Exposing Photo Manipulation from Shading and Shadows

Eric Kee, Columbia University
James F. O'Brien, University of California, Berkeley
and
Hany Farid, Dartmouth College

We describe a method for detecting physical inconsistencies in lighting from the shading and shadows in an image. This method imposes a multitude of shading- and shadow-based constraints on the projected location of a distant point light source. The consistency of a collection of such constraints is posed as a linear programming problem. A feasible solution indicates that the combination of shading and shadows is physically consistent, while a failure to find a solution provides evidence of photo tampering.

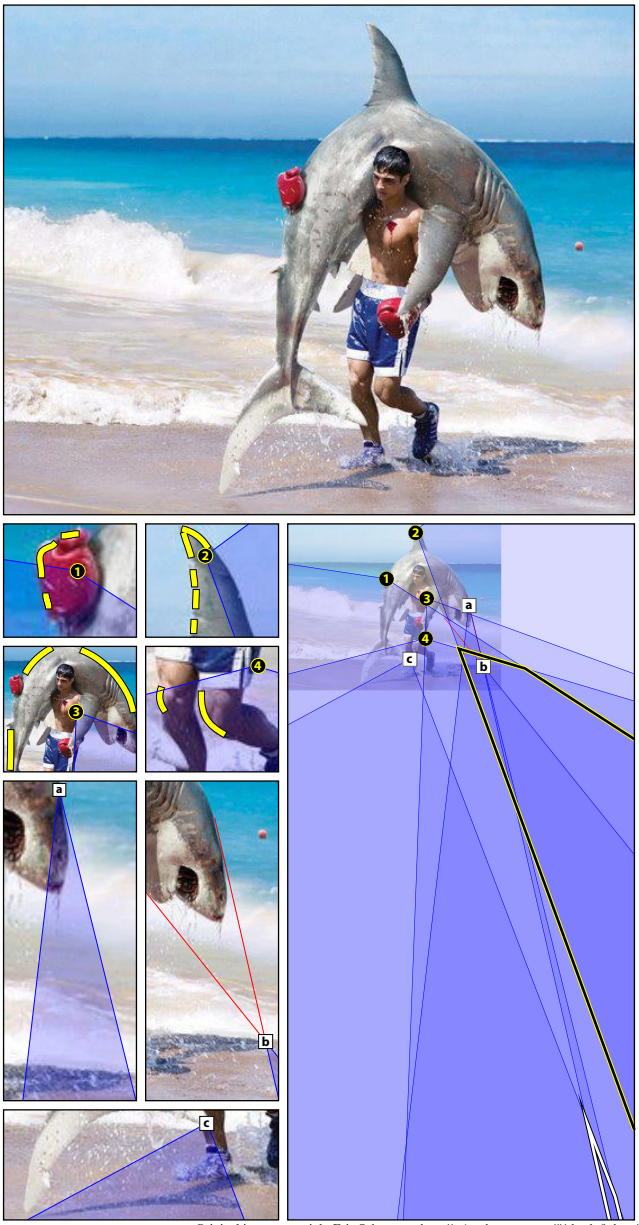
Categories and Subject Descriptors: I.2.10 [**Artificial Intelligence**]: Vision and Scene Understanding—*Scene Analysis*; I.3.6 [**Computer Graphics**]: Methodology and Techniques—*Image Forensics*; K.4.m [**Computing Milieux**]: Computers and Society—*Miscellaneous*

Additional Key Words and Phrases: Image forensics, photo manipulation, image manipulation, forgery detection, lighting, shadows

1. INTRODUCTION

Advances in photo editing software have made it possible to create visually compelling photographic forgeries. The ubiquity of these images has undermined our trust in photography, and impacts law enforcement, national security, the media, advertising, commerce, and more. The field of photo forensics has emerged to help restore trust in photography [Farid 2009; Rocha et al. 2011].

In the absence of an embedded watermark or signature, forensic methods assume that image manipulation will disrupt statistical, geometric, or physical properties in an image. To the extent that such perturbations can be quantified and detected, forgeries can be objectively invalidated. For example, format-specific analyses



Original image copyright Eric Sahrmann, *http://ericsahrmann.com/#/shark-fighter/*

Fig. 1. The projected location of the light source is constrained by shading along occluding contours ❶ – ❹ and by cast shadows [a] – [c]. The intersection of the shading constraints (outlined in black) is inconsistent with the intersection of the shadow constraints (outlined in white), revealing this image to be a fake (not surprisingly).

This paper has been accepted to ACM TOG for publication.
This work was supported by NSF grants IIS-0915462, IIS-1353155, and CNS-0708209, Intel's Science and Technology Center for Visual Computing, and by gifts from Adobe Systems, Microsoft, NVIDIA, and Pixar.
Authors' email addresses: erickee@cs.columbia.edu, job@berkeley.edu, farid@cs.dartmouth.edu

DOI 10.1145/XXXXXXX.YYYYYYY
http://doi.acm.org/10.1145/XXXXXXX.YYYYYYY

exploit artifacts that are introduced when a JPEG image is compressed multiple times [Popescu and Farid 2005a; Mahdian and Saic 2009; Kee et al. 2011]. Pixel-based analyses detect low-level correlations that are introduced by cloning [Fridrich et al. 2003], re-sizing [Popescu and Farid 2005a], or non-linear filtering [Lin et al. 2005]. Sensor-based analyses can detect inconsistencies in chromatic aberrations [Johnson and Farid 2006], color filter array interpolation [Popescu and Farid 2005b; Kirchner 2010], or sensor noise [Chen et al. 2008; Fridrich 2009]. Additionally, physically based analysis can detect inconsistencies in reflections [O'Brien and Farid 2012], lighting [Johnson and Farid 2007; Kee and Farid 2010; Johnson and Farid 2005], or shadows [Kee et al. 2013].

Any one of these techniques may be applicable to only a subset of images, and may even be vulnerable to counter attack. Working in concert, however, all of these techniques make it increasingly difficult and time consuming — but never impossible — for a forger to create a compelling fake.

We propose a new physically based forensic method that simultaneously analyses shading and shadows in a photo. Such methods are attractive because shading and shadows result from the 3-D interaction of lighting and geometry, and can therefore be difficult to accurately modify using 2-D photo editing software. In addition, these effects are often measurable even in low resolution and low quality images. It has been shown that the human visual system can be surprisingly bad at detecting even large physical inconsistencies in lighting and shadows [Jacobson and Werner 2004; Ostrovsky et al. 2005; Farid and Bravo 2010]. As a result, a forger is likely to leave behind inconsistent shading and shadows, and a visual inspection by a forensic analyst is likely to be subjective or unreliable. In contrast our proposed geometric analysis can objectively detect inconsistencies.

It has been well established that accurately estimating lighting from a single image is difficult. This estimation is made more difficult when an image is of low resolution and poor quality, and made even more complex when an image is a composite of multiple photos taken under varying lighting. In a forensic setting, however, our goal is not necessarily to estimate the location of the light, but rather to determine if the shading and shadows in a photo are consistent with a single light source. To this end, we seek only to constrain the projected location of the light source in the image plane. By combining a multitude of potentially weak constraints, we can determine if the shading and shadows in a photo are physically consistent.

Shown in Fig. 1 is an example of our shading and shadow analysis applied to a falsified photo. A collection of shading and shadow constraints restrict the projected location of the light source. The shading constraints are labeled ❶ – ❹, and the shadow constraints are labeled [a] – [c]. Magnified views of each constraint are shown in the bottom left. The intersection of the four shading constraints is outlined in black (bottom right panel). The intersection of the three shadow constraints is outlined in white (bottom right panel). Because these regions do not intersect, we can prove that there does not exist a single distant point light source that can explain the shading and shadows in this image. It can be objectively argued, therefore, that the shading and shadows in Fig. 1 are physically inconsistent, correctly revealing this photo to be a fake.

2. RELATED WORK

A large body of work has been established for the estimation of lighting from the shading on an object. Standard techniques often assume, as we do, a single distant point light source illuminating a Lambertian surface of constant reflectance. Within the vision literature, methods include shape-from-shading [Zheng and Chellappa 1991; Hougen and Ahuja 1993; Kim et al. 1998], and inverse lighting in which the 3-D geometry is assumed to be known [Marschner and Greenberg 1997; Zhang and Yang 2001; Hara et al. 2005; Zhou and Kambhamettu 2008]. Some methods take advantage of both shading and shadow cues, [Li et al. 2003; Wang and Samaras 2003], but, unlike our approach, require complete 3-D geometric models. Methods that do not require known 3-D geometry have used manually-approximated 3-D geometry [Karsch et al. 2011]. When 3-D models are not available or cannot be easily estimated, 2-D occluding contours can be used to estimate a 2-D subset of the 3-D light direction [Nillius and Eklundh 2001].

In the forensics literature, shading and shadows have been individually exploited to detect photo tampering [Johnson and Farid 2005; 2007; Zhang et al. 2009; Kee and Farid 2010; Liu et al. 2011; Riess and Angelopoulou 2010]. The combination of shading and shadow has been considered for estimating 2-D light direction [Johnson et al. 2008], and for estimating the location of light sources [Stork and Johnson 2006a; 2006b]. Their work, demonstrated in the analysis of paintings, assumes orthographic projection which can be an overly limiting assumption in many practical situations.

Our approach differs in three fundamental ways from these previous methods. First, instead of explicitly estimating lighting, we seek only to constrain the light source. This allows us to analyze a wide variety of images, even those in which there is insufficient information to uniquely estimate lighting. Second, we formulate the estimation of lighting from shading under a model of linear perspective. In previous work [Johnson and Farid 2005; 2007; Johnson et al. 2008; Stork and Johnson 2006a; 2006b], orthographic projection was assumed to simplify estimation of surface normals along occluding boundaries (by assuming that the z-component of the 3-D surface normal is zero). This assumption, however, leads to errors in the estimated lighting when objects are on opposing sides of an image, and particularly when the projected location of the light differs significantly between the two models. And third, because the analysis of cast shadows [Kee et al. 2013] implicitly assumes a model of perspective projection, our shading-based constraints can be combined with shadow constraints. In contrast to [Kee et al. 2013], this combination forces a would-be forger to consider and potentially correct both the shading and shadows, rather than the shadows alone. This makes the creation of a forgery more difficult and time consuming.

Our forensic technique has several advantages in that it makes minimal assumptions about the underlying scene geometry, makes relatively few assumptions about the scene (single distant point light source illuminating a Lambertian surface of constant reflectance), assumes a more plausible model of perspective projection, combines both shading and shadow cues within a single unified framework, and, perhaps most importantly, can be applied to a broader range of images in which it would be difficult or impossible to explicitly estimate lighting.

3. METHODS

In this section we formulate how the shading on an object constrains the location of a distant point light source in 3-D and in the projected 2-D image. We show that these 2-D shading constraints can be combined with 2-D shadow constraints within a linear framework to determine if the shading and shadows in an image are physically consistent with a single light source.

Throughout, the following notational conventions are used. Scalar values are lowercase letters (a), vectors are bold lowercase

letters ($\mathbf{v}$), and matrices are bold capital letters ($\mathbf{A}$). Italicized vectors denote quantities in the 2-D image plane ($\boldsymbol{v}$), while roman vectors denote quantities in the 3-D world or higher dimensions ($\mathbf{v}$). Individual elements of vectors or columns of matrices are referenced by subscripts: v_k is the k^{th} element of vector $\mathbf{v}$, and $\mathbf{a}_k$ is k^{th} column of matrix $\mathbf{A}$. Vectors that represent homogeneous quantities in the image plane are denoted as $\tilde{\boldsymbol{v}}$ and their Cartesian equivalent is simply $\boldsymbol{v}$.

3.1 Shading constraints (3-D)

Under an infinitely distant point light source, the radiance r of an image patch that corresponds to a surface with constant Lambertian reflectance and normal $\mathbf{n} = \begin{pmatrix}\mathrm{n}_x & \mathrm{n}_y & \mathrm{n}_z\end{pmatrix}^T$ is given by

$$r = \mathbf{n} \cdot \mathbf{s} + a, \tag{1}$$

where $\mathbf{s} = \begin{pmatrix}\mathrm{s}_x & \mathrm{s}_y & \mathrm{s}_z\end{pmatrix}^T$ specifies the direction to the light source, and the magnitude of $\mathbf{s}$ is proportional to the brightness of the light. The constant term a is the so-called ambient term which approximates indirect illumination. Note that this expression assumes that the angle between the surface normal and light is less than 90°. The components of this lighting model can be estimated from $k \geq 4$ patches with known surface normals. The equations for each surface normal and corresponding radiance are packed into the following linear system:

$$\begin{pmatrix}\mathbf{n}_1^T & 1\\ \mathbf{n}_2^T & 1\\ \vdots & \vdots\\ \mathbf{n}_k^T & 1\end{pmatrix}\begin{pmatrix}\mathbf{s}\\ a\end{pmatrix} = \mathbf{r} \tag{2}$$

$$\mathbf{N}\mathbf{e} = \mathbf{r}\ , \tag{3}$$

where $\mathbf{r}$ is a vector of observed radiance for each patch. The lighting parameters $\mathbf{e}$ can be estimated using standard least squares:

$$\mathbf{e}_0 = (\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{r}\ , \tag{4}$$

where $\mathbf{e}_0$ denotes the estimated parameters, and its first three components specify the brightness-scaled light direction, $\mathbf{s}_0$. The normalized light direction, $\mathbf{s}_0/||\mathbf{s}_0||$, can be visualized as a point on the surface of the unit sphere, Fig. 2(a).

When the matrix $\mathbf{N}$ is well conditioned, this estimation is straightforward. In practice, however, the matrix $\mathbf{N}$ may be poorly conditioned or rank deficient due to a limited number, or range, of known surface normals. In such situations it is not possible to determine an exact solution, $\mathbf{e}_0$, of the lighting parameters but it is still possible to constrain the solution to a space of possible solutions.

Consider the singular value decomposition (SVD) of the linear system in Equation (3):

$$\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T\mathbf{e} = \mathbf{r}\ . \tag{5}$$

This decomposition projects the lighting parameters $\mathbf{e}$ onto an orthonormal basis $\mathbf{V}$:

$$\mathbf{V} = \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4\end{pmatrix} = \begin{pmatrix}\mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 & \mathbf{w}_4\\ b_1 & b_2 & b_3 & b_4\end{pmatrix} \tag{6}$$

where $\mathbf{w}_i$ are 3×1 vectors and b_i are scalar values. Note that $\mathbf{w}_i$ and b_i form a basis for the brightness-scaled light direction $\mathbf{s}$ and ambient component a. After projection onto vectors $\mathbf{v}_i$, $\mathbf{e}$ is scaled by the diagonal matrix $\mathbf{\Lambda}$ of singular values λ_i, and then transformed into image radiance by the matrix $\mathbf{U}$. Using the convention that the singular values, $\lambda_1, \ldots, \lambda_4$, are sorted from largest to smallest, the well-conditioned dimensions of the solution space are given by the columns of $\mathbf{\Lambda}$ and $\mathbf{V}$ that correspond to singular values $\lambda_i/\lambda_1 \gg 0$. The lighting parameters $\mathbf{e}$ can be estimated within the subspace spanned by the well-conditioned vectors $\mathbf{v}_i$ of the singular value decomposition:

$$\mathbf{e}_0 = \mathbf{N}^+\mathbf{r}\ , \tag{7}$$

where $\mathbf{N}^+$ denotes the pseudo-inverse.

When one singular value is degenerate, $\lambda_4/\lambda_1 \approx 0$, possible lighting solutions are given by $\mathbf{e}_0 + \alpha\mathbf{v}_4$, where $\mathbf{e}_0$ is the solution to the linear system in Equation (7), α is a scalar, and $\mathbf{v}_4$ is the fourth column of matrix $\mathbf{V}$. This constrains the brightness-scaled direction of the light source $\mathbf{s}$ to a line in 3-D:

$$\mathbf{s} = \mathbf{s}_0 + \alpha\mathbf{w}_4\ , \tag{8}$$

where $\mathbf{s}_0$ is the first three components of $\mathbf{e}_0$ corresponding to the light direction, and $\mathbf{w}_4$ is the first three components of singular vector $\mathbf{v}_4$. This line passes through the point $\mathbf{s}_0$ and is parallel to $\mathbf{w}_4$. The corresponding set of normalized light directions, $\mathbf{s}/||\mathbf{s}||$, forms a semi-circular arc, which is the projection of the line onto the unit sphere, Fig. 2(b).

When two singular values are degenerate, $\lambda_4/\lambda_1 \approx 0$ and $\lambda_3/\lambda_1 \approx 0$, possible lighting solutions are given by $\mathbf{e}_0 + \alpha\mathbf{v}_4 + \beta\mathbf{v}_3$, where $\mathbf{e}_0$ is the solution to the linear system in Equation (7), and both α and β are scalars. This constrains the brightness-scaled direction of the light source $\mathbf{s}$ to a plane in 3-D:

$$\mathbf{s} = \mathbf{s}_0 + \alpha\mathbf{w}_4 + \beta\mathbf{w}_3\ , \tag{9}$$

which passes through $\mathbf{s}_0$ and is parallel to both $\mathbf{w}_4$ and $\mathbf{w}_3$. The corresponding set of normalized light directions forms a hemispherical surface, which is the projection of the plane onto the unit sphere, Fig. 2(c).

When three singular values are degenerate, $\lambda_4/\lambda_1 \approx 0$, $\lambda_3/\lambda_1 \approx 0$, and $\lambda_2/\lambda_1 \approx 0$, the null space of the linear system spans the full 3-D space, thus providing no constraint on the light direction $\mathbf{s}$.

3.2 Shading constraints (2-D)

In the previous section we described how the shading on an object can constrain the 3-D direction of a distant point light. In this section we develop a representation of these constraints that can be combined with constraints from shadows [Kee et al. 2013] (detailed in Section 3.4). Light directions in 3-D can be treated as points at infinity and projected into the 2-D image plane. Constraints on the 3-D lighting direction can, therefore, be expressed as constraints on the projected light location in the 2-D image plane.

For simplicity, but without loss of generality, let the surface normals $\mathbf{n}$ and direction of the light source $\mathbf{s}$ be specified in the camera coordinate system in which the camera faces down the $-z$ axis. Under an idealized perspective camera with focal length f and image center $\boldsymbol{c}$, the light source at infinity in direction $\mathbf{s}$ projects to a point in the image plane:

$$\tilde{\boldsymbol{s}} = \begin{pmatrix}f & 0 & -c_x\\ 0 & f & -c_y\\ 0 & 0 & -1\end{pmatrix}\mathbf{s} \tag{10}$$

$$= \mathbf{K}\mathbf{s} \tag{11}$$

where the image point $\tilde{\boldsymbol{s}}$ is specified in homogeneous coordinates, and $\mathbf{K}$ is the intrinsic matrix of the camera.

When the linear system, Equation (3), is fully conditioned, the direction of the light source, $\mathbf{s}$, can be unambiguously estimated,

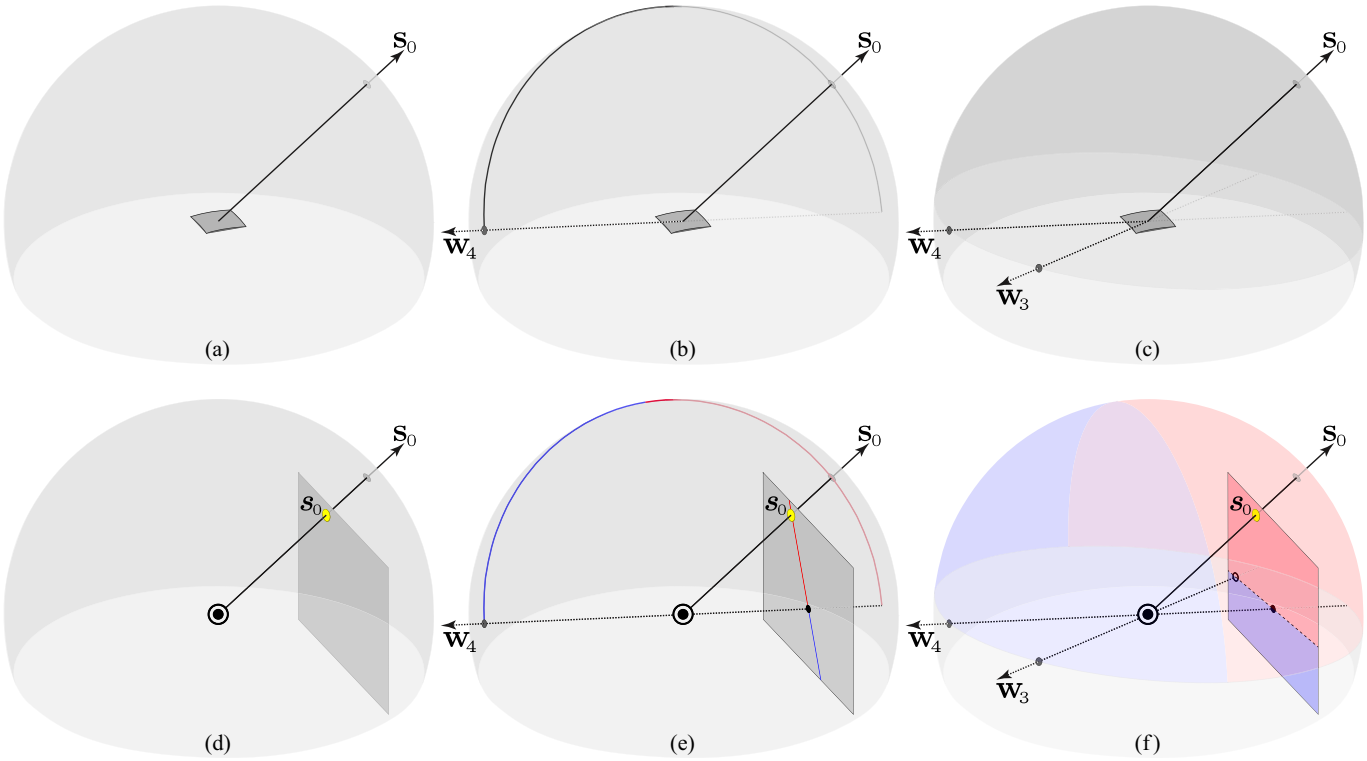


Fig. 2. Shading constraints in 3-D (top) and 2-D (bottom): (a) when the linear system in Equation (3) is fully conditioned, the 3-D direction of the light source, $\mathbf{s}_0$, can be uniquely determined; (b-c) if the linear system is not fully conditioned, then the 3-D direction of the light source is constrained to a semi-circular arc or a hemisphere, depending on the rank of the linear system ($\mathbf{w}_3$ and $\mathbf{w}_4$ are the vectors corresponding to the singular values of the linear system). Shown in panels (d)-(f) is the projection of these constraints onto the image plane, where the camera's center of projection is denoted by ◉.

$\mathbf{s} = \mathbf{s}_0$, Fig. 2(a). This constrains the projected location, $\tilde{\boldsymbol{s}}$, of the light source $\mathbf{s}$ to a single point in the image:

$$\tilde{\boldsymbol{s}} \;=\; \mathbf{K}\mathbf{s}_0 \;. \tag{12}$$

Shown in Fig. 2(d) is the projection, $\boldsymbol{s}_0$, of the 3-D direction, $\mathbf{s}_0$, onto a 2-D image plane (the camera center of projection is denoted by ◉).

When one singular value is degenerate, the brightness-scaled light direction is constrained to a line in 3-D, $\mathbf{s} = \mathbf{s}_0 + \alpha\mathbf{w}_4$, which corresponds to a half-circle of directions, Fig. 2(b). These directions project to a line in the image plane:

$$\tilde{\boldsymbol{s}} \;=\; \mathbf{K}(\mathbf{s}_0 + \alpha\mathbf{w}_4) \;. \tag{13}$$

Shown in Fig. 2(e) is the projection of the 3-D constraint (a half-circle of directions) onto a 2-D image plane. Note that light directions behind the camera's center of projection (blue arc) are projected into the lower part of the line (blue line segment), while light directions in front of the camera (red arc) are projected into the upper part of the line (red line segment). These two line segments meet at the image of vector $\mathbf{w}_4$, $\mathbf{K}\mathbf{w}_4$ (denoted by ● in the image plane). This line and its segments can be described by a position $\boldsymbol{p}$ and direction $\boldsymbol{d}$ in the image:

$$\tilde{\boldsymbol{p}} \;=\; \mathbf{K}\mathbf{w}_4 \tag{14}$$

$$\boldsymbol{d} \;=\; \text{sign}\,(\mathrm{s}_{0_z})(\boldsymbol{p} - \boldsymbol{s}_0) \;, \tag{15}$$

where $\boldsymbol{s}_0$ is the projected location of the least-squares solution of the light direction, $\mathbf{s}_0$, and s_{0_z} is its z-component. Direction $\boldsymbol{d}$ identifies the line segment onto which light directions in front of the camera project (see Appendix A for a derivation).

When two singular values are degenerate, the brightness-scaled light direction is constrained to a plane in 3-D, $\mathbf{s} = \mathbf{s}_0 + \alpha\mathbf{w}_4 + \beta\mathbf{w}_3$, which corresponds to a half-sphere of directions, Fig. 2(c). These directions project to a plane in the image:

$$\tilde{\boldsymbol{s}} \;=\; \mathbf{K}(\mathbf{s}_0 + \alpha\mathbf{w}_4 + \beta\mathbf{w}_3) \;. \tag{16}$$

Although this plane encompasses the entire image, it consists of two half-planes, Fig. 2(f): one for light directions in front of the camera (red), and one for directions behind the camera (blue). These half-planes are described by a position $\boldsymbol{p}$ and direction $\boldsymbol{d}$ in the image:

$$\tilde{\boldsymbol{p}} \;=\; \mathbf{K}\mathbf{w}_4 \tag{17}$$

$$\boldsymbol{d} \;=\; \text{sign}\left(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0\right)\left(\tilde{h}_1 \;\; \tilde{h}_2\right)^T \;, \tag{18}$$

where $\tilde{\boldsymbol{h}}$ is the homogeneous line through the image of $\mathbf{w}_4$ and $\mathbf{w}_3$:

$$\tilde{\boldsymbol{h}} \;=\; \mathbf{K}\mathbf{w}_4 \times \mathbf{K}\mathbf{w}_3 \;. \tag{19}$$

Direction $\boldsymbol{d}$ identifies the half-plane into which light directions in front of the camera project (see Appendix A for a derivation).

Finally, as noted previously, when three singular values are degenerate, the light direction is unconstrained in 3-D and thus provides no constraint on the projected location of the light in the image plane.

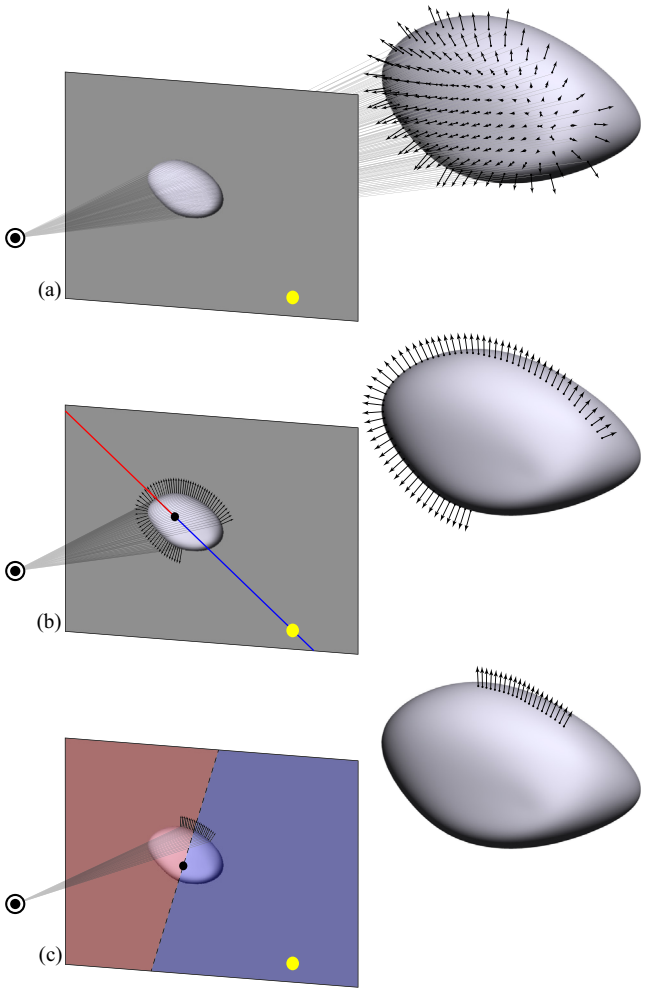


Fig. 3. Geometric configurations that lead to point (a), line (b), and half-plane (c) constraints on the projected location of the light source. Shown in each panel is the projection of a 3-D object (right) into the image plane (the camera's center of projection is denoted by ◉). A point constraint can be specified when a full 3-D model is available. Surface normals along an extended portion of an occluding contour provide a line constraint, and normals along a small portion of the occluding contour provide a half-plane constraint. (Note that, because the light is behind the camera, its projected location (yellow dot) is inverted under perspective projection.)

Fig. 3 illustrates how different configurations of surface normals can lead to the point, line, and half-plane constraints on the projected location of the light source. Shown in panel (a) is the case where a full 3-D model of the object is known, yielding a full rank matrix $\mathbf{N}$ in Equation (3) and a point constraint. Shown in panel (b) is the case where only approximately coplanar normals along an extended portion of the object's occluding boundary are known. These occluding normals yield one degenerate singular value in $\mathbf{N}$, producing a line constraint. Lastly, shown in panel (c) is a case where only normals along a small portion of the object's occluding boundary are specified, yielding two degenerate singular values in $\mathbf{N}$ and a half-plane constraint.

3.3 Satisfiability of multiple constraints

In a forensic setting, we would like to determine if the shading on objects in a photo are physically consistent with a single light source. We therefore combine the shading constraints from multiple objects in a scene into a single linear system to determine if they are collectively satisfiable. Consider first the case in which the light source is in front of the camera. When the light position can be constrained to a point in the image, Equation (12), the following equality constraint (in homogeneous coordinates) is placed on the location of the light source $\boldsymbol{s}$:

$$\tilde{\boldsymbol{s}} \;=\; \mathbf{K}\mathbf{s}_0 \;. \tag{20}$$

When the light position can be constrained to a line, Equation (13), the following constraint is placed on $\boldsymbol{s}$ (in cartesian coordinates):

$$\boldsymbol{s} \;=\; \alpha\boldsymbol{d} + \boldsymbol{p} \tag{21}$$

$$\alpha \;\geq\; 0 \;, \tag{22}$$

where $\boldsymbol{p}$ and $\boldsymbol{d}$ are given by Equations (14)-(15). Lastly, when the light position is constrained by a half-plane, Equation (16), the following inequality constraint is placed on $\boldsymbol{s}$:

$$\boldsymbol{s}\cdot\boldsymbol{d} - \boldsymbol{p}\cdot\boldsymbol{d} \;\geq\; 0 \;, \tag{23}$$

where $\boldsymbol{p}$ and $\boldsymbol{d}$ are given by Equations (17)-(19).

Each of these constraints are linear, and can therefore be combined into a linear program to check for mutual consistency. A viable solution to the linear program means that the shading in an image is consistent, while a failure to find a viable solution means that one or more of the constraints are inconsistent with a single light source.

When the light source is behind the camera, its projection into the image is inverted by linear perspective projection. To contend with this possibility, a second linear program is constructed in which all constraint inequalities are reversed. The shading is determined to be consistent if either linear program is satisfiable.

3.4 Shadow constraints (2-D)

The previous sections have described how the shading of an object can be used to constrain the projected location of a point light source. In addition, cast and attached shadows also constrain the projected location of a point light source [Kee et al. 2013]. These 2-D shadow constraints take the form of linear inequalities, and can therefore be combined with our shading constraints to further constrain the projected location of the light source. We briefly summarize the constraints provided by cast and attached shadows.

Cast shadows. Consider a ray that connects a point in a shadowed region to its corresponding point on the shadow-casting object, Fig. 4 (solid line). In the 3-D scene, this ray intersects the light source. In a 2-D image of the scene created under linear perspective, the projection of this ray remains a straight line that must connect the images of the shadow and object points, and must intersect the projected location of the light source. These constraints hold regardless of the geometry of the object and the surface onto which the shadows are cast, and hold for either an infinitely distant or local light. Cast shadows constrain the projected location of the light source $\boldsymbol{s}$ to a line:

$$\boldsymbol{s} \;=\; \alpha\boldsymbol{d} + \boldsymbol{p} \tag{24}$$

$$\alpha \;\geq\; 0 \;, \tag{25}$$

where $\boldsymbol{p}$ is a point in a shadow, and $\boldsymbol{d}$ is a vector connecting $\boldsymbol{p}$ to the corresponding point on the object. Notice that this constraint has

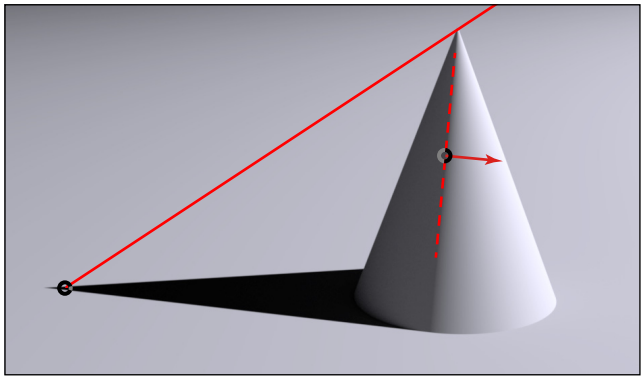


Fig. 4. A cast and attached shadow constraint. The projected location of the light source lies on a line that connects a point in the cone's cast shadow with its corresponding point on the cone. An attached shadow on the cone constrains the projected location of the light source to be in a half plane, specified by the orientation and polarity of the shadow terminator.

the same form as the line constraint provided by shading, Equations (21)-(22). As such, these constraints can be seamlessly combined with the shading constraints, Section 3.3.

Attached shadows. Attached shadows occur when objects occlude the light from themselves, leaving a portion of the object in shadow, Fig. 4 (dashed line and arrow). Points are in shadow if the surface normal makes an angle greater than 90 degrees with the direction toward the light. The surface contour that divides points in shadow from those receiving light is known as the terminator. Surface normals along this contour form a 90 degree angle with the direction to the light. Similar to a cast shadow, there is a correspondence between points in and out of shadow on either side of the terminator. This correspondence, however, can only be specified to within a half-plane because the light's elevation is ambiguous to within 180 degrees. Attached shadows therefore constrain the projected location of the light source $\boldsymbol{s}$ to a half-plane:

$$\boldsymbol{s} \cdot \boldsymbol{d} - \boldsymbol{p} \cdot \boldsymbol{d} \geq 0 \ , \tag{26}$$

where $\boldsymbol{p}$ is a point on a terminator, and $\boldsymbol{d}$ is a vector orthogonal to the terminator in the direction of the illuminated side of the attached shadow. Notice that this constraint has the same form as the half-plane constraint provided by shading, Equation (23). As such, these constraints can be seamlessly combined with the shading constraints, Section 3.3.

As with the shading constraints, when the light source is behind the camera, its projection into the image is inverted by linear perspective projection. To contend with this possibility, a second linear program is constructed in which all constraint inequalities are reversed (as with the shading constraints). The shadows are determined to be consistent if either linear program is satisfiable [Kee et al. 2013].

3.5 Uncertain shading constraints

To this point we have assumed an ideal system in which the shading and shadow constraints can be precisely and accurately specified. However, errors will be introduced into the estimated light direction by luminance non-linearities, imprecise normals (due to limited pixel resolution), poor signal-to-noise ratio (affected by surface albedo and exposure), and deviations from the assumptions of Lambertian reflectance and a distant point-source. It is therefore important to incorporate a model of uncertainty into the estimation of lighting. In a forensic setting, a full 3-D model of an object is not typically available, and we focus only on the more commonly occurring situation in which the lighting can only be constrained to a line or half-plane.

For simplicity, and because it yields linear constraints, we assume that uncertainty in the shading constraints affects only the constraint direction $\boldsymbol{d}$ in Equations (21) and (23) and that $\boldsymbol{p}$ is not significantly affected. This corresponds to the assumption that the dominant source of uncertainty is in the measured image radiance, rather than in the specified surface normals (this assumption proves reasonable for normals on occluding contours).

Shown in Fig. 5(a) and (c) are illustrations in which the estimated lighting constraints (parameterized by $\boldsymbol{p}$ and $\boldsymbol{d}$) do not correctly constrain the true projected location of the light source, $\mathbf{s}$. By introducing uncertainty in the estimated direction $\boldsymbol{d}$, these constraints take on a wedge-shape as shown in Fig. 5(b) and (d). In each case, the wedge-shaped region specifies a range of plausible projected light positions, and contends with uncertainty in the initial estimated lighting constraints.

Convex wedges, Fig. 5(b), can be described by the intersection of two half-planes, each of which are linear. Concave wedges, Fig. 5(d), however, are not linear, and preclude a linear solution to determine the satisfiability of constraints, Section 3.3.

This problem of concave constraints can be overcome by noting that, although the line and half-plane constraints are described in terms of degeneracy in the linear system, *both constraints can be constructed for a particular linear system.* That is, instead of constructing only one constraint for each object (either a line or half-plane constraint), we construct both constraints regardless of conditioning, and add angular uncertainty as dictated by the corresponding singular values of the linear system, Fig. 5(e). (More details on this construction are given below.) The constraints are then intersected to produce a combined constraint, Fig. 5(f). If this combined constraint specifies a convex wedge, it can be incorporated into the linear program, Section 3.3, otherwise, it can be omitted (thus providing no constraint).

The uncertainty of constraints, characterized as the width of the resulting wedge-shaped region, specify a confidence interval on the angle between the constraint direction $\boldsymbol{d}$ and the vector from $\boldsymbol{p}$ to the light source (up to an inversion of $\boldsymbol{d}$ due to lights that lie behind the camera). In part, this uncertainty depends upon measurable characteristics of the linear system, which includes the singular value associated with the constraint, the range of surface normals that are available, and the extent to which the estimated lighting parameters, $\mathbf{e}$, can explain the variation in the image intensity.

We therefore describe the width of the wedge-shaped region for line and half-plane constraints as functions $\Phi_{\text{L}}(\cdot)$ and $\Phi_{\text{H}}(\cdot)$ of three parameters κ, θ, and ρ. The first parameter, κ, is the condition of the line constraint, $\kappa_{\text{L}} = \lambda_3/\lambda_1$, or half-plane constraint, $\kappa_{\text{H}} = \lambda_2/\lambda_1$. The second parameter, θ, is the range of surface normals, measured to be independent of image resolution by constructing a histogram over the directions of the normals and computing the number of degrees spanned by non-empty bins. Lastly, the extent to which the lighting model can explain the observed image intensity is measured as the R-value, ρ, of the least-squares solution, Equation (7).

We used large-scale simulations to estimate the functions $\Phi_{\text{L}}(\cdot)$ and $\Phi_{\text{H}}(\cdot)$, as detailed in Appendix B. Random 3-D shapes are imaged under random point lights, and by cameras with random focal lengths and image centers. Common sources of uncertainty are included: a secondary point light is added to simulate inter-reflections and other effects, and the images are subjected to noise, random gamma corrections, and quantization. Appendix C details how the

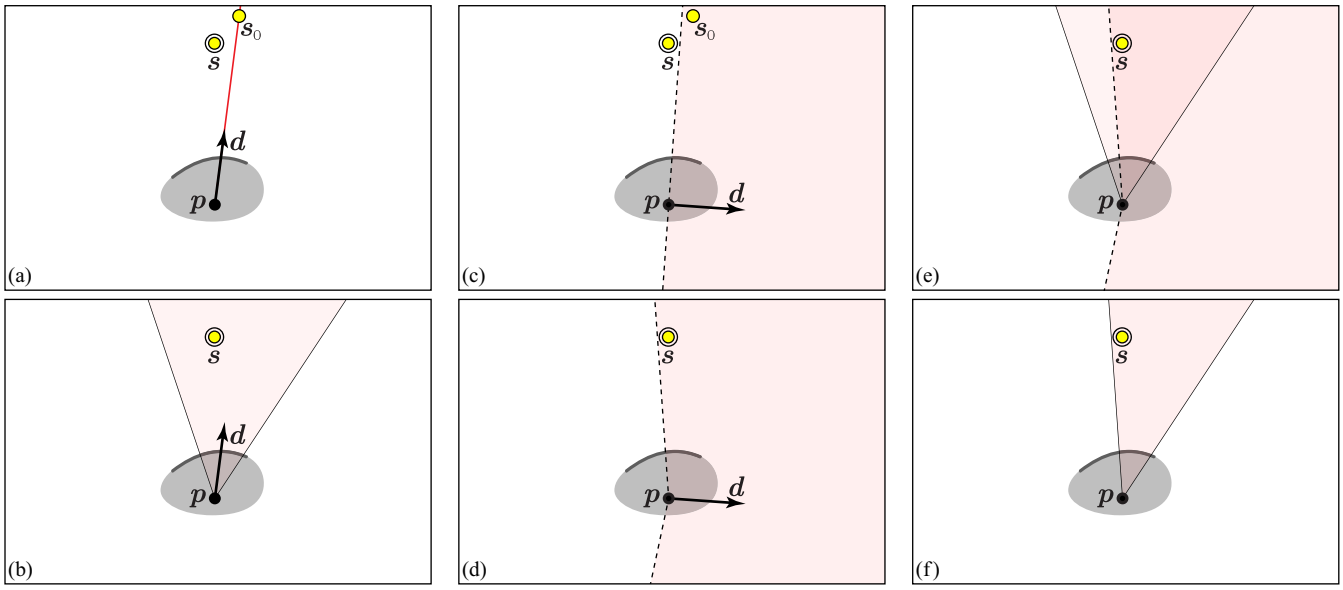


Fig. 5. Shown in each panel is an occluding contour along the top of an object. The projected location of the light source is denoted as $\mathbf{s}$ and the least squares estimate of the light source is denoted as $\mathbf{s}_0$. Shown are the following constraints, parameterized by a position $\boldsymbol{p}$ and direction $\boldsymbol{d}$: (a) a line constraint; (b) its corresponding convex wedge-shaped constraint; (c) a half-plane constraint; (d) its corresponding concave wedge-shaped constraint; (e) the pair of wedge-shaped constraints superimposed; and (f) the intersection of these two constraints yielding a single convex constraint.

objects are shaded. Random portions of the shaded objects are selected to compute the directions $\boldsymbol{d}$ of the line and half-plane constraints, and their angles ϕ to the light source. This simulation is repeated for many such samples ϕ, and each is assigned to its requisite quantized location in the domain of the functions $\Phi_{\text{L}}(\cdot)$ and $\Phi_{\text{H}}(\cdot)$. Those functions functions specify the angle below which a majority of the angles ϕ at each quantized location fall (e.g., 99.9%).

To build a constraint from an occluding contour, the position, direction, and width of the line and half-plane constraints are estimated, and their intersection computed. This produces a combined constraint with half-width, Φ_{C}, direction $\boldsymbol{d}_{\text{C}}$, and position $\boldsymbol{p}_{\text{C}} = \boldsymbol{p}$. Convex wedges ($\Phi_{\text{C}} \leq 90^{\circ}$) can be described by a pair of half-planes:

$$\boldsymbol{s} \cdot \boldsymbol{n}_{\text{C}}^{1} - \boldsymbol{p}_{\text{C}} \cdot \boldsymbol{n}_{\text{C}}^{1} \geq 0 \quad \text{and} \quad \boldsymbol{s} \cdot \boldsymbol{n}_{\text{C}}^{2} - \boldsymbol{p}_{\text{C}} \cdot \boldsymbol{n}_{\text{C}}^{2} \geq 0 \ , \tag{27}$$

where $\boldsymbol{s}$ is the unknown location of the light source. Vectors $\boldsymbol{n}_{\text{C}}^{i}$ are

$$\boldsymbol{n}_{\text{C}}^{1} = \mathbf{H}^{T}\,\mathbf{R}\boldsymbol{d}_{\text{C}} \quad \text{and} \quad \boldsymbol{n}_{\text{C}}^{2} = \mathbf{H}\mathbf{R}^{T}\,\boldsymbol{d}_{\text{C}} \ , \tag{28}$$

where $\mathbf{H}$ is a 90° rotation matrix, and matrix $\mathbf{R}$ specifies a rotation of Φ_{C} degrees. Lastly, concave wedges ($\Phi_{\text{C}} > 90^{\circ}$) cannot be described by linear constraints, and are omitted for simplicity.

3.6 Uncertain shadow constraints

It may be difficult to accurately specify a cast shadow constraint due to finite resolution, ambiguity in matching a cast shadow to its corresponding object, or when the shadow is indistinct. Unlike the shading constraints, modeling uncertainty in cast shadow constraints is more straightforward.

The line constraint for cast shadows can be relaxed to a wedge-shaped constraint by allowing a forensic analyst to select a point on a shadow and a range of possible corresponding points on the object that safely encompass the correct object-shadow pairing [Kee et al. 2013]. These relaxed constraints can be described by a pair of half-planes:

$$\boldsymbol{s} \cdot \boldsymbol{n}_{i}^{1} - \boldsymbol{p}_{i} \cdot \boldsymbol{n}_{i}^{1} \geq 0 \quad \text{and} \quad \boldsymbol{s} \cdot \boldsymbol{n}_{i}^{2} - \boldsymbol{p}_{i} \cdot \boldsymbol{n}_{i}^{2} \geq 0 \ , \tag{29}$$

where the normal vectors $\boldsymbol{n}_{i}$ are defined such that points within the wedge are on the positive side of both half-plane constraints. These wedge constraints are linear, and the satisfiability of a collection of these, along with other linear shading-based constraints, can be determined by linear programming, Section 3.3. Note that these constraints are identical to the wedge-shaped shading constraints, Equation (27). Lastly, an analyst may sometimes prefer to instead select a point on an object and a range of possible corresponding points in shadow (not vice versa). In this case the vectors $\boldsymbol{n}$ in Equation (29) are simply replaced by $-\boldsymbol{n}$.

As noted previously [Kee et al. 2013], we assume that an attached shadow constraint can be reliably specified due to the relative ease with which the terminator on an attached shadow can be specified. This assumption avoids non-linear (concave) constraints which would preclude a linear solution, Section 3.3.

3.7 Forensics

A photo may contain a variety of shaded objects along with cast and attached shadows. Shadow constraints are specified manually [Kee et al. 2013], while shading constraints are specified semi-automatically. In practice full 3-D geometric models of scene objects are rarely available, and determining surface normals from a single image is generally difficult and error prone. We therefore restrict ourselves to using points along occluding contours, for which normals can be reasonably estimated. The normals along occluding contours will generally not span the full 3-D space and therefore produce under-constrained systems corresponding to line and half-plane shading constraints.

An analyst specifies a shading constraint by first identifying a segment of an occluding contour having a constant Lambertian reflectance. The 2-D image normals, $\boldsymbol{n} = \begin{pmatrix} n_x & n_y \end{pmatrix}^T$, are computed from this occluding contour (a Bezier or other general-purpose curve can be fit to this contour, from which the 2-D normals can be analytically estimated). Under a model of linear perspective projection, the corresponding 3-D surface normal is given by:

$$\mathbf{n} = \begin{bmatrix} n_x \\ n_y \\ \frac{1}{f}\boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{c}) \end{bmatrix}, \tag{30}$$

where $\boldsymbol{x}$ is the 2-D image location of each corresponding 2-D normal, f is the camera's focal length, $\boldsymbol{c}$ is the image center, and it is assumed that $\mathbf{n}$ is subsequently scaled to unit length (see Appendix D for a derivation of this expression). These normals and the corresponding image radiance r along the occluding contour are used to estimate the lighting parameters $\mathbf{e}$, Equation (2). This estimation is straightforward if the camera focal length, f, and image center, $\boldsymbol{c}$, are known, or can be estimated. As we will describe below, however, the camera parameters need not be known.

The linear system used to estimate the lighting parameters $\mathbf{e}$, Equation (2), will typically be degenerate due to the limited range of normals provided by an occluding contour. This system is therefore solved by computing the pseudo-inverse, Equation (7). The SVD of the matrix of surface normals, $\mathbf{N} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$, is then computed. Using the estimated lighting parameters, $\mathbf{e}_0 = (\mathbf{s}_0 \;\; a_0)^T$, and the matrix $\mathbf{V}$, the position and direction of the line constraint, $\boldsymbol{p}_{\text{L}}$ and $\boldsymbol{d}_{\text{L}}$, are given by Equations (14)-(15), and the position and direction of the half-plane constraint, $\boldsymbol{p}_{\text{H}}$ and $\boldsymbol{d}_{\text{H}}$, are given by Equations (17)-(18).

Next, the line and half-plane constraints of an occluding contour are adjusted to contend with uncertainty in the estimated lighting parameters, Section 3.5. The line constraint takes on a wedge shape with position $\boldsymbol{p}_{\text{L}}$, direction $\boldsymbol{d}_{\text{L}}$, and angular half-width $\Phi_{\text{L}}(\kappa_{\text{L}}, \theta, \rho)$; the half-plane constraint also takes on a wedge shape with position $\boldsymbol{p}_{\text{H}}$, direction $\boldsymbol{d}_{\text{H}}$, and angular half-width $\Phi_{\text{H}}(\kappa_{\text{H}}, \theta, \rho)$. These two constraints are intersected to produce a single wedge-shaped region with position $\boldsymbol{p}_{\text{C}} = \boldsymbol{p}_{\text{L}} = \boldsymbol{p}_{\text{H}}$. Because both constituent wedges may be concave, the direction $\boldsymbol{d}_{\text{C}}$, and angular half-width, Φ_{C}, of the combined wedge is computed using a simple case-based algorithm. If the resulting wedge is concave, and hence non-linear, no constraint is specified.

Each convex constraint $(\boldsymbol{p}_c,\ \boldsymbol{d}_c,\ \Phi_{\text{C}})$ is added to a linear program by constructing a pair of half-planes, Equations (27)-(28). Similarly, each shadow constraint, specified by the analyst as either a wedge or half-plane constraint is added to the linear program. A solution to the linear program means that the shading and shadows in an image are consistent, while a failure to find a viable solution means that one or more of the constraints are inconsistent with a single light source. To contend with the situation when the light source is behind the camera, each constraint is reversed to yield a second linear program. Both of these systems are solved to determine if a viable light location exists that satisfies all constraints.

Lastly, recall that the camera focal length f and image center $\boldsymbol{c}$ must be known to estimate the 3-D surface normals, Equation (30). Because it can be difficult, or impossible, to estimate these parameters from a single image, we forgo their estimation, and instead search for camera parameters that yield a viable solution to the specified constraints. Specifically, we quantize the range of plausible focal lengths and image centers into a small set of values. The functions $\Phi_{\text{L}}(\cdot)$ and $\Phi_{\text{H}}(\cdot)$, which account for uncertainty in the estimated lighting parameters, are adjusted to include the additional uncertainty that is introduced by this quantization. Each quantized tuple of camera parameters is used to construct the required 3-D surface normals, and the satisfiability of the corresponding linear programs is checked. This approach is computationally efficient due to the simplicity of the linear formulation, and because a relatively coarse quantization of the focal length and image center has proven to be sufficient, Appendix B.

Fig. 6. A synthetic scene rendered with the physically based renderer Mitsuba [Jakob 2010]. This basic scene was used to test the reliability and specificity of the shading constraints.

4. RESULTS

We partition our results into two main sections. In the first section we report on a set of large-scale simulations used to validate the model of uncertainty employed by the shading-based constraints, Section 3.5. In particular, we report on the reliability and specificity of the shading constraints under a variety of assumptions that are consistent with, and that violate, our imaging assumptions. In the second section we apply our forensic technique to authentic and visually compelling forgeries.

4.1 Simulations

In the ideal situation, the wedge-shaped shading constraints will be just broad enough so as to always encompass the estimated light position, but narrow enough so as to reliably detect differences in lighting. We quantify these criteria in terms of the reliability and specificity of the wedge-shaped constraints. The reliability is measured as the probability that a constraint correctly encompasses the true light source. The specificity is measured as the probability that a random light position satisfies one or more constraints.

4.1.1 *Reliability.* Shown in Fig. 6 is a scene rendered with the physically based renderer Mitsuba [Jakob 2010]. This rock garden scene was rendered with either a point light, an area light subtending 5°, or a physically based daylight model [Preetham et al. 1999]. In each case, one of these distant light sources was placed in one of 49 possible directions, uniformly sampled over a hemisphere.

The scene was also rendered with the objects having different material properties. Each object's material was of constant reflectance and was either Lambertian, a rough diffuse material (modeling surfaces such as concrete and clay [Oren and Nayar 1994]), a rough plastic material (modeled by the physically-based Beck-

mann macrofacet distribution [Beckmann 1987]), or a woven cloth material [Irawan 2008].

The image rendering was configured to yield objects with an average resolution of 276 (default) or 136 square pixels, 1% (default) or 2% additive noise, and gamma correction in the range $[1.2,\ 1.8]$ (default) or $[1.5,\ 2.1]$.

Shading constraints were constructed by dividing the illuminated portion of the occluding contour on each rock into 20 segments, and selecting a random subset of between 1 and 20 segments. The corresponding 3-D occluding normals were estimated by selecting a focal length and image center within the precision assumed by the uncertainty functions, Φ_{L} and Φ_{H}, Appendix B. A total of $784,000$ constraints were specified for each of 9 types of scenes (3 lighting scenarios, 3 materials, and 3 rendering configurations).

The constraints built around the line and half-plane were designed to have an average reliability of 99.9%, Appendix B. These constraints are intersected to produce a combined constraint which therefore has an average reliability of $0.999^2 = 99.8\%$. We therefore expect that this percentage of constraints will encompass the projected location of the light source.

In the case of a distant point light source, Lambertian reflectance, a 276 square pixel resolution, 1% additive noise, and gamma in the range $[1.2,\ 1.8]$, the projected location of the light source was encompassed by 99.3% of the constraints. An area light yielded the same reliability, and a daylight model yielded a higher reliability of 99.9%. Reliability improves under the daylight model because it contributes a stronger ambient illumination, which decreases the signal-to-noise ratio. This decrease produces wider constraints, and improves reliability.

In the case of non-Lambertian surfaces illuminated by a distant point light source and the default remaining imaging parameters, rough diffuse objects, rough plastic, and cloth material yielded constraints with a reliability of 99.6%, 96.4%, and 99.2% respectively. Rough plastic led to a reduction in reliability by several percentage points because the roughened specular component biases the estimated light direction, but does not yield a corresponding increase in the constraint width.

Returning to a distant point light source and a Lambertian surface, a decrease in resolution to 136 square pixels yielded a reliability of 98.4%. An increase in the additive noise to 2% yielded a reliability of 99.8%. An increase in gamma into the range $[1.5,\ 2.1]$ yielded a reliability of 99.4%. As expected, the reduced resolution impacts the reliability. The increased gamma had little affect on reliability, and interestingly an increased level of noise improves reliability. This higher reliability is due to a reduced signal-to-noise ratio, which reduces the R-value of the least-squares solution and therefore yields a wider constraint.

To summarize, the overall reliability of the constraints was computed by combining constraints across all scene types (lights, reflectance models, and image degradations). This yields a mean reliability of 99.2%. This reliability does not depend on constraint width: the mean reliability for constraints of the same width was also 99.2% (with a minimum if 98.1%, and maximum of 99.8%).

4.1.2 *Specificity.* The width of each constraint determines its specificity, which is defined as the likelihood that a random light direction is satisfied by one, or more, constraints. We compute the specificity for a variable number and range of occluding contours from the scenes described in the preceding section, Fig. 6.

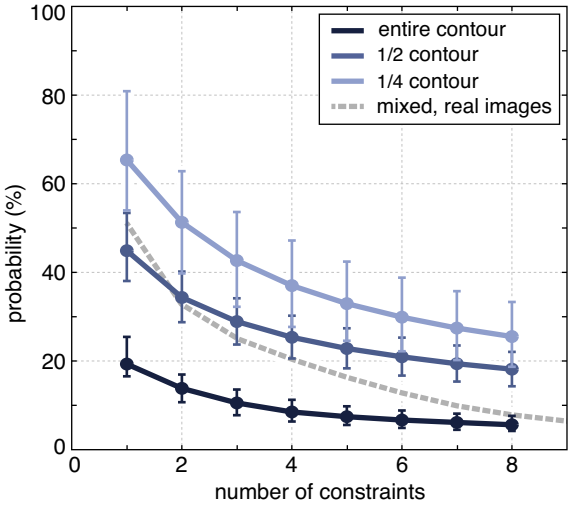


Fig. 7. The median probability that a randomly selected light direction will satisfy between one and eight shading constraints. Each curve corresponds to constraints built from the entire illuminated portion of the occluding contour (bottom), half of the contours (middle), or a quarter of the contours (top). The error bars identify the 25^{th} and 75^{th} quantile. Also shown (dashed gray) is this probability when using contours of mixed spatial extent from the authentic photos in Figs. 10, 11, 12, and 13 (see Section 4.2). The specificity of these real-world constraints matches closely with the rendered images.

Random 3-D light directions were drawn from a hemisphere, excluding elevations below 10°, and projected into the image.[1] When the entire illuminated portion of an occluding contour was used, the median constraint width was 30°, and was satisfied by 19% of random light directions. As constraints were added, up to a total of 8, this probability decreased to 6%, Fig. 7.

When one half of the contour was selected, the median constraint width was 75°, and was satisfied by 44% of random light directions. This probability decreased to 18% when all 8 available constraints were used, Fig. 7. When one quarter of the contour was selected, the median constraint width was 119°, and was satisfied by 65% of random light directions. This probability decreased to 26% when all 8 constraints were used. Overall, the specificity is directly proportional to the number and extent of constraints.

We also measured the decrease in specificity that results when the camera intrinsic parameters are unknown. Unknown camera parameters led to only a modest increase in the width of the constraints, and hence the specificity: the percentage of random light directions that satisfied constraints with unknown parameters increased by an average of only 1.8% over those with known parameters.

Lastly, recall that we exclude concave constraints as they cannot be incorporated into the linear program. In the above simulations, an average of only 1.2% of the constraints were concave.

4.1.3 *Shading vs. Shadows.* The addition of shading constraints allows an analyst to determine the consistency of regions that do not contain shadows, and also to detect mutually inconsistent combinations of shading and shadows. A forensic analysis of an image is therefore likely to consist of a combination of shading

[1]This measure of specificity was chosen to match [Kee et al. 2013], to which a comparison is made in Section 4.1.3.

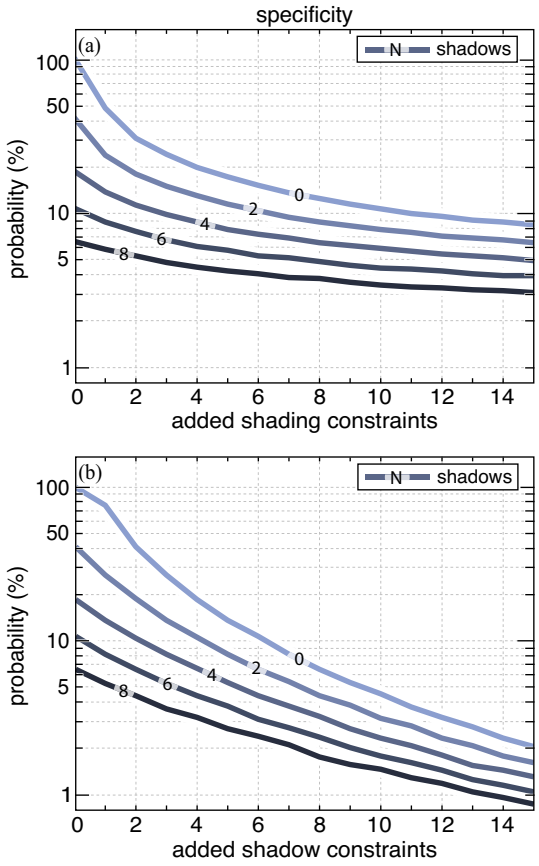


Fig. 8. The efficacy of combining shading and shadow constraints is shown in terms of specificity. Panel (a): specificity improves as shading constraints are added to $\{0, 2, 4, 6, 8\}$ shadow constraints. Panel (b): adding 8 or fewer additional shadow constraints instead of shading constraints yields similar improvement in specificity. Adding more than 8 shadow constraints yields greater specificity.

and shadow constraints. To assess the benefit afforded by this combination, we next compare and evaluate the usefulness of shading and shadow constraints in isolation and in combination.

Shown in Fig. 8(a) is the probability that a random light source will satisfy a combination of shading and shadow constraints drawn from the distribution found in the simulated scenes of Fig. 6, and [Kee et al. 2013]. Each colored curve corresponds to a fixed number of shadow constraints, and the horizontal axis corresponds to the number of additional shading constraints. Shown in Fig. 8(b) is a similar result in which shadow constraints are added, rather than shading constraints (horizontal axis). With a small number of constraints, these results show that the shading constraints are similar in specificity to the shadow constraints of [Kee et al. 2013]. For example, a random light source will satisfy 4 shading constraints with a probability of 20%, and for 4 shadow constraints the probability is a similar 19%. These results also show, however, that for a larger number of constraints (greater than 8), the shadow constraints are more specific than the shading constraints.

In a second set of simulations, we measured the probability that inconsistent lighting will be detected as a function of the angular distance between two light sources, and as a function of the number of shading and shadow constraints. Specifically, we simulated the creation of many forgeries, each by randomly selecting two light sources from among 49 directions, uniformly sampled on a hemisphere.[2] A set of m constraints was created to be consistent with the first light source, and a second set of n constraints was created to be consistent with the second light source, according to the distributions of constraint widths found in the simulated scenes. These two sets of constraints were then combined, and forgery was detected when the combined constraints were not satisfiable. The probability that these $m + n$ constraints were unsatisfiable was measured as a function of the median lighting difference, computed as the median angle between the projected directions of the two light sources at each constraint.

Shown in Fig. 9(a)-(f) are the results when using either $m = 5$ or $m = 15$ constraints from the first light source, and a varying number of n constraints from the second light source. In panels (a)-(b), the m constraints are from shading, and are combined with n shading constraints; in panels (c)-(d), the m constraints are from shadows, and are combined with n shading constraints; and in panels (e)-(f) the m constraints are from shadows, and are combined with n shadow constraints.

In general, shadow constraints are helpful for small lighting errors, while the shading constraints are slightly more effective overall. For example, using $m = 5$ shadow constraints, panel (c), rather than $m = 5$ shading constraints, panel (a), lighting errors below 30 degrees can be detected with slightly higher probability. In contrast, using $n = 4$ shading constraints, panels (a) and (c), a lighting inconsistency of 60 degrees can be detected with a probability of 80%; using shadows alone, panel (e), this same lighting inconsistency is detectable with probability 70%.

To summarize, the localization of a light source and the likelihood of detecting a lighting inconsistency improves as more constraints are added. Shading constraints offer an advantage of equal or better discrimination compared to shadow constraints. In part, this is because the shadow constraints consist of both cast and attached shadows. While cast shadows can be highly selective, attached shadows are less selective. Nevertheless, it will generally be advantageous to add as many constraints as possible to both localize the light source and increase the likelihood of finding inconsistent shading or shadow. Combining constraints from two distinct types of image information, shading and shadows, also offers the advantage that a forger hoping to avoid detection must now work to make both aspects of the image individually and mutually consistent.

4.2 Real world

Shown in the top left panel of Fig. 10 is an authentic photo. This scene was illuminated by a single light source located behind the camera and approximately 10 m from the scene. Shown in the top right panel are a variety of shading and shadow constraints, and magnified views are shown beneath.

Shading constraints (black dots) are computed from segments along occluding contours (outlined in yellow). These include cloth, ceramic, and organic materials. Only the reverse constraints are shown, corresponding to the case when the light is behind the camera. These shading constraints are physically consistent, as indicated by their feasible region (outlined in black).

Shadow constraints are specified by identifying correspondences between shadows and objects (white squares). Note that these constraints may be specified from shadow to object (labels **b**, **c**, **e**), or from object to shadow (labels **a**, **d**, **f**). As with the shading

[2]These 49 lights match [Kee et al. 2013], ensuring that the analysis faithfully compares the detection probability to our prior shadow-based work.

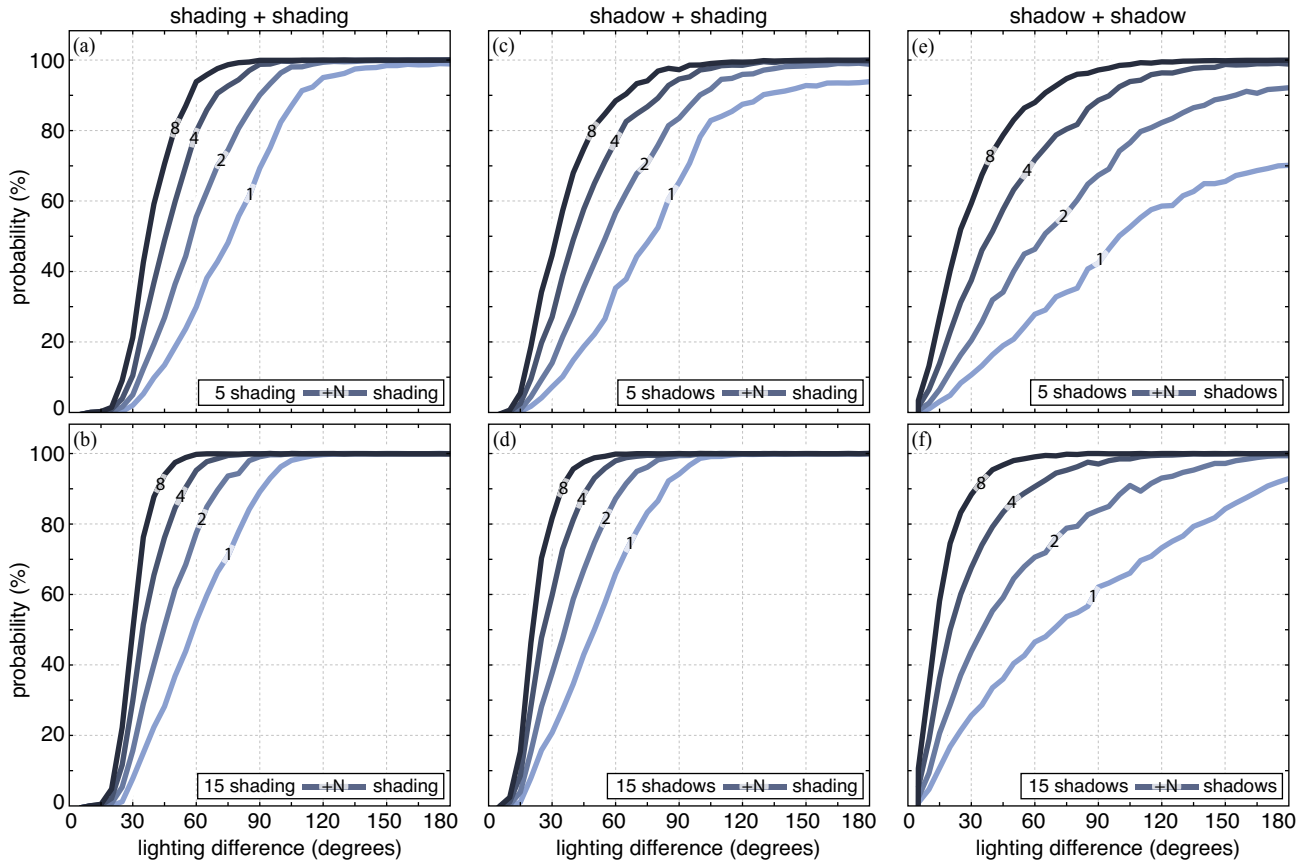


Fig. 9. The efficacy of combining shading and shadow constraints is shown in terms of detecting forgery. Forgery is detected by combining $m = \{5, 15\}$ constraints from a first light source, and n constraints from a second light source. Panels (a)-(b), 5 or 15 shading constraints are combined with n shading constraints. Panels (c)-(d), 5 or 15 shadow constraints are combined with n shading constraints. Panels (e)-(f), 5 or 15 shadow constraints are combined with n shadow constraints.

constraints, only the reverse constraints are shown. These shadow constraints are physically consistent, as indicated by their feasible region (outlined in white).

Because the individual feasible regions intersect, the shading and shadow constraints are mutually consistent. This common intersection of all shading and shadow constraints correctly suggests that the lighting in this scene is consistent with a single light source.

Shown in the bottom left panel of Fig. 10 is a composite image in which three objects (the squash and the green and orange shirts) have been replaced with versions photographed under different lighting. The same set of shading and shadows constraints were used as in the original version of this image. As shown in the lower right panel, however, the constraints are no longer satisfiable: the shadow constraints labeled **a** and **f** have a common intersection to the right, while the shading constraints labeled **8**, **9**, **0** have a common intersection to the left.

For the authentic photo in Fig. 10, a viable solution was found with a focal length $f = 60$ mm and image center $\boldsymbol{c} = (0, 0)$. For the composite photo, no combination of focal length and image center yields a viable solution. Although the shading inconsistencies on the modified objects are difficult to detect visually, our algorithm produces radically different constraints and easily identifies the forgery.

Shown in the top left panel of Fig. 11 is an authentic photo of a scene illuminated by the sun. Objects that produced distinct, pointed, shadows were placed in the scene to allow for precise localization of the light source. Shown in the top right panel are a variety of shading and shadow constraints; only the reverse constraints are shown. The available shading constraints are physically consistent, as indicated by their feasible region outlined in black. The shadow constraints are also physically consistent as indicated by their feasible region outlined in white. Furthermore, the shading and shadow constraints are mutually consistent because these two feasible regions intersect.

Shown in the bottom left panel of Fig. 11 is a composite image in which a pink ball has been added. The remaining cast shadows were then modified so as to be physically consistent with the inserted sphere. As seen in the lower right panel, the shading constraints are physically consistent. This is because the pink ball was photographed under lighting that was similar to the original scene. The shadow constraints are also physically consistent. Because these two feasible regions do not intersect, however, the shading and shadow constraints together are not mutually physically consistent.

Shown in top left panel of Fig. 12 is an authentic photo that at first glance may appear unlikely[3]. Segments of approximately con-

[3]Its authenticity is corroborated by additional photos on the source website.

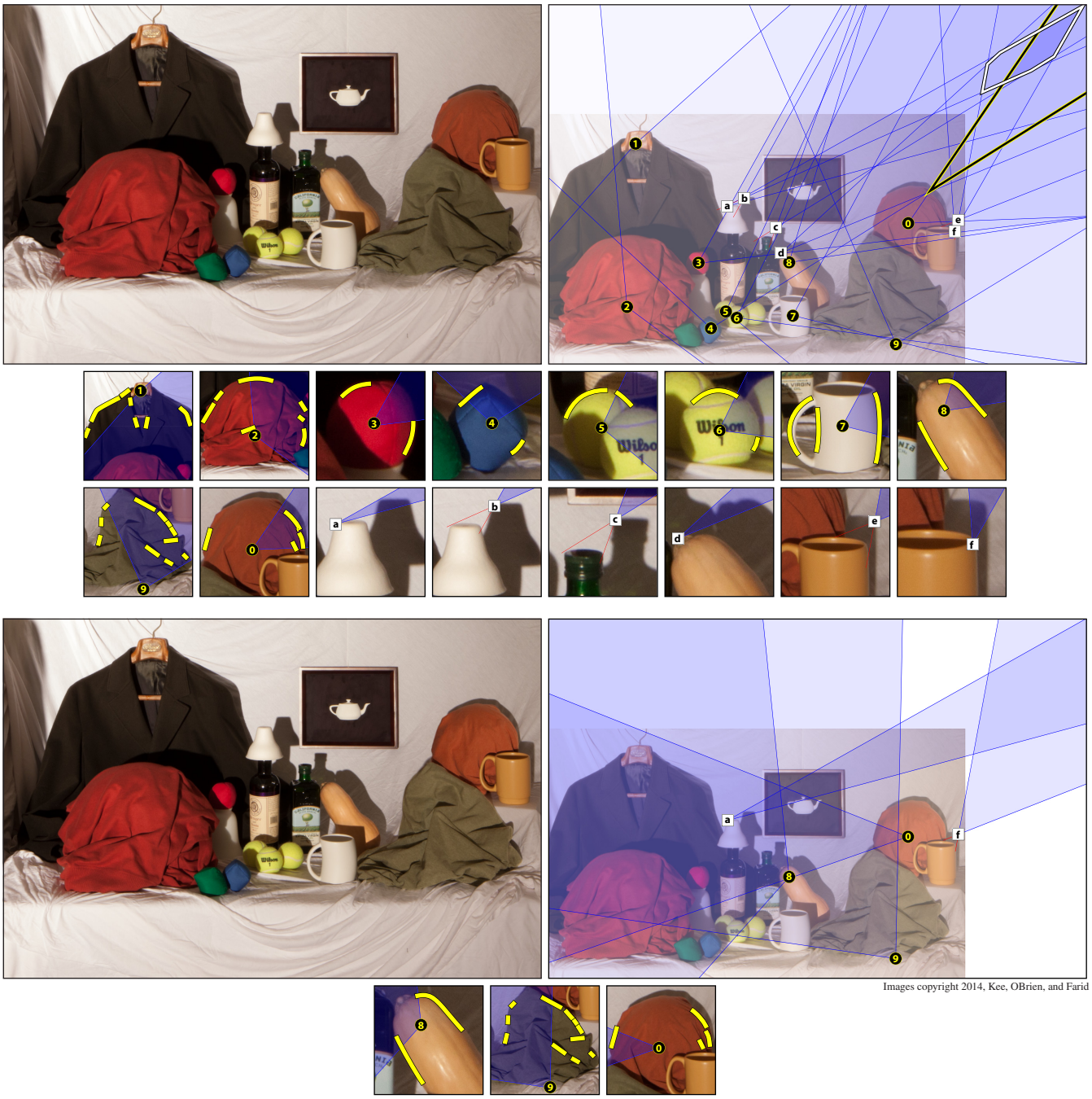


Fig. 10. A complete shading and shadow analysis of an authentic (top) and composite (bottom) photo. Top: shading constraints, denoted by black dots, are selected from a variety of occluding contours each outlined in yellow. The outlined black region is the common intersection of these constraints, showing that they are all mutually consistent. Shadow constraints are denoted by white squares. The outlined white region is the common intersection of these constraints, showing that the shadows are mutually consistent. Because the viable shading and shadow regions intersect, the lighting in this photo is physically consistent. Bottom: in this composite photo three objects with different lighting were inserted (the squash and the green and orange shirts). The five constraints shown in the lower right are inconsistent, revealing this photo to be a composite.

stant reflectance are selected along a variety of occluding contours, labeled ❶ – ❽. The outlined black region is the common intersection of these constraints. Shadow constraints, labeled [a] – [h] have a common intersection outlined in white. Because the viable shading and shadow regions intersect, the lighting in this photo is physically consistent.

Shown in Fig. 13 is another photo that may, at first glance, appear unlikely (there does in fact exist a putting green on the front lawn of

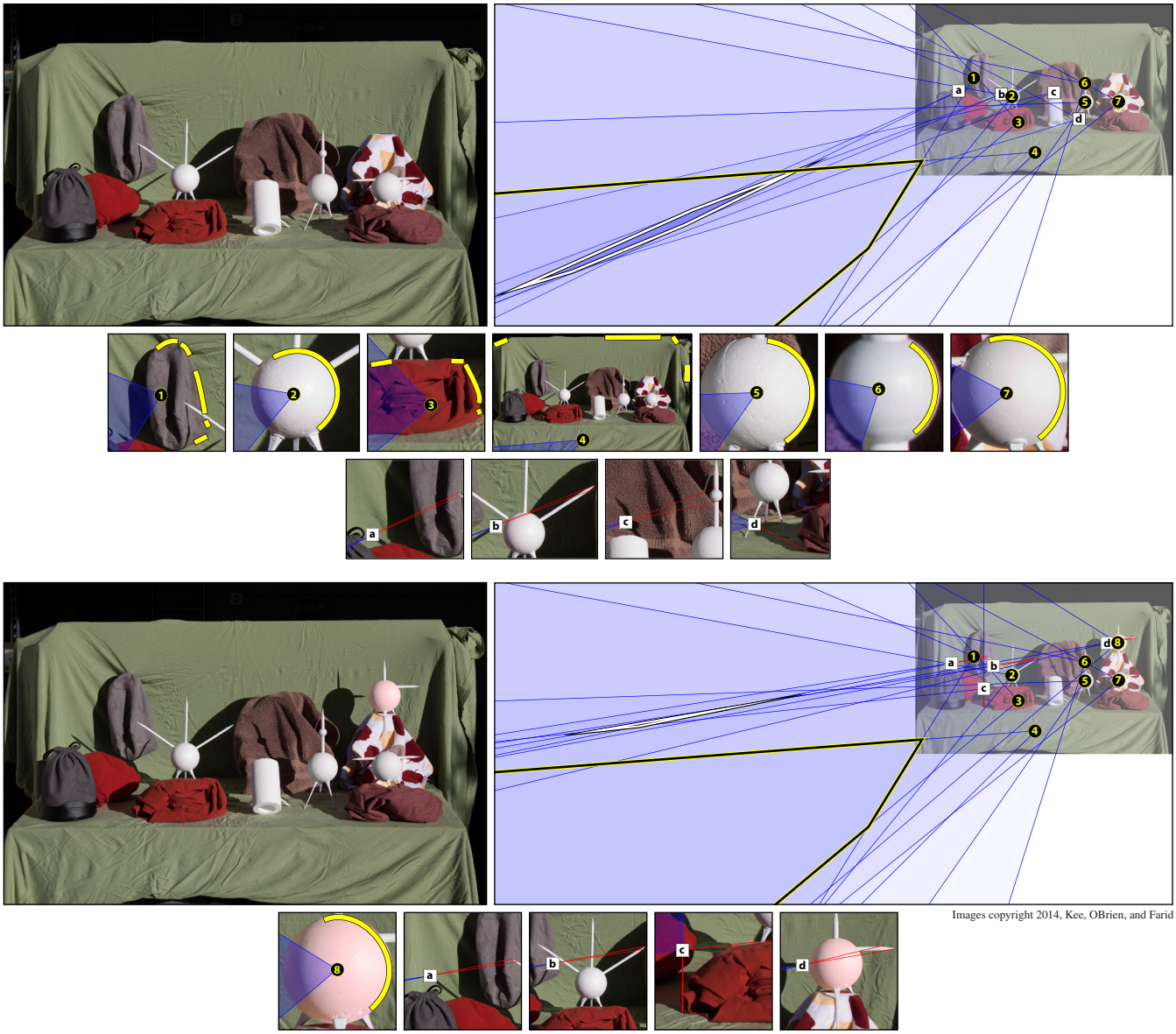


Fig. 11. A complete shading and shadow analysis of an authentic (top) and composite (bottom) photo. Top: shading constraints, denoted by black dots, are selected from a variety of occluding contours each outlined in yellow. The outlined black region is the common intersection of these constraints, showing that they are all mutually consistent. Shadow constraints are denoted by white squares. The outlined white region is the common intersection of these constraints, showing that the shadows are mutually consistent. Because the viable shading and shadow regions intersect, the lighting in this photo is physically consistent. Bottom: the pink ball in the upper right along with its cast shadow were inserted into the image. The remaining cast shadows were then digitally altered to be mutually consistent, as seen by the outlined white region. Despite the addition of the shading constraint from the pink ball, the shading on the objects is also mutually consistent, as seen by the outlined black region. The combination of shading and shadow constraints, however, is inconsistent: the viable region of the shading does not intersect the viable region of the shadows.

the White House). Segments are selected along the occluding contours of Vice President Biden, President Obama, and the bushes in the background, labels ❶ – ❼. These constraints intersect in the region outlined in black, suggesting that the shading in the foreground and background is physically consistent. The shadow constraints [a] – [i] are also consistent, and intersect in the region outlined in white. Because these viable shading and shadow regions intersect, the lighting in this photo is correctly determined to be consistent.

Shown in the top panel of Fig. 14 is a composite photo in which the man and boy were inserted from an unrelated photo. Despite the somewhat fantastic nature of the photo, the shading and shadows appear visually plausible. Shown in the lower left are a subset of shading and shadow constraints. Shading constraints are specified along the woman's shoulders ❶, the two hats ❷, ❹, and along the

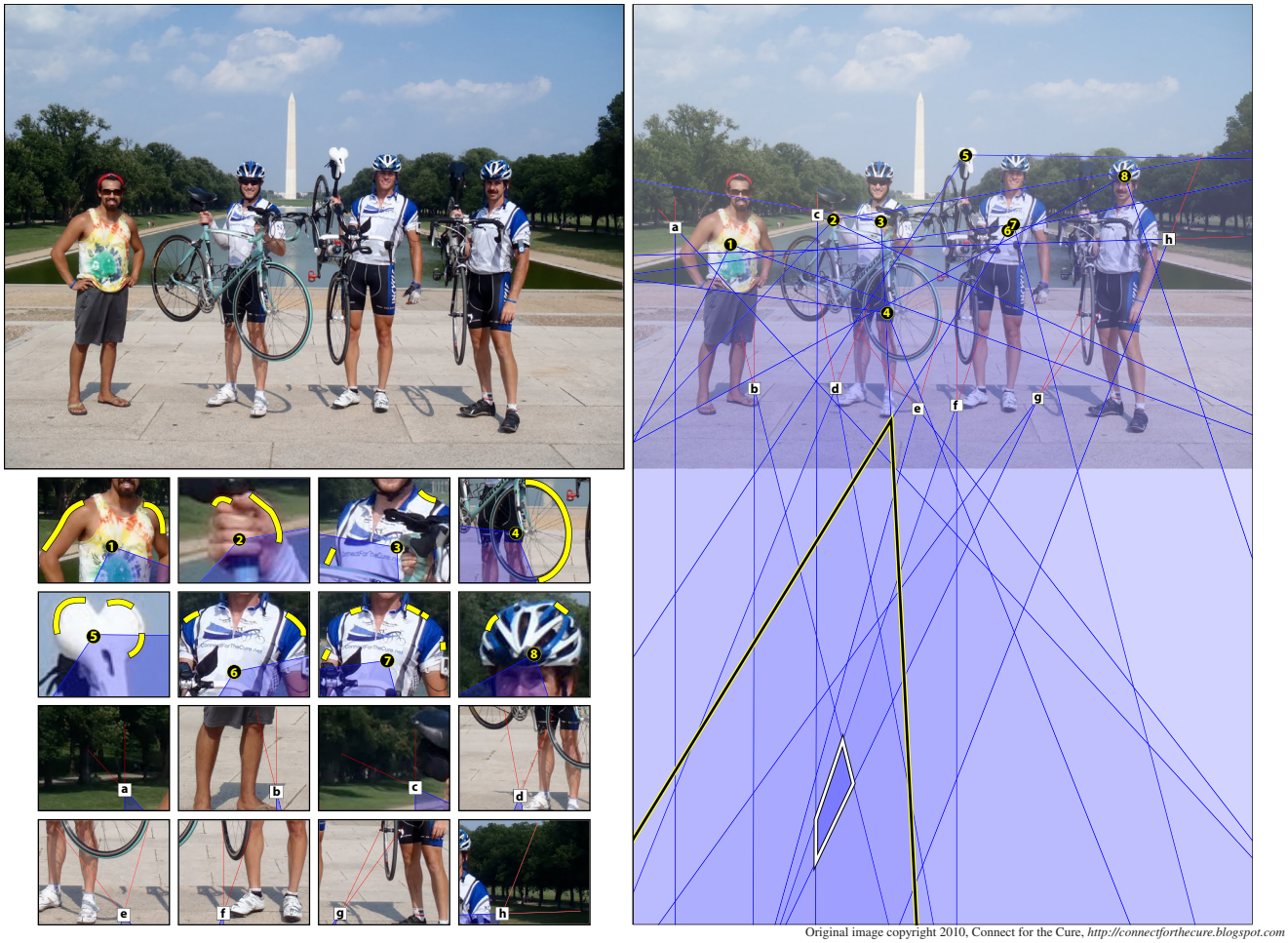


Fig. 12. An authentic photo and the result of a shading and shadow consistency check. The region of plausible light positions that satisfy the shading constraints is outlined in black, and for the shadows, in white. These regions intersect, correctly suggesting that the shading and shadow is physically consistent.

man's hands, ❸. Cast shadow constraints are labeled [a], [b], [d], [e], [f], and attached shadows are labeled [c], [g]. This set of shading and shadow constraints are not satisfiable. Shown in the lower right are two subsets of unsatisfiable constraints that separately entail both the man (top) and boy (bottom) There does not exist a light source in front of (shaded red) or behind the camera (shaded blue) that satisfies these constraints under any plausible camera.

Shown in Fig. 1 is a falsified photo that was downloaded from the Internet. A variety of shading constraints are specified along the occluding contour of the shark, labels ❶ – ❹, which are physically consistent due to their common intersection (black outline). Also specified are shadow constraints, [a] – [c], which are also physically consistent (white outline). However, the combination of these shading and shadow constraints is not physically consistent, and correctly suggest that the photo has been falsified.

Shown in Fig. 15 is a falsified image in which a 3-D model of the Stanford dragon has been inserted [Karsch et al. 2011]. This image was created by estimating the scene lighting and the camera parameters, and then rendering and inserting the dragon to match these parameters. A total of 9 shading constraints are specified along the occluding contours of the dragon, and on the snow in the foreground, labels ❶ – ❾. These shading constraints are physically consistent and intersect in the region outlined in black. An additional 6 shadow constraints were specified, labels [a] – [g], which are also physically plausible, and intersect in the region outlined in white. These two viable regions (black, white) overlap, thus suggesting that the shading and shadow are physically consistent. This illustrates a failure case for our algorithm, and provides some indication of the operations that might be needed to circumvent an analysis of shading and shadow.

The consistency of the constraints in Fig. 15 can also be compared to the ground-truth location of the light in the original scene (less the dragon). Notice that the shading and shadow constraints coincide with the shadow of the camera (lower left corner), which identifies the projected location of the light source. The camera's shadow and the projected light source are collocated because the light source, camera, and the shadow cast by the camera lie along a ray in the world that passes through (or near) the center of projection. All points along this ray, including the shadow, camera, and light source, must therefore project to the same point: the shadow of the camera. We can therefore verify that the shadow and shading constraints correctly encompass ground-truth location of the light source.

Next, we use the authentic photos (described above) to compare the specificity of real-world constraints with our simulated specificity, Section 4.1.2. Specifically, we collected constraints from each of the authentic photos in Figs. 10, 11, 12, and 13. The constraints in each photo were randomly combined into sets of size n, and the mean specificity of n constraints was computed. Shown in Fig. 7 (dashed gray) is this specificity, which matches closely with the specificity of constraints from the rendered images. This real-world specificity improves more quickly, however, because real-world constraints are computed from occluding contours of varying spatial extent.

To conclude, note that it is possible to incorrectly label an image as fake if an analyst chooses to analyze an object that violates our imaging assumptions. Shown in Fig. 16 is an analysis in which two objects violate our assumptions of constant Lambertian reflectance and a distant point light source. The constraints for these two objects (labeled ❼ and ❽) incorrectly suggest that the lighting is not consistent. This error occurs because the objects violate our assumptions: one is highly specular brushed metal, the other is a translucent glass sphere that is shadowed and is illuminated by an indirect bounce light. This example emphasizes the importance of having a knowledgeable analyst who understands how these tests should be applied. Our forensic method therefore shifts the debate over whether the shading and shadows look consistent to the more objective debate of whether particular objects satisfy the required reflectance and illumination conditions.

5. DISCUSSION

We have developed a forensic technique for measuring the consistency, or lack thereof, of shading and cast and attached shadows. Unlike previous approaches, instead of explicitly trying to estimate the lighting in a scene, we seek to extract and combine a multitude of partial constraints on the projected location of the light source. This approach allows us to contend with situations in which the lighting is under-constrained and cannot be explicitly estimated, a common problem in a forensic setting in which only a single image of uncertain origin is available.

Our analysis makes minimal assumptions: a scene containing a single distant point light source illuminating a Lambertian surface of constant reflectance and imaged under linear perspective. Although this precludes the analysis of scenes lit by multiple sources or diffuse area lights, it includes the common situation of outdoor scenes. In addition, we have shown that modest deviations from these assumptions do not severely hamper our technique. Should other violations of our assumptions (e.g., lens distortion) become an issue, they could be incorporated into our model of uncertainty.

We note that our model of uncertainty employs confidence regions, rather than explicit probability distributions. By intersecting confidence regions, we answer a forensic question: does there exist a light position that satisfies each constraint with probability greater than p? We therefore require an image to have a consistent interpretation among all constraints. Our use of probability differs from the more common framework in which distributions are multiplied, and maximum likelihood solutions are sought. That second approach is appropriate when one assumes that inconsistent constraints result from noise, and should be ignored.

Lastly, note that a talented forger could craft a forgery such that the shading and shadows are consistent by our measure. For example, a photo could be augmented with a synthetic object that is rendered with an approximation of the scene lighting [Karsch et al. 2011]. Or an object can be re-shaded and its shadows warped to conform with the lighting. While it is possible to falsify consistent shading and shadows, our approach does (at a minimum) make it more difficult and time consuming to create a convincing forgery.

REFERENCES

BECKMANN, P. 1987. *Scattering of Electromagnetic Waves from Rough Surfaces*. Number 0-89006-238-2. Artech House Publishers, New York, NY.

CHEN, M., FRIDRICH, J., GOLJAN, M., AND LUKAS, J. 2008. Determining image origin and integrity using sensor noise. *IEEE Transactions on Information Forensics and Security (TIFS) 3,* 1 (March), 74–90.

FARID, H. 2009. A survey of image forgery detection. *IEEE Signal Processing Magazine 2,* 26, 16–25.

FARID, H. AND BRAVO, M. J. 2010. Image forensic analyses that elude the human visual system. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 7541.

FRIDRICH, J. 2009. Digital image forensic using sensor noise. *IEEE Signal Processing Magazine 26,* 2, 26–37.

FRIDRICH, J., SOUKAL, D., AND LUKAS, J. 2003. Detection of copy move forgery in digital images. In *In Proceedings of Digital Forensic Research Workshop (DFRWS)*.

HARA, K., NISHINO, K., AND IKEUCHI, K. 2005. Light source position and reflectance estimation from a single view without the distant illumination assumption. *IEEE Trans. Pattern Anal. Mach. Intell. 27,* 4 (Apr.), 493–505.

HOUGEN, D. AND AHUJA, N. 1993. Estimation of the light source distribution and its use in integrated shape recovery from stereo and shading. In *Computer Vision, 1993. Proceedings., Fourth International Conference on*. 148 –155.

IRAWAN, P. 2008. Appearance of woven cloth. Ph.D. thesis, Cornell University, Ithaca, NY, USA.

JACOBSON, J. AND WERNER, S. 2004. Why cast shadows are expendable: Insensitivity of human observers and the inherent ambiguity of cast shadows in pictorial art. *Perception 33,* 11, 1369–1383.

JAKOB, W. 2010. Mitsuba renderer. http://www.mitsuba-renderer.org.

JOHNSON, M. AND FARID, H. 2007. Exposing digital forgeries in complex lighting environments. *IEEE Transactions on Information Forensics and Security 3,* 2, 450–461.

JOHNSON, M. K. AND FARID, H. 2005. Exposing digital forgeries by detecting inconsistencies in lighting. In *Proceedings of the 7th workshop on Multimedia and security*. MM&Sec '05. ACM, New York, NY, USA, 1–10.

JOHNSON, M. K. AND FARID, H. 2006. Exposing digital forgeries through chromatic aberration. In *Proceedings of the 8th workshop on multimedia and security*. MM&Sec'06. ACM, 48–55.

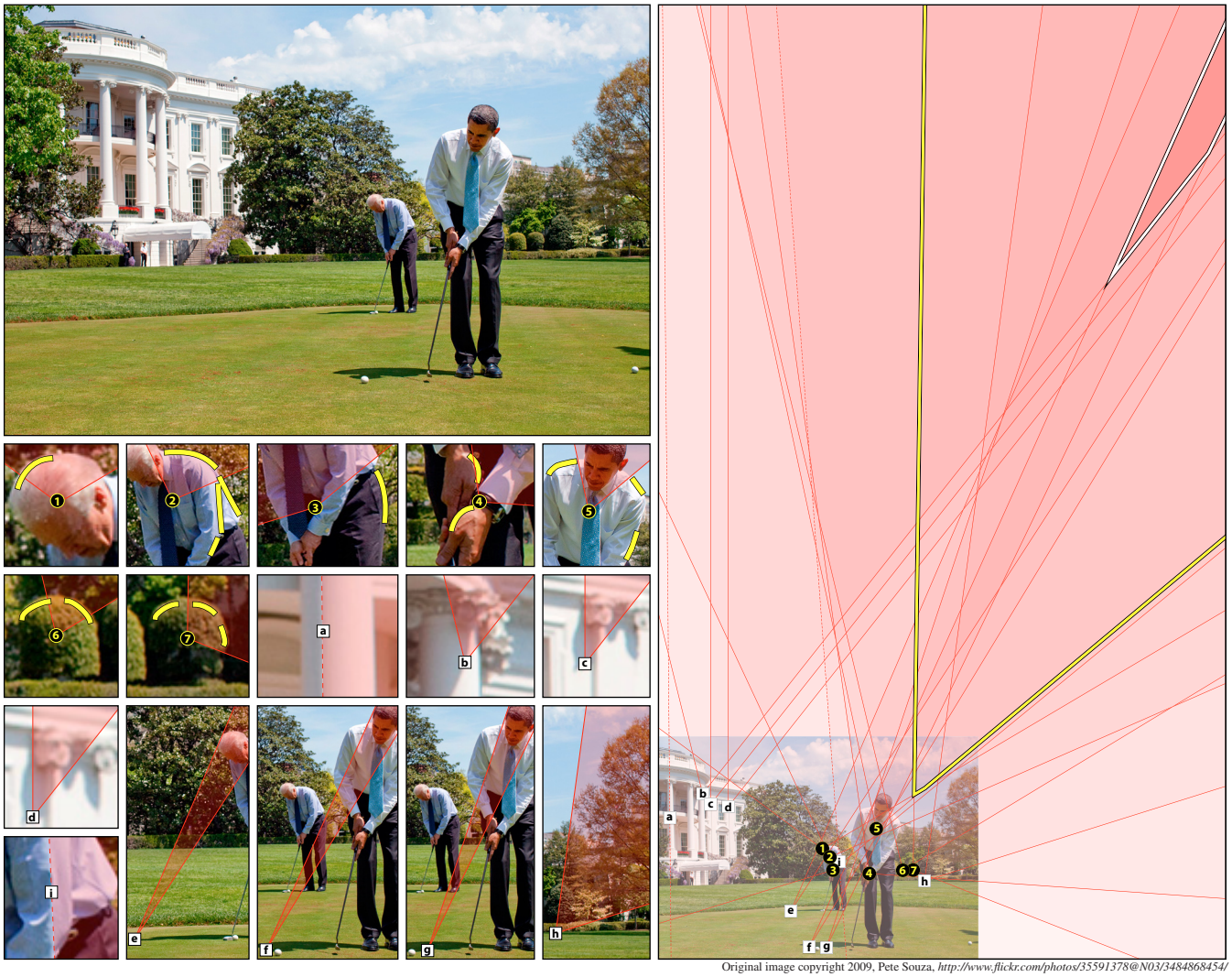
JOHNSON, M. K., STORK, D. G., BISWAS, S., AND FURUICHI, Y. 2008. Inferring illumination direction estimated from disparate sources in paintings: an investigation into Jan Vermeer's Girl with a Pearl Earring. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 6810.

KARSCH, K., HEDAU, V., FORSYTH, D., AND HOIEM, D. 2011. Rendering synthetic objects into legacy photographs. In *SIGGRAPH Asia*. 157:1–157:12.

KEE, E. AND FARID, H. 2010. Exposing digital forgeries from 3-D lighting environments. In *IEEE International Workshop on Information Forensics and Security (WIFS)*. 1–6.

KEE, E., JOHNSON, M. K., AND FARID, H. 2011. Digital image authentication from JPEG headers. *IEEE Transactions on Information Forensics and Security 6,* 3 (September), 1066–1075.

KEE, E., O'BRIEN, J., AND FARID, H. 2013. Exposing photo manipulation with inconsistent shadows. *ACM Trans. Graph. 32,* 3 (July), 28:1–28:12.



Original image copyright 2009, Pete Souza, *http://www.flickr.com/photos/35591378@N03/3484868454/*

Fig. 13. Our analysis finds that the shading and shadows in this photo of President Obama and Vice President Biden are physically consistent with a single light source. The region of plausible light positions that satisfy the shading constraints is outlined in black, and for the shadows, in white. These regions intersect, correctly suggesting that the shading and shadow is physically consistent.

KIM, C.-Y., PETROV, A. P., CHOH, H.-K., SEO, Y.-S., AND KWEON, I.-S. 1998. Illuminant direction and shape of a bump. *J. Opt. Soc. Am. A 15,* 9 (Sep), 2341–2350.

KIRCHNER, M. 2010. Efficient estimation of CFA pattern configuration in digital camera images. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 7541.

LI, Y., LIN, S., LU, H., AND SHUM, H.-Y. 2003. Multiple-cue illumination estimation in textured scenes. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. 1366 –1373 vol.2.

LIN, Z., WANG, R., TANG, X., AND SHUM, H.-Y. 2005. Detecting doctored images using camera response normality and consistency. *IEEE Conference on Computer Vision and Pattern Recognition 1*, 1087–1092.

LIU, Q., CAO, X., DENG, C., AND GUO, X. 2011. Identifying image composites through shadow matte consistency. *IEEE Transactions on Information Forensics and Security 6,* 3, 1111–1122.

MAHDIAN, B. AND SAIC, S. 2009. Detecting double compressed jpeg images. In *3rd International Conference on Crime Detection and Prevention (ICDP)*. 1–6.

MARSCHNER, S. R. AND GREENBERG, D. P. 1997. Inverse Lighting for Photography. In *In Proceedings of IS&T/SID Fifth Color Imaging Conference*. 262–265.

NILLIUS, P. AND EKLUNDH, J.-O. 2001. Automatic estimation of the projected light source direction. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. *Computer Vision and Pattern Recognition*.

O'BRIEN, J. F. AND FARID, H. 2012. Exposing photo manipulation with inconsistent reflections. *ACM Transactions on Graphics 31,* 1 (Jan.), 4:1–11. Presented at SIGGRAPH 2012.

OREN, M. AND NAYAR, S. 1994. Generalization of Lambert's Reflectance Model. In *ACM 21st Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*. 239–246.

OSTROVSKY, Y., CAVANAGH, P., AND SINHA, P. 2005. Perceiving illumination inconsistencies in scenes. *Perception 34*, 1301–1314.

Fig. 14. A composite photo in which the man and boy were inserted. Bottom left: shading constraints are labeled ❶ – ❹, cast shadow constraints specified from shadow to object object are labeled [a], [b], [d], cast shadows specified from object to shadow are labeled [e], [f], and attached shadows are labeled [c], [g]. Bottom right: two subsets of mutually inconsistent constraints revealing this image to be a composite with inconsistent lighting.

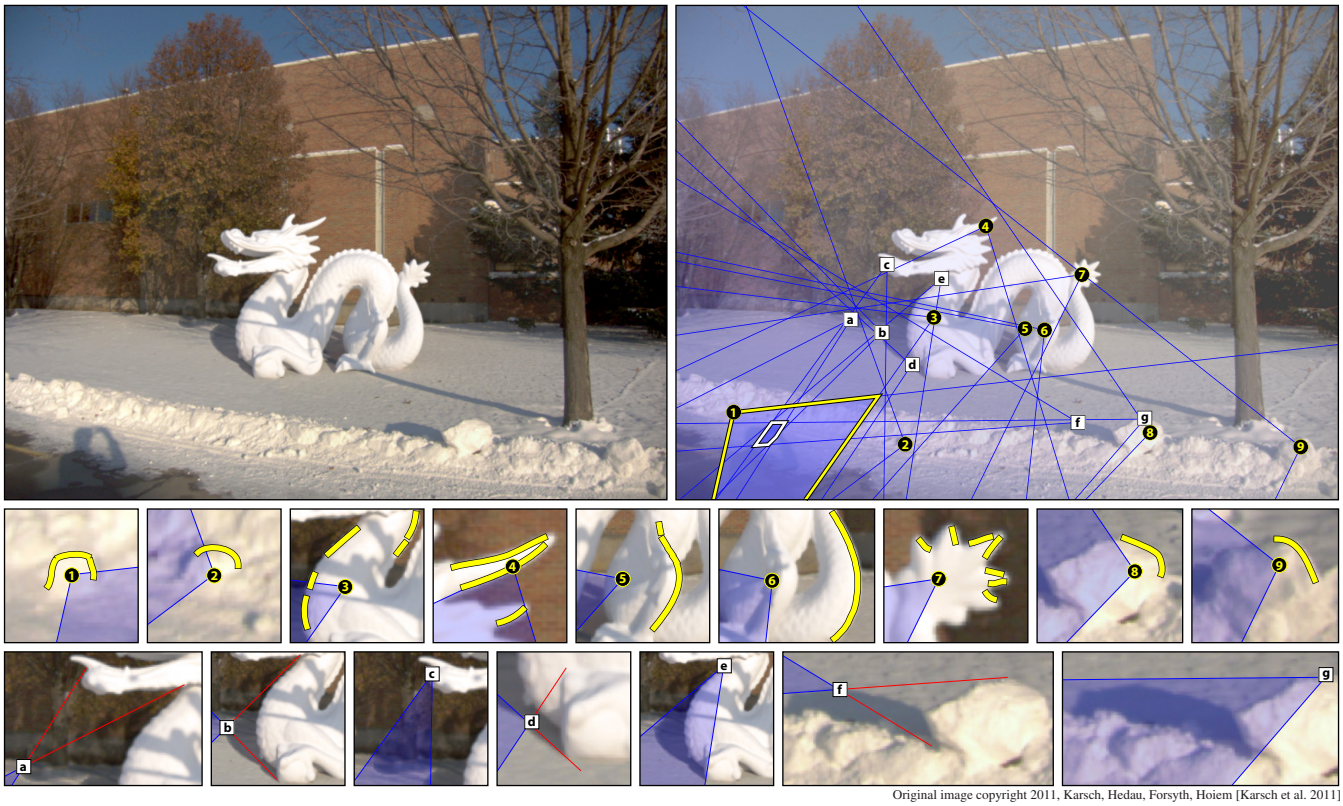


Fig. 15. Shown is a falsified image in which a 3-D model has been inserted [Karsch et al. 2011]. A total of 9 shading constraints are specified on the dragon and in the snow, labels ❶ – ❾, with a common intersection outlined in black. An additional 6 shadow constraints are specified, labels [a] – [g] have a common intersection outlined in white. Because these two viable regions overlap, we determine that the shading and shadows are physically consistent. This illustrates a failure case for our algorithm, and provides some indication of the operations that might be needed to circumvent an analysis of shading and shadow.

POPESCU, A. AND FARID, H. 2005a. Statistical tools for digital forensics. In *Information Hiding*, J. Fridrich, Ed. Lecture Notes in Computer Science, vol. 3200. Springer Berlin / Heidelberg, 395–407.

POPESCU, A. C. AND FARID, H. 2005b. Exposing digital forgeries in color filter array interpolated images. *IEEE Transactions on Signal Processing 53,* 10, 3948–3959.

PREETHAM, A., SHIRLEY, P., AND SMITS, B. 1999. A practical analytic model for daylight. In *Proceedings of the 26th annual conference on computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 91–100.

RIESS, C. AND ANGELOPOULOU, E. 2010. Scene illumination as an indicator of image manipulation. In *Information Hiding*, R. Böhme, P. W. Fong, and R. Safavi-Naini, Eds. Lecture Notes in Computer Science, vol. 6387. Springer Berlin Heidelberg, 66–80.

ROCHA, A., SCHEIRER, W., BOULT, T. E., AND GOLDENSTEIN, S. 2011. Vision of the unseen: Current trends and challenges in digital image and video forensics. *ACM Computing Surveys (CSUR) 43,* 4, 26:1–26:42.

STORK, D. AND JOHNSON, M. 2006a. Computer vision, image analysis, and master art: Part 2. *IEEE MultiMedia 13,* 4, 12–17.

STORK, D. AND JOHNSON, M. 2006b. Estimating the location of illuminants in realist master paintings computer image analysis addresses a debate in art history of the baroque. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*. Vol. 1. 255 –258.

WANG, Y. AND SAMARAS, D. 2003. Estimation of multiple directional light sources for synthesis of augmented reality images. *Graphical Models (Special Issue on Pacific Graphics) 65,* 4 (July), 185–205.

ZHANG, W., CAO, X., ZHANG, J., ZHU, J., AND WANG, P. 2009. Detecting photographic composites using shadows. In *IEEE International Conference on Multimedia and Expo*. 1042–1045.

ZHANG, Y. AND YANG, Y.-H. 2001. Multiple illuminant direction detection with application to image synthesis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 23,* 8 (August), 915 –920.

ZHENG, Q. AND CHELLAPPA, R. 1991. Estimation of illuminant direction, albedo, and shape from shading. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on*. 540 –545.

ZHOU, W. AND KAMBHAMETTU, C. 2008. A unified framework for scene illuminant estimation. *Image Vision Computing 26,* 3 (March), 415–429.

APPENDIX

A. LINE AND HALF-PLANE CONSTRAINTS

Line constraint

A line segment constraint is defined by a point $\tilde{\boldsymbol{p}}$ and direction $\boldsymbol{d}$:

$$\tilde{\boldsymbol{p}} = \mathbf{K}\mathbf{w}_4 \tag{31}$$

$$\boldsymbol{d} = \text{sign}\,(\text{s}_{0_z})(\boldsymbol{p} - \boldsymbol{s}_0) \ . \tag{32}$$

where $\mathbf{K}$ is the intrinsic camera matrix, $\mathbf{w}_4$ is the vector formed by the first three components of the singular vector $\mathbf{v}_4$ from Equa-

tion (5), and $\boldsymbol{s}_0$ is the least-squares estimate of the 3-D light direction as described in Section 3.1).

As can be seen in Fig. 2(e), the line segment corresponding to the forward constraint (illustrated in red) is simply the line that connects $\boldsymbol{p}$ and $\boldsymbol{s}_0$. Multiplying the direction $\boldsymbol{d}$ by $\text{sign}(\text{s}_{0_z})$ distinguishes between the forward (red) and backward (blue) constraint. That is, when the estimated light direction, $\mathbf{s}_0$, is in front of the camera, $\text{s}_{0_z} < 0$ (by convention) and $\boldsymbol{d}$ is defined as the line segment connecting $\tilde{\boldsymbol{p}}$ to $\boldsymbol{s}_0$. When the estimated light direction is behind the camera, then $\text{s}_{0_z} > 0$ and $\boldsymbol{d}$ is defined as the line segment connecting $\boldsymbol{s}_0$ to $\tilde{\boldsymbol{p}}$.

Half-plane constraint

A half-plane constraint is defined by a point $\tilde{\boldsymbol{p}}$ and direction $\boldsymbol{d}$:

$$\tilde{\boldsymbol{p}} = \mathbf{K}\mathbf{w}_4 \tag{33}$$

$$\boldsymbol{d} = \text{sign}(\tilde{\boldsymbol{h}} \cdot \tilde{\text{s}}_0) \begin{pmatrix} \tilde{h}_1 & \tilde{h}_2 \end{pmatrix}^T \quad , \tag{34}$$

where $\mathbf{K}$ is the intrinsic camera matrix, $\mathbf{w}_4$ is the vector formed by the first three components of the singular vector $\mathbf{v}_4$ from Equation (5), $\tilde{\boldsymbol{h}} = \mathbf{K}\mathbf{w}_4 \times \mathbf{K}\mathbf{w}_3$ is the homogeneous line that connects the images of singular vectors $\mathbf{w}_3$ and $\mathbf{w}_4$, and $\tilde{h}_1$ and $\tilde{h}_2$ are the first and second components of $\tilde{\boldsymbol{h}}$.

As shown in Fig.2(f), the line $\tilde{\boldsymbol{h}}$ defines the boundary between the forward (red) and reverse (blue) half-plane constraints. These correspond to the case when the light is in front of and behind the camera, respectively.

The line $\tilde{\boldsymbol{h}}$ can therefore be used to compute the half-plane direction, $\boldsymbol{d}$. Specifically, $\boldsymbol{d}$ is constructed from the first two components of the line, $\tilde{h}_1$ and $\tilde{h}_2$, which specify a vector that is orthogonal to the half-plane boundary. Multiplication by $\text{sign}(\tilde{\boldsymbol{h}} \cdot \tilde{\text{s}}_0)$ orients the direction $\boldsymbol{d}$ to identify the side of the half-plane into which light sources in front of the camera project.

To derive the factor $\text{sign}(\tilde{\boldsymbol{h}} \cdot \tilde{\text{s}}_0)$, note that a light $\boldsymbol{s}_0$ lies on the side ψ of the boundary:

$$\psi = \text{sign}\left(\tilde{\boldsymbol{h}}_{1,2} \cdot \boldsymbol{s}_0 + \tilde{h}_3\right) \tag{35}$$

$$= \text{sign}\left(\tilde{\boldsymbol{h}} \cdot \frac{\tilde{\boldsymbol{s}}_0}{-\text{s}_{0_z}}\right) \tag{36}$$

$$= -\text{sign}(\text{s}_{0_z})\text{sign}\left(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0\right) \quad , \tag{37}$$

where $\tilde{\boldsymbol{h}}_{1,2}$ is a vector of the first two components of $\tilde{\boldsymbol{h}}$, and the division by $-\text{s}_{0_z}$ normalizes the homogeneous vector $\tilde{\boldsymbol{s}}_0$. The negative is due to our assumption that the camera faces the $-z$ axis.

When the light is in front of the camera, it projects to side $\psi = \text{sign}(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0)$ because $\text{s}_{0_z} < 0$. If these forward lights project to side $\psi = -1$, then $\text{sign}(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0) = -1$ and we define $\boldsymbol{d}$ to point to the opposite side of the half-plane: $\boldsymbol{d} = \text{sign}(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0)\tilde{\boldsymbol{h}}_{1,2}$.

Similarly, when the light is behind the camera it projects to side $\psi = -\text{sign}(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0)$ because $\text{s}_{0_z} > 0$. If lights behind the camera project to side $\psi = 1$, then $\text{sign}(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0) = -1$ and we define $\boldsymbol{d}$ to point to the opposite side of the half-plane: $\boldsymbol{d} = \text{sign}(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0)\tilde{\boldsymbol{h}}_{1,2}$.

B. ESTIMATING UNCERTAINTY FUNCTIONS

The reliability of the estimated line and half-plane constraints, Section 3.2, is impacted by image noise, luminance non-linearities, resolution, and violations of the assumed imaging model. The functions $\Phi_{\text{L}}(\kappa_{\text{L}}, \theta, \rho)$ and $\Phi_{\text{H}}(\kappa_{\text{H}}, \theta, \rho)$, described in Section 3.5, model the uncertainty that results from these factors. The specific form of these functions is determined through a set of large-scale simulations, as described below.

The functions $\Phi_{\text{L}}(\kappa_{\text{L}}, \theta, \rho)$ and $\Phi_{\text{H}}(\kappa_{\text{H}}, \theta, \rho)$ were estimated by simulating images of occluding contours. Specifically, 200 random blobby shapes were generated and imaged under 50 random camera configurations, yielding $10,000$ distinct occluding contours. The focal lengths of the cameras were drawn uniformly at random between 18 mm and 200 mm. Each shape was placed randomly in the virtual scene so as to occupy approximately 235×235 pixels.

A total of 3 billion images of the occluding contours were created under differing lighting conditions. Contours were shaded by specifying a random surface albedo, ν in $[0.1,\ 0.5]$. Light directions, $\mathbf{s}$, were chosen randomly from the unit sphere, excluding directions more than 45° behind the contour. The contour was shaded according to the Lambertian reflectance equation, $r = \nu(\mathbf{s} \cdot \mathbf{n})$. To this primary illumination, a secondary light was added to simulate inter-reflections and other effects. Secondary lights were drawn randomly from the unit sphere, and their peak intensity on the occluding contour was, on average, 15% of the peak primary radiance (standard deviation 5%). An ambient illumination was also added, drawn from a normal distribution with mean 0 and standard deviation 0.1 (negative values were excluded). Lastly, the final rendered image was subjected to 1% additive Gaussian noise, gamma correction $r^{1/\gamma}$ with γ drawn uniformly in $[1.2,\ 1.8]$, and 8-bit quantization. See Appendix C for details.

The 3 billion images afforded a dense sampling of the wide range of line and half-plane constraints that result when image and lighting conditions vary, and when portions of the contour are obscured in an image. The illuminated portion of the contour in each image was divided into $10 - 20$ segments, and a random subset were used to build line and half-plane constraints. The 3-D normals on the contour were estimated, Equation (30), by selecting a random focal length, f, within δ_f log-units of the true focal length, and a random image center, $\boldsymbol{c}$, within $\delta_{\boldsymbol{c}}$ units of the true image center. (The values of δ_f and $\delta_{\boldsymbol{c}}$ are discussed below.) The true image center was selected uniformly at random within a circle with diameter 12% of the image height. The occluding normals were used to construct the linear system, and thus the line constraint, $(\boldsymbol{p}_{\text{L}},\ \boldsymbol{d}_{\text{L}})$, and half-plane constraint, $(\boldsymbol{p}_{\text{H}},\ \boldsymbol{d}_{\text{H}})$. The angle between the constraint directions $\boldsymbol{d}$ and the vector from $\boldsymbol{p}$ to the light source was then computed, up to an inversion of $\boldsymbol{d}$ for lights behind the camera. This produced 3 billion samples of these angles, ϕ_{L} and ϕ_{H}, from which the functions, Φ_{L} and Φ_{H}, were estimated.

The functions, $\Phi_{\text{L}}(\kappa_{\text{L}}, \theta, \rho)$ and $\Phi_{\text{H}}(\kappa_{\text{H}}, \theta, \rho)$, were estimated by quantizing the function domain and computing a confidence interval on the values ϕ_{L} and ϕ_{H} associated with each quantized location. The condition, $\kappa_{\text{L}} = \lambda_3/\lambda_1$ and $\kappa_{\text{H}} = \lambda_2/\lambda_1$, was quantized, in the log domain, into 24 levels between 0.001 and 0.7 (the minimum and maximum practical values that were encountered). The number of surface normal directions, θ, was measured by computing the angle of each 2-D image normal, constructing a histogram of these angles, and computing the number of degrees comprised by the non-empty bins. This number of directions, θ, was quantized into 36 levels (10° increments between 0° and 360°). The R-value, ρ, of the least squares solution, Equation (4), was quantized into 30 levels between 0.7 and 1. Note that these particular quantization ranges and levels are not essential, and were chosen to give reasonable resolution and intuitive quantization steps. The value of the functions, Φ_{L} and Φ_{H}, at each quantized position specified the angle below which 99.9% of the samples, ϕ_{L} and ϕ_{H}, fell.

Recall that the functions Φ_{L} and Φ_{H} are estimated by assuming that the camera focal length, f and image center $\boldsymbol{c}$ are known within δ_f and $\delta_{\boldsymbol{c}}$ units of their true values. The effect of these precisions

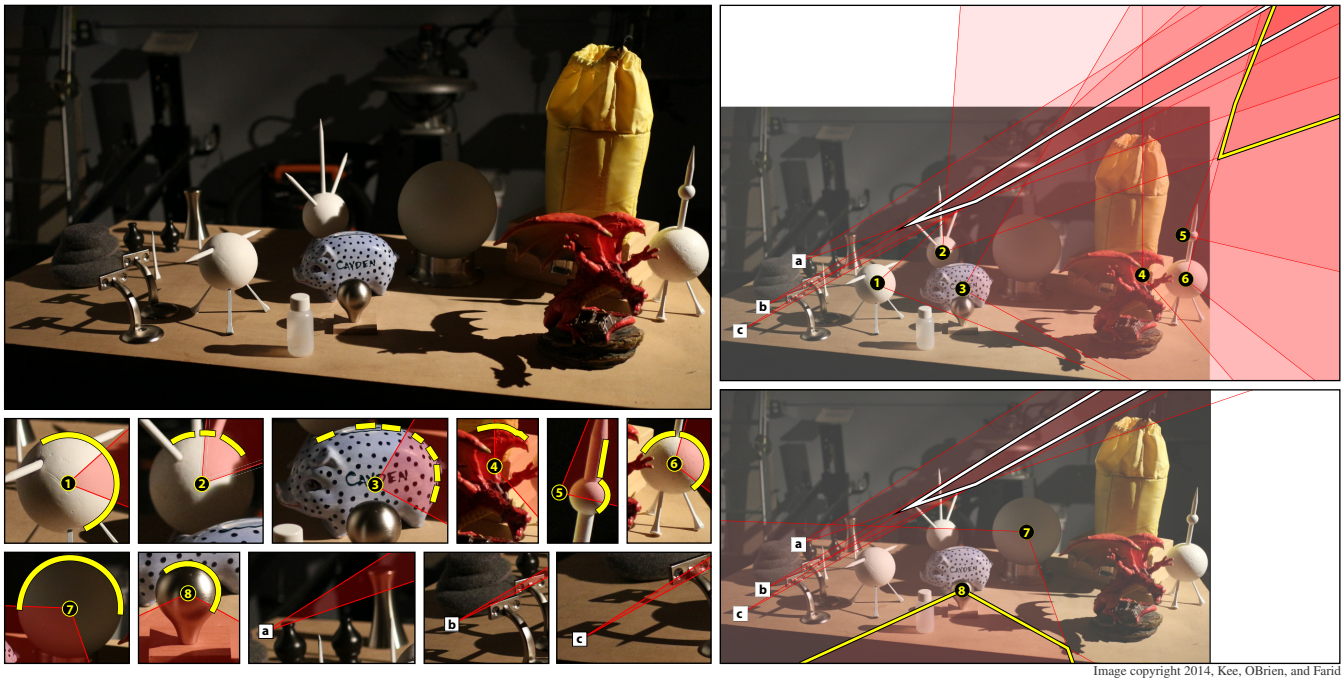


Fig. 16. Shown is an authentic image (top left) in which some objects violate the assumptions of constant Lambertian reflectance, and direct illumination by a point light. Shading constraints ❶-❻ (bottom left) are specified on objects that adhere to our assumptions, whereas ❼ and ❽ are not. Right: constraints ❶-❻ are consistent with shadows [a]-[c] (top). Constraints ❼ and ❽ are inconsistent with the shadows (bottom), but only because they violate our assumptions.

on the functions was tested by estimating the functions at varying values of δ_f and δ_c. Three precisions were selected, corresponding to fractions of the plausible range of the $\log_{10}$ focal length and of the image center. (Plausible focal lengths were in $[18,\ 200]$mm, and plausible image centers were within 12% of the image height.) The precisions were $1/6^{\text{th}}$, $1/3^{\text{rd}}$, and 1 (the entire range). The focal precisions were $\delta_f = \pm(0.087,\ 0.174,\ 0.522)$ log focal units. The image center precisions were $\delta_c = \pm(2\%,\ 4\%,\ 12\%)$.

The effect of the precision of the focal length and image center was tested by computing the mean angle in the functions, Fig. 17. When the focal length and image center are known within $1/6^{\text{th}}$ or $1/3^{\text{rd}}$ of the plausible range, this mean half-angle was the same: 33° for Φ_{L} and 96° for Φ_{H}. The mean half-angle of Φ_{L} increases significantly for a fraction of 1, suggesting that estimates of the focal length within $1/3^{\text{rd}}$ of the plausible range are sufficient, ($\delta_f = \pm 0.174$ and $\delta_c = \pm 4\%$). For example, this focal precision suggests that focal lengths within the following ranges will produce roughly the same shading constraint: $18-40$mm, $40-90$mm, and $90-200$mm.

Camera precision	$1/6^{\text{th}}$	$1/3^{\text{rd}}$	1
Mean (std-dev), Φ_{L}	33° (18°)	33° (18°)	58° (30°)
Mean (std-dev), Φ_{H}	96° (5°)	96° (5°)	96° (5°)

Fig. 17. The effect of varying precision in the focal length, f, and image center, $\boldsymbol{c}$, on the average value of the wedge functions, Φ_{L} and Φ_{H}. Precision is measured as a fraction of the plausible range of the log focal length, 18mm–200mm, and image center, $\pm 12\%$ of the image height. Estimates of f and $\boldsymbol{c}$ that are more precise than $1/3^{\text{rd}}$ do not affect average width.

C. SHADING OF OCCLUDING CONTOURS

The creation of 3 billion images of random occluding contours was simulated under varying lighting. A primary light $\mathbf{s}$, secondary light $\mathbf{s}_2$, and an ambient illumination a were randomly drawn for each of the $10,000$ distinct occluding contours, Appendix B. The radiance of the primary and secondary lights was:

$$r_1 = \nu(\mathbf{s}\cdot\mathbf{n}) \tag{38}$$

$$r_2 = \nu(\mathbf{s}_2\cdot\mathbf{n}) \ , \tag{39}$$

where $\mathbf{n}$ is the surface normal, ν is the surface albedo drawn uniformly and at random from the range $[0.1,\ 0.5]$, and where any radiances less than zero are clipped to zero. These light sources were combined as follows:

$$r_c = (1-\alpha)r_1 + \alpha r_2 \ , \tag{40}$$

where the contribution, $\alpha \in [0,\ 1]$, of the secondary light source accounted for only a fraction of the observed image intensity:

$$\frac{\alpha \max(r_2)}{(1-\alpha)\max(r_1)} = \tau \ , \tag{41}$$

with the $\max$ computed along the portion of the contour illuminated by the primary light source, and the fraction τ drawn from a normal distribution with mean 0.15 and standard deviation 0.05.

The ambient illumination, a, was added to the radiance, r_c

$$r = \frac{r_c + a}{0.5 + a} \ , \tag{42}$$

where 0.5 is the maximum possible radiance due to the primary and secondary lights and albedo, ν. This divisive normalization simulates a camera exposure that accommodates the brightest possible occluding contour. The final radiance, r, is bounded in $[0,\ 1]$.

The image of the occluding contour was computed by adding 1% Gaussian noise η, random gamma of the form $r^{1/\gamma}$ with γ drawn uniformly and at random in the range $[1.2,\ 1.8]$, and quantizing:

$$\hat{r} \ = \ Q\big((Q(r+\eta))^{1/\gamma}\big) \ , \tag{43}$$

where the function $Q(\cdot)$ quantizes the signal into 256 levels.

D. OCCLUDING SURFACE NORMALS

Occluding contours typically form at the boundary of objects in an image. The camera ray to points along these contours is tangent to the surface and orthogonal to the surface normal. As such, surface normals along occluding contours comprise only two degrees of freedom. Under orthographic projection, the z-component of these occluding surface normals is zero, and the x, y components are given by the image normal [Johnson and Farid 2005; Nillius and Eklundh 2001]. We, however, need to specify these normals under a model of perspective projection.

Under perspective projection, camera rays to points along an occluding contour span a hemisphere of directions, but are projected onto a plane. This planar projection distorts the shape of the occluding contour and obscures the 3-D orientation of the surface normals. These properties can be recovered by mapping the planar image onto the portion of the hemisphere that it subtends.

We map the planar image I to a spherical image, $\hat{I}$, of radius r:

$$\hat{I}(\theta,\phi,r) \ = \ I\big(\boldsymbol{x}(\theta,\phi,r)\big) \ , \tag{44}$$

where (θ,ϕ,r) is the spherical coordinate of a point in the spherical image. The corresponding point in the planar image is $\boldsymbol{x}$,

$$\boldsymbol{x}(\theta,\phi,r) \ = \ \frac{f}{-\mathrm{x}_z(\theta,\phi,r)} \begin{bmatrix} \mathrm{x}_x(\theta,\phi,r) \\ \mathrm{x}_y(\theta,\phi,r) \end{bmatrix} + \boldsymbol{c} \ , \tag{45}$$

under a camera[4] with focal length f and image center $\boldsymbol{c}$. The 3-D point $\mathbf{x}(\theta,\phi,r)$ — referenced above by subscripts x, y, z — is the Cartesian equivalent of the point on the sphere:

$$\mathbf{x}(\theta,\phi,r) \ = \ \begin{bmatrix} r\cos\theta\sin\phi \\ r\sin\theta\sin\phi \\ r\cos\phi \end{bmatrix} \ . \tag{46}$$

Points $\boldsymbol{x}$ are therefore

$$\boldsymbol{x}(\theta,\phi,r) \ = \ \frac{f}{-r\cos\phi} \begin{bmatrix} r\cos\theta\sin\phi \\ r\sin\theta\sin\phi \end{bmatrix} + \boldsymbol{c} \tag{47}$$

$$= \ -f \begin{bmatrix} \cos\theta\tan\phi \\ \sin\theta\tan\phi \end{bmatrix} + \boldsymbol{c} \ , \tag{48}$$

where the radius of the spherical image, r, does not affect the projection of $\mathbf{x}(\theta,\phi,r)$ to $\boldsymbol{x}(\theta,\phi,r)$.

Given the spherical image $\hat{I}$, computing the 3-D surface normal of an occluding contour under perspective projection is analogous to the orthographic case: the occluding normal is the direction in this spherical image that is orthogonal to the image of the contour. That direction can be derived from the gradient of $\hat{I}(\theta,\phi,r)$ (without sacrificing generality). The spherical gradient is

$$\nabla\hat{I}(\theta,\phi,r) \ = \ \frac{1}{r\sin\phi}\frac{\partial\hat{I}}{\partial\theta}\hat{\boldsymbol{\theta}} \ + \ \frac{1}{r}\frac{\partial\hat{I}}{\partial\phi}\hat{\boldsymbol{\phi}} \ + \ \frac{\partial\hat{I}}{\partial r}\hat{\boldsymbol{r}} \tag{49}$$

$$= \ \begin{bmatrix} \frac{\partial\hat{I}}{\partial\theta} & \frac{\partial\hat{I}}{\partial\phi} & \frac{\partial\hat{I}}{\partial r} \end{bmatrix} \begin{bmatrix} \frac{1}{r\sin\phi} & 0 & 0 \\ 0 & \frac{1}{r} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\theta}}^{\,T} \\ \hat{\boldsymbol{\phi}}^{\,T} \\ \hat{\boldsymbol{r}}^{\,T} \end{bmatrix}, \tag{50}$$

where the diagonal matrix normalizes the scale of the partial derivatives along the spherical basis directions $\hat{\boldsymbol{r}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$. The partials are given by the general chain rule:

$$\begin{bmatrix} \frac{\partial\hat{I}}{\partial\theta} & \frac{\partial\hat{I}}{\partial\phi} & \frac{\partial\hat{I}}{\partial r} \end{bmatrix} = \begin{bmatrix} \frac{\partial I}{\partial x_x} & \frac{\partial I}{\partial x_y} \end{bmatrix} \begin{bmatrix} \frac{\partial x_x}{\partial\theta} & \frac{\partial x_x}{\partial\phi} & \frac{\partial x_x}{\partial r} \\ \frac{\partial x_y}{\partial\theta} & \frac{\partial x_y}{\partial\phi} & \frac{\partial x_y}{\partial r} \end{bmatrix} \tag{51}$$

$$= \nabla I \begin{bmatrix} f\sin\theta\tan\phi & -f\cos\theta\sec^2\phi & 0 \\ -f\cos\theta\tan\phi & -f\sin\theta\sec^2\phi & 0 \end{bmatrix}. \tag{52}$$

Combining Equation (50) and Equation (52) and reducing gives

$$\nabla\hat{I}(\theta,\phi,r) = \nabla I \begin{bmatrix} \sin\theta & \cos\theta \\ -\cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & r/f \end{bmatrix} \begin{bmatrix} -\hat{\boldsymbol{\theta}}^{\,T} \\ -\hat{\boldsymbol{\phi}}^{\,T} \end{bmatrix}, \tag{53}$$

where the radial component $\hat{\boldsymbol{r}}$ does not appear because the radius of the spherical image does not affect the image. The negation of the bases accounts for the negation during projection, Equation (48).

Three transformations are therefore required to map the 2-D image gradient, ∇I, into a 3-D vector that is parallel to the 3-D occluding surface normal. First the 2-D gradient is rotated into the local 2-D coordinate system that is defined by the spherical basis vectors. The $\hat{\boldsymbol{\theta}}$ basis vector of this coordinate system is parallel to the image plane and is oriented in an axial direction around the image center. The $\hat{\boldsymbol{\phi}}$ basis vector accounts for variations in z, and its x, y components are oriented in a radial direction away from the image center. The radial component of the image gradient is next scaled by r/f, which corrects for the geometric distortion in the image of the local region. This distortion stretches the image when it is projected onto the image plane, which shortens the radial component of the 3-D gradient by f/r, the cosine of the angle between the image plane and the plane spanned by the spherical basis vectors. Lastly, this rectified 2-D image gradient is mapped into 3-D as a weighted combination of the 3-D basis vectors, $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$.

The spherical coordinates can be eliminated from Equation (53),

$$\nabla\hat{I}(\theta,\phi,r) \ = \ \nabla I \begin{bmatrix} 1 & 0 & (x_x - c_x)/f \\ 0 & 1 & (x_y - c_y)/f \end{bmatrix} \tag{54}$$

$$= \ \begin{bmatrix} \nabla I & \frac{1}{f}\nabla I\big(\boldsymbol{x}-\boldsymbol{c}\big) \end{bmatrix}, \tag{55}$$

and the resulting expression applies to any measure of the 2-D normal along an occluding contour in an image. The 3-D surface normal $\mathbf{n}$ of an occluding contour under perspective projection is

$$\mathbf{n} \ = \ \begin{bmatrix} n_x \\ n_y \\ \frac{1}{f}\boldsymbol{n}\cdot(\boldsymbol{x}-\boldsymbol{c}\,) \end{bmatrix} \ , \tag{56}$$

where $\boldsymbol{n} = \begin{pmatrix} n_x & n_y \end{pmatrix}^T$ is the 2-D normal of the occluding contour in the image, $\boldsymbol{x}$ is the position of the 2-D normal in the image, f is the focal length, and $\boldsymbol{c}$ is the image center.

Under orthographic projection, $f \to \infty$ and $\mathbf{n} \to \begin{pmatrix} n_x & n_y & 0 \end{pmatrix}^T$, as expected. Similarly, the z-component of $\mathbf{n}$ vanishes at the image center, $\boldsymbol{x} - \boldsymbol{c} = 0$. The direction of the x, y components of $\mathbf{n}$ are unchanged by projection. Note that this expression for the normal does not have unit length — even if $\boldsymbol{n}$ does — because the planar image projects onto the sphere with spatially varying density, and that modulates the gradient magnitude.

[4] This camera faces down the $-z$ axis by our convention, Equation (10).