diffuse, specular, and roughness parameters [Sato97]. Our pipeline is similar to theirs, except that so far we have only attempted to extract diffuse reflectances. To eliminate specular contributions, we discarded observations close to the mirror direction. This approach is robust, and since under a diffuse assumption we have redundant color observations for each point on the surface, the loss of data is inconsequential. Discarding pixels near the mirror direction also circumvents the problem of sensor saturation, although this could be solved using high dynamic range methods [Debevec97] at the cost of acquiring more images.

Mapping color onto the mesh. The first step in our pipeline was to subtract the images acquired with and without the spotlight as described in section 3, thereby compensating for ambient illumination. Next, the difference image was corrected for geometric distortion, chromatic aberration, and radiometric effects using the

[1] A theodolite is a surveying tool that locates points in 3-space by optically sighting them from two calibrated locations on the ground and measuring their direction angles from those locations. Using an optical theodolite with an accuracy of 0.5 mm, we measured several points on most of the statues we scanned.

Figure 9: A rendering of our full-resolution, merged model of Michelangelo's St. Matthew. The original dataset contained 104 scans, 800,000,000 polygons, and 4,000 color images. The model shown here contains 386,488,573 polygons. It still contains some holes, and we have not yet mapped our color data onto it. See figure 8 for closeup views.

calibration data described in section 2.6. We then needed to decide which vertices of our merged mesh saw the color camera and light source. For this purpose we used a hardware-accelerated polygon renderer. We rendered the mesh from both points of view, read back the depth buffers, and for each mesh vertex compared its distance from the camera and light source with the contents of the depth buffers at the appropriate pixel. If the vertex was occluded from the camera or light source, we discarded it. Although it is possible to miss thin occluders using this approach, marble statues generally do not contain any. If a vertex saw both the camera and the light, we projected it to the camera image and sampled the color found there.

Computing reflectance. Once we associated a color with a vertex, we performed an inverse lighting calculation to convert color to reflectance. This calculation is sometimes called "de-shading." Since the exit pupil of our fiber-optic cable's focusing assembly was small relative to its distance from the statue, we treated it as a point source. In this case, the irradiance E on the surface is

$$E = \frac{1}{r^2} I \cos\theta$$

where I is the radiant intensity of the point source, θ is the obliquity of the surface relative to the light, and r is the distance between them. Knowing the irradiance and the reflected radiance $L(\omega_r)$ in the direction ω_r towards the camera, and assuming ideal Lambertian reflection, the reflectance R is

$$R = \frac{L(\omega_r)}{E} = \frac{L(\omega_r)\, r^2}{I \cos\theta}$$

In our case, we did not know the radiant intensity of our light source, and, because we did not measure the absolute sensitivity of our camera, we knew reflected radiance only up to an unknown constant k. However, we knew from our calibration procedure the radiance (up to this same constant) reflected from a white card placed at the standoff distance. Assuming that the card is also ideal Lambertian, its reflectance R_c is

$$R_c = \frac{L_c(\omega_r)\, r^2}{I \cos\theta_c}$$

where $\theta_c = 0$ and $r = 112$ cm. The ratio of the reflectance of the statue surface to the reflectance of the white card is

$$\frac{R}{R_c} = \frac{L(\omega_r)}{L_c(\omega_r)\cos\theta}$$

where r and the unknowns I and k have canceled out. By separately determining the reflectance R_c of our white card relative to a reflectance standard such as Spectralon®, we could estimate the absolute reflectance of the statue surface.

Blending multiple observations. Each mesh vertex usually saw many color images. If the surface were ideal Lambertian, the computed reflectances would agree. However, our surfaces were not, and our observations included noise, miscalibration, and other errors. We therefore needed a rule for blending reflectances together. For this purpose, we computed a confidence for each reflectance based on the following factors:

- obliquity of the surface with respect to the light
- projected area of the surface with respect to the camera
- proximity to the mirror direction, to suppress highlights
- proximity to a silhouette edge with respect to the camera
- proximity to a silhouette edge with respect to the light
- proximity to the edge of the color image

Confidence was made lower near silhouettes to account for blur in the camera lens and penumbrae due to the non-zero extent of the light source. To prevent rapid changes in confidence from triggering sudden switches from one color image to another, we smoothed confidences among neighbors on the mesh. However, to remain conservative we never increased confidence, only decreased it. The last step was to sum the weighted reflectances at each mesh vertex.

Discussion. Although our color processing pipeline produced visually satisfactory results (see figure 10), there are several factors we did not consider. We treated the diffuse reflectance as ideal Lambertian, although it is not [Oren94]. We also ignored interreflections, which may be significant since our statues are light-colored [Yu99]. By modeling the effect of these interreflections, it may be possible to improve our estimate of surface shape [Nayar90]. Similarly, we ignored subsurface scattering [Dorsey99]. However, its contribution is probably minor on dirty, unpolished statues like the David, especially relative to its size.

Finally, in calculating irradiance at each point, we are employing an aggregate surface normal obtained from a 3D scan of the surface. Such de-shading calculations suffer from two problems. First, they are sensitive to noise in the surface normals; however, our use of redundant range images and a volumetric range merging algorithm reduces this noise somewhat. Second, geometric details too fine to scan will not be present in the normal field and will not enter the irradiance calculation. These details consequently manifest themselves as changes to the diffuse reflectance. As a result, care must be taken when using our reflectances for scientific analyses, for example to estimate marble properties. Interestingly, we may have acquired enough redundant color imagery to calculate a view-dependent reflectance texture [Dana99], permitting correct renderings from all viewpoints. This is a topic for future research.

5. Handling large datasets

One significant challenge we faced in this project was the size of our datasets, the largest of which was the David (see table 1). In our post-processing pipeline, we addressed this problem by using an efficient global alignment algorithm and a blocked range image merging algorithm. However, our scanning procedure posed additional challenges. In order to plan scans for a statue, we had to load its 3D model into memory. As each scan completed, we needed to add it quickly to the model. At the time of the project we knew of no modeling package into which we could load a 2-billion polygon model, nor any simplification algorithm that could be reasonably run on a mesh of this size (and we tried several). Therefore, we spent a lot of time writing code for handling large scanned models.

Range images versus polygon meshes. Our first technique was to store our data as range images instead of as polygon meshes. A range image is a 2D array $r(u, v)$ of range values r, some of which might be invalid due to occlusions of the laser or range camera. A range image is, of course, only a special case of a displacement map, a well-known object in computer graphics. Range images are efficient because the u and v coordinates are implicit and the r values have limited precision. To take advantage of this natural compression, we designed a file format containing an array of 16-bit range values and a header with enough information to map these range values to 3D points. To efficiently skip over missing range samples, we run-length encoded this array. If stored as uncompressed 3D vertex coordinates and index lists, the geometry of the David would occupy 36 gigabytes. Stored as run-length encoded

Figure 10: On the left is a photograph of Michelangelo's David. On the right is a rendering made from our model. Constructed at a resolution of 1.0 mm, the model is watertight and contains 4 million polygons. Its surface reflectance was computed from digitized color images as described in section 4.2. The photograph was taken under uncalibrated conditions, so its viewpoint and illumination differ slightly from those of the rendering. The raw data for this part of the statue was acquired at a resolution of 0.29 mm and contained 480 million polygons and 2,000 color images. Using one processor of an SGI Onyx2, it took 12 hours to align merge, and map color onto this model.

range images, it occupies only 2 gigabytes, a savings of 18:1 with no loss in information. More lossless compression could be obtained by using entropy coding, but decoding would be slower.

Range image pyramids. No matter how efficient our storage mechanism is, no current workstation can display 2 billion polygons in real time, so we needed strategies for working at reduced resolution. Fortunately, as long as we stored our models as range images, we could simplify them quickly by subsampling them, thereby creating range image pyramids. Depending on the task, we had several ways of constructing these pyramids, as shown in figure 11.

Lazy evaluation. To save space and time, we constructed range image pyramids from range images on demand, never storing them, and we constructed only those levels requested by the user. To display a range image, or to merge multiple range images as described in section 4.1, it must be converted to 3D points, then to a triangle mesh. We did this lazily as well. Fortunately, the conversion from range samples to 3D points can be done quickly using incremental arithmetic, and since the points appear in scanline order, they can be triangulated quickly by connecting adjacent points. Of course, care must be taken to avoid bridging depth discontinuities.

Viewer based on point rendering. If one only wants to view a 3D model, and not perform geometric operations on it, then it need not be represented polygonally. With this in mind we developed a viewer that combines a multiresolution hierarchy based on bounding spheres with a rendering system based on points [Rusinkiewicz00]. Our viewer preprocesses new meshes in seconds, launches quickly, maintains a constant frame rate regardless of object complexity, yields reasonable image quality during motion, and refines progressively if idle to a high final image quality. With modest hardware acceleration, our viewer permits real-time navigation of scanned models containing hundreds of millions of polygons.

The statue	
height without pedestal	517 cm
surface area	19 m^2
volume	2.2 m^3
weight	5,800 kg
Our raw dataset	
number of polygons	2 billion
number of color images	7,000
losslessly compressed size	32 GB
Other statistics	
total size of scanning team	22 people
staffing in the museum	3 people (on average)
time spent scanning	360 hours over 30 days
man-hours scanning	1,080
man-hours post-processing	1,500 (so far)

Table 1: Some statistics about our scan of Michelangelo's statue of David. The area, volume, and weight of the statue are estimated from our data.

6. Conclusions

We have described a system for digitizing the shape and color of large statues, and a demonstration of this system on the statues of Michelangelo. As computer scientists and technologists, our principal goal in pursuing this project was to push the state-of-the-art in 3D scanning. Our model of Michelangelo's David is two orders of magnitude larger than any existing scanned model. In trying to acquire and post-process this data, we were forced to invent new methods for representing, viewing, aligning, merging, and viewing large 3D models, methods that we have presented here and in related papers [Pulli99, Rusinkiewicz00].

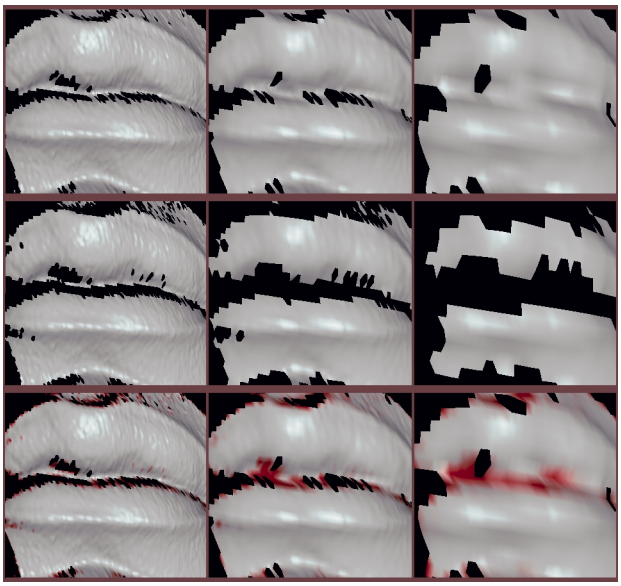


Figure 11: Range image pyramids of David's lips. The fastest way to construct a pyramid is to subsample the range image, without filtering, by successive factors of 2x in *u* and *v* (top row, left to right). To help us find holes, we can also delete any range sample if any of its 4 children at the next finer pyramid level is missing (middle row), making holes visible even at the coarsest resolution. Alternatively, we can redden the parents of missing children in proportion to the fraction of its children that are missing (bottom).

Although it is not the purpose of this paper to tell the story of the Digital Michelangelo Project, we faced many logistical problems that future digitizing projects may encounter, so it is worthwhile briefly enumerating them.

First and foremost, getting permission to scan Michelangelo's statues was a long, delicate, and occasionally painful process. Part of the pain was developing meaningful, equitable, and enforceable intellectual property agreements with the cultural institutions whose artistic patrimony we were digitizing. Since the goals of our project were scientific, our arrangement with the museums was simple and flexible: we are allowed to use and distribute our models and computer renderings for scientific use only. In the event we, or they, desire to use the models commercially, there will be further negotiations and probably the payment of royalties.

Second, we underestimated the difficulty of digitizing under field (non-laboratory) conditions. Shipping 4 tons of equipment to a foreign country, trucking it through narrow streets, and carrying it into historic buildings, was nerve-wracking and expensive. As soon as we moved our equipment into a museum, we became a liability to them - physically, logistically, and legally. During 5 months of scanning, we spent $50,000 hiring museum guards to watch over us, the statues, and the tourists. In the Accademia gallery, we found it necessary to remove the glass security barricades surrounding every statue, including the David. To minimize the time these statues were left unprotected, shifts were long and days off were few. Scanning during museum hours posed additional problems: bumped scanners, color images ruined by tourist flashbulbs, and a constant stream of questions (which we always answered).

Lastly, although most of our plans worked out, there were several disappointments. In particular, we were unable to scan the Pietà (in St. Peter's Basilica) because it was under restoration in preparation for the Year 2000 Jubilee. Scanning this statue may be impractical in any case; it is mounted high on a pedestal, it is sandwiched between an immovable altar in front and a niche behind, and it is highly polished. It may also be too geometrically complicated to scan using laser triangulation technology.

One of the tangible results of the Digital Michelangelo Project is a set of 3D geometric models, one per statue, architectural setting, or map fragment that we scanned. In the months (and maybe years) ahead we will process the data we have collected, and in the case of the Forma Urbis Romae (see figure 16) we will try to assemble the map. Our plan is to make these models freely available to the research community.

7. Acknowledgements

More people helped us than we can possibly acknowledge in this space. Scanning alongside us in the museums were Alana Chan, Kathryn Chinn, Jelena Curless, Unnur Gretarsdottir, Rahul Gupta, Wallace Huang, Dana Katter, Dan Perkel, Domi Piturro, Semira Rahemtulla, Alex Roetter, Marco Tarini, Maisie Tsui, and Daniel Wood. Backing us up in the Stanford Computer Graphics Laboratory à Firenze were John Gerth and Lisa Pacelle. Key contributors to the scanner design were Everett Reed, Jim Meek, Lloyd Addleman, and David Addleman of Cyberware. Our Italian collaborators (and occasional political brushfire fighters) included Cristina Acidini Luchinat, Licia Bertani, Barbara Caputo, Franca Falletti, Laura Ferrea, Eugenio La Rocca, Susanna Le Pera, Roberto Scopigno, and Anna Mura Somella. Our collaborators at the National Research Council of Canada included Marc Rioux, Guy Godin, Angelo Beraldin, and Luc Cournoyer. We received invaluable technical advice throughout the project from Paul Debevec and Matti Auvinen. Figure 13 was rendered by Henrik Wann Jensen. Our sponsors were Stanford University, Interval Research Corporation, and the Paul Allen Foundation for the Arts.

8. References

[Beraldin97] Beraldin, J.-A., Cournoyer, L., Rioux, M., Blais, F., El-Hakim, S.F., Godin, G., "Object model creation from multiple range images: acquisition, calibration, model building and verification," *Proc. 1st Int'l Conf. on 3-D Digital Imaging and Modeling*, IEEE, 1997, pp. 326-333.

[Beraldin99] Beraldin, J.-A., Blais, F., Cournoyer, L., Rioux, M., El-Hakim, S.F., Rodell, R., Bernier, F., Harrison, N., "Digital 3D imaging system for rapid response on remote sites," *Proc. 2nd Int'l Conf. on 3-D Digital Imaging and Modeling*, IEEE, 1999, pp. 34-43.

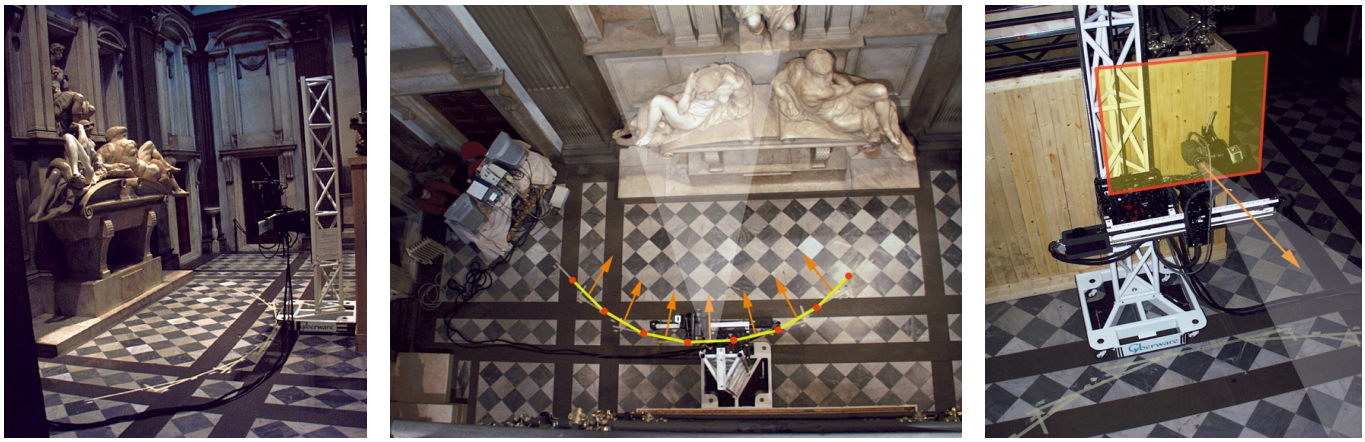
[Bergevin96] Bergevin, R., Soucy, M., Gagnon, H., Laurendeau, D., "Towards a general multi-view registration technique," *IEEE Trans. PAMI*, Vol. 18, No. 5, May, 1996, pp. 540-547.

[Bernardini00] Bernardini, F., Martin, I., Rushmeier, H., "High-Quality Texture Synthesis from Multiple Scans," IBM Research Report RC 21656(97598), February, 2000, IBM Research, Yorktown Heights, NY.

[Besl92] Besl, P., McKay, N., "A Method for Registration of 3-D Shapes," *IEEE Trans. PAMI*, Vol. 14, No. 2, February, 1992, pp. 239-256.

[ChenMed92] Chen, Y., Medioni, G., "Object modeling by registration of multiple range images," *Image and Vision Computing*, Vol. 10, No. 3, April, 1992, pp. 145-155.

[Rocchini99] Rocchini, C., Cignoni, P., Montani, C., Scopigno, R., "Multiple textures stitching and blending on 3D objects," *Proc. 10th Eurographics Rendering Workshop*, Springer-Verlag, 1999, pp. 119-130.



(a) (b) (c)

(d)

Figure 12: In addition to scanning Michelangelo's statue of Night, we also acquired a dense light field of it [Levoy96]. This particular statue is highly polished and thus optically interesting. (a) Our camera "swath" followed a 90° eye-level arc around the statue (taped out on the floor). (b) A plan view of the camera path. It consisted of 7 "light field slabs," shown here by yellow line segments. (c) Each slab was a planar array of 62 x 56 images, with images spaced 12.5mm apart on the plane. The total number of images was thus 24,304. (d) A representative image from each slab. Acquisition was performed using a Sony DKC-ST5 ultra-high resolution 3-CCD digital still camera mounted in place of our scan head. Shooting the light field took 35 hours over 4 consecutive nights. Keeping the illumination constant throughout the shoot was hard. Having a 3D model and dense light field of the same object will provide a unique opportunity for exploring hybrid image-based representations in which these two kinds of data are combined [Wood00].

[Curless96] Curless, B., Levoy, M., "A Volumetric Method for Building Complex Models from Range Images," Proc. SIGGRAPH '96, ACM, 1996, pp. 303-312.

[Dana99] Dana, K.J., Nayer, S.K., Ginneken, B.V, Koenderink, J.J., "Reflectance and Texture of Real-World Surfaces," *ACM Trans. on Graphics*, Vol. 18, No. 1, January 1999, pp. 1-34.

[Debevec97] Debevec, P.E., Malik, J., "Recovering high dynamic range radiance maps from photographs," Proc. SIGGRAPH '97, ACM, 1997, pp. 369-378.

[Dorsey99] Dorsey, J., Edelman, A., Jensen, H.-W., Legakis, J., Pedersen, H.-K., "Modeling and rendering of weathered stone," Proc. SIGGRAPH '99, ACM, 1999, pp. 225-234.

[Giovannini99] Giovannini, P., "Il 'San Matteo' di Michelangelo: analisi delle tracce di lavorazione, studio degli strumenti e osservazioni sulla tecnica di scultura," *Rivista dell'Opificio delle Pietre Dure*, Vol. 10, 1998, pp. 205-228.

[Hebert95] Hebert, M., Ikeuchi, K., Delingette, H., "A Spherical Representation for Recognition of Free-Form Surfaces," *IEEE Trans. PAMI*, Vol. 17, No. 7, July, 1995, pp. 681-690.

[Heikkilä97] Heikkilä J., Silvén O., "A four-step camera calibration procedure with implicit image correction," *Proc. CVPR '97*, IEEE, 1997, pp. 1106-1112.

[Johnson97] Johnson, A.E., Hebert, M., "Surface Registration by Matching Oriented Points," *Proc. 1st Int'l Conf. on 3D Digital Imaging and Modeling*, IEEE, 1997, pp. 121-128.

[Jokinen99] Jokinen, O., "Self-calibration of a light striping system by matching multiple 3-D profile maps," *Proc. 2nd Int'l Conf. on 3-D Digital Imaging and Modeling*, IEEE, 1999, pp. 180-190.

[Levoy96] Levoy, M., Hanrahan, P., "Light Field Rendering," Proc. SIGGRAPH '96, ACM, 1996, pp. 31-42.

[Miller94] Miller, G., "Efficient Algorithms for Local and Global Accessibility Shading," Proc. SIGGRAPH '94, ACM, 1994, *Computer Graphics* pp. 319-326.

[Nayar90] Nayar, S. K., Ikeuchi, K., Kanade, T., "Shape from Interreflections," *Proc. ICCV '90*, IEEE, 1990, pp. 2-11.

[Oren94] Oren, M., Nayar, S.K., "Generalization of Lambert's reflectance model," Proc. SIGGRAPH '94, ACM, 1994, pp. 239-246.

[Maver93] Maver, J., Bajcsy, R., "Occlusions as a Guide for Planning the Next View," *IEEE Trans. PAMI*, Vol. 15, No. 5, May, 1993, pp. 417-433.

[Pito96] Pito, R., "A sensor-based solution to the next best view problem," *Proc. ICPR '96*, 1996, pp. 941-945.

[Pulli97] Pulli, K., *Surface Reconstruction and Display from Range and Color Data*, Ph.D. dissertation, University of Washington, 1997.

[Pulli99] Pulli, K., "Multiview registration for large data sets," *Proc. 2nd Int'l Conf. on 3-D Digital Imaging and Modeling*, IEEE, 1999, pp. 160-168.

[Rushmeier97] Rushmeier, H., Taubin, G., Guéziec, A., "Applying shape from lighting variation to bump map capture," *Proc. 8th Eurographics Rendering Workshop*, Springer-Verlag, 1997, pp. 35-44.

[Rushmeier98] Rushmeier, H., Bernardini, F., Mittleman, J., Taubin, G., "Acquiring input for rendering at appropriate levels of detail: digitizing a Pietà," *Proc. 9th Eurographics Rendering Workshop*, Springer-Verlag, 1998, pp. 81-92.

[Rusinkiewicz00] Rusinkiewicz, S., Levoy, M., "Qsplat: a multiresolution point rendering system for large meshes," Proc. SIGGRAPH 2000 (these proceedings).

[Sato97] Sato, Y., Wheeler, M.D., Ikeuchi, K., "Object shape and reflectance modeling from observation," Proc. SIGGRAPH '97, ACM, 1997, pp. 379-387.

[Tolnay45] De Tolnay, C., "Michelangelo," Princeton University Press, 1945.

[Wood00] Wood, D., Azuma, D., Aldinger, K., Curless, B., Duchamp, T., Salesin, D., Stuetzle, W., "Surface Light Fields for 3D Photography," Proc. SIGGRAPH 2000 (these proceedings).

[Yu99] Yu, Y., Debevec, P., Malik, J., Hawkins, T., "Inverse global illumination: recovering reflectance models of real scenes from photographs," Proc. SIGGRAPH '99, ACM, 1999, pp. 215-224.

[Zhang99] Zhang, D., Hebert, M., "Harmonic maps and their applications in surface matching," *Proc. Computer Vision and Pattern Recognition (CVPR) '99*, IEEE, 1999, pp. 525-530.

Figure 13: A computer rendering made from a 2.0 mm, 8-million polygon model of Michelangelo's David. The raw 0.29 mm dataset contains 2 billion polygons. The veining and reflectance are artificial. The rendering includes physically correct subsurface scattering [Dorsey99], but with arbitrary parameters. The model contains some slightly misaligned scans and is not watertight. Improving this model, and computing correct reflectances for it, is an ongoing project.

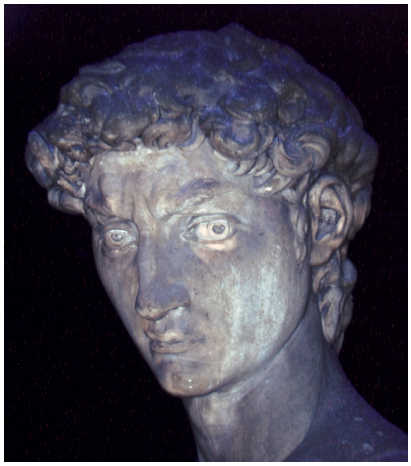


Figure 14: A photograph of David's head taken under ultraviolet light. Once mapped onto our 3D model, this data will show the location of waxes and other organic materials. This may help conservators plan future cleanings of the statue.

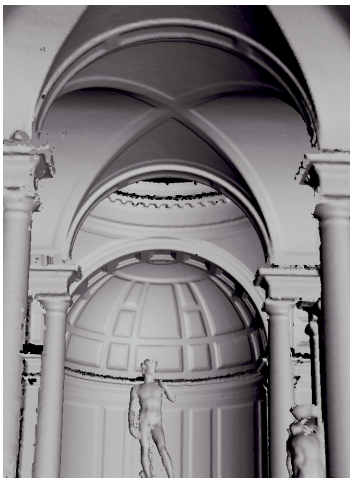


Figure 15: A computer rendering, with artificial reflectance, of a 15-million polygon model of the Accademia gallery in Florence. The data for this model was acquired using a Cyra time-of-flight laser scanner with a Z resolution of about 5 mm. We also acquired a model of the Medici Chapel, and we acquired color data at both sites.

Figure 16: A fragment of the Forma Urbis Romae. During our year in Italy, we scanned all 1,163 fragments of this ancient marble map. This fragment is 68 cm long. Our goal is to piece the map back together by developing compact digital signatures for the geometry of the fragments and searching among these signatures for matches.