3D LIGHTING-BASED IMAGE FORGERY DETECTION USING SHAPE-FROM-SHADING

Wei Fan[a,b], Kai Wang[a], François Cayre[a]†, and Zhang Xiong[b]*

[a]GIPSA-Lab, 11 rue des Mathématiques, F-38402 St-Martin d'Hères Cedex, France
[b]School of Computer Science and Engineering, Beihang University, Beijing 100191, P. R. China

ABSTRACT

This paper concentrates on lighting-based forensics. We first show how to fool the forgery detector based on 2D lighting coefficients using a simple counter-forensic strategy. This intermediary result advocates the use of more involved 3D lighting coefficients for forensics purposes. Such a research line means that we need at least an approximation of the 3D surface of the suspect object. Contrary to previous approaches that concentrated on particular kind of shapes (e.g. human faces), we propose a promising approach based on shape-from-shading. This new 3D lighting-based forensic method is more general as the 3D shape is learned from the picture itself. Furthermore, the results are in par with the less general state-of-the-art methods.

***Index Terms*—** Digital forensics, image forgery detection, complex lighting environment, spherical harmonics, shape-from-shading, counter-forensics

1. INTRODUCTION

With the increasing popularity and sophistication of photo manipulation software, our trust on the authenticity of digital images is decreasing. Doctored images can be easily found in our daily life, and have been used, for instance, in advertising, political and personal attacking, and forgery of scientific results. Accordingly, many image forensic techniques have been proposed during the last decade [1, 2], with the objective to faithfully detect image forgeries. Compared to the authentication based on digital watermarking, forensic techniques can assess the authenticity of an image in a passive and blind way, without resorting to previously embedded information (i.e. the watermark). These techniques make assumption that manipulating an image will probably disturb the intrinsic property, either geometrical, physical or statistical, of the authentic image. Therefore, inconsistencies in these properties over the image can be considered as an evidence of tampering.

In this paper, we concentrate on the physics-based image forgery detection that examines inconsistencies in lighting under complex natural illumination. In practice, it is very difficult to forge physically consistent lighting when splicing objects from different images, meanwhile experiments show that such inconsistencies may be difficult to perceive by human eyes [3]. Lighting-based forensics have been addressed by Johnson, Farid and Kee, under respectively simple directional lighting [4], 2D complex lighting [5] and 3D complex lighting [6]. The basic idea of the last two methods is to first recover the lighting environment, as represented by a group of spherical harmonics coefficients [7], and then compare the coefficients estimated from different parts of the image. Our new forensic method also follows this approach.

The work presented in this paper can be thought of as one iteration of lighting-based counter-forensics and counter-counter-forensics. We show shortcomings of a previous forensic method and demonstrate the possibility of developing a new lighting-based image forensic tool relying on the most recent results from shape-from-shading research [8]. Our contributions are summarized as follows:

- First, we show, through two simple examples, that the 2D lighting-based forensic method [5] is not completely reliable and may be vulnerable to counter-forensic attacks.
- Second, we use the shape-from-shading technique [8] for lighting environment estimation, which is new in the field of image forgery detection. Our motivation was to use 3D surface normals to estimate a more complete description of the lighting environment.
- Finaly, compared to the 3D lighting-based forensic method in [6], which relies on a predefined 3D model and is specific to human face images, our method does not need such a 3D model and seems more generic.

The remainder of this paper is organized as follows: Sec. 2 presents some background knowledge on lighting environment estimation, Sec. 3 depicts two simple examples to attack the 2D lighting-based forensic method in [5], Sec. 4 describes our new 3D lighting-based forensic tool, Sec. 5 shows some experimental results, and we draw conclusions in Sec. 6.

2. LIGHTING ESTIMATION

In order to model complex lighting environment, we assume that: (a) the lighting is distant; (b) the surfaces are convex and

*Wei Fan performed this work while at GIPSA-Lab on the grant from China Scholarship Council (No. 2011602067).
†This work was funded, in part, by French ANR Estampille.

Lambertian; (c) the surface reflectance is constant; and (d) the camera response is linear.

Denote $L(\omega)$ as the illumination function describing the intensity of the incident light from direction ω which is a unit vector. Let $R(\mathbf{n}, \omega)$ be the reflectance function of the surface, where $\mathbf{n}$ is the unit length surface normal vector. On the convex surface of a Lambertian object, we suppose there are no cast shadows or interreflections [7]. Hence, the irradiance is only due to the lighting environment, and it can be described as a convolution over the upper hemisphere $\Omega(\mathbf{n})$:

$$E(\mathbf{n}) = \int_{\Omega(\mathbf{n})} L(\omega) R(\mathbf{n}, \omega) d\omega. \tag{1}$$

A common way to approximate this function is using spherical harmonics to expand both the illumination function and the reflectance function to yield:

$$E(\mathbf{n}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \hat{A}_l L_{l,m} Y_{l,m}(\mathbf{n}), \tag{2}$$

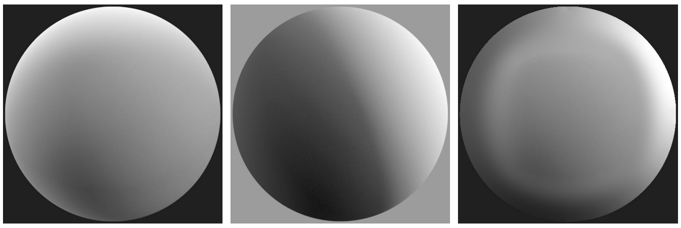
where $Y_{l,m}(\cdot)$ is the m^{th} spherical harmonic of order l, with $l \geq 0$ and $-l \leq m \leq l$. $L_{l,m}$ are the spherical harmonics coefficients representing the lighting environment. Constants $\hat{A}_l$ are the Lambertian reflectance coefficients, which decay rapidly when $l > 2$. Consequently with $l \leq 2$, $E(\mathbf{n})$ can be well approximated using only the first nine terms.

It is the surface diffuse albedo ρ, which is the multiplicative factor mapping the image irradiance to the intensity. Without loss of generality, we assume $\rho = 1$ for simplicity and $I(\mathbf{p}) = E(\mathbf{n_p})$ at the point $\mathbf{p}$ on a Lambertian surface. Thus the lighting coefficients are estimated up to an unknown factor. Given the estimated surface normals at $k > 9$ points on the surface of an object and their intensties, it is possible to estimate the nine 3D lighting coefficients by

$$\begin{bmatrix} \mathbf{i}_1^T \\ \vdots \\ \mathbf{i}_k^T \end{bmatrix} = \begin{bmatrix} \hat{A}_0 Y_{0,0}(\mathbf{n}_1) & \dots & \hat{A}_2 Y_{2,2}(\mathbf{n}_1) \\ \vdots & \ddots & \vdots \\ \hat{A}_0 Y_{0,0}(\mathbf{n}_k) & \dots & \hat{A}_2 Y_{2,2}(\mathbf{n}_k) \end{bmatrix} \begin{bmatrix} \mathbf{l}_{0,0}^T \\ \vdots \\ \mathbf{l}_{2,2}^T \end{bmatrix}. \tag{3}$$

Note $\mathbf{i} = [I^r \; I^g \; I^b]^T$ is the image intensity for RGB color images, and $\mathbf{l}_{l,m} = [L_{l,m}^r \; L_{l,m}^g \; L_{l,m}^b]^T$ is the vector containing the lighting coefficients corresponding to spherical harmonic $Y_{l,m}$ in red, green and blue channels, respectively.

Obviously, the estimation of the 3D lighting coefficients requests 3D surface normals. And without multiple images or known geometry, it is always difficult to satisfy this requirement [5]. Nevertheless, under the assumption of orthographic projection, the z-component of the surface normal is zero along the occluding contours of an object. Therefore, the spherical harmonics $Y_{1,0}$, $Y_{2,-1}$ and $Y_{2,1}$ are all zeros, and $Y_{2,0} = -\sqrt{5/16\pi}$ becomes a constant. We add the terms corresponding to spherical harmonics $Y_{0,0}$ and $Y_{2,0}$ together and factor $\hat{A}_0$ and $\hat{A}_2$ to the lighting coefficients. Denote $\hat{A}'_0 = Y'_{0,0} = 1$, we can estimate $L'_{0,0} = \sqrt{\pi/4}L_{0,0} - \sqrt{5\pi/256}L_{2,0}$. Hence, along the boundaries of an object, the five 2D lighting coefficients we are able to compute are $L'_{0,0}$, $L_{1,-1}$, $L_{1,1}$, $L_{2,-2}$, and $L_{2,2}$.



(a) GRC (b) GAL (c) Forgery

Fig. 1. Mapping (c) on (a) of the 2D lighting coefficients taken from (b).

Table 1. Pairwise 2D lighting differences

	Errors
GRC vs. GAL	0.2370
Forgery vs. GRC	0.2336
Forgery vs. GAL	2.6228×10^{-5}

Forgeries are detected by comparing the lighting coefficients estimated from different objects in an image. In [5], the authors proposed a distance measure between two lighting environments that is normalized to the interval $[0, 1]$. Here, we also use this measure to evaluate lighting differences.

3. COUNTER FORENSICS

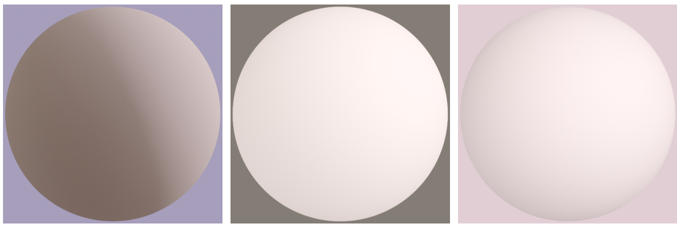
In [5], although the authors proposed the 3D lighting-based forensic model, due to the difficulty of 3D normal estimation, their main approach for forgery detection is still concentrated on 2D lighting-based forensic method. In this section, we introduce two counter-forensic methods to show how 2D lighting-based forensic method can be vulnerable.

3.1. Fooling the 2D Lighting-based Detector

We rewrite Eq. (3) in matrix form $\mathbf{I} = \mathbf{ML}$. The lighting coefficients are obtained as the least-squares solution to the system: $\mathbf{L} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{I}$. we can see that the estimation of lighting coefficients needs both the surface normals (determining $\mathbf{M}$) and the image intensities ($\mathbf{I}$).

Lighting-based forensics compare the lighting coefficients from different objects to decide whether the image is a forgery. The goal of counter-forensics is to fool the detector so that it obtains different lighting coefficients. For an object in the image, a simple strategy is to first keep the surface normals unchanged to yield the same $\mathbf{M}$; meanwhile, if we succeed in modifying the pixel values along the occluding contours, i.e. modifying $\mathbf{I}$, different lighting coefficients $\mathbf{L}$ can be generated.

The weakness of the 2D lighting-based forensic method we are targeting is that it uses only the information along the object boundaries. It should therefore be possible to create a



(a) GAL (b) GAL forgery-1 (c) GAL forgery-2

Fig. 2. Rendered spheres from actual and fake lighting environments.

Table 2. Comparison of estimated lighting coefficients

		R	G	B
Fig. 2-(a)	3D	6.3×10^{-4}	6.9×10^{-4}	6.6×10^{-4}
	2D	5.0×10^{-5}	4.8×10^{-5}	5.6×10^{-5}
Fig. 2-(b)	3D	0.1135	0.1233	0.1093
	2D	3.5×10^{-4}	3.8×10^{-4}	3.4×10^{-4}
Fig. 2-(c)	3D	0.2008	0.2059	0.1947
	2D	3.3×10^{-4}	3.7×10^{-4}	3.2×10^{-4}

fake picture by modifying the pixel values along the occluding contours, as long as the contour shape is kept the same in order to yield the same boundary normals.

We have tested this idea on the synthetic images rendered using the *pbrt* environment [9] with lighting probes maintained by Paul Debevec [10]. Fig. 1-(a) and -(b) are the rendered Lambertian spheres with lighting probes captured in Grace Cathedral, San Francisco (GRC) and Galileo's Tomb, Florence (GAL)[1]. In order to create the forgery of Fig. 1-(c), we doctor the boundary pixel values according to $\mathbf{I}_{forg} = \mathbf{M}_{grc}\mathbf{L}_{gal}$. We also use interpolation in the area between the boundary and the central part (which is kept the same as the original image of Fig. 1-(a)) to smooth the image. A visually more plausible forgery compared with Fig. 1-(c), can be obtained, by using advanced image processing algorithms.

According to the thresholds reported in [5], the results in Table 1 show that the image in Fig. 1-(c) successfully fools the 2D lighting-based forensic method, which mistakenly considers that Fig. 1-(b) and -(c) are from the same lighting environment. However, the truth is that the picture in Fig. 1-(c) is a forgery created from -(a) using the 2D lighting coefficients estimated from -(b). Note that Fig. 1 only shows the green channel of the images, but the modification in red/blue channels can be achieved similarly.

3.2. Building upon 3D Lighting Environments

The 2D lighting-based forensic method is only able to estimate five lighting coefficients: $L'_{0,0} = \sqrt{\pi/4}L_{0,0} - \sqrt{5\pi/256}L_{2,0}$, $L_{1,-1}$, $L_{1,1}$, $L_{2,-2}$ and $L_{2,2}$, which correspond to non-z-component-related spherical harmonics. Among the five 2D lighting coefficients, note that $L'_{0,0}$ is the combination of $L_{0,0}$ and $L_{2,0}$. Hence, if there are two lighting environments with similar $L_{0,0}$, $L_{2,0}$, $L_{1,-1}$, $L_{1,1}$, $L_{2,-2}$ and $L_{2,2}$ but different $L_{1,0}$, $L_{2,-1}$ and $L_{2,1}$, the 2D lighting-based forensic method would fail to distinguish them. After simulation, we verified that this weakness may be made use of by opponents to perform counter-forensic attacks.

Based on the ground truth of the nine 3D lighting coefficients of lighting environment GAL, we create two lighting environment forgeries. Fig. 2-(a) is the rendered sphere using the original lighting environment GAL, while Fig. 2-(b) and -(c) are rendered using GAL forgeries. In GAL forgery-1, the modification is applied to $L_{1,0}$, $L_{2,-1}$ and $L_{2,1}$, while one more coefficient $L_{2,0}$ is added to the modification in GAL forgery-2. And in both cases, $L_{0,0}$ is modified due to the fact that we should ensure that each value in the modified lighting environment should be positive. But it does not increase the lighting environment difference [5].

The figures in Table 2 are the estimation errors of RGB channels, in both the 3D and 2D cases, as compared with the ground truth of the GAL lighting environment. The small errors in the fourth and the sixth rows in Table 2 indicate that the 2D lighting-based forensic method is successfully attacked. 3D lighting-based forensics is a more reliable choice as it still can distinguish different lighting environments with similar 2D lighting coefficients.

4. IMPROVED LIGHTING-BASED FORENSICS

As explained in Sec. 3, it is relatively easy for opponents to fool the 2D lighting-based forensic method by modifying the boundary pixels. Moreover, the 2D method also fails to detect different 3D lighting environments that have similar 2D lighting coefficients.

A natural extension consists of using 3D lighting environments. Estimating a 3D lighting environment involves that some 3D information about the shape of the underlying object is available. Such an approach was already suggested by Kee and Farid [6]. They focused on detecting forgeries for human faces by matching a 3D model of a face on the image.

We propose a different approach by using shape-from-shading (SFS) [8] to estimate the 3D normals of the underlying object. Our goal is to produce a more general 3D lighting-based forensic tool that works on objects of arbitrary shape.

As shown above, the first nine spherical harmonics ($l \leq 2$) are either constant ($l = 0$), linear ($l = 1$) or quadratic ($l = 2$). As the 1st-order approximation of the Lambertian irradiance can capture up to 87.5% of the light energy [8], the image formation model can be simplified to a linear problem from Eq. (3):

$$\begin{bmatrix} \mathbf{i}_1^T \\ \vdots \\ \mathbf{i}_k^T \end{bmatrix} = \begin{bmatrix} \mathbf{n}_1^T & 1 \\ \vdots & \vdots \\ \mathbf{n}_k^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{a}^T \end{bmatrix}, \tag{4}$$

where $\mathbf{A} = \hat{A}_1[\mathbf{l}_{1,1} \;\; \mathbf{l}_{1,-1} \;\; \mathbf{l}_{1,0}]$ and $\mathbf{a} = \hat{A}_0\mathbf{l}_{0,0}$.

[1] http://www.pauldebevec.com/Probes

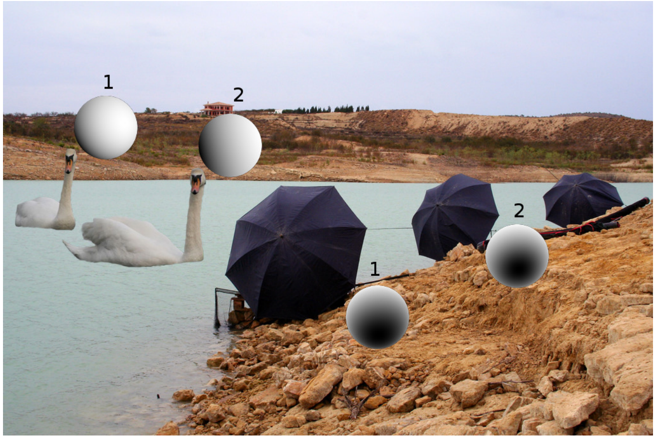


Fig. 4. A forgery with two swans inserted, and the lighting spheres from the estimated 3D lighting coefficients.

The main idea to recover the unit surface normal vector $\mathbf{n}^*$ is to solve the following quadratically constrained linear least-squares problem [8]:

$$\mathbf{n}^* = \arg\min_{\mathbf{n}} \|\mathbf{A}\mathbf{n} - \mathbf{b}\|^2, s.t. \|\mathbf{n}\| = 1, \tag{5}$$

where $\mathbf{b} = \mathbf{i} - \mathbf{a}$. Once the recovered surface normals are obtained, we can use the model of Eq. (3) to compute the nine 3D lighting coefficients.

The process of 3D lighting coefficients estimation based on SFS is enumerated as follows:

1. Use bicubic interpolation to coarsely estimate the surface normals $\{\mathbf{n}_0\}$ of the target object;
2. Estimate the linear/constant lighting coefficients $\mathbf{A}$ and $\mathbf{a}$ by Eq. (4);
3. Solve least-squares problem Eq. (5) at each point of the surface to recover the surface normals $\{\mathbf{n}_r\}$;
4. Add a smoothness constraint [8] to obtain the global optimum of the surface normals $\{\mathbf{n}_s\}$;
5. Recompute $\mathbf{A}$ and $\mathbf{a}$ for another iteration of surface recovery using $\{\mathbf{n}_s\}$ from Step 4, and repeat Steps 3-5;
6. According to Eq. (3), compute the 3D lighting coefficients by using $\{\mathbf{n}_s\}$.

5. RESULTS

The six images shown in the first row of Fig. 3 are: the lighting probe captured in a Eucalyptus Grove, UC Berkeley (EUC, also maintained by Paul Debevec [10]), the rendered Stanford bunny under EUC lighting environment, the RGB map of the recovered surface normal components, the z-component of the recovered surface normals, the spheres representing the actual lighting and the estimated 3D lighting coefficients (green channel). The three figures on the very right in the first row of Fig. 3 are the errors between the ground truth and the estimated 3D lighting coefficients in red, green and blue channels respectively from up to down. And the second/third rows of Fig. 3 are the results for the following lighting environments: Dining room of the Ennis-Brown House, Los Angeles, California (ENN), and Pisa courtyard nearing sunset, Italy (PIS)[2].

Table 3. Errors between object pairs of Fig. 4

	R	G	B
S-1 vs. S-2	0.0277	0.0356	0.0456
U-1 vs. U-2	0.0031	0.0035	0.0058
S-1 vs. U-1	0.4533	0.4432	0.3696
S-1 vs. U-2	0.4394	0.3902	0.3001
S-2 vs. U-1	0.4722	0.4801	0.4245
S-2 vs. U-2	0.4752	0.4314	0.3535

In the third and the fourth columns in Fig. 3, some noise appears in the RGB map and the z-component of the 3D surface normal estimates. This can be explained by the fact that the linear image formation model in Eq. (4) is only a rough approximation. And the crudely estimated lighting coefficients $\mathbf{A}$ and $\mathbf{a}$ therefore introduce surface normal recovery errors when solving Eq. (5). Although the 3D surface normal estimates are not perfect, but because $k \gg 9$ in Eq. (3), the matrix $\mathbf{M}$ is highly overdetermined and in practice the system can yield good results. We have tested our 3D lighting-based forensic method using 11 lighting probes to render Stanford bunny. In the red channel, the average estimate error is 0.0313 with a maximum of 0.0558 and a minimum of 0.0063. In the green channel, the average estimate error is 0.0283 with a maximum of 0.0580 and a minimum of 0.0094. And in the blue channel, the average estimate error is 0.0278 with a maximum of 0.0594 and a minimum of 0.0064. Compared with [6], for the synthetic images, we achieve better results even without a predefined 3D model.

Shown in Fig. 4 is a forgery from [5]. We extract the information from the bodies of the swans and the umbrellas to establish their 3D models and the estimated surface normals are then used for 3D lighting estimation. Four lighting spheres with estimated 3D lighting coefficients are also shown in Fig. 4. Qualitatively, in accordance with the results in [5], the lighting spheres between the swans and between the umbrellas are both very similar, while the differences of those between the swans and the umbrellas are quite obvious. In addition, the pairwise lighting differences are summarized in Table 3. Note that all the errors either between the swans or between the umbrellas are smaller than 0.05, similar to the differences between consistent lightings in the simulations of [6]. Based on the significant differences of the lighting environment between the swans and the umbrellas, we can conclude that the picture is a forgery. Besides, instead of only using boundary information of the objects to estimate five 2D lighting coefficients, we are able to estimate the nine

[2] http://gl.ict.usc.edu/Data/HighResProbes

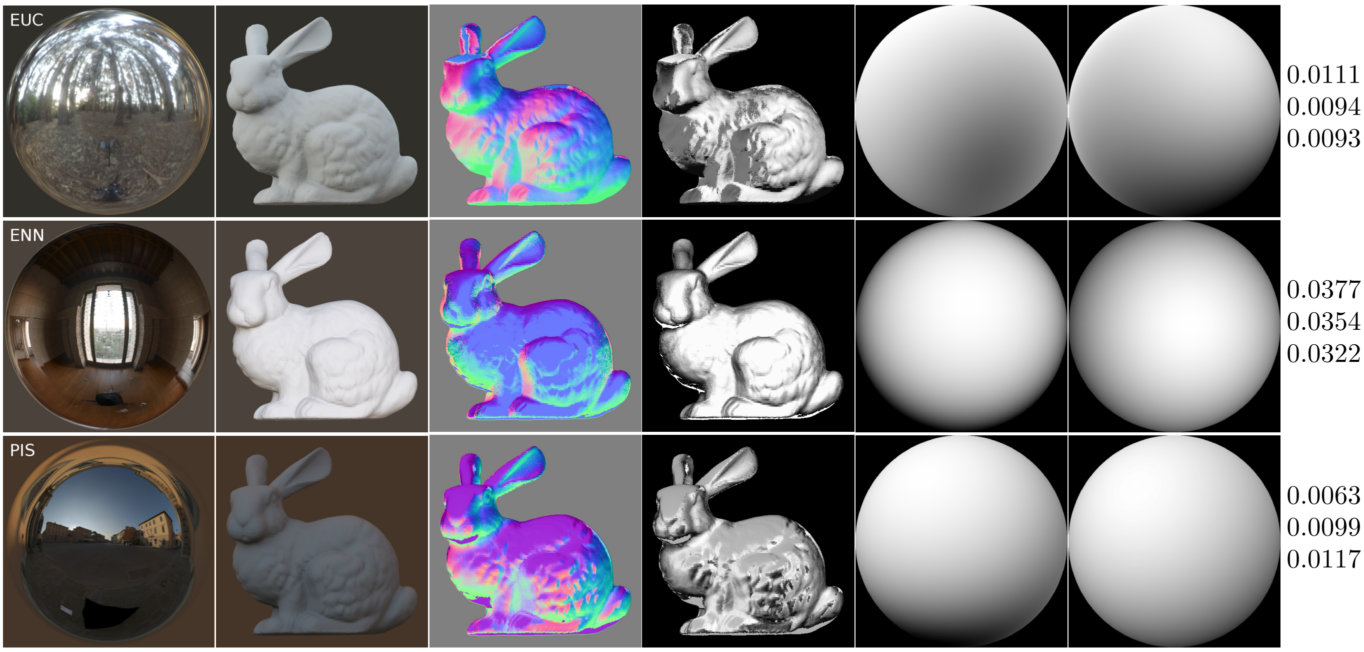


Fig. 3. From left to right are: the lighting probes, the rendered Stanford bunnies, the two results from SFS, the lighting spheres of the actual and the estimated lighting environments, and the figures of their lighting differences in red, green and blue channels respectively. The lighting environments from top to bottom are: EUC, ENN and PIS.

3D lighting coefficients, which is more reliable for lighting consistency comparisons.

6. CONCLUSIONS

We have presented that in lighting-based forensics, the original 2D lighting-based detector can be fooled by modifying the pixel intensities around the border of the inserted object. Therefore, we propose to use shape-from-shading to estimate 3D lighting coefficients in order to enhance the capabilities of the forgery detector. This has the potential to make lighting-based forensics more reliable and general.

The main issue with this new method is the estimation of the 3D shape of the object. At present a crude estimation of the shape seems sufficient for simple objects. Future work consists in improving the accuracy of shape recovery and investigating the effect for more complicated objects. We also plan to undertake comprehensive evaluation of the method by detecting more real world forgeries.

7. REFERENCES

[1] H. Farid, "A survey of image forgery detection," *IEEE Signal Process. Mag.*, vol. 26, no. 2, pp. 16–25, 2009.

[2] A. Rocha, W. Scheirer, T. Boult, and S. Goldenstein, "Vision of the unseen: Current trends and challenges in digital image and video forensics," *ACM Comput. Surv.*, vol. 43, no. 4, pp. 26:1–26:42, 2011.

[3] Y. Ostrovsky, P. Cavanagh, and P. Sinha, "Perceiving illumination inconsistencies in scenes," *Perception*, vol. 34, no. 11, pp. 1301–1314, 2005.

[4] M. K. Johnson and H. Farid, "Exposing digital forgeries by detecting inconsistencies in lighting," in *ACM Workshop on Multimedia & Security*, 2005, pp. 1–10.

[5] M. K. Johnson and H. Farid, "Exposing digital forgeries in complex lighting environments," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 450–461, 2007.

[6] E. Kee and H. Farid, "Exposing digital forgeries from 3-D lighting environments," in *IEEE Int. Conf. on Inf. Forensics Security*, 2010, pp. 1–6.

[7] R. Ramamoorthi and P. Hanrahan, "An efficient representation for irradiance environment maps," in *Proc. SIGGRAPH*, 2001, pp. 497–500.

[8] R. Huang and W. A. P. Smith, "Shape-from-shading under complex natural illumination," in *Proc. of IEEE Int. Conf. on Image Process.*, 2011, pp. 13–16.

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

[10] P. Debevec, "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *Proc. SIGGRAPH*, 1998, pp. 189–198.