Exposing Digital Forgeries in Complex Lighting Environments

Micah K. Johnson, *Student Member, IEEE,* and Hany Farid, *Member, IEEE*

Abstract

The availability of sophisticated digital imaging technology has given rise to digital forgeries that are increasing in sophistication and frequency. We describe a technique for exposing such fakes by detecting inconsistencies in lighting. We show how to approximate complex lighting environments with a low-dimensional model and, further, how to estimate the model's parameters from a single image. Inconsistencies in the lighting model are then used as evidence of tampering.

I. Introduction

The past few years has seen a considerable rise in the availability and sophistication of digital imaging technology (cameras, scanners, software) and their use in manipulating digital images. From the tabloid magazine to the fashion industry to main-stream media outlets to political campaigns to the photo hoaxes that land in our email in-boxes, doctored photographs are appearing with a growing frequency and sophistication. As a result, our trust in photographs is slowly being diminished.

The field of digital forensics has emerged over the past few years to combat this growing problem. Several techniques have been developed to detect various forms of digital tampering. Statistical techniques have been developed for detecting cloning [1], [2]; splicing [3]; re-sampling artifacts [4], [5]; color filter array aberrations [4]; and disturbances of a camera's sensor noise pattern [6]. Optical techniques have been developed to detect chromatic aberrations [7], and geometric techniques for rectifying perspectively distorted planar surfaces [8]. More recently two related approaches have been developed for detecting inconsistencies in lighting [9], [10]. Building specifically on this work, and more broadly on all of these forensic tools, we describe a new lighting-based digital forensic technique.

M. Johnson (kimo@cs.dartmouth.edu) and H. Farid (farid@cs.dartmouth.edu) are with the Department of Computer Science at Dartmouth College, 6211 Sudikoff Lab, Hanover NH 03755.

Fig. 1. A fake *Star* magazine cover showing the first author, Kimo, with actress Katie Holmes. Also shown is a magnified view of this forgery, and the original cover showing Holmes with actor Tom Cruise.

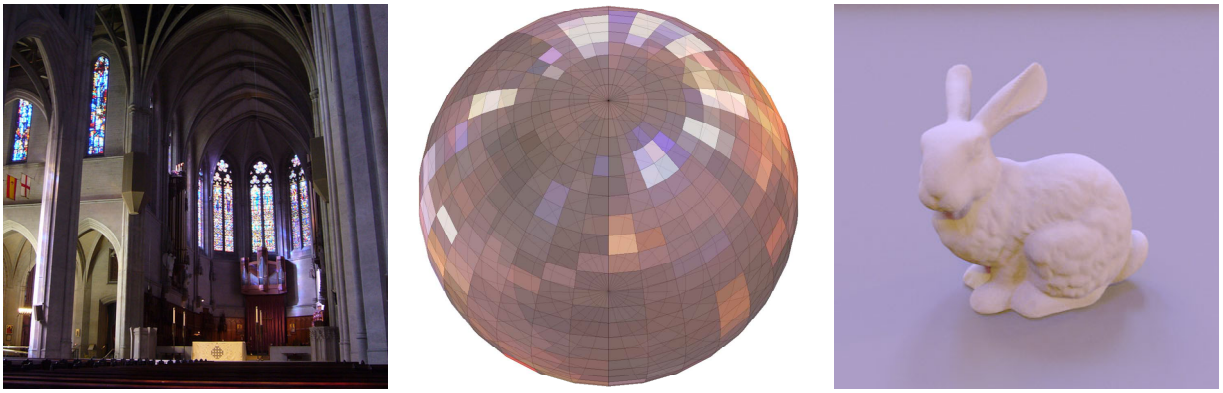

Fig. 2. Shown from left to right are an image taken inside Grace Cathedral in San Francisco, a sphere embodying the lighting environment in Grace Cathedral, and the Stanford bunny rendered under this lighting environment.

When creating a digital forgery from multiple images, it is often difficult to exactly match the lighting conditions. We have previously shown how to estimate the direction to a light source, and how inconsistencies in the illuminant direction can be used to detect tampering [9], [10]. This approach is appropriate when the lighting is dominated by a single light source, but is less appropriate in more complex lighting environments containing multiple light sources or non-directional lighting (e.g., the sky on a cloudy day). Shown in Fig. 1, for example, is a digital composite of "Katie" and "Kimo." At first glance, this composite is reasonably compelling. At a closer examination, however, the lighting on Kimo is seen to be strongly directional while the lighting on Katie is more diffuse. Here we describe how to quantify such complex lighting environments and how to use inconsistencies in lighting to detect tampering.

We leverage earlier work [11], [12] that shows that under some simplifying assumptions, arbitrarily complex lighting environments can be approximated with a low-dimensional model. We show how the parameters of a reduced version of this model can be estimated from a single image, and how this model can be used to detect consistencies and inconsistencies in an image. Results from a broad range of simulated and photographed images and visually plausible forgeries are presented.

II. Methods

The lighting of a scene can be complex—any number of lights can be placed in any number of positions, creating different lighting environments. In order to model such complex lighting, we assume that the lighting is distant and that surfaces in the scene are convex and Lambertian. To use this model in a forensic setting, we also assume that the surface reflectance is constant and that the camera response is linear.

A. Representing lighting environments

Under the assumption of distant lighting, an arbitrary lighting environment can be expressed as a non-negative function on the sphere, $L(\vec{V})$, where $\vec{V}$ is a unit vector in Cartesian coordinates and the value of $L(\vec{V})$ is the intensity of the incident light along direction $\vec{V}$, Fig. 2. If the object being illuminated is convex, the irradiance (light received) at any point on the surface is due to only the lighting environment; i.e., there are no cast shadows or interreflections [11]. As a result, the irradiance, $E(\vec{N})$, can be parametrized by the unit length surface normal $\vec{N}$ and written as a convolution of the reflectance function of the surface, $R(\vec{V},\vec{N})$, with the lighting environment $L(\vec{V})$:

$$E(\vec{N}) \;=\; \int_{\Omega} L(\vec{V})\, R(\vec{V},\vec{N})\, d\Omega, \tag{1}$$

where Ω represents the surface of the sphere and $d\Omega$ is an area differential on the sphere. For a Lambertian surface, the reflectance function is a clamped cosine:

$$R(\vec{V},\vec{N}) \;=\; \max\left(\vec{V}\cdot\vec{N}, 0\right), \tag{2}$$

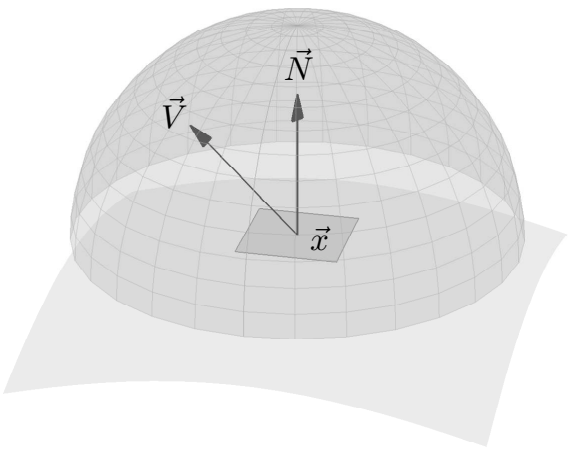

Fig. 3. The irradiance (light received) at a point $\vec{x}$ is determined by integrating the amount of incoming light from all directions $\vec{V}$ in the hemisphere about the surface normal $\vec{N}$.

which is either the cosine of the angle between vectors $\vec{V}$ and $\vec{N}$, or zero when the angle is greater than 90 degrees. This reflectance function effectively limits the integration in Equation (1) to the hemisphere about the surface normal $\vec{N}$, Fig. 3. In addition, while we have assumed no cast shadows, Equation (2) explicitly models *attached* shadows, i.e., shadows due to surface normals facing away from the direction $\vec{V}$.

The convolution in Equation (1) can be simplified by expressing both the lighting environment and the reflectance function in terms of spherical harmonics. Spherical harmonics form an orthonormal basis for piecewise continuous functions on the sphere and are analogous to the Fourier basis on the line or plane. The first three orders of spherical harmonics are defined in Appendix A and shown in Fig. 4.

The lighting environment expanded in terms of spherical harmonics is:

$$L(\vec{V}) \;=\; \sum_{n=0}^{\infty} \sum_{m=-n}^{n} l_{n,m} Y_{n,m}(\vec{V}), \tag{3}$$

where $Y_{n,m}(\cdot)$ is the m^{th} spherical harmonic of order n, and $l_{n,m}$ is the corresponding coefficient of the lighting environment. Similarly, the reflectance function for Lambertian surfaces, $R(\vec{V},\vec{N})$, can be expanded in terms of spherical harmonics, and due to its symmetry about the surface normal, only harmonics with $m = 0$ appear in the expansion:

$$R(\vec{V},\vec{N}) \;=\; \sum_{n=0}^{\infty} r_n Y_{n,0}\left(\begin{pmatrix} 0 & 0 & \vec{V}\cdot\vec{N} \end{pmatrix}^T \right). \tag{4}$$

Note that for $m = 0$, the spherical harmonic $Y_{n,0}(\cdot)$ depends only on the z-component of its argument.

Convolutions of functions on the sphere become products when represented in terms of spherical harmonics [11], [12]. As a result, the irradiance, Equation (1), takes the form:

$$E(\vec{N}) \;=\; \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \hat{r}_n l_{n,m} Y_{n,m}(\vec{N}), \tag{5}$$

where

$$\hat{r}_n \;=\; \sqrt{\frac{4\pi}{2n+1}}\, r_n. \tag{6}$$

The key observation in [11] and [12] was that the coefficients $\hat{r}_n$ for a Lambertian reflectance function decay rapidly, and thus the infinite sum in Equation (5) can be well approximated by the first nine terms:

$$E(\vec{N}) \;\approx\; \sum_{n=0}^{2} \sum_{m=-n}^{n} \hat{r}_n l_{n,m} Y_{n,m}(\vec{N}). \tag{7}$$

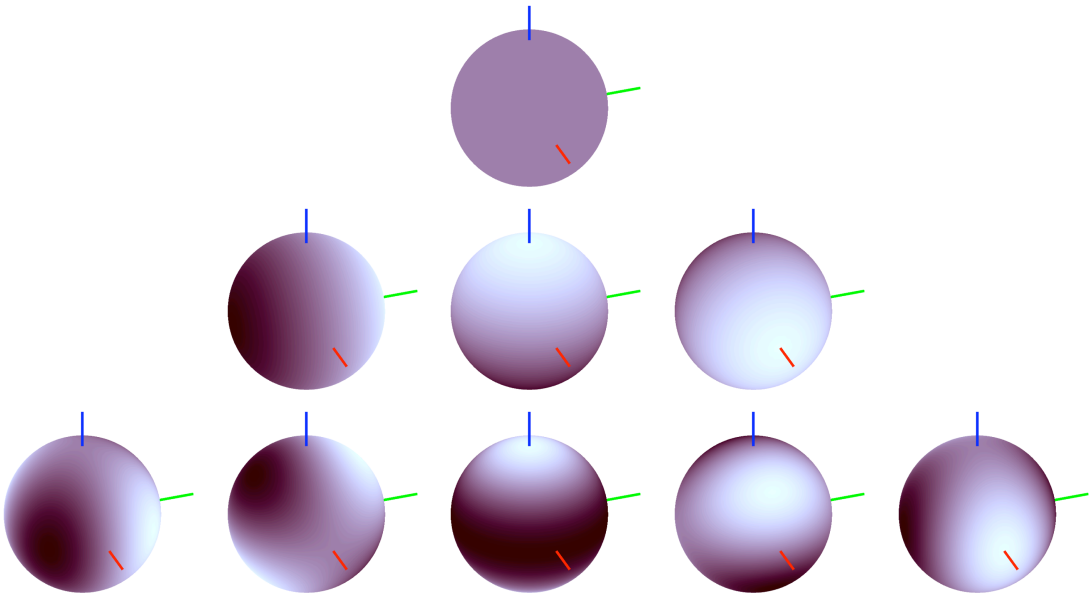

Fig. 4. The first three orders of spherical harmonics as functions on the sphere. Shown from top to bottom are the order zero spherical harmonic, $Y_{0,0}(\cdot)$; the three order one spherical harmonics, $Y_{1,m}(\cdot)$; and the five order two spherical harmonics, $Y_{2,m}(\cdot)$.

Since the constants $\hat{r}_n$ are known for a Lambertian reflectance function, the irradiance of a convex Lambertian surface under arbitrary distant lighting can be well modeled by the nine lighting environment coefficients $l_{n,m}$ up to order two.

B. From irradiance to intensity

Irradiance describes the total amount of light reaching a point on a surface. For a Lambertian surface, the reflected light, or radiosity, is proportional to the irradiance by a reflectance term ρ. In addition, Lambertian surfaces emit light uniformly in all directions, so the amount of light received by a viewer (i.e., camera) is independent of the view direction.

A camera maps its received light to intensity through a camera response function $f(\cdot)$. Assuming the reflectance term ρ is constant across the surface, the measured intensity at a point $\vec{x}$ in the image is given by [13]:

$$I(\vec{x}) \quad = \quad f(\rho t E(\vec{N}(\vec{x}))). \tag{8}$$

where $E(\cdot)$ is the irradiance, $\vec{N}(\vec{x})$ is the surface normal at point $\vec{x}$, and t is the exposure time. For simplicity, we assume a linear camera response, and thus the intensity is related to the irradiance by an unknown multiplicative factor, which is assumed to have unit value—this assumption implies that the lighting coefficients can only be estimated to within an unknown scale factor. Under these assumptions, the relationship between image intensity and irradiance is simply:

$$I(\vec{x}) \quad = \quad E(\vec{N}(\vec{x})). \tag{9}$$

C. Estimating lighting environments

Since, under our assumptions, the intensity is equal to irradiance, Equation (9) can be written in terms of spherical harmonics by expanding Equation (7):

$$\begin{aligned} I(\vec{x}) \quad &= \quad l_{0,0}\pi Y_{0,0}(\vec{N}) + l_{1,-1}\tfrac{2\pi}{3} Y_{1,-1}(\vec{N}) + l_{1,0}\tfrac{2\pi}{3} Y_{1,0}(\vec{N}) + l_{1,1}\tfrac{2\pi}{3} Y_{1,1}(\vec{N}) \\ &+ \quad l_{2,-2}\tfrac{\pi}{4} Y_{2,-2}(\vec{N}) + l_{2,-1}\tfrac{\pi}{4} Y_{2,-1}(\vec{N}) + l_{2,0}\tfrac{\pi}{4} Y_{2,0}(\vec{N}) \\ &+ \quad l_{2,1}\tfrac{\pi}{4} Y_{2,1}(\vec{N}) + l_{2,2}\tfrac{\pi}{4} Y_{2,2}(\vec{N}). \end{aligned} \tag{10}$$

Note that this expression is linear in the nine lighting environment coefficients, $l_{0,0}$ to $l_{2,2}$. As such, given 3-D surface normals at $p \geq 9$ points on the surface of an object, the lighting environment coefficients can be estimated as the least-squares solution to the following system of linear equations:

$$\begin{pmatrix} \pi Y_{0,0}(\vec{N}(\vec{x}_1)) & \frac{2\pi}{3} Y_{1,-1}(\vec{N}(\vec{x}_1)) & \dots & \frac{\pi}{4} Y_{2,2}(\vec{N}(\vec{x}_1)) \\ \pi Y_{0,0}(\vec{N}(\vec{x}_2)) & \frac{2\pi}{3} Y_{1,-1}(\vec{N}(\vec{x}_2)) & \dots & \frac{\pi}{4} Y_{2,2}(\vec{N}(\vec{x}_2)) \\ \vdots & \vdots & \ddots & \vdots \\ \pi Y_{0,0}(\vec{N}(\vec{x}_p)) & \frac{2\pi}{3} Y_{1,-1}(\vec{N}(\vec{x}_p)) & \dots & \frac{\pi}{4} Y_{2,2}(\vec{N}(\vec{x}_p)) \end{pmatrix} \begin{pmatrix} l_{0,0} \\ l_{1,-1} \\ \vdots \\ l_{2,2} \end{pmatrix} = \begin{pmatrix} I(\vec{x}_1) \\ I(\vec{x}_2) \\ \vdots \\ I(\vec{x}_p) \end{pmatrix} \tag{11}$$

$$M\vec{v} = \vec{b}, \tag{12}$$

where M is the matrix containing the sampled spherical harmonics, $\vec{v}$ is the vector of unknown lighting environment coefficients, and $\vec{b}$ is the vector of intensities at p points. The least-squares solution to this system is:

$$\vec{v} = \left(M^T M\right)^{-1} M^T \vec{b}. \tag{13}$$

This solution requires 3-D surface normals from at least nine points on the surface of an object. Without multiple images or known geometry, however, this requirement may be difficult to satisfy from an arbitrary image.

As in [9] and [14], we observe that under an assumption of orthographic projection, the z-component of the surface normal is zero along the occluding contour of an object. Therefore, the intensity profile along an occluding contour simplifies to:

$$I(\vec{x}) = A + l_{1,-1}\tfrac{2\pi}{3} Y_{1,-1}(\vec{N}) + l_{1,1}\tfrac{2\pi}{3} Y_{1,1}(\vec{N}) + l_{2,-2}\tfrac{\pi}{4} Y_{2,-2}(\vec{N}) + l_{2,2}\tfrac{\pi}{4} Y_{2,2}(\vec{N}), \tag{14}$$

where

$$A = l_{0,0}\tfrac{\pi}{2\sqrt{\pi}} - l_{2,0}\tfrac{\pi}{16}\sqrt{\tfrac{5}{\pi}}. \tag{15}$$

Note that the functions $Y_{i,j}(\cdot)$ depend only on the x and y components of the surface normal $\vec{N}$, Appendix A. That is, the five lighting coefficients can be estimated from only 2-D surface normals, which are relatively simple to estimate from a single image.[1] In addition, Equation (14) is still linear in its now five lighting environment coefficients, which can be estimated as the least-squares solution to:

$$\begin{pmatrix} 1 & \frac{2\pi}{3} Y_{1,-1}(\vec{N}(\vec{x}_1)) & \frac{2\pi}{3} Y_{1,1}(\vec{N}(\vec{x}_1)) & \frac{\pi}{4} Y_{2,-2}(\vec{N}(\vec{x}_1)) & \frac{\pi}{4} Y_{2,2}(\vec{N}(\vec{x}_1)) \\ 1 & \frac{2\pi}{3} Y_{1,-1}(\vec{N}(\vec{x}_2)) & \frac{2\pi}{3} Y_{1,1}(\vec{N}(\vec{x}_2)) & \frac{\pi}{4} Y_{2,-2}(\vec{N}(\vec{x}_2)) & \frac{\pi}{4} Y_{2,2}(\vec{N}(\vec{x}_2)) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{2\pi}{3} Y_{1,-1}(\vec{N}(\vec{x}_p)) & \frac{2\pi}{3} Y_{1,1}(\vec{N}(\vec{x}_p)) & \frac{\pi}{4} Y_{2,-2}(\vec{N}(\vec{x}_p)) & \frac{\pi}{4} Y_{2,2}(\vec{N}(\vec{x}_p)) \end{pmatrix} \begin{pmatrix} A \\ l_{1,-1} \\ l_{1,1} \\ l_{2,-2} \\ l_{2,2} \end{pmatrix} = \begin{pmatrix} I(\vec{x}_1) \\ I(\vec{x}_2) \\ \vdots \\ I(\vec{x}_p) \end{pmatrix} \tag{16}$$

$$M\vec{v} = \vec{b}, \tag{17}$$

which has the same least-squares solution as before:

$$\vec{v} = \left(M^T M\right)^{-1} M^T \vec{b}. \tag{18}$$

Note that this solution only provides five of the nine lighting environment coefficients. We will show, however, that this subset of coefficients is still sufficiently descriptive for forensic analysis.

When analyzing the occluding contours of objects in real images, it is often the case that the range of surface normals is limited, leading to an ill-conditioned matrix M. This limitation can arise from many sources, including occlusion or object geometry. As a result, small amounts of noise in either the surface normals or the measured intensities can cause large variations in the estimate of the lighting environment vector $\vec{v}$. To better condition the estimate, an error function $E(\vec{v})$ is defined that combines the least-squares error of the original linear system with a regularization term:

$$E(\vec{v}) = \|M\vec{v} - \vec{b}\|^2 + \lambda\|C\vec{v}\|^2, \tag{19}$$

[1]The 2-D surface normal is the gradient vector of an implicit curve fit to the edge of an object.

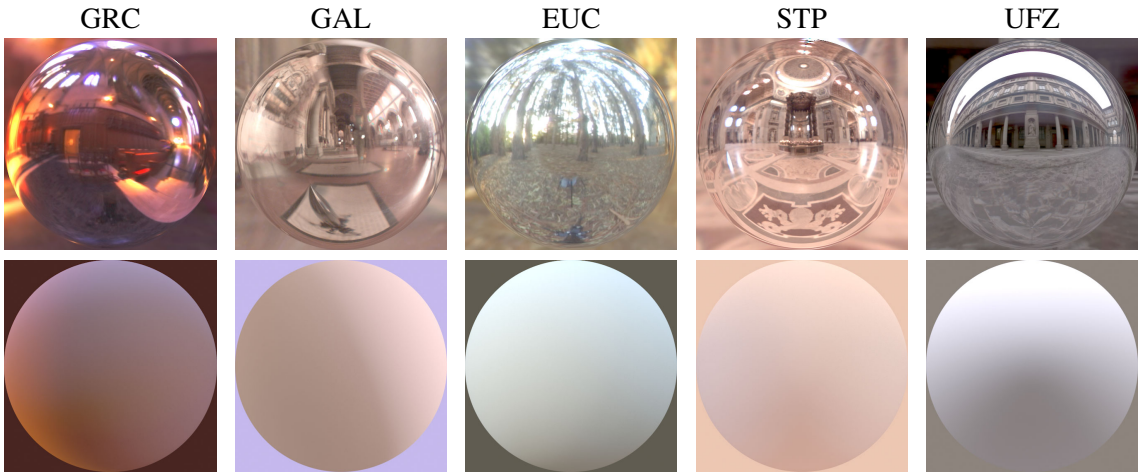

Fig. 5. Shown along the top row are five light probes from different lighting environments, from which lighting coefficients are computed (Table I). Shown in the bottom row are Lambertian spheres rendered from these coefficients.

where λ is a scalar, and the matrix C is diagonal with $\begin{pmatrix}1 & 2 & 2 & 3 & 3\end{pmatrix}$ on the diagonal. The matrix C is designed to dampen the effects of higher order harmonics and is motivated by the observation that the average power of spherical harmonic coefficients for natural lighting environments decreases with increasing harmonic order [15]. For the full lighting model when 3-D surface normals are available, Equation (12), the matrix C has $\begin{pmatrix}1 & 2 & 2 & 2 & 3 & 3 & 3 & 3 & 3\end{pmatrix}$ on the diagonal.

The error function to be minimized, Equation (19), is a least-squares problem with a Tikhonov regularization [16]. The analytic minimum is found by differentiating with respect to $\vec{v}$:

$$\begin{aligned}\frac{dE(\vec{v})}{d\vec{v}} &= 2M^TM\vec{v} - 2M^T\vec{b} + 2\lambda C^TC\vec{v} \\ &= 2(M^TM + \lambda C^TC)\vec{v} - 2M^T\vec{b},\end{aligned} \tag{20}$$

setting the result equal to zero, and solving for $\vec{v}$:

$$\vec{v} = (M^TM + \lambda C^TC)^{-1}M^T\vec{b}. \tag{21}$$

In practice, we have found that the conditioned estimate in Equation (21) is appropriate if less than 180° of surface normals are available along the occluding contour. If more than 180° of surface normals are available, the least-squares estimate, Equation (18), can be used, though both estimates will give similar results for small values of λ.

D. Comparing lighting environments

The estimated coefficient vector $\vec{v}$, Equation (21), is a low-order approximation of the lighting environment. For forensics purposes, we would like to differentiate between lighting environments based on these coefficients. Intuitively, coefficients from objects in different lighting environments should be distinguishable, while coefficients from objects in the same lighting environment should be similar. In addition, measurable differences in sets of coefficients should be mostly due to differences in the lighting environment and not to other factors such as object color or image exposure. Taking these issues into consideration, we propose an error measure between two estimated lighting environments.

Let $\vec{v}_1$ and $\vec{v}_2$ be two vectors of lighting environment coefficients. From these coefficients, the irradiance profile along a circle (2-D) or a sphere (3-D) is synthesized, from which the error is computed. The irradiance profiles corresponding to $\vec{v}_1$ and $\vec{v}_2$ are given by:

$$\vec{x}_1 = M\vec{v}_1, \tag{22}$$

$$\vec{x}_2 = M\vec{v}_2, \tag{23}$$

TABLE I
LIGHTING ENVIRONMENT COEFFICIENTS AND ESTIMATION ERRORS FROM DIFFERENT LIGHTING ENVIRONMENTS

	$l_{0,0}$	$l_{1,-1}$	$l_{1,0}$	$l_{1,1}$	$l_{2,-2}$	$l_{2,-1}$	$l_{2,0}$	$l_{2,1}$	$l_{2,2}$	3-D ($\times 10^{-4}$)	2-D ($\times 10^{-4}$)
GRC	0.44	0.35	-0.18	-0.06	-0.05	-0.22	-0.10	0.21	-0.05	2.6	2.7
GAL	0.76	0.34	-0.19	0.54	0.50	-0.10	-0.27	-0.14	0.42	4.4	0.8
EUC	0.43	0.36	0.03	-0.10	-0.06	-0.01	-0.13	-0.05	-0.00	0.2	0.1
STP	0.26	0.14	-0.01	0.02	0.01	-0.03	-0.08	0.00	-0.03	6.4	1.7
UFZ	0.31	0.37	-0.00	-0.01	-0.02	-0.01	-0.27	0.00	-0.24	2.5	1.4

where the matrix M is of the form in Equation (12) (for 3-D normals) or Equation (17) (for 2-D normals). After subtracting the mean, the correlation between these zero-meaned profiles is:

$$\text{corr}(\vec{x}_1, \vec{x}_2) \quad = \quad \frac{\vec{x}_1^T \vec{x}_2}{\|\vec{x}_1\| \|\vec{x}_2\|} \tag{24}$$

In practice, this correlation can be computed directly from the lighting environment coefficients:

$$\text{corr}(\vec{v}_1, \vec{v}_2) \quad = \quad \frac{\vec{v}_1^T Q \vec{v}_2}{\sqrt{\vec{v}_1^T Q \vec{v}_1} \sqrt{\vec{v}_2^T Q \vec{v}_2}}, \tag{25}$$

where the matrix Q for both the 2-D and 3-D cases is derived in Appendix C.

By design, this correlation is invariant to both additive and multiplicative factors on the irradiance profiles $\vec{x}_1$ and $\vec{x}_2$. Recall that our coefficient vectors $\vec{v}_1$ and $\vec{v}_2$ are estimated to within an unknown multiplicative factor. In addition, different exposure times under a nonlinear camera response function can introduce an additive bias, Appendix B. The correlation is, therefore, invariant to these factors and produces values in the interval $[-1, 1]$. The final error is then given by:

$$D(\vec{v}_1, \vec{v}_2) \quad = \quad \frac{1}{2} \left(1 - \text{corr}(\vec{v}_1, \vec{v}_2)\right), \tag{26}$$

with values in the range $[0, 1]$.

III. RESULTS

We tested our technique for estimating lighting environment coefficients on synthetically generated images and real images of natural lighting environments. The synthetic images were rendered using the `pbrt` environment [17] with data from a gallery of light probe images maintained by Paul Debevec [18]. The natural images were obtained in two different ways. For the first set, we photographed a known target in a variety of lighting conditions. For the second set, we downloaded twenty images from Flickr, a popular image sharing website. Results from four visually plausible forgeries are also presented [19].

For all images, the lighting environment coefficients were estimated from only the green channel. It may be the case that for some objects, other color channels could provide more reliable estimates. For these cases, a different color channel could be used, but it is not advisable to compare estimates across color channels since complex lighting environments often have different coefficients for each channel.

A. Simulation

Lighting environments can be captured by a variety of methods, such as photographing a mirror sphere [18], or through panoramic photography techniques. These methods produce high dynamic range images, known as light probe images, that represent the lighting environment function $L(\vec{V})$. The spherical harmonic coefficients are computed by integrating the lighting environment function $L(\vec{V})$ against the corresponding spherical harmonic basis function [20]:

$$l_{n,m} \quad = \quad \int_\Omega L(\vec{V}) Y_{n,m}(\vec{V}) \, d\Omega. \tag{27}$$

Shown in Table I are nine lighting environment coefficients computed from five different light probe images. The light probes, Fig. 5, were captured in the following locations: Grace Cathedral, San Francisco (GRC); Galileo's Tomb, Florence (GAL); a Eucalyptus Grove, UC Berkeley (EUC); St. Peter's Basilica, Rome (STP); and the Uffizi Gallery, Florence (UFZ).[2]

[2]Light probe images ©1998, 1999 Paul Debevec, available at http://www.debevec.org/Probes.

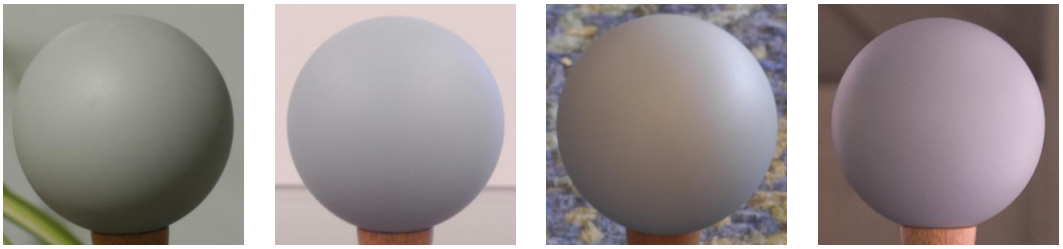

Fig. 6. Diffuse spheres photographed in four different lighting environments.

These lighting environment coefficients were used to render a Lambertian sphere in each of the five lighting environments, Fig. 5. Using the known geometry of these spheres, the lighting environment coefficients were estimated in two different ways: with 3-D surface normals from the visible side of the sphere, and with 2-D surface normals along the occluding contour. In both cases, the regularization term λ in Equation (21) was set to 0.01.

The estimation errors are reported in the last two columns of Table I. For the 2-D case, the errors are computed between the five estimated coefficients and the corresponding subset of actual coefficients ($l_{0,0}$, $l_{1,-1}$, $l_{1,1}$, $l_{2,-2}$, $l_{2,2}$). Overall, the errors are less than 0.001; for comparison, the average error between all ten pairs of different lighting environments is 0.13 with a minimum of 0.015.

B. Spheres

To test our ability to discriminate between lighting environments in real images, we photographed a diffuse sphere in 28 different locations with a 6.3 mega-pixel Nikon D-100 digital camera set to capture in high-quality JPEG mode. The focal length was set to 70 mm, the f-stop was fixed at $f/8$, and the shutter speed was varied to capture two or three exposures per location. In total, there were 68 images, four of which are shown in Fig. 6.

For each image, the Adobe Photoshop "Quick Selection Tool" was used to locate the occluding contour of the sphere from which both 2-D and 3-D surface normals could be estimated. The 3-D surface normals were used to estimate the full set of nine lighting environment coefficients and the 2-D surface normals along the occluding contour were used to estimate five coefficients. For both cases, the regularization term λ in Equation (21) was set to 0.01.

For each pair of images, the error, Equation (26), between the estimated coefficients was computed. In total, there were 2278 image pairs: 52 pairs were different exposures from the same location, and 2226 pairs were captured in different locations. The errors for all pairs for both models (3-D and 2-D) are shown in Fig. 7. In both plots, the 52 image pairs from the same location are plotted first (blue '+'), sorted by error. The 2226 pairs from different locations are plotted next (red '$\cdot$'). Note that the axes are scaled logarithmically.

For the 3-D case, the minimum error between an image pair from different locations is 0.0027 and the maximum error between an image pair from the same location is 0.0023. Therefore, the two sets of data, same location versus different location, are separated by a threshold of 0.0025.

For the 2-D case, thirteen image pairs (0.6%) fell below 0.0025. These image pairs correspond to lighting environments that are indistinguishable based on the five coefficient model. For example, two of these indistinguishable lighting environment pairs are shown in Fig. 8. In each plot, the red (dashed) and blue (dotted) lines are from different lighting environments, where the 2-D error between these environments is less than 0.0025. Both plots illustrate that different lighting environments can create similar intensity profiles, and low order approximations of these profiles will be unable to capture the differences. Therefore, while large errors indicate different lighting environments, small errors can only indicate indistinguishable lighting environments.

C. Photographs

To be useful in a forensic setting, lighting estimates from objects in the same lighting environment should be robust to differences in color and material type, as well as to geometric differences, since arbitrary objects may not have the full range of surface normals available. To test our algorithm under these conditions, we downloaded twenty images of multiple objects in natural lighting environments from Flickr [19], Fig. 9.

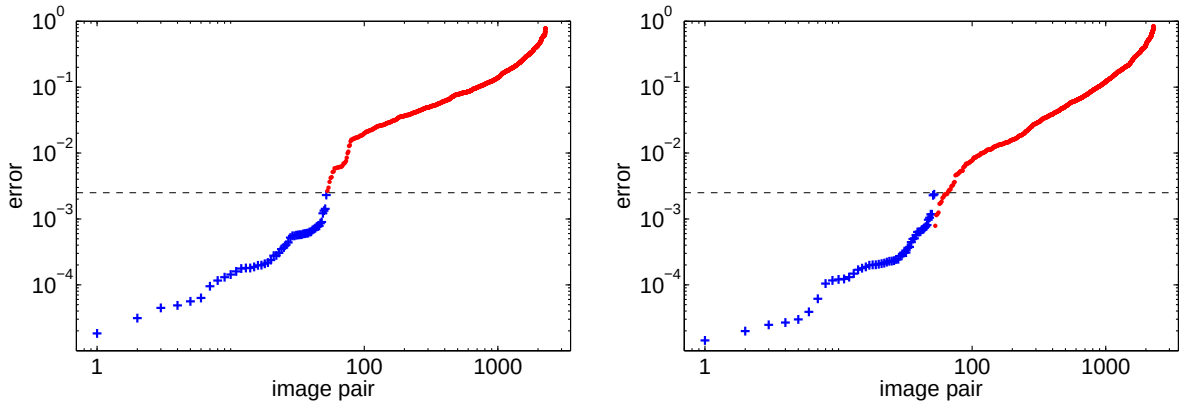

Fig. 7. Errors between image pairs corresponding to the same (blue '+') and different (red '·') locations using the full 9-parameter model with 3-D surface normals (left) and using the 5-parameter model with 2-D surface normals (right). Both the horizontal and vertical axes are scaled logarithmically.

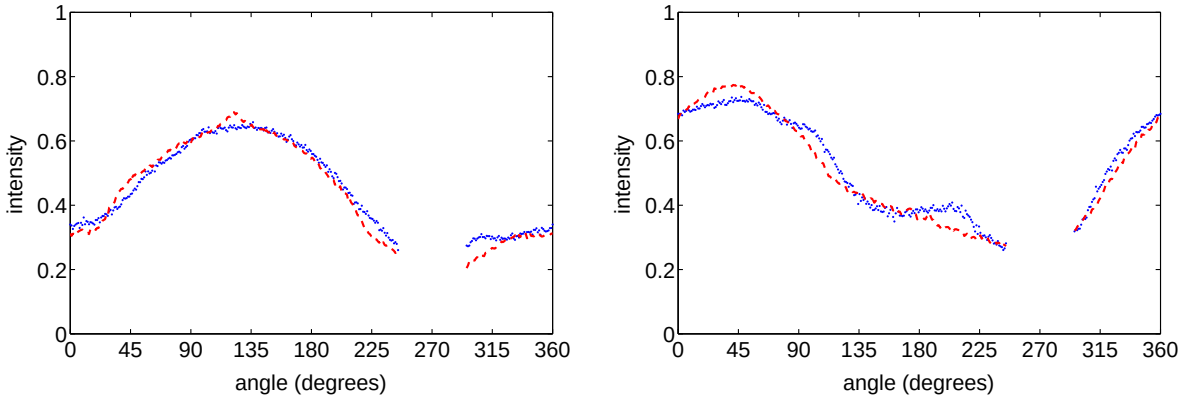

Fig. 8. Shown in each panel are intensity profiles from a pair of spheres in indistinguishable lighting environments. In each case, the error between the red dashed and blue dotted profiles is below the threshold of 0.0025. (The gap in the profiles corresponds to the sphere's mounting stand for which no intensity values are available.)

In each image, occluding contours of two to four objects were specified using a semi-automated approach. A coarse contour was defined by painting along the edge of the object using Adobe Photoshop. Each stroke was then automatically divided into quadratic segments, or regions, which were fit to nearby points with large gradients. The analyzed regions for all images are shown in Fig. 10. Analytic surface normals and intensities along the occluding contour were measured from the regions. With the 2-D surface normals and intensities, the five lighting environment coefficients were estimated, Equation (21). The regularization term λ in Equation (21) was increased to 0.1, which is larger than in the simulation due to an increasing sensitivity to noise (see Section III-D).

Across all twenty images, there were 49 pairs of objects from the same image and 1329 pairs of objects from different images. For each pair of objects, the error between the estimated coefficients was computed. For objects in the same image, the average error was 0.009 with a standard deviation of 0.007 and a maximum error of 0.027. For comparison, between objects in different images the average error was 0.295 with a standard deviation of 0.273. There were, however, 196 pairs of objects (15%) from different images that fell below 0.027. The lighting environments in these images (e.g., the two police images, the trees and skiers images, etc.) were indistinguishable using the five coefficient model.

For objects from the same image, the pair with the maximum error of 0.027 is the basketball and basketball player. The sweaty skin of the basketball player is somewhat shiny, a violation of the Lambertian assumption. In addition, the shoulders and arms of the basketball player provide only a limited extent of surface normals, making the linear system somewhat ill-conditioned. In contrast, the objects from the same image with the minimum error of 0.0001 are the left and right pumpkins on the bench. Both pumpkins provide a large extent of surface normals,

Fig. 9. Twenty images of multiple objects in natural lighting environments, see also Fig. 10.

Fig. 10. Superimposed on each image from Fig. 9 are the contours from which the surface normals and intensity values are extracted to form the matrix M and the corresponding vector $\vec{b}$, Equation (17).

over 200°, and the surfaces are fairly diffuse. Since the surfaces fit the assumptions and the linear systems are well-conditioned, the error between the estimated coefficients is small.

D. Sensitivity

In this section, we explore the sensitivity of the estimate to surface normal extent in the presence of additive noise and JPEG compression. Random lighting environments were generated by picking coefficients according to a unit-variance and zero-mean Gaussian distribution. To simulate natural lighting, the coefficients at order n were scaled so that the average power was proportional to $1/n^2$ [15]. From each lighting environment, we rendered images of

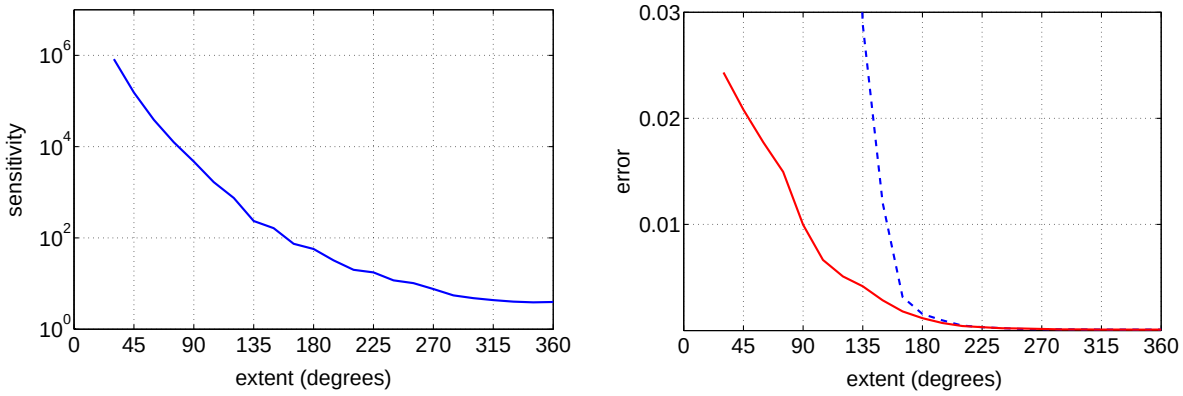

Fig. 11. Shown on the left is the sensitivity, Equation (28), of the least-squares problem, Equation (6), as a function of the surface normal extent (note that the vertical axis is scaled logarithmically). Shown on the right is the average error between the estimated and actual lighting environment vectors as a function of surface normal extent. Each data point corresponds to the error averaged over 2000 random lighting environments. The dashed blue curve corresponds to the unconditioned solution, Equation (17), and is largely unstable for a surface normal extent less than $180°$. The solid red curve corresponds to the conditioned solution, Equation (19), and is substantially more stable.

spheres and added Gaussian noise with standard deviation equal to 5% of the intensity range of the image. From each image, the coefficient vector $\vec{v}$ was estimated, Equation (21), with $\lambda = 0.01$. The surface normals were limited to a specified extent, from 30 to 360 degrees, about the primary illuminant direction. The surface normal extent affects the stability of the estimate $\vec{v}$, which can be formalized by computing the sensitivity of $\vec{v}$ to perturbations in M [21]:

$$\kappa(M) + \frac{\kappa(M)^2 \tan\theta}{\eta}, \tag{28}$$

where $\kappa(M) = \sigma_{\max}/\sigma_{\min}$ is the condition number of the matrix M (ratio of the largest to smallest singular value), and θ and η are:

$$\theta = \cos^{-1}\left(\frac{\|M\vec{v}\|}{\|\vec{b}\|}\right), \tag{29}$$

$$\eta = \sigma_{\max}\, \|\vec{v}\|/\|M\vec{v}\|. \tag{30}$$

As shown in the left panel of Fig. 11, the sensitivity, Equation (28), increases dramatically as the extent of surface normals decreases, indicating potential instability of the estimate $\vec{v}$. Shown in the right panel of Fig. 11 is the error averaged over 2000 random environments per surface normal extent for both the conditioned (solid red) and unconditioned systems (dashed blue), Equations (18) and (21). Note that the conditioned system provides considerably more accurate results when the surface normal extent is below $180°$.

The sensitivity to JPEG compression was also tested. As above, we generated random lighting environments and rendered images of spheres in these environments. These images were then saved with a JPEG quality between 5 and 100 (in a range of $[0, 100]$). The lighting environment coefficients were estimated from surface normals spanning a range of $135°$. For JPEG quality of 5, the average error over 2000 random trials is 0.03. For a quality between 10–35, the average error is 0.01; for a quality between 40–65, the average error is 0.005; and for a quality between 70–100, the average error is 0.002. Note that for JPEG quality between 40–100, these error are comparable or less than the errors introduced from additive noise, Fig. 11.

E. Forgeries

We created three forgeries by mixing and matching several of the images in Fig. 9, and downloaded one forgery from *Worth1000*, a Photoshop contest website [22]. These forgeries are shown in Fig. 12 and Fig. 13.

Regions along the occluding contour of two to four objects in each image were selected for analysis. These regions are superimposed on the images in the right column of Fig. 12 and Fig. 13. Surface normals and intensities along these occluding contour were extracted, from which the five lighting environment coefficients were estimated, Equation (21), with the regularization term $\lambda = 0.1$.

TABLE II
ERRORS BETWEEN PAIRS OF OBJECTS IN THE FORGERIES OF FIG. 12 AND FIG. 13

Police		Umbrellas		Soldiers		Snoop Dogg	
pair	error	pair	error	pair	error	pair	error
1, 2	0.006	1, 2	0.010	2, 3	0.002		
3, 4	0.004	3, 4	0.004				
1, 3	0.047	1, 3	0.152	1, 2	0.109	1, 2	0.388
1, 4	0.033	1, 4	0.194	1, 3	0.138		
2, 3	0.076	2, 3	0.229				
2, 4	0.054	2, 4	0.277				

Shown in each figure is a sphere rendered with the estimated coefficients. These spheres qualitatively show discrepancies between the lighting. The calculated errors between object pairs are summarized in Table II. For all pairs of objects originally in the same lighting environment (above the horizontal line), the average error is 0.005 with maximum error of 0.01. For pairs of objects from different lighting environments (below the horizontal line), the average error is 0.15 with a minimum error of 0.03.

IV. DISCUSSION

When creating a composite of two or more people, it is often difficult to exactly match the lighting, even if the lighting seems perceptually consistent. The reason for this is that complex lighting environments (multiple light sources, diffuse lighting, directional lighting) give rise to complex and subtle lighting gradients and shading effects in the image. Under certain simplifying assumptions (distant light sources and diffuse surfaces), arbitrary lighting environments can be modeled with a 9-dimensional model. This model approximates the lighting with a linear combination of spherical harmonics. We have shown how to approximate a simplified 5-dimensional version of this model from a single image, and how to stabilize the model estimation in the presence of noise. Inconsistencies in the lighting model across an image are then used as evidence of tampering.

We showed the efficacy of this approach on a broad range of simulated images, photographic images, and visually plausible forgeries. In each case, the model parameters can be well approximated, from which differences in lighting can typically be detected. There are, however, instances when different lighting environments give rise to similar model coefficients – in these cases the lighting differences are indistinguishable.

The ability to estimate complex lighting environments was motivated by our earlier work in which we showed how to detect inconsistencies in the direction to an illuminating light source [9], [10]. Our current work generalizes this approach by allowing us to estimate more complex models of lighting and in fact can be adapted to estimate the direction to a single light source. Specifically, by considering only the two first-order spherical harmonics, $Y_{1,-1}(\cdot)$ and $Y_{1,1}(\cdot)$, the direction to a light source can be estimated as $\tan^{-1}(l_{1,-1}/l_{1,1})$.

While any forensic tool is vulnerable to counter-measures, the precise matching of lighting in an image can be difficult, although certainly not impossible. And we have seen that even if different parts of an image are photographed under somewhat different lighting conditions, these differences can become indistinguishable in our lower-order model. We expect, nevertheless, that this work and a growing body of forensic tools will make it increasingly more difficult to create convincing forgeries.

ACKNOWLEDGMENTS

This work was supported by a Guggenheim Fellowship, a gift from Adobe Systems, Inc., a gift from Microsoft, Inc., a grant from the United States Air Force (FA8750-06-C-0011), and by the Institute for Security Technology Studies at Dartmouth College under grant 2005-DD-BX-1091 from the Bureau of Justice Assistance and Award Number 2006-CS-001-000001 from the U.S. Department of Homeland Security. Points of view or opinions in this document are those of the author and do not represent the official position or policies of the U.S. Department of Justice, the U.S. Department of Homeland Security, or any other sponsor. Thanks to the photographers on Flickr who granted permission for the use of their images: Bram Appleton, Charro Badger, Kenneth Barrett, Zack Brewer, David Bruce, Jim Ferguson, Marcus Frieze, Tommy Hemmert Jørgensen, Stig Nygaard, Nathan Phillips, and Clyde Robinson, Jr.

Fig. 12. Shown on the left are three forgeries: the ducks, swans, and football coach were each added into their respective images. Shown on the right are the analyzed regions superimposed in white, and spheres rendered from the estimated lighting coefficients (see also Table II).

APPENDIX A

The spherical harmonics of orders zero to two in terms of the Cartesian coordinates of the surface normal $\vec{N} = \begin{pmatrix} x & y & z \end{pmatrix}$:

$$
\begin{aligned}
Y_{0,0}(\vec{N}) &= \tfrac{1}{\sqrt{4\pi}} & Y_{1,-1}(\vec{N}) &= \sqrt{\tfrac{3}{4\pi}}y & Y_{1,0}(\vec{N}) &= \sqrt{\tfrac{3}{4\pi}}z \\
Y_{1,1}(\vec{N}) &= \sqrt{\tfrac{3}{4\pi}}x & Y_{2,-2}(\vec{N}) &= 3\sqrt{\tfrac{5}{12\pi}}xy & Y_{2,-1}(\vec{N}) &= 3\sqrt{\tfrac{5}{12\pi}}yz \\
Y_{2,0}(\vec{N}) &= \tfrac{1}{2}\sqrt{\tfrac{5}{4\pi}}(3z^2-1) & Y_{2,1}(\vec{N}) &= 3\sqrt{\tfrac{5}{12\pi}}xz & Y_{2,2}(\vec{N}) &= \tfrac{3}{2}\sqrt{\tfrac{5}{12\pi}}(x^2-y^2)
\end{aligned}
$$

APPENDIX B

In this appendix, we derive a relationship between image intensity, irradiance, and exposure time under a nonlinear camera response function.

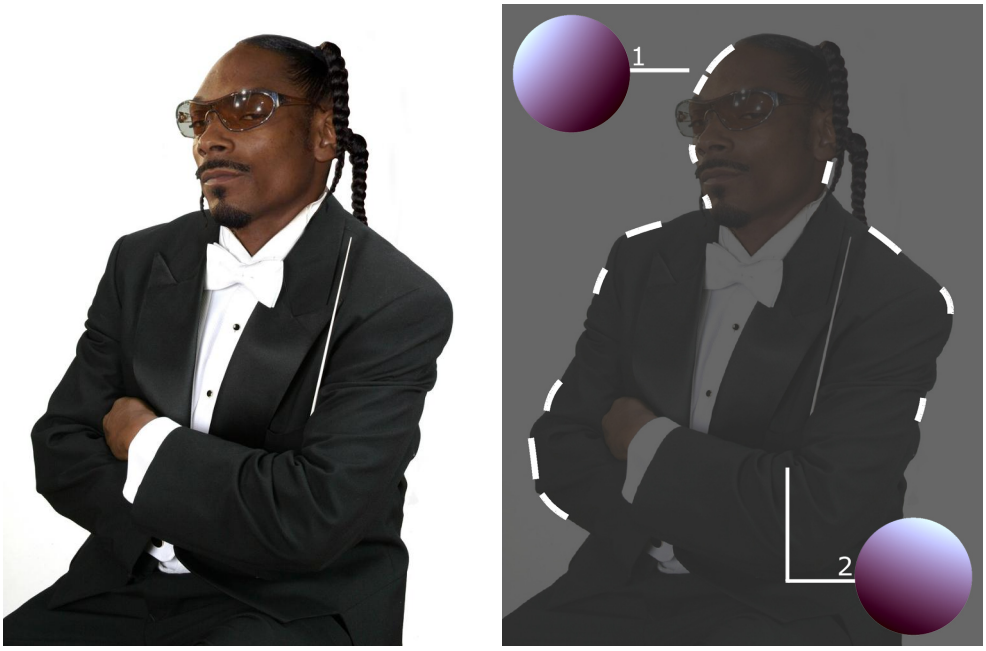

Fig. 13. Shown on the left is a forgery where the head of rapper Snoop Dogg has been placed on the body of an orchestra conductor. Shown on the right are the analyzed regions superimposed in white, and spheres rendered from the estimated lighting coefficients (see also Table II).

Consider a nonlinear camera response function $f(xt)$ that maps the product of irradiance x and exposure time t to intensity. In addition, suppose there is an object in an arbitrary lighting environment and the minimum and maximum irradiance values for the object are x_1 and x_2. Let t_1 and t_2 be two different exposure times, and without loss of generality, we assume $t_1 = 1$ and $t_2 > t_1$. The intensities for the first exposure, $t_1 = 1$, can be approximated by a truncated Taylor series expanded about the midpoint of irradiance values for the object:

$$\begin{aligned} f(xt_1) &\approx f(m_1) + f'(m_1)(xt_1 - m_1) && (31) \\ &= f(m_1) + f'(m_1)(x - m_1), && (32) \end{aligned}$$

where $m_1 = (x_1 + x_2)/2$. Similarly, the intensities for the second exposure, $f(xt_2)$, can be approximated by a truncated Taylor series expanded about the midpoint of the scaled irradiance values:

$$f(xt_2) \approx f(m_2) + f'(m_2)(xt_2 - m_2), \quad (33)$$

where $m_2 = (t_2x_1 + t_2x_2)/2 = t_2m_1$.

From Equations (32) and (33), the relationship between the intensities due to a change in exposure is given by:

$$\begin{aligned} f(xt_2) &\approx f(m_2) + f'(m_2)(xt_2 - m_2) \\ &= f(m_2) + f'(m_2)(xt_2 - m_2) + \left[\frac{f'(m_2)}{f'(m_1)}t_2 f(xt_1) - \frac{f'(m_2)}{f'(m_1)}t_2 f(xt_1)\right] \\ &= f(m_2) + f'(m_2)(xt_2 - m_2) + \alpha f(xt_1) - \alpha f(m_1) - f'(m_2)t_2(x - m_1) \\ &= f(m_2) + \alpha f(xt_1) - \alpha f(m_1) \\ &= \alpha f(xt_1) + \beta, \quad (34) \end{aligned}$$

where $\alpha = f'(m_2)t_2/f'(m_1)$ and $\beta = f(m_2) - \alpha f(m_1)$. Therefore, a change of exposure time under a nonlinear camera response function can be approximated by a linear change in intensity values, where the multiplicative term α and the additive term β both depend on the exposure time t_2.

This relationship can be seen intuitively in Fig. 14. Shown in the figure is the camera response function $f(\cdot)$ and its first-order approximations for two different exposures about the points m_1 and m_2. Note that approximations to $f(\cdot)$ at these two points are related by a multiplicative factor (slope) and an additive factor (offset).

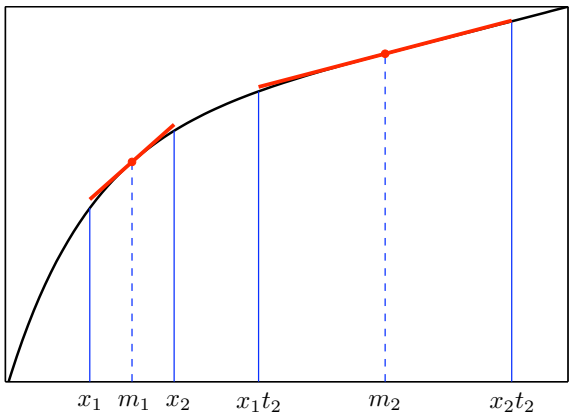

Fig. 14. Shown are the camera response function $f(\cdot)$ and its first-order approximations for two different exposures about the points m_1 and m_2.

APPENDIX C

In this appendix, we derive a compact expression for the error between two lighting environment vectors $\vec{v}_1$ and $\vec{v}_2$. We begin with the 2-D error, estimated using matrix M from Equation (17).

The average value of an irradiance profile, Equation (23), can be derived by integrating the $m_{i,j}$ functions in Equation (14) around the unit circle. Each of these functions integrates to zero, thus the average value is simply the ambient term A. From this observation, the zero-mean irradiance profile for vector $\vec{v}$ is given by:

$$\vec{x} \quad = \quad \hat{M}\vec{v}, \tag{35}$$

where $\hat{M}$ is the matrix M from Equation (17) with the first column of ones replaced with zeros. The numerator of Equation (25) can then be rewritten as an inner product of irradiance profiles:

$$\vec{x}_1^T \vec{x}_2 = (\hat{M}\vec{v}_1)^T(\hat{M}\vec{v}_2) = \vec{v}_1^T(\hat{M}^T\hat{M})\vec{v}_2 = \vec{v}_1^T Q_2 \vec{v}_2. \tag{36}$$

The terms of the matrix Q_2 are derived by integrating products of pairs of functions $m_{i,j}$ around the unit circle and normalizing by $\frac{1}{2\pi}$. Since these functions are orthogonal, the off-diagonal terms of the matrix Q_2 are zero. The terms on the diagonal are $\left(0 \quad \frac{\pi}{6} \quad \frac{\pi}{6} \quad \frac{15\pi}{512} \quad \frac{15\pi}{512}\right)$.

For the 3-D error, we limit the correlation to the visible hemisphere by restricting the bounds of the integration to values where $z \geq 0$. Since the coefficient vectors $\vec{v}_1$ and $\vec{v}_2$ are estimated from surface normals that face the camera, irradiance estimates for surface normals facing away from the camera (i.e., behind the object) are often numerically unstable. Restricting the integration to the visible hemisphere reduces the effect of this instability.

On the hemisphere, the average value of the irradiance profile is derived by integrating Equation (10) and normalizing by $\frac{1}{2\pi}$ (2π steradians of solid angle on the hemisphere):

$$\frac{1}{2\pi}\int_{\Omega_{z\geq 0}} E(\vec{N})\, d\Omega \quad = \quad l_{0,0}\frac{\pi}{2\sqrt{\pi}} + l_{1,0}\frac{\pi}{6}\sqrt{\frac{3}{\pi}}. \tag{37}$$

The zero-mean irradiance profile over the hemisphere using M from Equation (12) is therefore:

$$\vec{x} = M\vec{v} - B\vec{v} = (M - B)\vec{v}, \tag{38}$$

where B is a matrix the same size as M with $\frac{\pi}{2\sqrt{\pi}}$ in column 1 and $\frac{\pi}{6}\sqrt{\frac{3}{\pi}}$ in column 3. Following the derivation of Equation (36), a similar expression can be derived for the correlation between the zero-mean irradiance profiles on the hemisphere:

$$\vec{x}_1^T \vec{x}_2 = ((M - B)\vec{v}_1)^T((M - B)\vec{v}_2) = \vec{v}_1^T(M - B)^T(M - B)\vec{v}_2 = \vec{v}_1^T Q_3 \vec{v}_2. \tag{39}$$

The matrix Q_3 can be expanded as

$$Q_3 \quad = \quad M^T M - M^T B - B^T M + B^T B. \tag{40}$$

the components of Q_3 determined by symbolic integration software to be:

$$Q_3 = \begin{pmatrix} 0 & & & & & & & & \\ & \frac{\pi}{9} & & & & \frac{\sqrt{5}\pi}{64} & & & \\ & & \frac{\pi}{36} & & & & \frac{\pi\sqrt{5}}{64\sqrt{3}} & & \\ & & & \frac{\pi}{9} & & & & \frac{\sqrt{5}\pi}{64} & \\ & & & & \frac{\pi}{64} & & & & \\ & \frac{\sqrt{5}\pi}{64} & & & & \frac{\pi}{64} & & & \\ & & \frac{\pi\sqrt{5}}{64\sqrt{3}} & & & & \frac{\pi}{64} & & \\ & & & \frac{\sqrt{5}\pi}{64} & & & & \frac{\pi}{64} & \\ & & & & & & & & \frac{\pi}{64} \end{pmatrix}.$$

REFERENCES

[1] J. Fridrich, D. Soukal, and J. Lukáš, "Detection of copy-move forgery in digital images," in *Proceedings of Digital Forensic Research Workshop*, August 2003.

[2] A. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions," Department of Computer Science, Dartmouth College, Tech. Rep. TR2004-515, 2004.

[3] T.-T. Ng and S.-F. Chang, "A model for image splicing," in *IEEE International Conference on Image Processing (ICIP)*, Singapore, October 2004.

[4] A. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of re-sampling," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 758–767, 2005.

[5] İ. Avcıbaş, S. Bayram, N. Memon, B. Sankur, and M. Ramkumar, "A classifier design for detecting image manipulations," in *2004 International Conference on Image Processing, ICIP '04*, vol. 4, 2004, pp. 2645–2648.

[6] J. Lukáš, J. Fridrich, and M. Goljan, "Detecting digital image forgeries using sensor pattern noise," in *Proceedings of the SPIE*, vol. 6072, 2006.

[7] M. K. Johnson and H. Farid, "Exposing digital forgeries through chromatic aberration," in *ACM Multimedia and Security Workshop*, Geneva, Switzerland, 2006.

[8] ——, "Metric measurements on a plane from a single image," Department of Computer Science, Dartmouth College, Tech. Rep. TR2006-579, 2006.

[9] ——, "Exposing digital forgeries by detecting inconsistencies in lighting," in *ACM Multimedia and Security Workshop*, 2005.

[10] ——, "Exposing digital forgeries through specular highlights on the eye," in *9th International Workshop on Information Hiding*, Saint Malo, France, 2007.

[11] R. Ramamoorthi and P. Hanrahan, "On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object," *Journal of the Optical Society of America A*, vol. 18, pp. 2448–2559, 2001.

[12] R. Basri and D. W. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 218–233, 2003.

[13] P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, 1997, pp. 369–378.

[14] P. Nillius and J.-O. Eklundh, "Automatic estimation of the projected light source direction," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001.

[15] R. O. Dror, A. S. Willsky, and E. H. Adelson, "Statistical characterization of real-world illumination," *Journal of Vision*, vol. 4, no. 9, pp. 821–837, 2004.

[16] G. H. Golub, P. C. Hansen, and D. P. O'Leary, "Tikhonov regularization and total least squares," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 1, pp. 185–194, 1999.

[17] M. Pharr and G. Humphreys, *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann, 2004.

[18] P. Debevec, "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*. ACM Press, 1998, pp. 189–198.

[19] [Online]. Available: http://www.flickr.com

[20] R. Ramamoorthi and P. Hanrahan, "An efficient representation for irradiance environment maps," in *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM Press, 2001, pp. 497–500.

[21] L. N. Trefethen and D. Bau, III, *Numerical Linear Algebra*. SIAM, 1997.

[22] [Online]. Available: http://www.worth1000.com