Making One Object Look Like Another: Controlling Appearance Using a Projector-Camera System*

Michael D. Grossberg, Harish Peri, Shree K. Nayar, and Peter N. Belhumeur

Department of Computer Science, Columbia University
New York, New York 10027
E-mail: { mdog, hsp7, nayar, belhumeur}@cs.columbia.edu

Abstract

We present a method for controlling the appearance of an arbitrary 3D object using a projector and a camera. Our goal is to make one object look like another by projecting a carefully determined compensation image *onto the object. The determination of the appropriate compensation image requires accounting for spatial variation in the object's reflectance, the effects of environmental lighting, and the spectral responses, spatially varying fall-offs, and non-linear responses in the projector-camera system. Addressing each of these effects, we present a compensation method which calls for the estimation of only a small number of parameters, as part of a novel off-line radiometric calibration. This calibration is accomplished by projecting and acquiring a minimal set of 6 images, irrespective of the object. Results of the calibration are then used on-line to compensate each input image prior to projection. Several experimental results are shown that demonstrate the ability of this method to control the appearance of everyday objects. Our method has direct applications in several areas including smart environments, product design and presentation, adaptive camouflages, interactive education and entertainment.*

1 Introduction

An object's color and texture are critical to its appearance. The ability to control these attributes has powerful implications. For instance, objects can be camouflaged by blending them into the background. Existing patterns or colors on products can be made invisible in order to preview new designs. We know that changing an object's intrinsic color or texture cannot be accomplished without physically modifying the object. However, the object's apparent color and texture can be controlled using the illumination. Precise control of an object's illumination can be achieved using existing projection display technology.

Projection display technology has dramatically advanced in the past decade. Projectors are able to display images with high spatial resolution and dynamic range. Innovations in manufacturing have resulted in projectors which are both compact and inexpensive. All these developments make it practical to use projectors as components of complex display systems. Such systems integrate projectors, cameras, and computational devices for a wide range of applications. For example, multiple projector systems have been used to create large seamless high resolution displays [8, 6, 4, 7], to produce high quality images that have several component images [9], to create immersive environments [2, 14], and to eliminate shadows cast on a screen [1]. Projectors have also been used to change the appearance of a 3D object. For instance, a Lambertian white object can be made to appear colored or textured [15]. If the observer's position is known, a matte white object can be made to appear specular or even transparent [5].

The goal of this work is to provide a method to *control* the appearance of a colored and textured 3D object. We achieve this with a system that uses a projector and a camera. The user specifies the object's desired appearance from the camera's viewpoint. Success in controlling this appearance is defined by how similar the image acquired by the camera is to the desired appearance. To generate the desired appearance we project an appropriate image onto the object. Any image projected onto the object is modulated by the spatially varying reflectance properties of the object's surface. Humans are very sensitive to such modulations. Hence, in determining the image to project, we must account for the contribution of the spatially varying reflectance to the appearance of the object. A major challenge is to efficiently account for this contribution at every pixel, along with the other components that affect the measured appearance. These components are the environmental lighting on the scene, and the spectral responses, spatially varying fall-offs, and non-linearities of the projector-camera system.

In [18], a method was presented that color corrects an image when it is projected on a flat screen with homogeneous color. Controlling the appearance of a surface that has spatially varying color and texture requires a more sophisticated radiometric model. A model that takes us all the way from a display image to the image captured by the camera was presented in [12]. This model was used to correct for markings or imperfections on a flat projection screen or surface.

This paper describes an improved radiometric model which is dramatically less complex. While the previous model in [12] requires over 500MB to store the parameters, our model requires just 7.7MB. This is critical when using the model for real-time compensation because this data must be accessed for each image displayed. We also provide a simple and effective off-line radiometric calibration method that recovers the model without knowing the radiometric parameters of the individual components of the system.[1] A major contribution of our work

*This research was conducted at the Columbia Vision and Graphics Center in the Computer Science Department at Columbia University. It was funded by an NSF ITR Grant (IIS-00-85864).

[1] Note the ability to control the appearance of a surface using a projector

Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)
1063-6919/04 $20.00 © 2004 **IEEE**

is to show that this calibration can be done very efficiently and only requires projecting and capturing 6 images. This number is the minimum needed and is a significant improvement over the 260 images required in [12].

Results of the radiometric calibration are used on-line to compensate each display image prior to projection. The compensation step is simple and hence can be done at frame-rate. By displaying compensated images on the object, we can control the object's appearance as desired. We show several experimental results that demonstrate our method's ability to control the appearance of brightly colored and highly textured 3D objects.

Our method has direct applications in many areas. Smart environments can be enhanced by allowing the display of readable text on colored and textured everyday objects. Our method can be applied to product design, allowing designers to control the appearance of existing products in order to preview new designs. Classes can become more interactive and communicative. By making existing objects such as globes or posters take on different, dynamically changing appearances, it is possible to present educational material in a more compelling manner.

2 Model of a Projector-Camera System

Fig. 1 shows the dataflow pipeline for our projector-camera system. The projector displays a user-selected image onto the screen. The resulting appearance is captured by the camera. Note that the camera side of this system (bottom part of the pipeline) is used only during off-line calibration. After calibration, our system computes the display image $\mathcal{I}$ which, when projected on the scene, will produce a desired camera image $\mathcal{M}$.

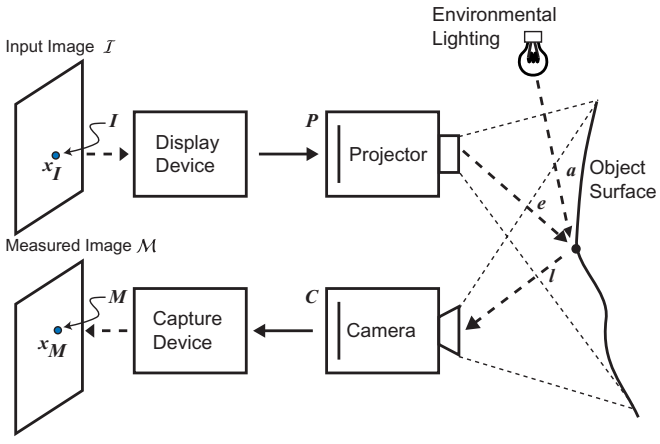


Figure 1: The dataflow pipeline for a projector-camera system.

The projector-camera system we have used is shown in Fig. 2. The projector is a Sony SVGA VPL-CS5. It has a native resolution of 800×600 pixels. We have constructed a variety of textured objects and backgrounds (an example scene is shown in Fig. 2) for testing our algorithms. The camera we have used is a Sony DXC 950 Power HAD model with a resolution of 640×480 pixels. Images are sent to the projector via a ATI Radeon VE display card and images from the camera are captured using a Matrox Meteor II frame-grabber. Our algorithms are run on a Dell Precision 330 computer with a Pentium P4 (1.8 GHz) processor and 1 Gb of memory.

depends on the dynamic range of the projector and the surface reflectance having a significant diffuse component.

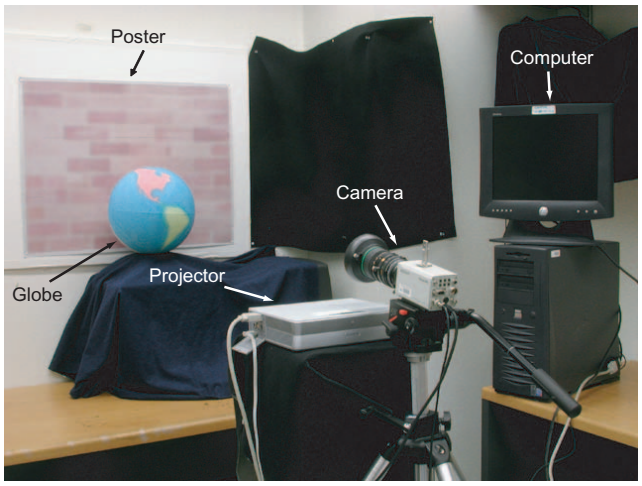


Figure 2: The projector-camera system we have used in our experiments. Images are displayed using a Sony SVGA VPL-CS5 projector and captured using a Sony DXC 950 Power HAD camera. The calibration and compensation algorithms are run on a Dell Precision 330 computer that uses an ATI Radeon VE card to output images and a Matrox Meteor II frame-grabber to capture images.

2.1 Geometric Mapping

In this work, our focus is on the radiometric aspect of controlling appearance. However, to achieve this we need to know the geometric mapping between points in the projector plane $\mathbf{x}_I$ and the image plane $\mathbf{x}_M$, as shown in Fig. 1. We note that a projector-camera system can be designed such that the geometric mapping between the displayed and acquired images is fixed and is unaffected by the location or the shape of the scene. This is achieved by making the optics of the projection and the imaging systems coaxial.[2]

Although our system is not a coaxial one, the camera is placed close to the projector. Hence, the geometric mapping between the camera and projector can be modeled with a piecewise 2D mapping. To determine this 2D mapping, we first obtain a sampling of corresponding points by projecting 1024 uniformly spaced square patches on the scene and acquire the corresponding images. The patches are efficiently scanned using binary coding; the correspondences for the centers of all 1024 patches are obtained using just 10 projected images.

We model the mapping between the two coordinate frames using piecewise second-order polynomials. We divide the camera and projector planes into small regions. In each region, we

[2]For instance, the same lens can be used by the projector and the camera by means of a beam-splitter placed behind the lens. The use of coaxial optics has the added benefit that all points that are visible to the projector are also visible to the camera; there is no possibility of occlusion.

fit a single second-order polynomial using the corresponding points previously obtained. This piecewise approach produces very accurate results as it can accommodate for geometric distortions that may be caused by the optics of the projector and the camera as well as parallax due to the shape of the scene. The final geometric mappings (both ways) between the projector and the camera are stored as look-up tables; each point in one domain is used as an index to obtain the corresponding point in the other. Details are described in [13].

2.2 Radiometric Model

Having obtained a geometric mapping between the projector and camera image planes, we focus on the radiometric aspect of the problem. Again, consider the projector-camera pipeline in Fig. 1. Note that each of the devices in our system has its own unknown, non-linear response. Since the process of radiometric compensation requires us to invert these responses, we will assume that the individual responses are monotonic. This is a reasonable assumption as all the devices are expected to increase in output with input.

We begin with a model for a single point in the scene. The projector and camera may have multiple color channels. To begin, we assume that the projector has only a single channel denoted by K. The input brightness value I_K (scalar value) is passed to the display device. It is mapped by the radiometric response function[3] p_K of the electronics of the projector to a projector brightness

$$P_K = p_K(I_K)\,. \tag{1}$$

This projector brightness is modulated by the spectral response $w_K(\lambda)$ of the projector channel, where λ is wavelength. The projector illuminates the scene point to contribute a scene irradiance of

$$e_K(\lambda) = P_K\, w_K(\lambda)\,. \tag{2}$$

The total irradiance on the scene point is given by the sum of the projector illumination and the environmental lighting. Let the irradiance due to the environmental lighting be $a(\lambda)$. Let the spectral reflectance of the irradiated scene point be $s(\lambda)$ in the viewing direction of the camera. Then, the radiance of the scene point in the direction of the camera can be written as

$$l_K(\lambda) = \big(a(\lambda) \;+\; P_K\, w_K(\lambda)\big)\, s(\lambda)\,. \tag{3}$$

Now, let us assume that the radiance of a scene point is being measured by a camera with a single spectral channel J with the spectral response $q_J(\lambda)$. Then, the irradiance measured by the camera's sensor is

$$C_J = \int \big(a(\lambda) \;+ P_K\, w_K(\lambda)\big)\, s(\lambda)\, q_J(\lambda)\, d\lambda\,. \tag{4}$$

This irradiance is processed by the electronics of the camera to produce a camera output. The camera output is mapped to the final measured brightness by the capture device (framegrabber). The relationship between the camera irradiance C_J and the measured brightness M_J is described by a non-linear response m_j,

$$M_J = m_J(C_J)\,. \tag{5}$$

The recovery of the response m_J is a well studied problem and can be accomplished from a small number of images of an arbitrary scene obtained at different exposures [10, 3, 11]. Applying the inverse response m_J^{-1} to M_J linearizes measured brightness with respect to C_J. In the remainder of this paper, we will assume that the measured brightness is linear in camera irradiance. Note that we are assuming the model to be independent at each point. Thus we must assume that inter-reflection in the scene is negligible.

The above expressions, together, model the whole pipeline in Fig. 1, from input projector brightness to the measured camera brightness. Since this model only describes a single channel, we must generalize for the case of multiple color channels. It is important to note that the spectral responses of the projector and camera channels can be arbitrary and are unknown to us. Let us assume that the projector and the camera each have three color channels (RGB). Then, we can extend the above radiometric model and write it compactly using vectors and matrices as

$$\mathbf{C} \;=\; \mathbf{V}\,\mathbf{P} + \mathbf{F}\,, \tag{6}$$

where:

$$\mathbf{C} = \begin{bmatrix} C_R \\ C_G \\ C_B \end{bmatrix},$$

$$\mathbf{V} = \begin{bmatrix} V_{RR} & V_{RG} & V_{RB} \\ V_{GR} & V_{GG} & V_{GB} \\ V_{BR} & V_{BG} & V_{BB} \end{bmatrix}, \mathbf{P} = \begin{bmatrix} P_R \\ P_G \\ P_B \end{bmatrix}, \mathbf{F} = \begin{bmatrix} F_R \\ F_G \\ F_B \end{bmatrix},$$

$$V_{KJ} = \int w_K(\lambda)\, s(\lambda)\, q_J(\lambda)\, d\lambda\,,$$

$$F_J = \int a(\lambda)\, s(\lambda)\, q_J(\lambda)\, d\lambda\,,$$

$$P_K = p_K(\, I_K\,)\,.$$

The matrix $\mathbf{V}$ is referred to as the color mixing matrix. The matrix captures all the couplings between the projector and camera channels and their interactions with the spectral reflectance. Note that even if the scene were an ideal white, and the camera and projector had channels with identical spectra we cannot assume $\mathbf{V}$ is diagonal. This is because the R, G, and B channels typically have broad and overlapping spectral curves in both cameras and projectors. The contribution of environmental lighting in the scene (independent of the projector) to camera irradiance is the vector $\mathbf{F}$. Note that a projector will also output some light even for a black display image. Since this offset can be considered independent of the input value to the projector, we absorb this offset into the environmental lighting term $\mathbf{F}$.

The above model nicely separates brightness non-linearities of

[3] We will use small letters to denote functions.

the system from the spectral characteristics of the system.[4] This model is valid at each point $\mathbf{x}_I$ in the projector plane. We note that the color mixing matrix $\mathbf{V}_{\mathbf{x}_I}$ and the environmental lighting component $\mathbf{F}_{\mathbf{x}_I}$ spatially vary with $\mathbf{x}_I$ but do not depend on the vector of input projector brightness values $\mathbf{I} = (I_R, I_G, I_B)$. We make the crucial observation that for a typical projector the non-linear projector response function $\mathbf{p}(\mathbf{I}) = (p_R(I_R), p_G(I_G), p_B(I_B))$, is the same for all points. Therefore, the model in Eq. (6) can be simplified to

$$\mathbf{C} = \mathbf{V}_{\mathbf{x}_I}\mathbf{p}(\mathbf{I}) + \mathbf{F}_{\mathbf{x}_I}. \tag{7}$$

Since the projector response function $\mathbf{p}$ is the same at every point, our radiometric model can be represented with a small number of parameters. The model is minimal in the sense that reducing the number of parameters further would require assumptions on the scene or the environmental lighting. In this model, $\mathbf{p}$ can be parameterized using splines, or simply represented with lookup tables. With a lookup table with 16-bit precision, this requires just over 390K for all three channels. The 3x3 matrix $\mathbf{V}_{\mathbf{x}_I}$ and the 3-vector $\mathbf{F}_{\mathbf{x}_I}$ together require 12 values per pixel. At a resolution of 640x480 this only requires 7.7MB at 16-bit precision to store all the parameters of the model. This modest memory requirement makes real-time application of the model possible.

3 Efficient Recovery of Model Parameters

The model parameters in Eq. (7) depend on the scene. When the scene changes, these parameters must be recalibrated. Thus, a simple and efficient calibration method is required. The efficiency of the calibration process is fundamentally limited by the number of images needed to recover the model parameters. In this section, we show that we can accurately estimate these paramters by acquiring just 6 images.

3.1 Decoupling the Color Channels

The first step in recovering the model parameters in Eq. (7) is to find the values of the matrix $\mathbf{V}_{\mathbf{x}_I}$ at each pixel $\mathbf{x}_I$. We omit the subscript $\mathbf{x}_I$ for the remainder of this section, but it is important to keep in mind that $\mathbf{V}$ varies spatially. It turns out to be convenient to treat the recovery of diagonal entries V_{KK} of $\mathbf{V}$ separately. Let $\mathbf{D}$ be the diagonal matrix with diagonal entries V_{KK}. Then we define a matrix $\tilde{\mathbf{V}} = \mathbf{V}\mathbf{D}^{-1}$ so that $\tilde{V}_{KK} = 1$. The entries of the matrix $\tilde{\mathbf{V}}$ encode the mixing between unlike projected and measured color channels.

To determine $\tilde{\mathbf{V}}$ we change the input projector brightness of one channel while keeping the others fixed. For example, consider the case where we only change the red channel of the input brightness:

$$\mathbf{I}^{(1)} = \begin{bmatrix} {I_R}^{(1)} \\ {I_G}^{(1)} \\ {I_B}^{(1)} \end{bmatrix}, \ \mathbf{I}^{(2)} = \begin{bmatrix} {I_R}^{(2)} \\ {I_G}^{(1)} \\ {I_B}^{(1)} \end{bmatrix}. \tag{8}$$

From Eq. (6), we have

$$\begin{aligned} \begin{bmatrix} {C_R}^{(1)} \\ {C_G}^{(1)} \\ {C_B}^{(1)} \end{bmatrix} &= \tilde{\mathbf{V}}\mathbf{D} \begin{bmatrix} {P_R}^{(1)} \\ {P_G}^{(1)} \\ {P_B}^{(1)} \end{bmatrix} + \mathbf{F}, \\ \begin{bmatrix} {C_R}^{(2)} \\ {C_G}^{(2)} \\ {C_B}^{(2)} \end{bmatrix} &= \tilde{\mathbf{V}}\mathbf{D} \begin{bmatrix} {P_R}^{(2)} \\ {P_G}^{(1)} \\ {P_B}^{(1)} \end{bmatrix} + \mathbf{F}. \end{aligned} \tag{9}$$

Since we have changed only the red channel of the input, the corresponding changes in the three channels are simply:

$$\begin{aligned} \Delta C_R &= \tilde{V}_{RR} V_{RR} \Delta P_R, \\ \Delta C_G &= \tilde{V}_{GR} V_{RR} \Delta P_R, \\ \Delta C_B &= \tilde{V}_{BR} V_{RR} \Delta P_R. \end{aligned} \tag{10}$$

Since $\tilde{V}_{RR} = 1$, we have,

$$\tilde{V}_{RG} = \frac{\Delta C_G}{\Delta C_R} \text{ and } \tilde{V}_{RB} = \frac{\Delta C_B}{\Delta C_R}. \tag{11}$$

Similarly, the unknown elements of $\tilde{\mathbf{V}}$ corresponding to the green and blue channels are obtained by changing the input brightness for each of those channels while keeping the others fixed. We can assume that each of the channels is changed with respect to a single base image. The 6 parameters in $\tilde{\mathbf{V}}$ require exactly 4 images to estimate.[5] In our experiments, we used one image with a low input brightness value. For the other three images, we changed the input brightness by a significant amount in each respective channel.

The recovery of $\tilde{\mathbf{V}}$ thus allows us to decouple the color channels by multiplying Eq. (7) by $\tilde{\mathbf{V}}^{-1}$. This gives

$$\tilde{\mathbf{C}} = \mathbf{D}\,\mathbf{p}(\mathbf{I}) + \tilde{\mathbf{F}}, \tag{12}$$

where $\tilde{\mathbf{C}} = \tilde{\mathbf{V}}^{-1}\mathbf{C}$ and $\tilde{\mathbf{F}} = \tilde{\mathbf{V}}^{-1}\mathbf{F}$. The key point is that the rows of Eq. (12) decouple into an independent equation for each channel

$$\tilde{C}_K = V_{KK}\, p_K(I_K) + \tilde{F}_K. \tag{13}$$

Thus, at each pixel, $\tilde{C}_K$ depends on the input brightness I_K for the channel K and is independent of the other channels..

3.2 Recovering the Projector Non-Linearity

The relationship between input projector brightness and camera irradiance can be determined at each pixel, and for each channel, by exhaustively capturing one image for each possible input projector brightness $I_K = 1$ to 255. This is highly inefficient given that, for each channel, the projector non-linearity is the same at every pixel. Consider a perfectly uniform white projection surface and a projector and camera with

[4] In our experiments, we have used a DLP projector. It is known that a DLP projector adds a "white" component that is a function (possibly non-linear) of the RGB color values (see [17, 19] for details). For this case, the color mixing is more complex. However, our experimental results indicate that the above model works as a reasonable approximation.

[5] The 3 channels per image give 12 measurements at each pixel. We note that recovering $\tilde{\mathbf{V}}$ is invariant to the additive factor $\mathbf{F}$ (3 parameters) and rescaling by $\mathbf{D}$ (3 parameters). Hence at each pixel, the images give 12-6=6 constraints.

perfectly uniform optics. In this special case the matrix $\tilde{\mathbf{V}}$ is the same for all pixels. We could perform the calibration of the projector non-linearity using just one calibration image. For example, let $\mathcal{U}$ be the calibration image, shown in Fig. 3(a), that consists of 20x20 pixel squares going from possible input brightness black (gray-level 0) to white (gray-level 255) in steps of one. Once we decouple the channels using $\tilde{\mathbf{V}}^{-1}$, we can determine camera irradiance as a function of input projector brightness by regression on each channel independently.

In our case, however, the optics are not uniform and the scene has varying surface reflectance and environmental lighting. To handle these variations, we construct an invariant of the projector illumination. Let $\mathcal{S}$ and $\mathcal{T}$ be uniform input images where all channels are equal to gray-levels S and T, respectively. Let $\tilde{C}_{K,\mathcal{S}}$, $\tilde{C}_{K,\mathcal{T}}$, $\tilde{C}_{K,\mathcal{U}}$, be the corresponding decoupled camera irradiances for the images $\mathcal{S}$, $\mathcal{T}$ and $\mathcal{U}$ at $\mathbf{x}_I$. It follows from Eq. (13) that we can define a value N_K given by

$$N_K = \frac{(\tilde{C}_{K,\mathcal{U}} - \tilde{C}_{K,\mathcal{S}})}{(\tilde{C}_{K,\mathcal{T}} - \tilde{C}_{K,\mathcal{S}})} \quad (14)$$

$$= \frac{(p_K(I_K) - p_K(S))}{(p_K(T) - p_K(S))}, \quad (15)$$

where I_K is the input projector brightness value of the image $\mathcal{U}$ at $\mathbf{x}_I$. Eq. (14) shows that N_K is invariant to any simultaneous scaling of all the camera irradiances $\tilde{C}_{K,\mathcal{U}}$, $\tilde{C}_{K,\mathcal{S}}$, $\tilde{C}_{K,\mathcal{T}}$. It is also invariant to simultaneous addition by a constant to all the camera irradiances. At a pixel, the contribution of reflectance, environmental lighting, and the fall-offs of the projector-camera system all scale, or add constants to, the camera irradiances. From Eq. (15) we see that the value N_K only depends on the value I_K in the image $\mathcal{U}$ since the projector response p_K is the same at every point. Since p_K is monotonic, it follows from Eq. (15) that there is a monotonic function ρ_K such that $I_K = \rho_K(N_K)$. We therefore can determine ρ_K by regression of the values N_K vs. I_K for all pixels.

In our experiment, we choose brightness values $S = 85$ and $T = 140$. Fig. 3(b) shows a sample poster we used to test our three-image calibration. The image in Fig. 3(c) shows the camera measurement of the spatially varying image projected on the poster. In Fig. 3(d) is an image whose intensities are the normalized values of the invariant N_K. This image should look similar to Fig. 3(a) since the values of the invariant are projector input values re-mapped by ρ_k^{-1} for each channel.

From Fig. 3(d), it is clear that some values of N_K depend on the reflectance values from the poster shown in Fig. 3(b). This is explained by the fact that the invariance does not hold at points where a pixel is too dark to be measured or saturates in one of the channels. Invariance also fails at points where there is a strong variation in either the object's reflectance or illumination.[6] We remove these points from consideration using a mask (see Fig. 4(a)) indicating the valid pixels. The mask is zero at pixels that are saturated or black in the captured images $\tilde{C}_{K,\mathcal{S}}$, $\tilde{C}_{K,\mathcal{T}}$, $\tilde{C}_{K,\mathcal{U}}$. In addition a mask pixel is zero if the corresponding point in the input image $\mathcal{U}$ (varying projector illumination) or the image acquired when projecting $\mathcal{T}$ (varying reflectance) has large variance in a neighborhood. All other mask pixels are one. We also exclude pixels where the estimate of $\tilde{\mathbf{V}}$ is corrupted by saturation.

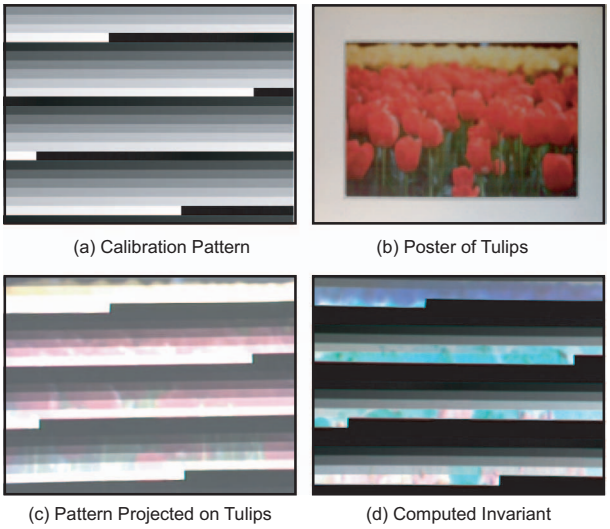


Figure 3: (a) A spatially varying calibration image used to recover the projector non-linearity. (b) A poster of tulips we used to test our three-image calibration method. The spatially varying image from (a) is projected on the poster and results in (c) when captured by the camera. From this an invariant image (d) is computed using Eq. (14) which depends on the projector illumination in (a). Note the invariance fails to be independent of the poster reflectance and environmental lighting at some points. This is due to saturated, dark and high variance pixels.

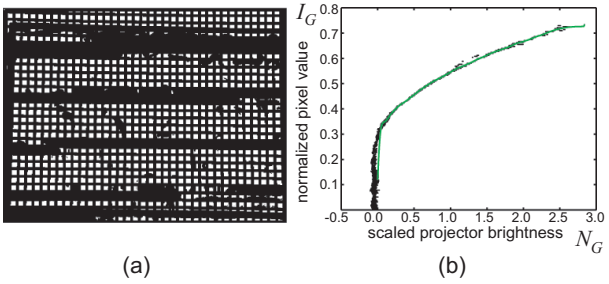


Figure 4: (a) Mask used to exclude points which are either saturated or have too much local variance. Black mask points were not used in determining the non-linearity of the display system. (b) Plot of projector input brightness vs. an invariant of projector output (for the Green channel).

Fig. 4(b) shows a scatter plot of pairs of values I_G and N_G for the green channel $(K = G)$. The plots for the other channels look similar. The green curve shown in Fig. 4(b) was obtained by non-parametric regression. The regression was performed at 1000 sample points to obtain ρ_K. As is clear from Fig. 4(b), almost all of our data lies very close to a smooth curve. We note that we recover $\tilde{\mathbf{V}}$ with 4 images and ρ with 3, requiring

[6] Pixels near an edge in the acquired image represent an average of values. This breaks the invariant. The camera and the projector also do not behave ideally with respect to step edges. For instance, there are slight misalignments of the color channels. Manufacturers may also introduce mild sharpening filters to enhance typical images. These factors produce artifacts at edges.

a total of 7 images. However, the darker constant image $\mathcal{S}$ can be used for both sets of images. Hence, we need just 6 images for calibration.

4 Computing the Compensation Image

Let $\rho = (\rho_R, \rho_G, \rho_B)$. We write the formula for the vector of RGB input values $\mathbf{I} = (I_R, I_G, I_B)$ required to achieve the camera vector of RGB camera irradiance values $\mathbf{C}$ at a point $\mathbf{x}_I$ in vector form as

$$\mathbf{I} = \rho \left((\tilde{\mathbf{V}}^{-1}\mathbf{C} - \tilde{\mathbf{C}}_{\mathcal{S}})./(\tilde{\mathbf{C}}_{\mathcal{T}} - \tilde{\mathbf{C}}_{\mathcal{S}}) \right) \qquad (16)$$

where the notation ./ denotes component-wise division, and $\tilde{\mathbf{C}}_{\mathcal{S}}$ and $\tilde{\mathbf{C}}_{\mathcal{T}}$ are the measured camera values for the projected images $\mathcal{S}$ and $\mathcal{T}$. We note that our on-line compensation method does not use the camera; we directly compute the compensation image from Eq. (16). This is critical to making real-time compensation possible.

5 Experimental Results

We have evaluated our algorithm on the colorful poster shown in Fig. 3(a). In Fig. 5(a) we see the results of projecting flat-gray uncompensated input images on the poster for three levels of gray (100,150,200). After performing our calibration, we used Eq. (16) to determine the corresponding compensation images shown in Fig. 5(b). These were projected on the poster producing a near uniform gray appearance. At points where the measured camera brightness saturates or at strong edges on the poster, the correction fails as is seen in the result for the brightest image (gray-level 200).[7]

The remainder of the small artifacts are due to the relatively low resolution camera we used as well as color abberations introduced by the projector at strong edges.

In general, however, the errors are quite small as can be seen from Table. 6, which shows maximum and RMS errors. We compared our results against those achieved from correction based on the calibration algorithm of [12] that uses 260 images. The errors were comparable. This shows that the compactness of our model and the efficiency of our calibration algorithm does not compromise performance.

We also applied our correction to achieve more complex (non-uniform) desired appearances on the tulip poster. For example, we would like to change our tulip poster to look like a poster of the woman's face shown Fig. 7(a). Simply projecting this image on the poster results in the appearance shown in Fig. 7(b). The result has an undesirable mix of appearances. In addition, the captured image is very dark. The compensation image in Fig. 7 (c) both increases the brightness and corrects for the spatially varying reflectance. The new appearance shown in Fig. 7(d) is very close to the desired appearance despite the strong colors on the tulips poster.

[7] The edge artifacts mentioned in Sec. 3.1 as well as camera's resolution place limits on the frequency of surface modulations a system can correct. This limitation required us to use smoothed reflectance patterns.

(a) Uniform gray (values 100,150,200) projected on flat poster

(b) Compensation images for uniform gray (values 100,150,200)

(c) Result of projection of compensation images on poster

Figure 5: Experimental results showing control of the appearance of the poster in Fig. 3(b). Despite the strong colors in the poster we are able to give it a near uniform gray appearance using our method.

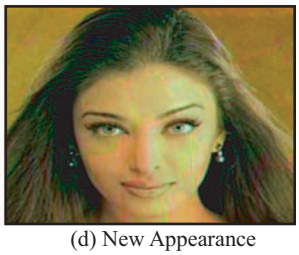
Projected Brightness R, G, B	Max. Error R,G,B		RMS Error R,G,B	
	Uncompensated	Compensated	Uncompensated	Compensated
50, 50, 50	25,26,25	8,6,7	23,24,23	4,3,4
100, 100, 100	66,70,69	12,8,10	59,66,64	3,2,3
150, 150, 150	77,89,83	13,13,14	42,74,66	6,3,3
200, 200, 200	53,81,62	31,23,31	36,47,27	8,4,4

Figure 6: Table showing numerical errors with and without compensation. The errors remain quite small as long as the compensation remains within the dynamic range that the projector can produce.

(a) Desired Appearance

(b) Uncompensated

(c) Compensation Image



(d) New Appearance

Figure 7: Experimental results showing the tulip poster in Fig. 3(b) transformed into a woman's face. When the desired image (a) is simply projected on the poster, the image (b) is both dark and modulated by the colorful flowers. By instead projecting the compensation image (c), we obtain an image (d) which is very close to the desired image.

Our method works equally well when changing the appearance of 3D objects. For example, the cube in Fig. 8(a) has the appearance of a cardboard computer box. We applied our radiometric calibration and correction to make it appear as a shaded neutral gray box, as shown in Fig. 8(b). We see in Fig. 8(c) that the original appearance cannot be changed by simply projecting the desired appearance. The result of projecting the compensation image in Fig. 8(d) gives the box the appearance that is shown Fig. 8(e), which is close to the desired appearance. Since we used a matte object, the box has the desired appearance from other viewing directions, as can be seen in Fig. 8(f). In Fig. 9(a)-(d), we changed the apparent pattern on the computer box to make it appear like that of a completely different brand.

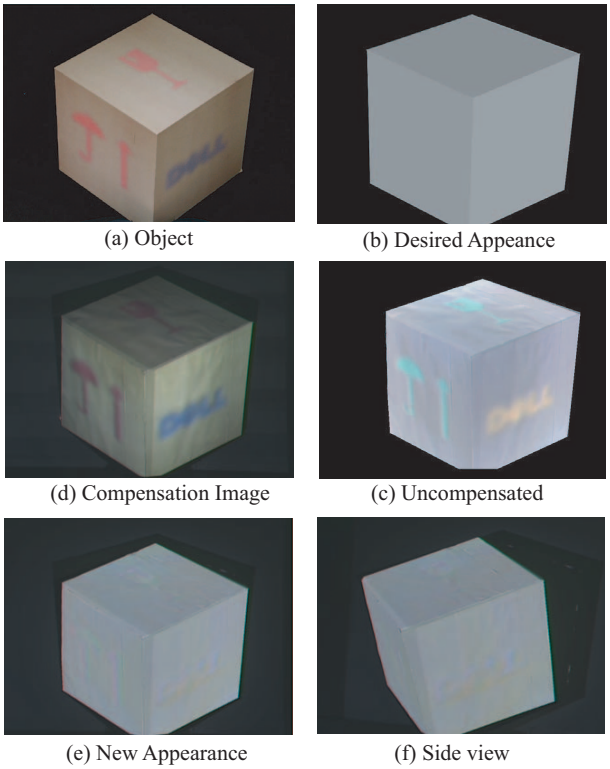
The image in Fig. 10(a) shows a brightly colored globe. In Fig. 10(b)-(f), we see the results of changing the appearance of the globe to that of a shaded gray sphere. This demonstrates that our method is effective at handling curved objects.

We also used our method to control the appearance of the entire scene shown in Fig. 10(a) including the brick background and the globe. In Fig. 11(a), the desired appearance is of a soccer ball with soccer players in the background. As seen in Fig. 11(b), despite the strong pattern of the ball, the intrinsic appearance of the object is clearly visible in the uncompensated image while the background is too dark. Upon compensation, the scene takes on the desired appearance. As we change the view, the appearance is maintained although the shadow cast by the sphere shows the limitations of using only one projector.

In [16], a video was used to give specially prepared white painted objects apparent motion. Using our method, we can transform an ordinary scene of colored and textured objects into a dynamically changing scene. For example, we give the globe in Fig. 10 the appearance of Jupiter rotating in space. Fig. 12(a) shows one frame of the desired video. Fig. 12(b) shows one frame of the uncompensated video. Both the incorrect shading as well as the globe's strong color can be clearly seen. After compensation, the same video frames have the correct new appearance for all views as shown in Fig 12(c,d). Real-time video correction is practical since the efficiency of our model enables correction at frame-rate.

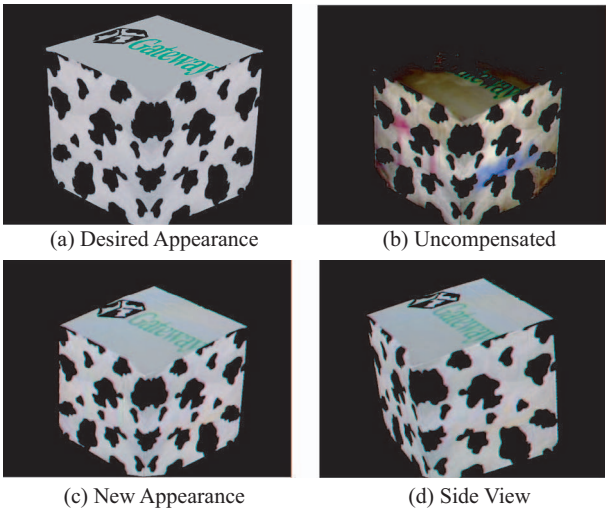
6 Conclusion

We have demonstrated the ability to control the appearance of 3D objects using a projector-camera system. This is done by compensating for the intrinsic colors and textures of the objects. Our compensation technique is based on a novel radiometric model and an efficient off-line algorithm that computes the parameters of the model. We have applied our compensation method to a variety of colored and textured 3D objects. The method works for scenes with both planar and curved surfaces. Since it is very efficient, it can also be used to temporally vary the appearance of an object to emulate novel dynamic scenes. The ability to control the appearance of everyday objects makes the projector a more versatile device that can be used to create experiences that are more realistic, compelling and communicative.



(a) Object (b) Desired Appeance (d) Compensation Image (c) Uncompensated (e) New Appearance (f) Side view

Figure 8: Experimental results on controlling the appearance of a computer box with symbols and lettering. The algorithm automatically compensates for both the albedo and shading of the 3D object.



(a) Desired Appearance (b) Uncompensated (c) New Appearance (d) Side View

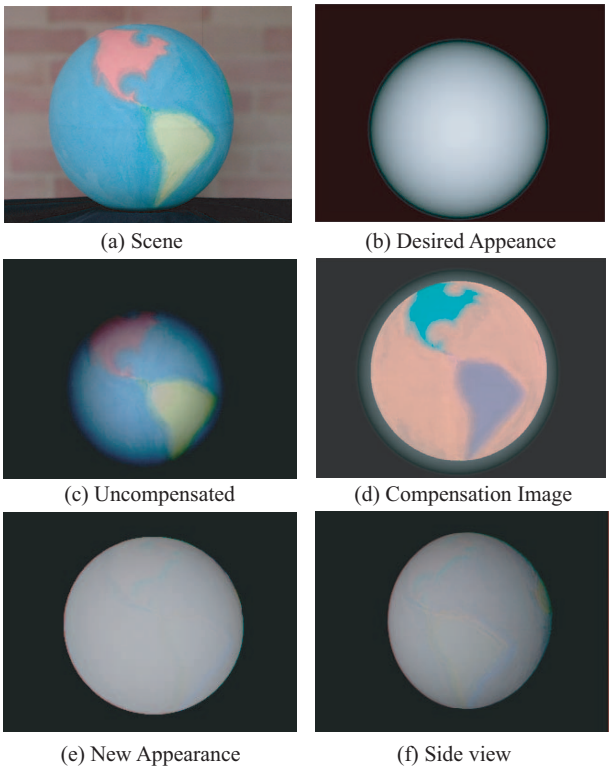
Figure 9: We change the apparent brand and design of the computer box from Fig.8(a). This allows the preview a new box design.

References

[1] T.J. Cham, J.M. Rehg, R. Sukthankar, and G. Sukthankar. Shadow elimination and occluder light suppression for multi-projector displays. In *CVPR03*, pages II: 513–520, 2003.

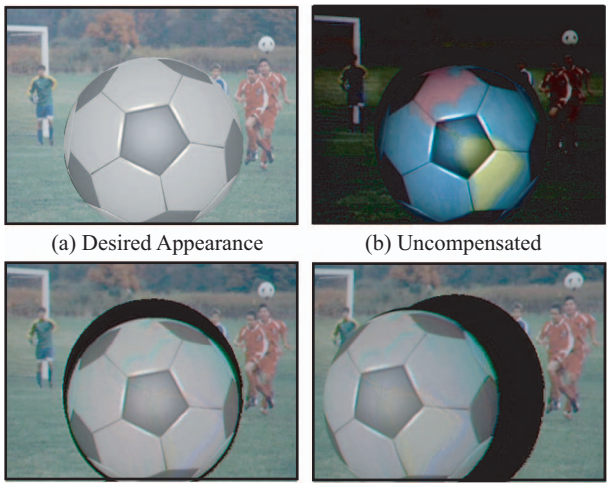
[2] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti. Surround-screen projection-based virtual reality: The design and implementation of the cave. In *Proc. of SIGGRAPH*, pages 135–142, Anaheim, CA, 1993.

[3] P. Debevec and J. Malik. Recovering High Dynamic Range Radiance Maps from Photographs. *Proc. of ACM SIGGRAPH 1997*, pages 369–



(a) Scene (b) Desired Appeance (c) Uncompensated (d) Compensation Image (e) New Appearance (f) Side view

Figure 10: Experimental results on controlling the appearance of a brightly colored globe. The compensation gives the desired appearance from all views.



(a) Desired Appearance (b) Uncompensated (c) New Appearance (c) Side View

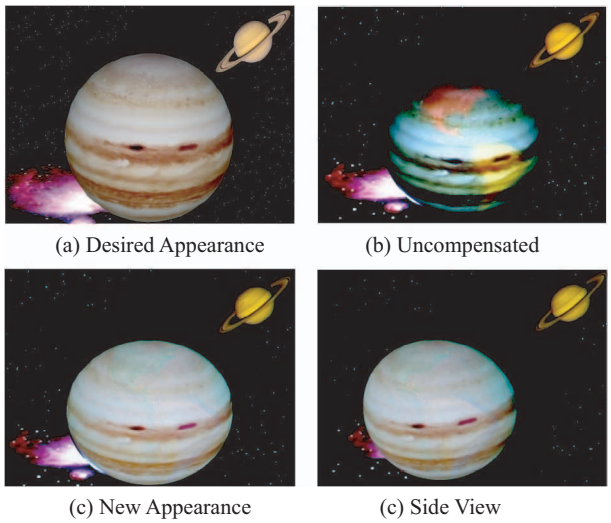
Figure 11: Controlling the appearance of the scene in Fig. 10(a). Here, both the globe and the brick background are changed so that the scene appears to be a soccer ball in front of a playing field.

378, 1997.

[4] G. Humphreys and P. Hanrahan. A distributed graphics system for large tiled displays. In *Proc. IEEE Visualization*, pages 215–223, 1999.

[5] M. Inami, N. Kawakami, D. Sekiguchi, Y. Yanagida, T. Maeda, and S. Tachi. Visuo-haptic display using head-mounted projector. In *Proc. of the IEEE VR*, pages 233–240, 2000.

[6] P. Inova. Seamless video display, #4,974,073. US Patent, 1988.



(a) Desired Appearance (b) Uncompensated (c) New Appearance (c) Side View

Figure 12: Experimental results from a video showing the globe with brick background in Fig. 10(a) transformed into a rotating Jupiter in space.

[7] K. Li and Y. Chen. Optical blending for multi-projector display wall system. In *IEEE Proc. 12th Lasers and Electro-Optics Society*, volume 1, pages 281–282, 1999.

[8] P. Lyon. Edge-blending multiple projection displays on a dome surface to form continuous wide angle fields-of-view. In *Proc. of 7th I/ITEC*, pages 203–209, 1985.

[9] A. Majumder and G. Welch. Computer graphics optique: Optical superposition of projected computer graphics. In *Fifth Immersive Projection Technology Workshop*, May 2001.

[10] S. Mann and R. Picard. Being 'Undigital' with Digital Cameras: Extending Dynamic Range by Combining Differently Exposed Pictures. *Proc. of IST's 48th Annual Conference*, pages 442–448, May 1995.

[11] T. Mitsunaga and S. K. Nayar. Radiometric Self Calibration. In *Proc. of Computer Vision and Pattern Recognition '99*, volume 1, pages 374–380, June 1999.

[12] S. Nayar, H. Peri, M. Grossberg, and P. Belhumeur. A projection system with radiometric compensation for screen imperfections. In *IEEE International Workshop on Projector-Camera Systems*, Oct. 2003.

[13] S. K. Nayar, H. Peri, M. Grossberg, and P. Belhumeur. A projector with automatic radiometric screen compensation. Technical report, Department of Computer Science, Columbia University, New York, 2003.

[14] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. *Proc. of SIGGRAPH*, pages 179–188, 1998.

[15] R. Raskar, G. Welch, K-L. Low, and D. Bandyopadhyay. Shader lamps: Animating real objects with image based illumination. In *Eurographics Workshop on Rendering*, pages 89–102, 2001.

[16] R. Raskar, R. Ziegler, and T. Willwacher. Cartoon dioramas in motion. In *Inter. Symp. on Non-Photorealistic Animation and Rendering (NPAR)*, pages 7–12, 2002.

[17] M. Stone. Color balancing experimental projection displays. In *Proc. of the 9th Color Imaging Conference, Scottsdale, AZ*, number 5-9, pages 342–347, Nov. 2001.

[18] M. Tsukada, C. Funayama, and J. Tajima. Devlopment of projector color reproduction adapted to wall color. In *Proc. Color Forum Japan 2002*, pages 103–106, 2002.

[19] G. Wallace, H. Chen, and K. Li. Color gamut matching for tiled display walls. In *Proceedings of the workshop on Virtual environments 2003*, pages 293–302. ACM Press, 2003.