A System for Reconstructing Integrated Texture Maps for Large Structures

Chen Xu Athinodoros Georghiades Holly Rushmeier Julie Dorsey

Department of Computer Science
Yale University
New Haven, CT 06520-8285

Abstract

We consider the problem of creating integrated texture maps of large structures scanned with a time-of-flight laser scanner and imaged with a digital camera. The key issue in creating integrated textures is correcting for the spatially varying illumination across the structure. In most cases, the illumination cannot be controlled, and dense spatial estimates of illumination are not possible. We present a system for processing multiple color images into an integrated texture that makes use of the laser scanner return intensity and the captured geometry, together with color balancing and mapping of illumination-corrected images onto the target geometry after filtering into two spatial frequency bands.

1. Introduction

Systems combining laser range scanners and color digital cameras have become popular for obtaining 3D digital models. Models consisting of both geometry and relightable textures (i.e., textures representing reflectances, rather than reflected light) can be used to synthesize realistic images of large objects, or structures, in new settings. Processing a set of range and color images into an accurate, integrated model is still problematic. Color images record intensities that are a function of surface reflectance. However, knowledge of the surface geometry and illumination are needed to extract the reflectance. In this paper, we present a system for reconstructing a seamless texture map representing a spatially dense estimate of diffuse surface reflectance for objects that are too large to capture with controlled lighting.

In previous work, high quality results have been demonstrated for spatially varying bidirectional reflectance distribution functions (BRDF) from systems with a high degree of control over the view and lighting conditions [1]. Typically, tight control over conditions can only be maintained when capturing relatively small objects, on the order of a couple of meters or less in diameter. For scans of large structures outdoors or of interior spaces in use, precise control of illumination is not possible.

Several approaches have been developed to form seamless, integrated texture maps. Techniques include using estimates of illumination that are either spatially or spectrally sparse, calibrating capture devices to obtain response curves not provided by hardware vendors, blending variations in tone and luminance over low spatial frequencies, detecting shadows or highlights in images using ratios of image pixel values, color balancing overlapping images, and using spot measurements of materials. Our system combines these ideas to produce a texture-mapped model from digital images and a time-of-flight range scanner. In addition to combining ideas from other methods, our texture processing system has the following unique features: 1) Segmentation of captured images using laser scanner return intensity into regions of different spectral illumination before color balancing to account for complex lighting variations; 2) Adjustment of the integrated texture map to match local measurements of reflectance.

We begin by reviewing previous work in synthesizing texture maps from multiple digital images. We then describe our data acquisition. Next, we describe the main pipeline of our system that adjusts color images using laser scanner return intensity, color balances adjusted images, and combines them using two passes with different spatial frequencies. We end by demonstrating our results for a large outdoor sculpture and for a vaulted interior space.

2. Previous Work

We assume as a starting point for creating a seamless texture a geometric mesh and color images that have been geometrically registered with respect to it. We also assume that the transfer function of the camera is known, so that the image values can be assumed to be adjusted to be linearly related to the light from the scene. A summary of some of the methods for combining range images and registering 2D images can be found in [2]. Given the input, the key elements in forming the texture are correcting the color images for illumination effects and then combining the results to eliminate the appearance of image-to-image variations.

2.1. Illumination Correction

The red, green, and blue values of a color image are the result of the interaction of the incident illumination, object

geometry, object reflectance and the camera transfer function. When illumination is known reliably, parameters for a surface reflectance function can be estimated using the image values [1].

When illumination cannot be controlled, an alternative is to attempt to measure it at sparse (or even single) spatial locations. Yu and Malik [3] estimated illumination from sun, sky and environment from photographs. Given these estimates and a simple geometric model of a building, "pseudo reflectances" were estimated for the building surface. Debevec et al. [4] measure incident illumination from high-dynamic-range images of spheres of known reflectance. Assuming each surface reflectance is a linear combination of a small number of measured reflectances, a global illumination solution is run for a structure with the measured illumination as the light-source term. The coefficients of the linear combination of reflectances for each location are iteratively adjusted until the global illumination solution matches the captured color images of the structure.

Rather than measure spatially sparse estimates of illumination for multiple wavelengths and directions, an alternative is to use the return intensity provided by many scanners for each measured position. Baribeau et al. [5] used return intensity for lasers at red, green and blue wavelength to estimate surface reflectance. Umeda et al. [6] modified this approach to use the return intensity at a single wavelength. After a calibration step to establish the non-obvious relationship between the return intensity and albedo for their triangulation scanner, the spatial map of the return intensity was used to adjust one channel—in their case the red channel, for a laser scanner operating with a red laser—of a color image that captured surface variations at a much higher spatial resolution. Assuming that the color balance of the capture camera adjusted for the spectrum of the incident light, the ratio used to adjust the red channel from illuminated to reflectance estimate was also used to adjust the green and blue channels.

Sagawa et al. used corrected return intensity from a time-of-flight (TOF) laser scanner as a monochromatic estimate of reflectance at each vertex in a model [7]. Images of the rendered reflectance per vertex model were used to register color images to the model, but not to modify those color images. In using the return intensity from a TOF scanner, the relationship between depth, albedo and orientation needs to be carefully calibrated. As demonstrated in studies to evaluate TOF scanner accuracy, such as [8] and [9], the processing of the return signal to produce an accurate depth produces a return intensity that is not predicted by models of light reflection.

One problem with assuming that the camera color balance corrects the incident illumination to white is that the color of incident illumination spatially varies. Kawakami et al. [10] recently studied this problem for outdoor scenes. They developed a method that attempts to correct chromaticity variations in images due to illumination chromaticity variations using only image data. A complete model accounting for these chromaticity variations was developed, rather than just applying the ratio for one channel to the other two.

A final alternative is not to correct for illumination at all, but to rely on methods of adjusting overlapping images for consistency to produce an acceptable texture map [11, 12, 13].

2.2. Combining Images

Even if the original color images are adjusted to remove illumination effects, image-to-image variations will persist. Rushmeier and Bernardini [14] proposed a global color balance scheme to adjust all overlapping images to make corresponding points match as closely as possible. Agathos and Fisher [11] proposed a similar approach, but allowed crosstalk between the color channels.

Uniform adjustments across images improve balance, but seams may be apparent, where different parts of the object are texture mapped from different sources. Bernardini et al. [15] avoid such seams by blending all possible color images that project on a point. This requires precise alignment of images to avoid ghosting artifacts.

An alternative adopted by Callieri et al. [12] is to compute a color difference correction that would equalize the difference between neighboring patches, and then spatially diffuse that correction through the texture map. The idea of combining a low-spatial frequency correction factor with texture containing high spatial frequency features is similar to Burt-Adelson image mosaicking [16]. Inspired by the work of Burt and Adelson, Baumberg [17] developed a method of combining low and high spatial frequency bands of color images to average out color variations, while capturing detailed high-frequency features without ghosting effects.

Troccoli and Allen [13] introduced a method for the adjustment of overlapping color images projected on a single geometry, so that they are consistent with each other. This is performed without removing all illumination effects, but rather using relighting. To adjust an image B to appear to be illuminated as image A, the ratio of the image values for different surface orientations is learned in the overlap region and then applied to the rest of B.

Another approach to enforcing consistency across seams is to view texture map construction as a classification of each pixel as one of a small number of material types [1, 4, 18]. Pixels that might be adjusted to slightly different reflectance values in different images are likely to be assigned the same material. This approach works well when images are well aligned and the number of materials is small.

In our system, similar to the work by Umeda et al. we use the return intensity to adjust the color image values. This allows us to deal with illumination levels that vary across the

scene. Recognizing that the color of illumination can vary within an image, we introduce a novel technique to use the return intensities to segment the images into regions that can have separate color adjustments. Using the previous method by Rushmeier and Bernardini, we perform a global color balance on the segmented images. We adopt the approach used by Baumberg combining the corrected color images in a way that more closely follows the Burt-Adelson approach for 2D color imaging. Finally, we introduce a novel step that uses of spot measurements to adjust the texture.

3. Data Acquistion

Our method for integrating texture maps assumes a known geometric mesh, aligned color images, and a model for how scanner return intensities relate to surface albedo. In this section, we describe how we acquire and process the preliminary data.

Mesh and Image Capture and Processing: To obtain the mesh and aligned images, we use existing hardware and software. We obtain range images with a Cyrax 2500 TOF range scanner. We align and integrate the scans into a single triangle mesh using the MeshLab [19] software, and segment the mesh into patches for a texture atlas. We obtain color images with an Olympus C8080WZ digital camera. For high dynamic range (HDR) images, we acquire images at multiple exposures. We process these into a single image using HDRShop [20]. We align the color images to the geometry by picking corresponding points on the 2D images and 3D model, and then computing the camera model using Tsai's algorithm [21].

Relating the Return Intensity to Albedo: Our method requires a model relating the return intensity to the diffuse albedo of the scanned surface at the laser light frequency (532 nm for the Cyrax). We treat the Cyrax 2500 scanner as a "black box," and perform controlled experiments to determine a model of the return intensity with respect to the distance of the reflecting surface, z, the angle of incidence of the laser beam, θ, and the diffuse albedo α. By constraining this model to be invertible, we can use it to infer the albedo.

To determine the dependence with respect to albedo, we scan the GretagMacbeth ColorChecker chart at different depths and orientations. The chart provides calibrated reflectance values. From these, we can build a model for the return intensity, R. Although the parametric model we determine is for the particular scanner we used, other scanners could be used within our framework.

From our measurements we have noted that the return intensity peaked at around $17 - 20m$, as also observed by Sagawa et al. [7], who used an earlier Cyrax model. We also observed a relatively weak dependence to the incidence angle, θ. Unlike Sagawa, et al. [7], we explicitly measure the dependence of the return intensity with respect to the cosine of the incident angle.

In designing our return intensity model, we accounted for the above observations, along with the requirement that the model be invertible. We use the following parametric model:

$$R = f(\alpha, z, \theta) = f(\alpha, z)g(\theta), \tag{1}$$

where,

$$g(\theta) = a_1 \cos\theta + a_2, \tag{2}$$

with $a_1 + a_2 = 1$, and

$$f(\alpha, z) = p_1(z)\alpha + p_2(z), \tag{3}$$

where $p_1(z)$ and $p_2(z)$ are quadratic functions of depth, z, each with three unknown parameters.

We first determine the parameters of $g(\theta)$, using measurements at a fixed distance of a cylindrical surface with constant albedo. The parameters a_1 and a_2 are determined using least squares and then are normalized such that their sum is 1. Subsequently, we can determine the six parameters describing $p_1(z)$ and $p_2(z)$. We scanned the MacBeth chart at four different depths ($5m$, $9m$, $17m$, and $33m$), and at four incidence angles ($0°$, $30°$, $45°$, and $60°$). The average, measured return intensity of each of the 24 patches of the MacBeth chart at a particular distance and incidence angle provides a single constraint to the following equation:

$$p_1(z)\alpha + p_2(z) = \frac{R}{g(\theta)}, \tag{4}$$

where $g(\theta)$ is known, having determined its parameters before. Given the 384 measurements, we can then reliably determine the six parameters of $p_1(z)$ and $p_2(z)$, again using least-squares. With this model at hand, we can estimate the albedo of the surface, which can be used to perform color correction of the captured photographs.

4. Proposed System

Our system for reconstructing an integrated texture map consists of four steps: 1) "Unlight" the color images using the albedo estimated from the return intensity; 2) Balance the colors between segmented versions of the unlit images; 3) Combine all of the color balanced images, and 4) Adjust for spot measurements of reflectance.

4.1. "Unlighting" Captured Images

With the model presented in Section 3, we can estimate the albedos for the sampling points in the range images from the return intensity. Similar to [6], our goal is to adjust the green channel (corresponding to the laser wavelength of 532 nm) to match these albedo estimates. The resolution of the range images is much lower than the color images. In our system, the resolution of range images will be from 300x300 to 1000x1000, and for color images it is 3264x2448. Our approach is to calculate the ratio of the low-frequency part

of the green channel of the HDR images to the estimated albedo, and use this ratio to correct all color channels.

Specifically, we begin by estimating the albedo, α, at each point in the range image using Eq. 4. Next, we compute a low-frequency version of the green channel of the HDR images by Gaussian filtering and store the result as $\bar{g}$. We then project the points from the range image into the HDR image. Because the projection of points in the range image is not evenly distributed on the HDR image plane, for each HDR image pixel, we interpolate values of α from the range image. We use the 3D point projections within a window the size of the Gaussian filter we used on the green channel, and weigh the values using inverse pixel distance. Given the values of α and $\bar{g}$ at each pixel, we form the ratio $\bar{g}/\alpha$. At this step, we just simply multiply all the color channels by this ratio for each pixel. Unlike [6], we save the ratio for every pixel for further processing.

4.2. Color Balancing

To adjust all overlapping images to have the same color and luminance, we use a variation of the color balance method described in [14]. We densely sample the model to find corresponding points in the various images. We form a set of linear equations for an adjustment value for the red, green and blue channels in each image. Our dense sampling gives us an overconstrained system, and a least squares solution is found.

Because of the variation of the spectrum of light in the scenes we consider, a single adjustment value per image is not adequate. We note that the ratios computed in 4.1 indicate the variation of the overall light level in the scene. Regions of high illumination will generally be more influenced by the spectrum of the direct light source. Regions of low illumination will be more influenced by the spectrum of indirect light. With the recognition that the light level indicated in the ratio image indicates regions of potentially different illumination spectra, we segment each image based on the ratio image values. This segmentation is performed using the k-means clustering algorithm on the ratio image values. Rather than running the color balance on full images, we run the color balance procedure treating each image segment independently.

4.3. Combining Images

After color-balancing the return-intensity adjusted images, we need to determine the texture of the individual, near-flat, geometric patches to form the complete texture map of the object. Each individual patch may be visible from multiple camera views, and hence would project to multiple adjusted images. To combine the overlapping images we use a technique similar to [17]

The color-balanced, adjusted images are decomposed into two frequency bands, one low-pass and the other high-pass. This is achieved using Gaussian filtering, where the width is of the same order as the misalignment between the color images and their corresponding scans. In determining the final texture of a patch, we first construct its low-pass and high-pass versions and then simply combine them. In our approach, the former is the weighted average of the low-pass versions of the color images in which a patch projects. The weights are a combined function of the angle of the normal of a point on the geometry with respect to the camera viewing direction, the distance of that point from the camera center, and its distance from an occluding boundary. The weighted averaging of low-pass images prevents the appearance of seams. The high-pass version of the patch texture is determined by choosing the high-pass texture colors from the image with the highest associated weight. Combining the "best" high-pass texture with the weighted averaged low-pass colors, leads to a global texture map with detailed features, but with minimal ghosting and very few seams.

4.4. Color Adjustment to Local Measurement

Our aim is to make the texture as close to the true diffuse color as possible. Our images to this point have been processed as though the camera color balance was accurate over most of each image. This is generally not the case. We can measure the diffuse color with an optical device. In our system we use the XRiteColor Digital SwatchBook. For each sample, the SwatchBook can output a reflectance spectrum at wavelengths from 400nm to 700nm in 10nm steps. We convert this spectrum into RGB color space. The number of points we can measure is quite limited however, and possibly there will be fairly large area that we cannot reach and measure. Our problem is to use limited sample data to adjust the entire texture.

We note that after processing there are residual irregularities in the texture that are functions of both location and the specific color value. To adjust the texture values then we account for the distance from each location measurement in both physical 3D space and in color space. We compute a correction for each texture pixel as a weighted average of the spot measurements, where the weights are the inverse of the product of Euclidean distance and difference in the color space. We calculate the ratio at each color channel, and to prevent noise, we calculate the color difference using Gaussian-filtered texture.

5. Results and Conclusion

We present results for two large objects captured under challenging lighting conditions. Figure 1 shows data for an outdoor sculpture of a dinosaur. Nine photographs were obtained under natural lighting on different days. Sources of lighting variations include direct sun, self-shadows, self-interreflections between the sculpture and its pedestal, shadowing by a neighboring building, skylight and reflections from the ground which was snow covered in one case, and

was green grass in another. The figure shows captured geometry, return intensities, color images, return intensity adjusted color images, the albedo ratio images, and color balanced images. The top row of Fig. 3 shows the effect of the within image color balancing. The camera color white balance did not adequately adjust for the blue tone of the illumination. Since we could not access the sculpture with the X-Rite device, final texture is adjusted based on images of the sculpture with a grey scale card. Renderings of the final model are shown in Fig. 6.

Figure 2 shows data for an indoor Guastivino vault. When we captured the 23 HDR photographs used (5 exposures per image), the vault it was illuminated by small halogen lights mounted a few inches from the vaulting, fluorescent light from neighboring hallways and natural daylight from several windows. The vault ceiling geometry is shown in Fig. 2a. Figure 2b shows the same series of processing steps as shown for the dinosaur in Figure 1. The second row of Fig. 3 shows the effect of the within image color balancing for the vault. Figure 4 shows the result of the two-frequency band image combination, compared to using all images or choosing one only. The red tint of the processed unlit images show that the camera color balance did a poor job adjusting to white. Figure 5 shows the locations of spot measurements obtained with the X-Rite device, and the true color is shown on the measurement locations. Renderings of the final model are shown in Fig. 7. Some color variations persist. The overly green results in the areas near the halogen lights are the result of the HDR images failing to capture the complete color content.

Conclusion: We have described and demonstrated a system for reconstructing seamless texture maps for large structures scanned under uncontrolled illumination. We use the scanner return intensity to estimate albedo, and segment images into areas of different illumination. We use a two-spatial frequency approach to combine the images, and then adjust the resulting texture map with spot measurements of reflectance.

We thank the Peabody Museum at Yale for permission to scan the Torosaurus sculpture used our examples. This work was funded by NSF grant CCF-0528204.

References

[1] Hendrik P. A. Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel, "Image-based reconstruction of spatial appearance and geometric detail," *ACM Trans. Graph.*, vol. 22, no. 2, pp. 234–257, 2003.

[2] F. Bernardini and H. Rushmeier, "The 3D model acquisition pipeline," *Computer Graphics Forum*, vol. 21, no. 2, pp. 149–172, 2002.

[3] Yizhou Yu and Jitendra Malik, "Recovering photometric properties of architectural scenes from photographs," in *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, 1998, pp. 207–217.

[4] Paul Debevec et al., "Estimating surface reflectance properties of a complex scene under captured natural illumination," Tech. Rep., USC ICT Technical Report ICT-TR-06.2004, 2004.

[5] R. Baribeau, M. Rioux, and G. Godin, "Color reflectance modeling using a polychromatic laser range sensor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 263–269, 1992.

[6] Kazunori Umeda, Megumi Shinozaki, Guy Godin, and Marc Rioux, "Correction of color information of a 3D model using a range intensity image," in *3DIM '05: Proceedings of the Fifth International Conference on 3-D Digital Imaging and Modeling*, 2005, pp. 229–236.

[7] Ryusuke Sagawa, Ko Nishino, and Katsushi Ikeuchi, "Adaptively merging large-scale range data with reflectance properties," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 392–405, 2005.

[8] Geraldine Cheok, Stefan Leigh, and Andrew Rukhin, "Calibration experiments of a laser scanner," Tech. Rep., NISTIR 6922, 2002.

[9] J. Clark and S. Robson, "Accuracy of measurements made with a Cyrax 2500 laser scanner against surfaces of known colour," in *XXth ISPRS Congress, Istanbul Turkey*, 2004, pp. 1031–6.

[10] Rei Kawakami, Katsushi Ikeuchi, and Robby T. Tan, "Consistent surface color for texturing large objects in outdoor scenes," in *ICCV*, 2005, pp. 1200–1207.

[11] Alexander Agathos and Robert B. Fisher, "Colour texture fusion of multiple range images.," in *3DIM*, 2003, pp. 139–146.

[12] Marco Callieri, Paolo Cignoni, Claudio Rocchini, and Roberto Scopigno, "Weaver, an automatic texture builder," in *3DPVT*, 2002, pp. 562–567.

[13] Alejandro Troccoli and Peter K. Allen, "Relighting acquired models of outdoor scenes," in *3DIM '05: Proceedings of the Fifth International Conference on 3-D Digital Imaging and Modeling*, 2005, pp. 245–252.

[14] Holly E. Rushmeier and Fausto Bernardini, "Computing consistent normals and colors from photometric data," in *3DIM*, 1999, pp. 99–108.

[15] Fausto Bernardini, Ioana M. Martin, and Holly E. Rushmeier, "High-quality texture reconstruction from multiple scans," *IEEE Trans. Vis. Comput. Graph.*, vol. 7, no. 4, pp. 318–332, 2001.

[16] Peter J. Burt and Edward H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Trans. Graph.*, vol. 2, no. 4, pp. 217–236, 1983.

[17] A. Baumberg, "Blending images for texturing 3D models," in *British Machine Vision Conference*, 2002, pp. 404–413.

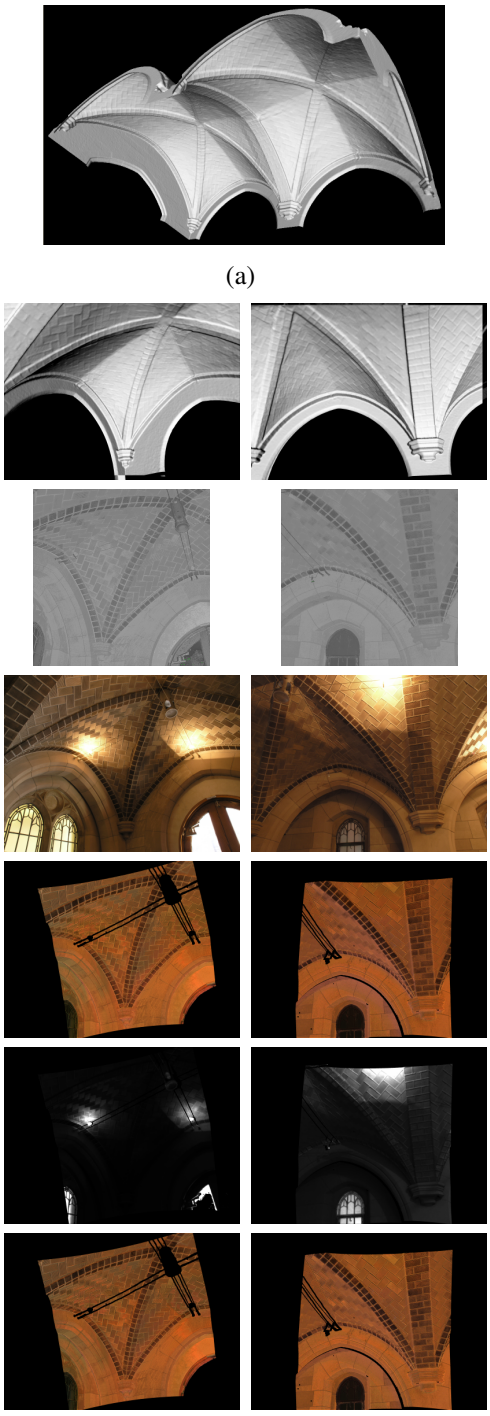
[18] Hiroki Unten and Katsushi Ikeuchi, "Color alignment in texture mapping of images under point light source and general lighting condition," in *CVPR (1)*, 2004, pp. 234–239.

[19] Visual Computing Lab ISTI-CNR, "Meshlab," http://meshlab.sourceforge.net/.

[20] "HDR shop," http://gl.ict.usc.edu/HDRShop/.

[21] R.Y. Tsai, "An efficient and accurate camera calibration technique for 3-D machine vision," in *Computer Vision and Pattern Recognition*, 1986, pp. 364–374.

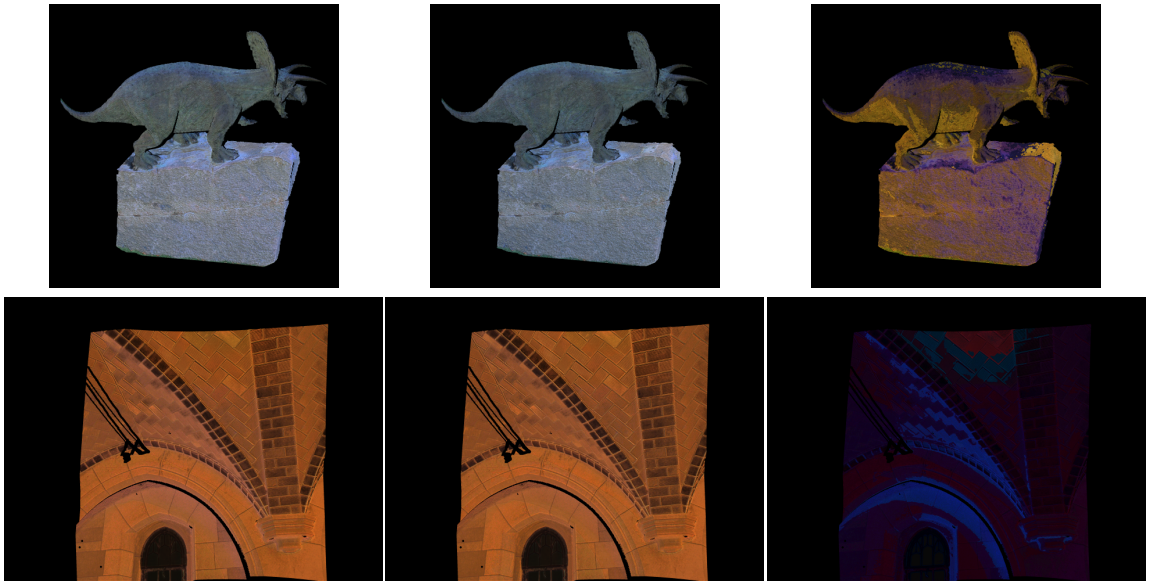
Figure 1. ROW 1: Scanned geometry; ROW 2: Laser scanner return intensity (green implies invalid region); ROW 3: Captured images; ROW 4: Corrected images using return intensity; ROW 5: Ratio of captured images to return intensity; ROW 6: Color-corrected images using our segmentation-based algorithm.



(a)

(b)

Figure 2. (a) The complete geometry model of a vaulted ceiling. (b) ROW 1: Scanned geometry; ROW 2: Laser scanner return intensity; ROW 3: Captured images; ROW 4: Corrected images using return intensity; ROW 5: Ratio of captured images to return intensity; ROW 6: Color-corrected images using our segmentation-based algorithm.



(a) (b) (c)

Figure 3. Color correction of the return-intensity-adjusted images of the dinosaur sculpture (top row), and of the vaulted ceiling (bottom row). (a) The color-corrected images, where only one set of color coefficients was used for the whole image. (b) The result using our approach, where each image is segmented into seven regions depending on the ratio of the original color image and the scanner return intensity. Note that the colors in these images are more uniform, indicating that the effect of the color of the illuminant has been reduced. (c) The difference between the left and middle column images. Note how our approach adjusts the colors differently in regions of different illumination. (The image differences have been scaled by 8 times in order to be more easily visible. They are of the order of 10% of the dynamic range of the images.)

(a) (b) (c)

Figure 4. Comparison of our two-frequency-band, texture-mapping approach to other image mixing methods: (a) Texture of part of the vaulted ceiling created with our approach, where color-corrected images are separated into two frequency bands. The low-pass version of the images is mixed using weighted averaging to form the low-frequency component of the final texture map. The high-frequency component of the texture map is formed by the high-pass version of the image with the highest weight. Note how the resulting texture created by our approach preserves details without blurring or ghosting and without creating seams. (b) The output texture created using weighted-averaging to mix the original color-corrected images. Note the relatively blurred texture and the "ghosting" effects. (c) The resulting texture when it is simply created by choosing the original image with the highest weight. Note the seams and breaks in its appearance.

Figure 5. An image showing where on the surface of the vaulted ceiling the diffuse color measurements were made using the XRiteColor Digital SwatchBook. The dots on the surface show both the position and the measured color at different parts of the geometry. Note the differences between the measured and texture-map colors (as also indicated in the inset, which shows how a measurement was made). We use these spot measurements to correct the final texture map.

Figure 6. Three rendered images of the color-corrected, texture-mapped dinosaur sculpture. (Note that the black regions on the object are due to a lack of image or geometric information. Those regions were inaccessible during the data acquisition process.)

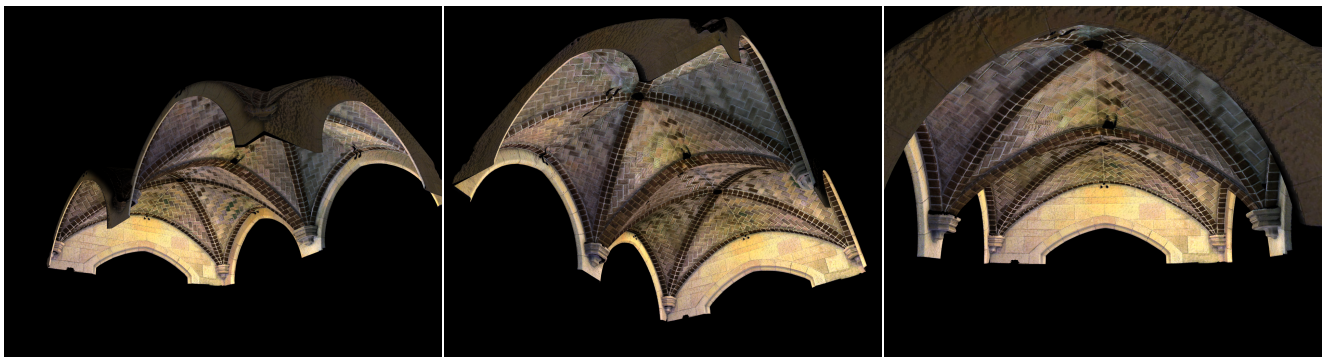


Figure 7. Three rendered images of the color-corrected, texture-mapped, vaulted ceiling. (Note that the black regions on the object are due to a lack of image or geometric information. Those regions were inaccessible during the data acquisition process.)