In Search of Illumination Invariants*

Hansen F. Chen
Department of Physics
Yale University
New Haven, CT 06520-8120
hansen.chen@yale.edu

Peter N. Belhumeur†
Departments of EE and CS
Yale University
New Haven, CT 06520-8267
belhumeur@yale.edu

David W. Jacobs
NEC Research Institute
Princeton, NJ 08540
dwj@research.nj.nec.com

Abstract

We consider the problem of determining functions of an image of an object that are insensitive to illumination changes. We first show that for an object with Lambertian reflectance there are no discriminative functions that are invariant to illumination. We do this by showing that given any two images, one can construct a single Lambertian object that can produce both images under two very simple lighting conditions. This result leads us to adopt a probabilistic approach in which we analytically determine a probability distribution for the image gradient as a function of the surface's geometry and reflectance. Our distribution reveals that the direction of the image gradient is insensitive to changes in illumination direction. We verify this empirically by constructing a distribution for the image gradient from more than 20 million samples of gradients in a database of 1,280 images of 20 inanimate objects taken under varying lighting conditions. Using this distribution, we develop an illumination insensitive measure of image comparison and test it on the problem of face recognition.

*This paper is based on "Comparing Images Under Variable Illumination," by Jacobs, Belhumeur, and Basri, which appeared in the IEEE Conference on Computer Vision and Pattern Recognition, June 1998, and on "In Search of Illumination Invariants," by Chen, Belhumeur, and Jacobs, which appeared in the IEEE Conference on Computer Vision and Pattern Recognition, June 2000.

†Belhumeur was supported by Presidential Early Career Award for Scientists and Engineers IIS-9703134, ARO DAAH04-95-1-0494, NSF KDI-9980058, NIH RO1-EY 12691-01.

1 Introduction

Changes in viewpoint and illumination can dramatically alter the appearance of an object. To generalize effectively, image-based recognition systems must use methods of image comparison that work in spite of these changes. We focus here solely on changes in illumination and ask: Are there discriminative illumination invariants? If not, are there local image measurements that are at least insensitive to illumination changes?

We show that even for objects with Lambertian reflectance [28], there are no discriminative functions of images of objects that are invariant to illumination. This differs from earlier findings in that we do not assume a homogeneous BRDF [27, 5], coplanarity [36], or consider invariants based on multiple images [48]. To do this, we show that for any two images — whether or not they are of the same object — there is always a family of surfaces, albedo patterns, and light sources that could have produced them.

This result suggests that in comparing images for recognition, alignment, or tracking, one must resort to probabilistic measures of comparison. (Here we do not assume we have access to multiple training images of the same object under varying illumination. If we did, we could extract invariants [48] or construct representations for modeling the illumination variation [43, 20, 3, 18].) We develop a measure of image comparison by considering illumination to be a random variable which gives rise to the apparent randomness in local image measurements. Specifically, we derive the probability distribution of the image gradient of a point on a surface as determined by the differential geometric and reflectance properties at that point. Our distribution reveals that the direction of the image gradient (which is perpendicular to the flow field [5]) is insensitive to changes in illumination direction. Using the image gradient distribution, we then develop an illumination insensitive measure of image comparison.

Finally, we verify our qualitative theoretical arguments by empirically constructing both our distribution for the image gradient and our illumination insensitive measure of image comparison. We do this using a database of 1,280 images of 20 objects taken under varying illumination direction. We then test our illumination insensitive measure on the problem of face recognition, and compare

our performance on 450 images of 10 individuals to that of other existing methods.

2 Background

In recent years there has been much work on object recognition by image comparison. In these methods, an object is not described in terms of its 3-D properties, but rather in terms of the 2-D images that it produces. One approach to appearance-based recognition is to sample an object's possible images, and then to compare, in a lower dimensional image subspace, a novel image to the set of sampled images, using pattern recognition techniques such as nearest neighbors. Turk and Pentland [44] (inspired by the work of Kirby and Sirovich [26]) suggest such an approach, using principal component analysis to compactly represent the training images. Murase and Nayar [35] suggest accounting for lighting and viewpoint variation with such an approach by explicitly sampling images of the object under all possible viewing conditions. The advantage of these methods is that they do not have to derive 3-D structure, or to extrapolate to all possible images of an object based on a small set of samples. The disadvantage is that the set of images an object can produce is extremely large, since there are so many ways that viewing conditions can vary. Murase and Nayar, for example, are only able to deal with one rotational degree of freedom in the object, and one positional degree of freedom in a point light source. In principal their approach can handle greater variability, but the number of images an object produces grows exponentially with the number of degrees of freedom in the viewing conditions.

To build flexible recognition systems that account for all the variability of the image formation process, it seems necessary to find ways to generalize from images that are few relative to this variability. The most powerful form of generalization is based on invariants. An invariant is a property of an object that shows up in every image of that object. Geometric invariants have been used for the recognition of restricted classes of objects, (e.g., planar [29] or symmetric [16]) that have identifiable local features. But it has been shown that invariants do not exist for general classes of 3-D objects ([7, 9, 33]). Photometric invariants have also been used for planar objects ([36, 42]). Moses and Ullman have shown that photometric invariants do not exist, even for Lambertian objects, if the

"objects" are allowed to consist of mosaics of possibly disconnected planar patches. In this paper, we show the stronger result that discriminative invariants do not exist even when these surface normals are required to form surfaces. Both geometric and photometric invariants can also exist when one has access to multiple images of an object at recognition time. Such techniques are used for structure-from-motion (e.g., [13]) or photometric stereo (e.g., [48, 10, 37]).

In the absence of invariants for recognition, one can attempt to use a small number of images of an object to characterize the complete set of images that the object might produce. Ullman and Basri [45] take this approach, showing that in many cases each of the 2-D images produced by an object from different viewpoints is a linear combination of a small number of basis images. The spirit of this work is to use images of an object to predict what other images the object might produce, without ever explicitly reconstructing the object's 3-D properties. Jacobs [23] determines the most compact possible representations of these sets of images, while Faugeras and Robert [14] presents methods for extrapolating the images that an object produces under perspective projection. Shashua [43] and Moses [31] take an approach similar to that of Ullman and Basri to show that the set of intensity images produced by an object under variable lighting conditions can be predicted linearly from a small number of basis images. Belhumeur and Kriegman [3] show how to extend this work to account for multiple or diffuse light sources, and self-shadowing. In all this work, it is assumed that training images provide sufficient prior information to completely characterize the entire set of images that an object can produce. This information may be strictly weaker than having a precise 3-D description of the object, because various approaches may not attempt to recover elements of the structures that are altered by a 3-D affine transformation (e.g., [45, 43]) or a bas-relief transformation (e.g., [4]). However, it is clear that much of the structure of an object must be recovered in order to predict all images that it can produce, and that this may be difficult to do based on a small set of unconstrained training images. For example, Jacobs, Belhumeur, and Basri [24] show that under arbitrary, diffuse lighting conditions, it is not possible to exactly reconstruct the albedos of an object, even when its structure is precisely known.

Finally, one may wish to perform recognition under variable viewing conditions when there

is simply not enough prior information to fully characterize the effect of this variation. In this case, one may wish to use image comparision methods that are insensitive to changes in viewing condition, even if they are not completely invariant to changes. These are sometimes called *quasi-invariants*. A number of geometric quasi-invariants have been proposed (e.g., [7, 51]); they are sometimes relaxed versions of invariant properties. Related photometric methods have long been used; one justification for edge-based recognition methods is that for polyhedral objects edge position is insensitive to illumination and viewpoint changes. (However for smooth surfaces, edge position is sensitive to illumination changes.) More recently, Rao and Ballard [39], for example, recognize objects using descriptors built from the multi-scale output of filters. These are insensitive to modest changes in viewpoint and illumination (for insensitive comparisons of color see, e.g., [17]). Irani and Anandan [22] suggest comparing images using the squared output of directional derivative filters to gain insensitivity to changes in sensing modality.

After proving that illumination invariant properties do not exist for 3-D objects, we also develop a probabilistic semi-invariant measure, namely, the direction of the image gradient, that is insensitive to lighting variation. Our method is based on a geometric analysis of the effects of lighting variation on the image gradients produced by a 3-D object. We compare this to empirical data, and produce and test a new image comparison method.

3 Do Illumination Invariants Exist?

Given two images, are they created by two distinct objects, or the same object under different illuminations? While it seems possible that one can always, or at least sometimes, distinguish with certainty between these two scenarios, in this section we show the contrary is true.

To analyze this question, we study the existence of discriminative illumination invariants: functions of an image which are invariant to illumination on the object but vary with the object identity. Formally, let $\mathcal{O}$ be some set of rigid objects including their optical properties; $\mathcal{S}$ be certain illumination conditions; and $\mathcal{I}$ be the set of all images. Function $Q : \mathcal{O} \times \mathcal{S} \to \mathcal{I}$ gives the image $I \in \mathcal{I}$ of object $o \in \mathcal{O}$ under illumination $s \in \mathcal{S}$, i.e., $I = Q(o, s)$.

We adopt the following definitions:

Definition 1 *A function μ on $\mathcal{I}$ is invariant to illumination* $\iff$ $\mu(Q(o,s)) = \mu(Q(o,l))$, $\forall s, l \in \mathcal{S}, o \in \mathcal{O}$.

Definition 2 *An illumination invariant μ is nondiscriminative for object set $\mathcal{O}$* $\iff$ $\mu(I) = \mu(J)$, $\forall I \neq J$, *where I and J are in the range of Q. An illumination invariant is discriminative iff it is not nondiscriminative.*

This definition implies μ does not depend on o for nondiscriminative invariants.

Lemma 3.1 *There are no discriminative illumination invariants for $\mathcal{O}$ if for any two images I and J in the range of Q, there is always an object $o \in \mathcal{O}$ which, under some pair of lighting conditions in $\mathcal{S}$, could have produced both image I and J.*

Proof. The proof follows immediately from the above definitions. □

The existence of discriminative invariants depends on the specific set of illumination and objects. It is obvious that for a large enough class of illuminations, there are no such invariants. For example, a movie projector can create arbitrary images by projecting carefully designed patterns on any given object with nonzero reflectance. The larger the set of illumination conditions we admit, the smaller the set of invariants will be. Likewise, for purely specular convex surfaces, (e.g., mirrors) it is again obvious that there are no discriminative illumination invariants. However these cases are clearly extremes in illumination and reflectance. Next, we consider in detail the case of Lambertian surfaces illuminated by point light sources distant from the object. What is surprising is that even for this simple case there are no discriminative illumination invariants. We first prove the claim ignoring the effects of interreflection, then show that our result holds when interreflection is taken into account. Moreover, our simulations show that their additive effect on the image intensity is often negligible when considered in the context of this problem [8]. The fact that, statistically the typical albedo of

an object is less than 0.15 [40, 41], lends empirical support for the technique, letting the albedo of the object approach zero, employed in the proof of Theorem 3.2 below.

Theorem 3.1 *Discounting interreflection, given two arbitrary image functions I and J, and two arbitrary point light sources $\vec{s}$ and $\vec{l}$ at infinity, and if the projection on the image plane of the sum $\vec{s}+\vec{l}$ is nonzero, there exists a family of smooth Lambertian surface f such that I is the image of f under $\vec{s}$ and J is that of f under $\vec{l}$.*

Proof. This is just a recapitulation of the Lemmas 3.5 and 3.2 that are to follow. □

Corollary 3.1 *Discounting interreflection, all illumination invariants for objects with Lambertian reflectance under point light sources at infinity are nondiscriminative.*

Proof. Let μ be an illumination invariant. Given two arbitrary images I, J in the range of Q, by Theorem 3.1, there always exists a Lambertian surface of an object o and two light sources at infinity s and l, such that I and J are images of o under s and l, respectively, or $I = Q(o, s)$ and $J = Q(o, l)$. By Lemma 3.1, μ must be a nondiscriminative invariant. □

Note that for the above propositions we require objects to be composed of surfaces, not freely floating planar facets in space as in [32].

We prove the much stronger Theorem 3.1 rather than the weaker Corollary 3.1 directly, since some researchers believe humans are able to determine the direction of the light source when viewing an image [38, 6, 21, 25, 30, 46, 49, 50, 52]. We show even under this more stringent condition, there is still no way to tell with certainty if the given two images are generated from two different surfaces.

If we assume in addition that the invariant function is continuous with respect to the image functions, the above result can be extended to Lambertian surfaces with interreflection.

Theorem 3.2 *Taking interreflection into account, all* continuous *illumination invariants for objects with Lambertian reflectance under point light sources at infinity are nondiscriminative.*

Proof. See appendix. □

In this section we outline the proof of Theorem 3.1. The basic strategy of our proof is to write two equations that describe the formation of the images, with the object's shape and albedo as unknowns. After eliminating albedo, we arrive at a single first order linear partial differential equation in shape. We then show that this PDE always has a unique solution, up to a set of initial conditions that determine a family of solutions. Using this object shape, albedo is uniquely solved for linearly, up to a scale factor. Finally, we must show that the solutions we arrive at do not contain shadows, which justifies our use of a simple linear description of image formation.

We begin the outline of our proof of Theorem 3.1 with the following setup. Set the camera optical axis as the z-axis. An object o is viewed from the direction $(0, 0, 1)$. The set $\mathcal{O}$ (more accurately the visible surface of $\mathcal{O}$) is defined to be the set of pairs (f, α), where f is the graph $(x, y, z = f(x, y))$ of a smooth function $z = f(x, y)$, and $\alpha(x, y)$ is a nonnegative function called albedo, both defined on the prescribed unit square. Here, the set of point light sources at infinity $\mathcal{S}$ is represented by the set of 3-D vectors $\vec{s} = (s_x, s_y, s_z)$ in the opposite direction of the light rays, with $||\vec{s}||$ equal to the powers of the sources. There are two regions on the surface $(x, y, f(x, y))$: one that is illuminated by the light source, and the other called shadow region (SR hereafter) that is either facing away and creates so called attached shadows, or obscured from the light source by other parts of the surface. Suppose an image I is generated by f under point light sources $\vec{s} = (s_x, s_y, s_z)$. Excluding interreflection, the image intensity I is described by

$$I(x, y) = \begin{cases} \alpha(x, y)\ \vec{s} \cdot \hat{n}(x, y) & \text{if } (x, y, f(x, y)) \notin \text{SR} \\ 0 & \text{otherwise} \end{cases}, \tag{1}$$

where $\alpha(x, y)$ is the albedo of object o, $\vec{n} = (-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1)$ is a normal vector to the surface f of o, and $\hat{n}$ is the unit vector for $\vec{n}$. Henceforth, we will adopt the following convention: if $\vec{r}$ is a vector, then $\hat{r} = \vec{r}/\|\vec{r}\|$.

There is, in principle, no more restriction other than nonnegativity for a function to describe the image of some object. By painting the albedo proportional to the given image function on a

flat surface, and placing the surface under any point light source with appropriate power, one can always obtain any prescribed image. However, for convenience of later discussion, we define

Definition 3 *I is called an image function, or $I \in \mathcal{I}$, if I is C^1 and positive on a compact set. C^1 on the boundary means there is a tangent plane approached by interior and boundary points.*

The restriction is not severe in that any real valued piecewise continuous function on a compact set can be approached pointwise by a sequence of continuous functions which can in turn be approached uniformly by a sequence of analytic functions (Stone-Weierstrass approximation theorem). We may directly introduce discontinuities. However, its existence and the specific geometry of the set on which I is discontinuous will greatly complicate the ensuing technical investigation without altering the essential result[1].

Our goal is to show that for any two $I, J \in \mathcal{I}$, there always exists a smooth solution on the unit square to the partial differential equations (PDE's) given in Eq.'s 2 and 3, which in turn leads to Theorem 3.1.

We construct a surface according to the local differential description, or the upper part of Eq. 1, as if there is no effect of cast shadows.

$$I(x,y) = \alpha(x,y)\, \vec{s} \cdot \hat{n}(x,y), \tag{2}$$

$$J(x,y) = \alpha(x,y)\, \vec{l} \cdot \hat{n}(x,y) \tag{3}$$

We then show that if the boundary conditions are chosen appropriately, the surface thus constructed actually casts no shadows, and therefore renders the lower condition of Eq. 1 moot:

Lemma 3.2 *Let Ω be a compact subset of $\mathbb{R}^2$. Suppose I, $J \in \mathcal{I}(\Omega)$ (image functions on Ω). There is a family of C^1 surfaces f that satisfies Eq. 2 and 3 has no attached or cast shadows.*

To prove the above lemma, we make the following construction which leads to three additional

[1] If there are discontinuities, if the set of discontinuities obey certain smoothness conditions, the set of $(\vec{l}, \vec{s})$ that admits continuous surfaces may not be the whole 6-D space, but will be of nonzero measure

lemmas. (The trusting reader may want to skip ahead to Section 4.) All the proofs are relegated to the Appendix.

We cross multiply the two sides of Eq. 2 with those of Eq. 3. We then divide the resulting equation by α (α cannot be zero as neither I nor J assumes zero value anywhere) to obtain the following first order linear partial differential equation:

$$(I\vec{l} - J\vec{s}) \cdot \vec{n} = 0. \tag{4}$$

(Wolff and Angelopoulou [47] and Fan and Wolff [12] also derive this equation, and then use it for stereo matching and photometric stereo). Once this equation is solved, we can substitute the result into our original equation to obtain a linear equation in albedo. When the magnitude of the light is unknown, this equation determines albedo uniquely, up to a scale factor. To obtain valid albedos we must scale the magnitude of the light so that all albedos are less than one.

To solve Equation 4 we use the method of *characteristic curves*. In this method, one performs a change of variables to obtain a PDE in one variable. Solving this PDE tells us the surface height along a curve on the object. The complete equation will have a unique solution, up to an initial condition, if every point is on a single characteristic curves, so that solving these curves separately provides exactly one height for every point on the surface.

As a simple example of this method, consider the case of $l = (0, 0, 1)$ and $s = (1, 0, 0)$. Then we have

$$I\frac{\partial f}{\partial x} - J = 0.$$

Our equation reduces to one with a partial derivative. The characteristic curves in this case are horizontal lines across the image. The value of $f(x, y)$ is given by

$$f(x, y) = f(0, y) + \int_0^x \frac{I(w, y)}{J(w, y)} dw$$

where we have no obvious source of knowledge available to provide the initial condition $f(0, y)$.

The above construction provides one pair of point light sources that when provided with an appropriate boundary condition gives the required smooth surface. However, here we intend to

show the existence of a surface even when the point light sources are given for each of the two images.

Generally, the characteristic curves $\vec{r}(t)$ of Eq. 4 satisfy

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = I\vec{l} - J\vec{s}. \tag{5}$$

Obviously, we only need to solve for the characteristic curves in the XY plane. The z-component can be solved by simply integrating the z-component of the right hand side above along the XY trajectory. The surface is constructed by weaving the characteristic curves together by choosing continuous boundary values on part of the boundary of the unit square.

We claim that for a vector field uniformly bounded from below, as is the case here, there is a characteristic curve in the XY-plane through each point in D_2 with both ends lying on the boundary ∂D_2.

Lemma 3.3 *Let Ω be a compact subset of $\mathbb{R}^2$. Suppose I, $J \in \mathcal{I}(\Omega)$ (image functions on Ω), and the projection on the image plane of the sum $\vec{s}+\vec{l}$ is nonzero. Let $\vec{\rho}$ denote points in the XY-plane. Through an arbitrary point q in Ω, there exists a unique $\zeta_q \in R$ and a unique characteristic curve projection on the XY-plane $\vec{\rho}(t)$, $t \in [0, \zeta_q]$ in Ω homeomorphic to $[0, \zeta_q]$, such that $\vec{\rho}(0), \vec{\rho}(\zeta_q) \in \partial\Omega$.*

Lemma 3.3 shows the global existence of characteristic curves through any point in the unit square.

Lemma 3.4 *Let the hypothesis of Lemma 3.3 be satisfied. The XY-plane characteristic curve is C^1 with respect to t and its initial point.*

The uniqueness of the characteristic curves provided by Lemma 3.4 removes the ambiguity, enabling the later construction of a smooth surface instead of a multivalued function. It also guarantees the needed smoothness for the characteristic curves, although not yet for the surface.

The smooth surface can then be fabricated by carefully assigning smooth boundary conditions.

Lemma 3.5 *Let the hypothesis of Lemma 3.3 be satisfied. There is a C^1 surface f on Ω satisfying Eq. 4.*

The above construction from the local description would not satisfy Eq. 1 if the surfaces cast shadows onto themselves. Our task would be made much simpler if there is a surface without attached or cast shadows. Fortunately, Lemma 3.2 assures that there indeed exist such surfaces provided that the boundary condition is judiciously chosen.

To summarize, we have shown that given any two images, and any two point light sources, we can construct a Lambertian surface that will generate each image, in the absence of interreflections. In the appendix, we show that the effects of interreflections can be made arbitrarily small. This, in turn, shows that no discriminative illumination invariants exist.

4 A Probabilistic Model

In the previous section, we concluded that for Lambertian and purely specular surfaces there are no illumination invariants. Thus, we must settle for something less: a weaker, probabilistic form of invariance. To this end, we will show that even if the direction of the light source is uniformly distributed, the direction of the image gradient is not. We first analytically construct the probability distribution for the image gradient as a function of the surface curvature and reflectance. We will verify this empirically by constructing a distribution for the image gradient from a database of real images.

Suppose $(x, y, f(x, y))$ is a smooth, i.e. C^k or analytic, surface. Let (u, v) be coordinates on the surface such that $(x(u,\cdot), y(u,\cdot), f[x(u,\cdot), y(u,\cdot)])$ as a function of u only and $(x(\cdot,v), y(\cdot,v), f[x(\cdot,v), y(\cdot,v)])$ as a function of v only are lines of curvature, with u, v being the length of the respective lines. For subsequent development, let κ_u and κ_v be the principal curvatures of the surface in principal directions $\hat{u}$ and $\hat{v}$, respectively. We set up a local Cartesian coordinate system by adopting the tangents of the lines of curvature along with the surface normal as the third axis and call this the u-v-n coordinate system.

The BRDF (bidirectional reflectance distribution function) α is a function of (u, v), the direction of the incident light $\hat{i}$, and the direction of outgoing light $\hat{o}$ in this local coordinate system. The radiance from a point on a smooth surface in the camera direction $\hat{c}$ (in the u-v-n coordinate system)

is then

$$\begin{aligned} L(u,v,\hat{c}) &= \int_{\Omega} \alpha[u,v,\hat{i}(u,v,\hat{s}),\hat{o}(u,v,\hat{c})]\,\hat{n}(u,v)\cdot\vec{s}\,\mathrm{d}\hat{s} \\ &= \hat{n}\cdot\int_{\Omega}\alpha\vec{s}\,\mathrm{d}\hat{s} \end{aligned} \tag{6}$$

where Ω denotes the solid angle of light seen at the point of concern.

We analyze the influence of the differential geometric and reflectance properties by examining the scene radiance under a single light source at infinity. The scene radiance in Eq. 6 becomes

$$L(u,v,\hat{c}) = \alpha[u,v,\hat{i}(u,v,\hat{s}),\hat{o}(u,v,\hat{c})]\;\vec{s}\cdot\hat{n}(u,v). \tag{7}$$

Assume the light sources are spherically symmetrically distributed. Suppose we have a patch of a Lambertian surface with constant albedo and principal curvatures $|k_u| \neq |k_v|$. Then the gradient of the scene radiance most likely lies in the direction in which the magnitude of the curvature is maximal. Note that in the special case of cylindrical objects, e.g., coffee mugs, animal limbs, telephone poles, etc., the image intensity gradient lies orthogonal to the cylinder axis and tangent to the surface, and the isophotes run parallel to the cylinder axis. As a second example, let us consider a planar patch of surface with nonhomogeneous BRDF. For any BRDF, the direction of the gradient of the scene radiance always points in the direction of the spatial gradient of the BRDF. These observations suggest that if the light source directions are distributed uniformly, the distribution of the gradient of scene radiance will be biased by the underlying geometry and reflectance.

We shall expound these observations in two steps. First, we derive the relation between the gradient of the scene radiance $\vec{\nabla}L$ and the surface's local geometry and reflectance. Second, we impose a spherically symmetric probability distribution on the light source $\vec{s}$ and determine the resulting distribution on the gradient of the scene radiance $\vec{\nabla}L$. (Note that for simplicity we are doing our analysis in the tangent plane of the surface, whereas the image records the projection of scene radiance from the surface down to the x-y plane. Thus, our analysis ignores the effects of projections.)

In the following derivation, the operator $\vec{\nabla}$ is understood to be the gradient taken in the tangent plane, or u-v plane, at a point on the surface. A short calculation shows that the gradient of the scene radiance

$$\vec{\nabla} L = \alpha(\vec{s} \cdot \vec{\nabla})\hat{n} + (\vec{\nabla}\alpha)\vec{s} \cdot \hat{n}. \tag{8}$$

Eq. 8 highlights two factors that determine the gradient of the scene radiance. The first term, called the geometric gradient, is the contribution from geometric changes; the second term, called the reflectance gradient, is the contribution from changes in the BRDF.

Consider the geometric gradient term:

$$(\vec{s} \cdot \vec{\nabla})\hat{n} = \hat{u}\kappa_u s_u + \hat{v}\kappa_v s_v \tag{9}$$

where κ_u and κ_v are the two principal curvatures, and s_u and s_v are the u and v components of the light source in the u-v-n coordinate system.

Consider next the reflectance gradient term:

$$\vec{\nabla}\alpha = \hat{u}\frac{\partial \alpha}{\partial u} + \hat{v}\frac{\partial \alpha}{\partial v} \tag{10}$$

where, after a little calculation,

$$\begin{aligned}
\frac{\partial \alpha}{\partial u} &= \vec{\nabla}_{\hat{c}}\alpha \cdot \frac{\partial \hat{c}}{\partial u} + \vec{\nabla}_{\hat{s}}\alpha \cdot \frac{\partial \hat{s}}{\partial u} + \hat{u} \cdot \vec{\nabla}\alpha \\
&= \kappa_u \hat{v} \cdot \left(\vec{\nabla}_{\hat{c}}\alpha \times \hat{c} + \vec{\nabla}_{\hat{s}}\alpha \times \hat{s}\right) + \hat{u} \cdot \vec{\nabla}\alpha, \\
\frac{\partial \alpha}{\partial v} &= \kappa_v \hat{u} \cdot \left(\vec{\nabla}_{\hat{c}}\alpha \times \hat{c} + \vec{\nabla}_{\hat{s}}\alpha \times \hat{s}\right) + \hat{v} \cdot \vec{\nabla}\alpha
\end{aligned}$$

where $\vec{\nabla}_{\hat{c}}$ is the gradient taken with respect to $\hat{c}$ and $\vec{\nabla}_{\hat{s}}$ is the gradient taken with respect to $\hat{s}$.

We combine the geometric and reflectance gradients to get an overall expression for $\vec{\nabla} L$:

$$\vec{\nabla} L = \underbrace{(\hat{u}\ \kappa_u s_u + \hat{v}\ \kappa_v s_v)}_{\text{geometric gradient}} + \underbrace{\left[(\kappa_u\ \hat{u}\ \hat{v} + \kappa_v\ \hat{v}\ \hat{u}) \cdot (\vec{\nabla}_{\hat{c}}\alpha \times \hat{c} + \vec{\nabla}_{\hat{s}}\alpha \times \hat{s}) + \vec{\nabla}\alpha\right]\ \vec{s} \cdot \hat{n}}_{\text{reflectance gradient}}. \tag{11}$$

(Note that $\hat{u}\hat{v}$ and $\hat{v}\hat{u}$ are tensor products. The associative rule applies here.) The first term in the square bracket, although classified as part of the reflectance gradient, is induced by the rotation of

the u-v-n coordinate system along the lines of curvature. If the BRDF is Lambertian, the above expression becomes

$$\vec{\nabla} L = \underbrace{(\hat{u}\ \kappa_u s_u + \hat{v}\ \kappa_v s_v)}_{\text{geometric}} + \underbrace{(\vec{\nabla}\alpha)\vec{s}\cdot\hat{n}}_{\text{reflectance}}\ . \tag{12}$$

Now that we have the expression for the gradient of the scene radiance in terms of the differential geometric and reflectance properties of the patch of surface, we can determine the distribution for scene radiance by imposing a distribution on light sources. We consider light sources above the tangent plane satisfying $s_n > 0$. We felt it important for this discussion to choose a distribution that does not favor any particular direction, or a distribution that depends on the magnitude $|\vec{s}|$ alone, even though it may prove useful to assume otherwise under some particular circumstances. In addition, we assume the components of the source s_u, s_v, and s_n are distributed independently. With these assumptions, it can be shown that the probability density for $\vec{s}$ is given by

$$\rho_s(\vec{s}) = \frac{1}{(\sqrt{2\pi}\sigma)^3} e^{-\frac{1}{2\sigma^2}(s_u^2+s_v^2+s_n^2)}, \quad s_n \in [0,\infty). \tag{13}$$

Note again this distribution is uniform over the direction of the source. We subsequently deduce the corresponding probability density function of $\vec{\nabla} L$ as follows:

General Case:

The geometry and reflectance factors are mixed. The distribution in u-v coordinate is

$$\begin{aligned} & \rho_{u,v}(u,v) \\ = & \frac{1}{\kappa_u\kappa_v}\int_0^\infty \rho_s\Big(\Big(\frac{u-as_n}{\kappa_u}\Big)^2 + \Big(\frac{v-bs_n}{\kappa_v}\Big)^2 + s_n^2\Big)\,\mathrm{d}s_n \\ = & \frac{1}{\pi^{\frac{3}{2}}\sigma^2\kappa_u\kappa_v c}\, e^{-\frac{1}{2\sigma^2}\left(\left(1-\frac{a^2}{c^2}\right)\xi^2 - 2\frac{ab}{c^2}\xi\zeta + \left(1-\frac{b^2}{c^2}\right)\zeta^2\right)} \int_{-\frac{a\xi+b\zeta}{\sqrt{2}\sigma c}}^\infty e^{-\tau^2}\mathrm{d}\tau. \end{aligned} \tag{14}$$

where $a = \frac{\partial\alpha}{\partial u}$, $b = \frac{\partial\alpha}{\partial v}$, $c = \sqrt{\Big(\frac{a}{\kappa_u}\Big)^2 + \Big(\frac{b}{\kappa_v}\Big)^2 + 1}$, $\xi = \frac{u}{\kappa_u}$, and $\zeta = \frac{v}{\kappa_v}$.

We pay particular attention to two special cases which we feel are the determining factors for the distribution of the image gradient.

Special Case I: Non-zero Curvatures with Constant Albedo

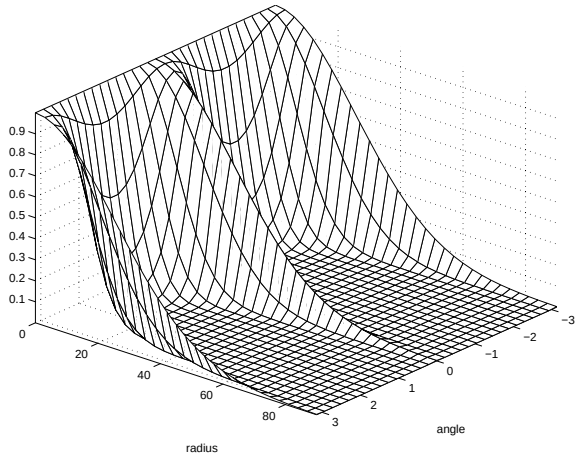


Figure 1: Distribution $\rho_{u,v}$ for constant albedo as described in Eq. 15 in polar coordinates (r, φ).

If the surface patch has spatially homogeneous reflectance (constant albedo), then $\frac{\partial \alpha}{\partial u} = \frac{\partial \alpha}{\partial v} = 0$. In coordinate system u-v-n, the probability density function for $\vec{\nabla} L$ is then

$$\rho_{u,v}(u,v) = \frac{1}{\pi^{\frac{3}{2}} \sigma^2 \kappa_u \kappa_v} \, e^{-\frac{1}{2\sigma^2}\left(\left(\frac{u}{\kappa_u}\right)^2 + \left(\frac{v}{\kappa_v}\right)^2\right)} \, . \tag{15}$$

Note that the level curves of this function are concentric ellipses and that there is a ridge along the larger of the curvature directions. In polar coordinates (r, φ), where $r = \sqrt{u^2 + v^2}$ and $\tan \varphi = \frac{v}{u}$, there are two equal ridges, at $\varphi = 0$ and $\varphi = \pi$, along the direction of r-axis whenever $\frac{|\kappa_u|}{|\kappa_v|} \neq 1$ as shown in Figure 1. The ridges grow sharper as the ratio deviates farther away from 1. The existence of these ridges implies that the gradient is more likely to point in the larger principal curvature direction. As stated previously, if one of the principal curvatures, say κ_v, is 0, such as in the case of a cylindrical surface, the image gradient always points along the direction of the nonzero curvature κ_u.

Special Case II: Zero Curvatures with Nonconstant Abedo

Let $\kappa_u = \kappa_v = 0$. The two curvatures are the same, so the orientation of the u-v coordinate system is ambiguous. We may simply choose the direction of the albedo gradient as the positive u axis. The probability density function for the gradient is then simply

$$\rho_{u,v}(u,v) = \frac{\sqrt{2}}{\pi\sigma} \, e^{-\frac{1}{\sigma^2}\left(\frac{u}{\alpha_u}\right)^2} \delta(v) \Theta(u), \tag{16}$$

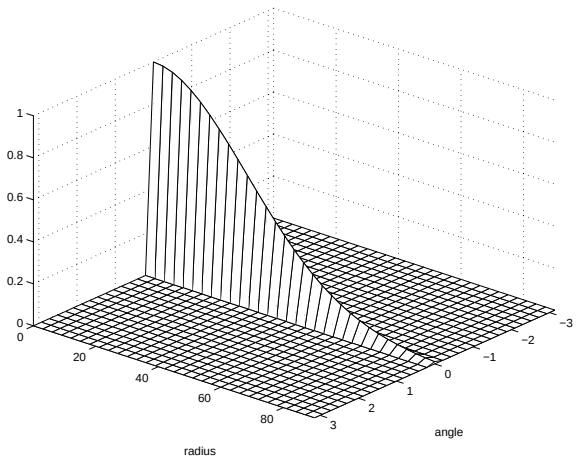


Figure 2: Distribution $\rho_{u,v}$ for a planar surface as described in Eq. 16 in polar coordinates (r, φ).

where $\alpha_u = \frac{\partial \alpha}{\partial u}$, δ is the Dirac delta function, and Θ is the Heaviside theta function (or step function). The distribution in the u-v coordinate system is one thin slice along the positive u axis. In polar coordinates (r, φ), there is only one slice at $\varphi = 0$ along r as depicted in Figure 2. Therefore, the image gradient is always directed along the albedo gradient, regardless of the direction of the light source.

We surmise that the above two special cases dominate the natural occurrence of images. In most natural images, the gradient in intensity will be largely due to albedo change (case II), and the geometric influence comes into play only when the material is almost optically homogeneous (case I). We conjecture that the linear combination of these two cases is enough to describe the distribution of the image gradient.

We want to develop an illumination insensitive measure for image comparison encompassing the two determining factors described above. If we are given the directions and magnitudes of principal curvature along with the variance of the BRDF, then we can use the density function $\rho_{u,v}$ to determine, in a probabilistic sense, how faithful any image is to the given values. The problem we are interested in is slightly different: we want to compare two images and determine the likelihood that they have been produced by the same object. For this problem, the magnitudes and directions of surface curvature and the BRDF are unknown. We must then look at the joint density for

gradients of scene radiance as given in two images, (as they are the only observables amongst these previously introduced quantities) and integrate out the unknown unobservable quantities, i.e., the magnitude and direction of the principal curvatures and the reflectance parameters.

Furthermore, we do not know the variation in the strength of the source (the standard deviation σ in Eq. 13). This variation in the strength of the light source is, of course, embedded in — but not determining — the variation in the magnitude of the gradient. Because of this, we elect to integrate out the magnitude of the image gradient as well, distilling the discriminative power of the distribution. On the other hand, the magnitude of the gradient is not completely proportional to the power of the light source. As the image intensity decreases as in the shadow region, the variation in angle φ of $\rho_{u,v}(r, \varphi)$ decreases, as can be deduced from Eq. 14, and is clearly depicted in Figure 1. Moreover, the signal to noise ratio decreases as well. There is thus distinction between $\rho_{u,v}$ at high image intensity and low image intensity. We loose some information in collapsing the dimension of the probability density $\rho_{u,v}$. We believe nevertheless that under usual circumstances the information thus lost is minute and we gain much in the resulting simplicity and computational speed. The experimental result in Section 5 confirms our intuition.

The gradients are observed in a fixed x-y coordinate system. Given the angle γ between $\hat{x}$ and $\hat{u}$, the principal curvatures denoted by κ's, and the albedo gradient $\vec{\nabla}\alpha$, the probability density of observing a gradient with magnitude r and angle φ from $\hat{x}$ is $\rho_r(r, \varphi|\gamma, \kappa, \vec{\nabla}\alpha) = r\ \rho_{(u,v)}(r\cos(\varphi - \gamma), r\sin(\varphi - \gamma))$. (The scaling of ρ by r comes from the Jacobian in converting from Cartesian to polar coordinates.) Noting that the angular dependence is only on the difference of the angles $\varphi - \gamma$, we rewrite the density as $\rho_r(r, \varphi - \gamma|\kappa, \vec{\nabla}\alpha)$. Now, the *joint* probability density of observing two scene radiance gradients (r_1, φ_1) and (r_2, φ_2) under two independent and identically distributed light sources is

$$\rho(r_1, \varphi_1, r_2, \varphi_2) = \int \rho_r(r_1, \varphi_1 - \gamma|\kappa, \vec{\beta})\rho_r(r_2, \varphi_2 - \gamma|\kappa, \vec{\beta})\ \mathrm{d}P(\gamma, \kappa, \vec{\beta}), \tag{17}$$

where $\vec{\beta} = \vec{\nabla}_{(u,v)}\alpha$, $P(\gamma, \kappa, \vec{\beta})$ is the probability measure on the unobservable random variables, and the integration is over the whole sample space. Furthermore if azimuthal symmetry holds for P,

then the density ρ above can be rewritten as a function of three variables

$$\rho(r_1, \varphi = \varphi_1 - \varphi_2, r_2) = \int \int_{\gamma=-\pi}^{\pi} \rho_r(r_1, (\varphi_1 - \varphi_2) - \gamma|\kappa, \vec{\beta}) \rho_r(r_2, -\gamma|\kappa, \vec{\beta}) \, \mathrm{d}\gamma \mathrm{d}P(\kappa, \vec{\beta}). \tag{18}$$

Azimuthal symmetry is intrinsic for any set of reasonably random image samples: the unrestrained relative rotation of the objects and the camera along the optical axis of the lens will almost surely render the azimuthal angle indistinguishable. Eq. 18 implies that the joint density depends on the magnitude of the angle between two image gradients or the absolute value of $\varphi = \varphi_2 - \varphi_1$. It is thus an even function with respect to φ.

In keeping with our earlier assumption that the density function $\rho_{u,v}$ is the linear combination of the special cases I and II, we evaluate the joint density $\rho(r_1, \varphi, r_2|\kappa, \vec{\beta})$ as such a linear combination. Let ρ_1 and ρ_2 denote the distribution in polar coordinate in the special case I and II, respectively. The joint density for special case I is then

$$\begin{aligned} & p_1(r_1, \varphi, r_2|\kappa) \\ = & \int_{-\pi}^{\pi} \rho_1(r_1, \gamma|\kappa) \rho_1(r_2, \gamma + \varphi|\kappa) \, \mathrm{d}\gamma \\ = & \frac{2}{(\pi\sigma^2\kappa_u\kappa_v)^2} \, e^{-\frac{1}{4\sigma^2}(r_1^2 + r_2^2)\left(\frac{1}{\kappa_1^2} + \frac{1}{\kappa_2^2}\right)} \, I_0\left(\frac{1}{(2\sigma)^2} \left|\frac{1}{\kappa_1^2} - \frac{1}{\kappa_2^2}\right| \sqrt{r_1^4 + r_2^4 + 2r_1^2 r_2^2 \cos(2\varphi)}\right), \end{aligned} \tag{19}$$

where I_0 is the 0'th hyperbolic Bessel function. The joint density for special case II is

$$\begin{aligned} p_2(r_1, \varphi, r_2) &= \int_{-\pi}^{\pi} \rho_2(r_1, \gamma|\vec{\beta}) \rho_1(r_2, \gamma + \varphi|\vec{\beta}) \, \mathrm{d}\gamma \\ &= \frac{2}{\pi^2\sigma^2} e^{-\frac{r_1^2 + r_2^2}{(\sigma\alpha_u^2)}} \delta(\varphi). \end{aligned} \tag{20}$$

The joint density is then the linear combination of Eq.s 19 and 20,

$$\rho(r_1, \varphi, r_2|\kappa, \vec{\beta}) = w_1 p_1 + w_2 p_2, \tag{21}$$

where $w_1 \geq 0$ and $w_2 \geq 0$ are the weights for the two cases, and $w_1 + w_2 = 1$.

Figure 3(a) shows a graph of the joint probability density function $\rho(r_1, \varphi, r_2|\kappa, \vec{\beta})$. For this graph we chose $\sigma = 1.7$, $\kappa_u = 4$, $\kappa_v = 1$, $a_u = 6$, $w_1 = 0.22$, and $w_2 = 0.78$. We chose these

values simply to illustrate the nature of the shape of the density. Different values will of course yield different densities, but their underlying shapes remain qualitatively the same.

Since the magnitude r of an image gradient is directly proportional to the power of the light source, and we consider only the local effect, the magnitude is greatly influenced by the light source power and thus not as sensitive to the underlying geometry and photometry as the angle φ. Therefore, if we desire an even simpler illumination insensitive measure, we can integrate ρ over r_1 and r_2 and obtain

$$\rho_\varphi(\varphi) = \int\int \rho(r_1, \varphi, r_2)\, \mathrm{d}r_1 \mathrm{d}r_2. \tag{22}$$

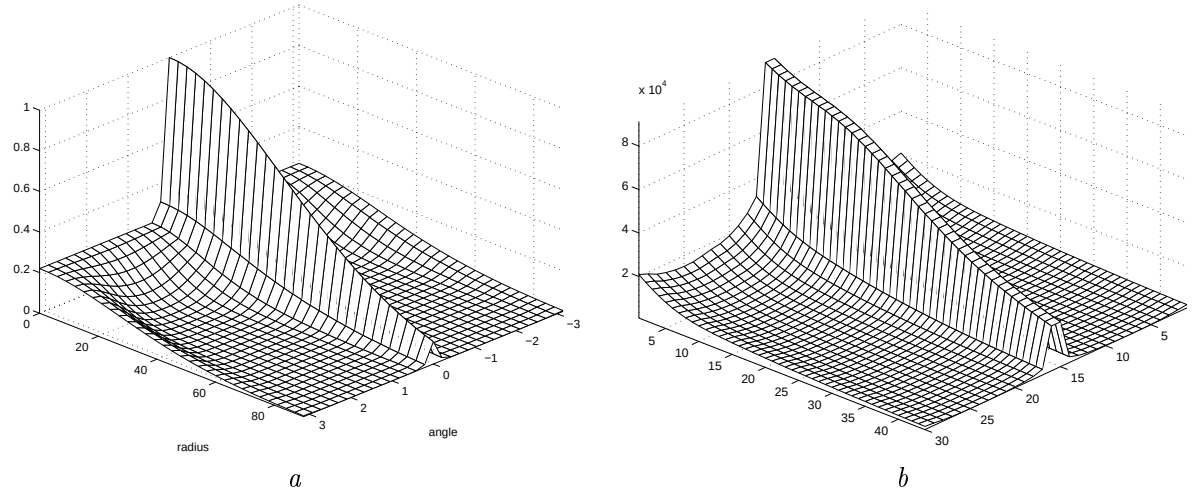
Clearly, ρ_φ inherits from $\rho(r_1, \varphi, r_2)$ the property of having a unique maximum at $\varphi = 0$. On the other hand, the original measure $\rho(r_1, \varphi, r_2)$ may have more discriminatory power, especially for the dark regions of the images, or the small magnitudes r_1 and r_2.

The joint distributions $\rho_\varphi(\varphi)$ and $\rho(r_1, \varphi, r_2)$ are the illumination insensitive measures we seek. Under each of these distributions, we would expect high probability to be assigned to two images of the same object (differing only in illumination, not viewpoint) and low probability assigned to two images of different objects. Yet, we are hampered by our ignorance of the probability distribution for the unobservables (curvatures and albedo gradients) and thus, cannot perform the integration in Eq. 18. We note that all of the above analysis is done in the tangent plane of the surface, and consequently ignores the effect of the projection of the radiance from the surface onto the image plane. However, it will become apparent in the next section that this effect is small.

To circumvent this difficulty, and to offer a practical solution, in the next section we will construct empirically the distribution in Eq. 18 from real images of objects under varying illumination. We will also compare the empirical data against the theoretical construct.

5 Empirical Joint Density

Using a geodesic dome (see Figure 4) with 64 photographic flashes, we gathered a database of 1,280 images of 20 objects (see Figure 5). The 64 flashes are positioned on the dome to cover slightly more than a hemisphere of directions. The objects included folded cloth, a computer keyboard, cups, an



a *b*

Figure 3: (a) Joint probability densities calculated from the theoretical model of the two image gradients $\rho(r_1, \varphi, r_2 = 50)$, expressed as a function of the magnitude of one gradient and the angle between the two, with the other's magnitude set to 50. (b) Empirical data of the same function as in (a).

umbrella, plants, a styrofoam mannequin, among others[2]. A stationary camera, positioned at the center of the hemisphere covered by the flashed light, captured 64 images, each illuminated by one of the 64 lights flashing in quick succession, for each stationary object. The 64 images of a telephone, one of the objects, is shown in Figure 6. We estimate the density $\rho(r_1, \phi, r_2)$ in Eq. 18 directly from a histogram of the image gradients. Eq. 18 is used with the except that the integral becomes discrete summation and the summation is taken over the pixel coordinates and objects instead of the curvatures and albedo gradient. Since the image function is scaled by the power of the light source, we only need to collect the images under light sources with the same power, which is the case for the flashs on the dome. The image gradient distribution under light sources with spherically symmetric distribution is then attained by integrating the distribution $\rho(r_1, \phi, r_2 | s_1 = 1, s_2 = 1)$ under unit light source power with respect to the power distribution

$$\rho(r_1, \varphi, r_2) = \frac{1}{2\pi\sigma^6} \int_0^\infty \int_0^\infty \rho(\frac{r_1}{s_1}, \varphi, \frac{r_2}{s_2} | s_1 = 1, s_2 = 1) e^{\frac{1}{2\sigma^2}(s_1^2 + s_2^2)} s_1^2 s_2^2 \, \mathrm{d}s_1 \mathrm{d}s_2, \tag{23}$$

As we have discussed in Section 4, if convenience or simplicity is desired, $\rho_\varphi(\varphi)$ is another

[2]All databases used in this paper are available for download from subdirectories "hrlfaces", "yaleAselected," and "objects" at ftp://Plucky.cs.yale.edu/FTPRoot.

Figure 4: Geodesic dome with 64 flashes used to capture data base images.

excellent illumination insensitive measure. The distribution $\rho_\varphi(\varphi)$ in angle alone can be gathered directly from the image database under a uniform single power lighting distribution. We see this by substituting the above expression for $\rho(r_1, \varphi, r_2)$, the angular distribution

$$\begin{aligned}
\rho_\varphi(\varphi) &= \int\int \rho(r_1, \varphi, r_2)\, \mathrm{d}r_1 \mathrm{d}r_2 \\
&= \int\int \rho(\eta_1, \varphi, \eta_2 | s_1 = 1, s_2 = 1) E[s]^2 \mathrm{d}\eta_1 \mathrm{d}\eta_2 \\
&= \frac{1}{2\pi\sigma^6} \int\int e^{\frac{1}{2\sigma^2}(s_1^2 + s_2^2)} s_1^3 s_2^3 \, \mathrm{d}s_1 \mathrm{d}s_2 \int\int \rho(\eta_1, \varphi, \eta_2 | s_1 = 1, s_2 = 1) \mathrm{d}\eta_1 \mathrm{d}\eta_2 \\
&= \frac{2\sigma^2}{\pi} \int\int \rho(\eta_1, \varphi, \eta_2 | s_1 = 1, s_2 = 1)\, \mathrm{d}\eta_1 \mathrm{d}\eta_2. \qquad (24)
\end{aligned}$$

where all the integrations are over the positive real axis, and E is the expectation value with respect to the radial distribution of the light power λ. We perform a change of variables from (r, λ) to (η, λ) to obtain the second equation. As shown in the second line, Eq. 24 is valid for any radial power distribution.

A slice of the joint probability density $\rho(r_1, \varphi, r_2)$ for a fixed r_2 is shown in Figure 3(b). As

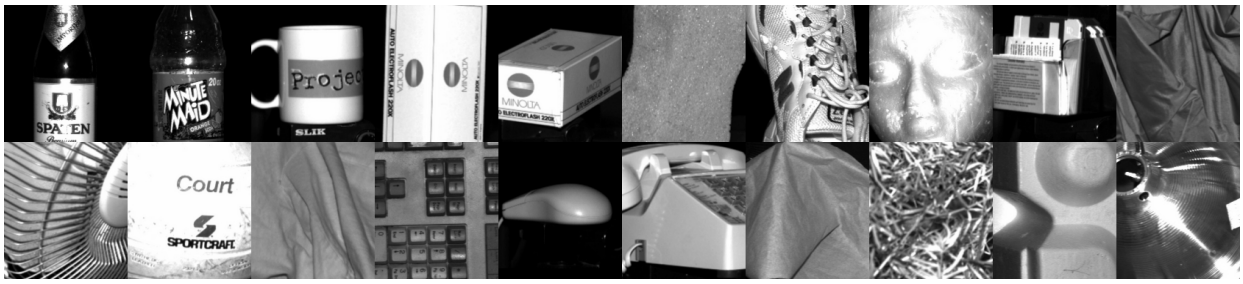


Figure 5: 20 objects in the database used to collect the empirical distribution of image gradient.

Figure 6: A telephone, one of the 20 objects, under the illumination of 64 flashes.

expected in Section 4, the distribution exhibits most of the features in the theoretical model depicted in Figure 3(a). It is symmetric with respect to the plane $\varphi = 0$. There is a prominent ridge at $\varphi = 0$ and it is the unique global maximum of $\rho(r_1, \phi, r_2)$ on the line of arbitrarily fixed r_1, r_2.

As anticipated in Section 4, the linear combination of the two special cases captures the dominant properties of the joint image gradient distribution, and the effect of projecting the tangent plane to the imaging plane can be safely ignored. This shows that the statistical regularity of the scene radiance gradient does reflect the intrinsic geometric and reflectance properties of surfaces and that this regularity can be exploited. In Section 6, we will demonstrate this on the problem of face recognition under varying illumination.

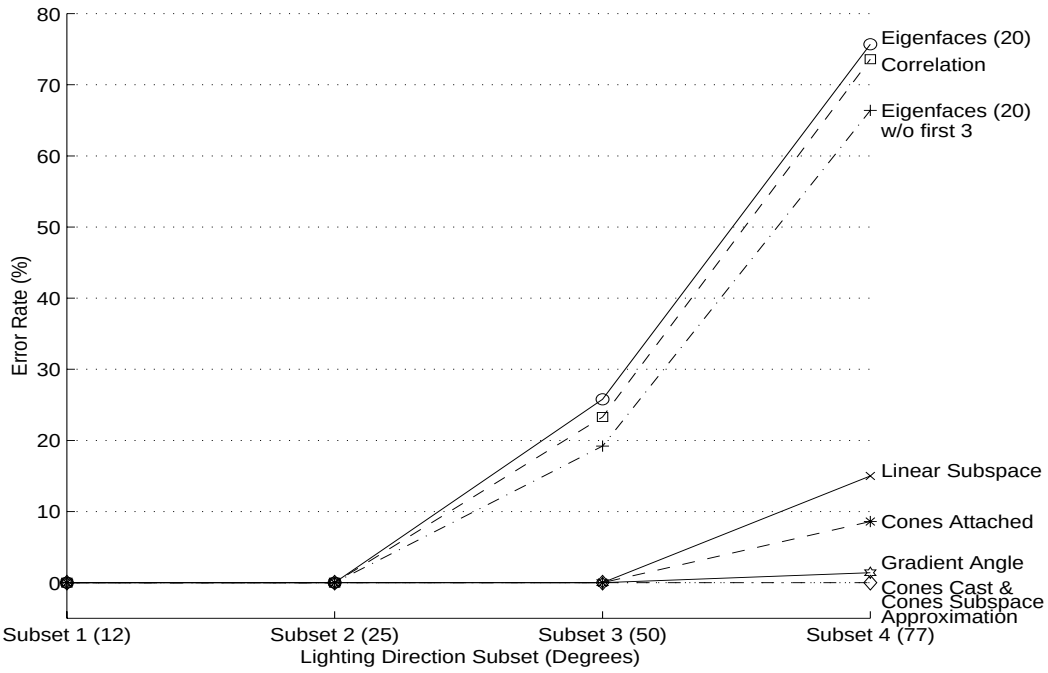
6 Application to Face Recognition

Given an object o which generates images I and J under two different lighting conditions, the joint probability of observing the gradients of I and J is assumed to satisfy

$$\begin{aligned} P(I,J) &= \prod_{i \in M} \rho(\vec{\nabla} I_i, \vec{\nabla} J_i) \\ &= \prod_{i \in M} \rho_\varphi(\varphi(i)), \end{aligned}$$

where M is the set of pixel indices and φ is the angle between the two gradient vectors. We treat the points on the surface as being independent, ignoring correlations arising from spatial proximity. From a Bayesian perspective, this is equivalent to assuming that when the images come from two different objects, the difference in the gradient direction will be uniformly distributed. We use density function ρ_φ as the illumination insensitive measure for the sake of simplicity.

We apply this scheme to face recognition. Given a test image I of a face, we compute $P(I,J)$ for every training image using an empirically collected probability database as described in Section 5. The one training image having the highest P value is deemed the most likely to have come from the same face as the test image I. Figure 7 shows the result of a face recognition test and compares it to those of other methods. 450 images of 10 faces each under 45 different lighting conditions are used. One image of each face under frontal illumination is taken as a training image. The recognition test is then performed for the remaining 440 images. The results are grouped into 4 subsets according to the lighting angle with respect to the frontal or camera axis. The first subset covers lighting angles $0^\circ - 12^\circ$, the second covers $12^\circ - 25^\circ$, the third $25^\circ - 50^\circ$, and the fourth $50^\circ - 77^\circ$. We compared



Comparison of Recognition Methods				
Method	Error Rate (%) vs. Illumination			
	Subset 1	Subset 2	Subset 3	Subset 4
Correlation	0.0	0.0	11.7	65.0
Eigenfaces	0.0	0.0	16.7	69.3
Eigenfaces w/o 1st 3	0.0	0.0	3.3	57.9
Linear subspace	0.0	0.0	1.7	12.9
Cones-attached	0.0	0.0	0.8	9.3
Gradient Angle	0.0	0.0	0.0	1.4
Cones-cast	0.0	0.0	0.0	0.0

Figure 7: Each of the methods except Gradient Angle, which has only one training image, is trained on images with near frontal illumination (Subsets 1). This graph shows the error rates under more extreme light source conditions.

our method with those tested and reported in [19]. The image gradient method clearly outperforms all other methods except Cone Cast.

We also perform the test using the probability density $\rho(r_1, \varphi, r_2)$ including the gradient magnitudes instead of ρ_φ with only the gradient angle, and the error rates are on a par with those above. It confirms our earlier assertion in Section 4 that not much information is lost when the magnitudes

Subset 1

Subset 2

Subset 3

Subset 4

Figure 8: Images of one of the 10 individuals under the 4 subsets of lighting.

are integrated out. Nevertheless, we hasten to reassert that when the image intensity is extremely low as in the shadow region and the shadow region is large, we expect that the higher dimensional density would be more discriminating.

It should be noted that all the methods other than Gradient Angle use all of Subset 1 for training, and so by definition have zero error rates for Subset 1. In contrast, our method uses only one frontal image in Subset 1 for each individual. The result of our method is still better than the others with the exception of Cone Cast. Also note that Cone Cast is the Illumination Cone method

in [18] in which Subset 1 images are used to construct person specific illumination representations. Our method, however, has the advantage of being much simpler and faster than Cone Cast using only local image comparisons.

It is worthwhile to note that the probability distribution used to perform the test is gathered from images of objects rather than human faces. The distributions collected from different categories of objects or faces are remarkably similar. This is expected from the analysis in Section 4.

7 Conclusion

This paper presents two results: first, illumination invariants do not exist for Lambertian surfaces; and second, the angle (or direction) of the image gradient is insensitive to changes in illumination direction. The latter statement is consistent with the conclusion in [1] that linear filters for image comparison do not exist, since the gradient angle is a nonlinear function of the image. However, we cannot conclude that image edges are good measures of image comparison under varying illumination — in fact the contrary is true. Most edge detection methods are highly sensitive to the magnitude of the image gradient. As we can see from Eq. 12, the magnitude of the gradient varies drastically with the change in the direction of the light source. The distributions in Figure 3 also shows the slow variation of the density function with respect to the magnitude of the gradient. Therefore the gradient magnitude is a poor indicator of the underlying surface geometry and photometry. Nevertheless, when combined with the gradient angle, the magnitude may extract more information from extremely low intensity images.

Our assumption about pixel independence implicit in Eq. 25 is obvious incorrect — the light sources are fixed for all pixels in a given image, and neighboring surface points tend to have similar geometric and reflectance properties. We hope to remedy the crudeness of our independence assumption by exploring how spatial dependencies in pixel values [2] can be exploited to produce even more discriminating measures of the images.

8 Appendix

The propositions in Section 3 are proved here.

The following observation on the characteristic curve Eq. 5 is crucial to the proofs of the lemmas.

Observation 8.1 *Let q be an arbitrary point in R^3, S_q the plane containing q parallel to $\vec{l}$ and $\vec{s}$, and $\vec{r}_q$ the characteristic curve passing through q. It is clear that if $\vec{r}_q$ exists, it remains in the plane S_q. Moreover, the characteristic vector at every point resides in the planar cone $C = \{a\vec{l} - b\vec{s},\ a, b \geq 0\}$ generated by $\vec{l}$, $-\vec{s}$, and the negative of C. Consequently the curve $\vec{r}_q$ is also confined in the two-sided cone $q + C \cup q - C$. Specifically, the projection $\vec{\rho}_p$ of $\vec{r}_q$ onto the XY-plane is contained in a cone C_p, the projection of C on the XY-plane.*

We also need another lemma, Whitney's Extension Theorem [15], to extend a C^1 function onto a larger set in the proof of the lemmas in the main text.

Lemma 8.1 *Let A be a closed set of points a in R^m at which the values and derivatives of a desired C^1 function are prescribed by linear polynomials $P_a : R^m \to R$. For each compact subset C of A and $\delta > 0$, let $\xi(C, \delta)$ be the supremum of the numbers $\frac{|P_a(b) - P_b(b)|}{|a-b|}$, $\|DP_a(b) - DP_b(b)\|$, where D is the differentiation operator, over all $a, b \in C$ with $0 < |a - b| \leq \delta$. If the prescribed data satisfy the coherence condition that $\lim_{\delta \to 0} \xi(C, \delta) = 0$ for each compact subset C of A, then there exists a C^1 funciton g satisfying $g(a) = P_a(a),\ Dg(a) = DP_a(a),\ \forall a \in A$, and $\inf_{x \in R^m} g(x) = \inf_{a \in A} P_a(a)$, $\sup_{x \in X} = \sup_{a \in A} P_a(a)$.*

Proof of Lemma 3.3.

We shall first caution the reader that this lemma concerns the global existence of a solution as opposed to the local existence. We intend to turn the ordinary equation into an integral equation and solve the integral equation with iterations. However, the interval of integration should be fixed and the trajectory from each iteration should all reach the boundary of the domain Ω. Thus the vector field should be extended to a larger domain containing Ω to satisfy these requirements.

Proof. Without loss of generality, we assume that the given point is at the origin and Ω is bounded

by $R > 0$. Since I and J are positive and continuous on a compact set, $M' > I, J > c$ for some constants $M' > c > 0$. Recall (Observation 8.1) that the characteristic vectors in the XY-plane are confined in the cone C_p. We denote by $\vec{h}$ the vector bisecting the angle subtended by $\vec{l_p}$ and $-\vec{s_p}$. As I and J are both bounded from below by $c > 0$, the projection of the vector field on $\vec{h}$ is bounded from below by $c(l_p + s_p)\cos\theta$, where θ is half the angle between $\vec{l_p}$ and $\vec{s_p}$, the projections of $\vec{l}$ and $\vec{s}$ on the XY plane respectively.

If $\theta < \frac{\pi}{2}$, any curve starting from the origin with these velocity (or tangential) vectors will leave Ω in time $T = \frac{R}{c\cos\theta}$. If $\theta = \frac{\pi}{2}$, $I\vec{l_p} - J\vec{s_p}$ is on a straight line. The exit time is then simply bounded by $T = \frac{R}{c}$. Let $\max_{\vec{\rho}\in\Omega} |I(\vec{\rho})\vec{l_p} - J(\vec{\rho})\vec{s_p}| = M$. By setting $P_{\vec{a}}(\vec{\rho}) = I(\vec{a}) + DI(\vec{a})(\vec{\rho})$, and the same for $J(\vec{\rho})$, Lemma 8.1 (Whitney's Extension Theorem) allows us to extend the functions $I, J \in C^1(\Omega)$ to C^1 functions, again denoted by I and J, onto the closed ball B of radius MT, such that both newly constructed I and J maintain their original values on Ω, and assume the same respective maxima and minima on B as on Ω.

We are now in a position to recursively construct a sequence of functions $\{\vec{\rho}_n\}_{n=0}^{\infty} : [0, T] \to B$. Let

$$\vec{\rho}_0(t) = 0. \tag{25}$$

Suppose $\vec{\rho}_n$ is defined on $[0, T]$. Define

$$\vec{\rho}_{n+1}(t) = \int_0^t \left[I(\vec{\rho}_n(\tau))\vec{l_p} - J(\vec{\rho}_n(\tau))\vec{s_p} \right] \mathrm{d}\tau, \ t \in [0, T]. \tag{26}$$

Clearly $\vec{\rho}_n(t) \in B, \ \forall t \in [0, T], \ \forall n \in \mathbf{N}$ or $\{\vec{\rho}\}_{n=0}^{\infty}$ is uniformly bounded by B. Since $I, J \in C^1$ on the convex and compact set B, by the mean value theorem, $|(I(\vec{\rho}_2)\vec{l_p} - J(\vec{\rho}_2)\vec{s_p}) - (I(\vec{\rho}_1)\vec{l_p} - J(\vec{\rho}_1)\vec{s_p})| < \lambda|\vec{\rho}_2 - \vec{\rho}_1|, \forall \vec{\rho}_1, \vec{\rho}_2 \in B$ for some constant $\lambda > 0$. This leads to

$$|\vec{\rho}_{n+1}(t) - \vec{\rho}_n(t)| < \lambda \int_0^t |\vec{\rho}_n(\tau) - \vec{\rho}_{n-1}(\tau)| \, \mathrm{d}\tau. \tag{27}$$

Apply the bound $|I\vec{l} - J\vec{s}| < M$ to Inequality 27,

$$|\vec{\rho}_1(t)| < Mt. \tag{28}$$

We obtain by induction,

$$|\vec{\rho}_{n+1}(t) - \vec{\rho}_n(t)| < M\frac{\lambda^n t^{n+1}}{(n+1)!},\ \forall t \in [0,t]. \tag{29}$$

The series $\sum_{n=0}^{\infty}(\vec{\rho}_{n+1}(t) - \vec{\rho}_n(t))$ is majorized by that of $\frac{M}{\lambda}e^{\lambda t}$. Therefore the sequence $\vec{\rho}_n$ tends uniformly to a continuous function $\vec{\rho}$. Since B is closed, $\vec{\rho}(t) \in B$, $\forall t \in [0,T]$.

$I\vec{l_p} - J\vec{s_p}$ is continuous on the compact set B, and is therefore uniformly continuous. Together with the uniform convergence of $\{\vec{\rho}_n\}_{n=0}^{\infty}$, it enables us to take the limit $n \to \infty$ on both sides of Eq. 26, and substitute $\vec{\rho}_n$ with $\vec{\rho}$ in the integrand.

Let $\zeta = \inf_{t\in[0,T]}\{t : \vec{\rho}(t) \in B - Int(\Omega)\}$. Since the image of the continuous function $\vec{\rho}$ on the compact set $[0,T]$ is also compact and therefore closed, its intersection with the closed set $B - Int(\Omega)$ is closed. Again by continuity of $\vec{\rho}$ $\{t : \vec{\rho}(t) \in B - Int(\Omega)\}$ is closed, and thus $\vec{\rho}(\zeta) \in B - Int(\Omega)$. Consequently $\vec{\rho}(\zeta) \in \partial\Omega$ and $\vec{\rho}(t) \in Int(\Omega)$, $\forall t \in [0,\zeta)$. Thus $\vec{\rho}(t),\ t \in [0,\zeta]$ is a solution of Eq. 5 on Ω.

The uniqueness of the characteristic curve through the origin on Ω can be shown following the classical argument. Suppose there are two solutions to Eq. 5, $\vec{\rho}_1$ and $\vec{\rho}_2$ with the same initial points $\vec{\rho}_1(0) = \vec{\rho}_2(0) = 0$ (They may exist only in the interior of Ω without reaching the boundary.). Suppose they are both defined up to t. We have

$$\begin{aligned} |\vec{\rho}_2(t) - \vec{\rho}_1(t)| &\leq \int_0^t \left|(I[\vec{\rho}_2(\tau)]\vec{l_p} - J[\vec{\rho}_2(\tau)]\vec{s_p}) - (I[\vec{\rho}_1(\tau)]\vec{l_p} - J[\vec{\rho}_1(\tau)]\vec{s_p})\right| \mathrm{d}\tau \\ &\leq \lambda\int_0^t |\vec{\rho}_2(\tau) - \vec{\rho}_1(\tau)|\,\mathrm{d}\tau, \end{aligned}$$

or, what is equivalent,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-\lambda t}\int_0^t |\vec{\rho}_2(t) - \vec{\rho}_1(t)|\,\mathrm{d}\tau\right) \leq 0. \tag{30}$$

It is obvious from the initial condition $|\vec{\rho}_2(0) - \vec{\rho}_1(0)| = 0$, that $|\vec{\rho}_2(t) - \vec{\rho}_1(t)| = 0$ for every defined t. Consequently, there is a unique solution emanating from the origin and evolving positively in time. It lands on the boundary $\partial\Omega$ at $t = \zeta$.

As for the other part of the curve in the negative direction, replace t with $-t$, every step of the above argument carries through. □

Proof of Lemma 3.4. The proof for these global properties is the same as that for the corresponding local ones in many classic texts (e.g., [11]). □

Proof of Lemma 3.5.

Lemma 3.3 guarantees that every point in Ω is an outgrowth from the boundary. A surface covering the whole domain Ω can be constructed by growing the characteristic curves from boundary curves. However the boundary conditions should not lead to multi-valued function f. To avoid this difficulty, we simply expend the vector field to a triangle containing Ω, and assign boundary values to one side of the triangle which will not lead to multi-valuedness. The required surface is simply the restriction of the surface constructed on the triangle to Ω.

Proof. Let $\vec{c} = \frac{1}{2}(\vec{l_p} - \vec{s_p})$. Draw a straight line Γ perpendicular to $\vec{c}$ such that Ω is on the side of Γ that $\vec{c}$ points to. Draw two lines parallel to $\vec{l_p}$ and $-\vec{s_p}$ respectively, so that together with Γ they form a triangle enclosing the compact set Ω. Then restrict Γ to denote the closed segment that is a side of the triangle (If $\vec{l_p}$ and $-\vec{s_p}$ are parallel, we can either consider the triangle as having one apex at infinity, or cut off the two parallel sides and form a rectangle, so long as it contains Ω.). We again use Whitney's Extension Theorem, Lemma 8.1, and extend I and J on Ω to two C^1 functions on the closed triangle, retaining their original minima and maxima.

Let the compact set Ω in the premise of Lemma 3.3 be the triangle constructed above. It then follows from Lemma 3.3 and Observation 8.1, and the fact that Γ is not parallel to any characteristic direction of the PDE, that through any point $\vec{\rho} \in \Omega$ runs a unique characteristic line emanating from a unique point on Γ and reaching the other two sides of the triangle. Parameterize Γ by its length s from one of its two ends. Denote the characteristic curve emanating from Γ by $\vec{\rho}(s,t)$. By Lemma 3.4, $(x(s,t), y(s,t)) = \vec{\rho}(s,t)$ establishes a C^1 diffeomorphic coordinate chart for the triangle (a two dimensional manifold).

Let $f_0(s)$ be an arbitrary C^1 function on Γ and let

$$f(s,t) = \int_0^t (I[\vec{\rho}(s,\tau)]l_z - J[\vec{\rho}(s,\tau)]s_z)\ \mathrm{d}\tau \tag{31}$$

It is obvious that $f(s,t) \in C^1$, and therefore $\vec{r}(s,t) = (x(s,t), y(s,t), f(s,t))$ is a C^1 manifold. By the diffeomorphism $(x,y) = \vec{\rho}(s,t)$, f could be written as a function of (x,y) and the manifold is the graph of f. The desired surface is the manifold restricted to Ω. □

Proof of Lemma 3.2.

As is mentioned in Section 3, for Eq.'s 2 and 3 to adequately provide the global description of the required surface, it is necessary that the surface constructed should have no attached or cast shadows. Otherwise Eq. 1 instead should be used. Since all the characteristic curves stay confined in a planar cone described in Observation 8.1, no point on the surface will attach a shadow to itself or cast one on another point so long as the boundary curve assigned to the side of the triangle satisfies certain condition.

Proof. Construct the same triangle as in Lemma 3.5. Let $f_0(s)$ be the initial condition on Γ as defined in the proof of Lemma 3.5. By Lemma 3.5, the graph can also be denoted as $\vec{r}(s, t=0)$. In addition, $\vec{r}(s,0)$ is such that

$$-\vec{s} \times \vec{l} \cdot \frac{\partial \vec{r}(s,0)}{\partial s} > 0. \tag{32}$$

Point a casts a shadow on point b if and only if a and b are on the same line parallel to either $\vec{s}$ or $\vec{l}$. By the way we construct the surface, it is a manifold parameterized by (s,t) via a diffeomorphism $\vec{r}(s,t)$. Let $a = \vec{r}(s_1,t_1)$ and $b = \vec{r}(s_2,t_2)$, we distinguish two cases: 1) $s_1 = s_2$; 2) $s_1 \neq s_2$.

In case of 1),

$$\vec{r}(s_1,t_2) - \vec{r}(s_1,t_1) = \int_{t_1}^{t_2} \left(I[\vec{\rho}(s_1,t)]\vec{l} - J[\vec{\rho}(s_1,t)]\vec{s}\right)\ \mathrm{d}\tau. \tag{33}$$

Since $J[\vec{\rho}(s,t)] > 0$ and continuous, $-\int_{t_1}^{t_2} J[\vec{\rho}(s_1,t)]\ \mathrm{d}\tau < 0$ and thus $\vec{r}(s_1,t_2) - \vec{r}(s_1,t_1)$ is not parallel to $\vec{l}$. By the same token, $\vec{r}(s_1,t_2) - \vec{r}(s_1,t_1)$ is not parallel to $\vec{s}$ either.

For case 2), we show that $\vec{r}(s_1,t_2)-\vec{r}(s_1,t_1)$ is not in the plane of $\vec{l}$ and $\vec{s}$.

$$\begin{aligned} -\vec{s}\times\vec{l}\cdot\vec{r}(s,t) &= -\vec{s}\times\vec{l}\cdot\left[\int_0^t\left(I[\vec{\rho}(s,\tau)]\vec{l}-J[\vec{\rho}(s,\tau)]\vec{s}\right)\,\mathrm{d}\tau+\vec{r}(s,0)\right] \\ &= -\vec{s}\times\vec{l}\cdot\left(a(s,t)\vec{l}-b(s,t)\vec{s}+\vec{r}(s,0)\right) \\ &= -\vec{s}\times\vec{l}\cdot\vec{r}(s,0). \end{aligned} \tag{34}$$

$a(s,t)$ and $b(s,t)$ stand for the scalar functions of vectors $\vec{l}$ and $\vec{s}$. Without loss of generality, we assume $s_2>s_1$.

$$-\vec{s}\times\vec{l}\cdot(\vec{r}(s_2,t_2)-\vec{r}(s_1,t_1))=-\vec{s}\times\vec{l}\cdot(\vec{r}(s_2,0)-\vec{r}(s_1,0))>0, \tag{35}$$

where the last inequality comes from integrating Inequality 32 over s from s_1 to s_2 and the claim is verified.

Divide Inequality 34 by s_2-s_1 and let $s_2\to s_1$, we obtain

$$-\vec{s}\times\vec{l}\cdot\frac{\partial\vec{r}(s,t)}{\partial s}>0. \tag{36}$$

Then for the surface normal vector $\frac{\partial\vec{r}(s,t)}{\partial s}\times\frac{\partial\vec{r}(s,t)}{\partial t}$, the sign of the inner product is always

$$\frac{\partial\vec{r}(s,t)}{\partial s}\times\frac{\partial\vec{r}(s,t)}{\partial t}\cdot\vec{s}=\frac{\partial\vec{r}(s,t)}{\partial s}\cdot\left(I\vec{l}-J\vec{s}\right)\times\vec{s}>0, \tag{37}$$

since $I(x,y)$ and $J(x,y)$ are positive. The same inequality holds for the inner product of the normal vector and $\vec{l}$. In other words, there is no attached shadow. □

Proof of Theorem 3.2.

We have shown that in the absence of interreflections we can construct an object and two light sources that produce the two images I and J. We now show that with interreflections, there will still be no continuous discriminative invariants. To do this we show that if we scale up the magnitude of the light sources, while scaling down the albedos by an inverse amount, we can suppress the effect of interreflections to the infinitesimal and thus approximate the given images with those generated by the surface with arbitrary precision.

Proof. An image $I(x,y)$ generated by an interreflecting surface $f(x,y)$ on a compact set $(x,y) \in \Omega$ can be described by the differo-integral equation

$$I = I_0 + \alpha S[f]I, \tag{38}$$

where I is the image function, and I_0 is the intensity of the first reflection off the surface of the light directly from the light source. Linear operator $S[f]$ is a function of the continuous surface $f \in C(\Omega) \cap$ piecewise $C^1(\Omega)$,

$$S[f]I(\vec{\rho}) = -\int_\Omega \mathrm{d}^2\rho'\, \delta(\vec{\rho}, \vec{\rho}\,')\frac{\hat{n}\cdot\hat{r}\;\hat{n}'\cdot\hat{r}}{r^2} I(\vec{\rho}\,'), \tag{39}$$

where $\vec{\rho} = (x,y)$, $\vec{\rho}\,' = (x',y')$, $\vec{\rho}, \vec{\rho}\,' \in \Omega$. $\mathrm{d}^2\rho'$ stands for area element on the surface. $\vec{r} = (x'-x, y'-y, z'-z)$, $z = f(\vec{\rho}) = f(x,y)$, and $z' = f(\vec{\rho}\,') = f(x',y')$. $r = |\vec{r}|$, $\hat{r} = \vec{r}/r$. $\hat{n}$ and $\hat{n}'$ are the surface normal at points $(x,y,f(x,y))$ and $(x',y',f(x',y'))$, respectively. $\delta(\vec{\rho}, \vec{\rho}\,')$ is an index function equal to 1 if the line segment connecting $\vec{\rho}$ and $\vec{\rho}\,'$ does not intersect the surface $(x,y,f(x,y))$, and 0 if the line segment does.

It is clear by conservation of energy, the maximum norm of the operator $\|S[f]\|_{L^\infty} < 1$, and $0 < \alpha_m \stackrel{\text{def}}{=} \max_{\vec{\rho}\in\Omega} \alpha(\vec{\rho}) < 1$. The solution of Eq. 38 is then readily obtained by Neumann series, i.e.,

$$I = \sum_{n=0}^{\infty} \alpha^n S^n I_0, \tag{40}$$

where S^n denotes the n'th operation of the operator S on I_0 and α^n is the n'th multiplication. Thus

$$\|I - I_0\|_{L^\infty} < \alpha_m \frac{\|I_0\|_{L^\infty}}{1 - \alpha_m \|S\|_{L^\infty}}. \tag{41}$$

Given two image functions I_0 and J_0 and two arbitrary light source direction $\hat{s}$ and $\hat{l}$, there are, according to Theorem 3.1, a surface f and an albedo α, such that $I_0 = \alpha\vec{s}\cdot\hat{n} = a\tilde{\alpha}\hat{s}\cdot\hat{n}$ and $J_0 = \alpha\vec{l}\cdot\hat{n} = b\tilde{\alpha}\hat{l}\cdot\hat{n}$ (and there is no cast or attached shadow on f), where $a \stackrel{\text{def}}{=} \alpha_m|\vec{s}|$, $b \stackrel{\text{def}}{=} \alpha_m|\vec{l}|$, and $\tilde{\alpha} \stackrel{\text{def}}{=} \frac{\alpha}{\alpha_m}$.

Define I and J, as at the beginning of the proof, to be the images generated by the surface f with interreflection under $\vec{s}$ and $\vec{l}$, respectively. Let $s \stackrel{\text{def}}{=} |\vec{s}| = \frac{a}{\alpha_m}$ and $l \stackrel{\text{def}}{=} |\vec{l}| = \frac{b}{\alpha_m}$, $I(\alpha_m)$ and

$J(\alpha_m)$ are then functions of $0 < \alpha_m < 1$. I_0 and J_0, on the other hand, are independent of α_m. It then follows from Inequality 41 that $I(\alpha_m) \to I_0$ and $J(\alpha_m) \to J_0$, as $\alpha_m \to 0^+$.

Suppose μ is a continuous illumination invariant. By the definition of illumination invariant, $\mu(I(\alpha_m)) = \mu(J(\alpha_m))$, since I and J are generated from the same surface and albedo. By the continuity of μ, $\mu(I_0) = \lim_{\alpha_m \to 0^+} \mu(I(\alpha_m)) = \lim_{\alpha_m \to 0^+} \mu(J(\alpha_m)) = \mu(J_0)$. Hence the invariant μ is nondiscriminative. □

Acknowledgements

We are grateful to Alan Yuille, David Kriegman, David Forsyth, Mike Langer, Steven Zucker, Athinodoros Georghiades, Jonas August, Patrick Huggins, Melissa Koudelka and Todd Zickler for their help and advice in writing this paper.

References

[1] Y. Adini, Y. Moses, S. Ullman. "Face Recognition: The Problem of Compensating for Changes in Illumination Direction", *IEEE Trans.PAMI*, Vol.19, No.7: 721–732, July 1997.

[2] J. August and S. Zucker, "The curve indicator random field: curve organization via edge correlation," in *Perceptual Organization for Artificial Vision Systems*, K. Boyer and S. Sarkar, Eds., pp. 265–288, Kluwer Academic, Boston, 2000.

[3] P. Belhumeur, D. Kriegman. "What is the Set of Images of an Object Under All Possible Lighting Conditions?", *IEEE CVPR 96*: 270–277, 1996.

[4] P. Belhumeur, D. Kriegman, and A. Yuille, 1997. "The Bas-Relief Ambiguity" *CVPR*:1060–1066.

[5] P. Breton, S. Zucker. "Shadows and Shading Flow Fields", *CVPR 96*: 782-789, 1996.

[6] M. J. Brooks and B. K. P. Horn, "Shape and Source from Shading", *Proc. of the 9th Int. Joint Conf. on Artificial Intelligence*: 932–936, 1985.

[7] J. Burns, R. Weiss, E. Riseman. "The Non-Existence of General-Case View-Invariants", *Geometric Invariance in Computer Vision*, edited by J. Mundy, A. Zisserman, MIT Press, Cambridge, 1992.

[8] H. Chen, P. Belhumeur. "Illumination Invariance". Technical Report, Center for Computational Vision, Computer Science Dept., Yale University, Dec., 1999.

[9] D. Clemens, D. Jacobs. "Space and Time Bounds on Model Indexing", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.13, No.10: 1007–1018,1991.

[10] E. Coleman, R. Jain. "Obtaining 3-Dimensional Shape of Textured and Specular Surfaces Using Four-Source Photometry", *CGIP 82*, Vol.18, No.4: 309–328, 1982.

[11] Earl A. Coddington, Norman Levinson. *Theory of Ordinary Differential Equations*, Krieger Publishing Co. Malabar, Fl, 1984.

[12] J. Fan, L. Wolff. "Surface Curvature and Shape Reconstruction from Unknown Multiple Illumination and Integrability", *Computer Vision and Image Understanding*, Vol.65, No.2: 347–359, 1997.

[13] Faugeras, O., 1992, "What can be Seen in Three Dimensions with an Uncalibrated Stereo Rig?" *Second European Conference on Computer Vision*:563-578.

[14] O. Faugeras, L. Robert. "What Can Two Images Tell Us about a Third One?", *Int. J. of Comp. Vis.*, Vol. 18, No.1: 5–19, 1996.

[15] H. Federer, "Geometric Measure Theory", Springer-Verlag, New York, 1969.

[16] Forsyth, D., Mundy, J., Zisserman, A., and Rothwell, C., 1992. "Recognising Rotationally Symmetric Surfaces from their Outlines," *European Conf. on Comp. Vis.*:639–647.

[17] B. Funt, G. Finlayson. "Color Constant Color Indexing", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.17, No.5: 522-529, 1995.

[18] A. Georghiades, D. Kriegman, P. Belhumeur. "Illumination Cones for Recognition Under Variable Lighting: Faces", *CVPR 98*: 52–59, 1998.

[19] A. Georghiades, P. Belhumeur ,D. Kriegman. "From Few to Many: Generative Models for Recognition Under Variable Pose and Illumination", *Int. Conf. on Automatic Face and Gesture Recognition 2000*, 2000.

[20] P. Hallinan. "A Low-Dimensional Representation of Human Faces for Arbitrary Lighting Conditions", *CVPR 94*: 995–999, 1994.

[21] D. R. Hougen and N. Ahuja, "Estimation of the Light Source Distribution and Its Use in Integrated Shape Recovery from Stereo and Shading", *IEEE 4th Int. Conf. on Computer Vision*, Berlin, Germany: 148–155, 1993.

[22] M. Irani, and P. Anandan, "Robust multi-sensor image alignment", ICCV 98, 959-966.

[23] D. Jacobs. "Matching 3-D Models to 2-D Images", *Int. J. of Comp. Vis.*, Vol.21, No.1-2: 123–153, 1997.

[24] D. Jacobs, P. Belhumeur, R. Basri. "Comparing Images Under Variable Illumination", *CVPR 98*: 610-617, 1998.

[25] C-Y Kim et al. "Illuminant Direction and Shape of a Bump", *J. Opt. Soc. Am.* **15**: 2341–2350, 1998.

[26] M. Kirby, and L. Sirovich, 1990, "The application of the Karhunen-Loeve procedure for the characterization of human faces", *IEEE transactions on Pattern Analysis and Machine Intelligence*, **12**(1):103-108.

[27] J. J. Koenderink, A. J. van Doorn. "Photometric Invariants related to solid Shape", *Optica Acta*, Vol. 27, No.7: 981-996, 1980.

[28] J. Lambert. "Photometria Sive de Mensura et Gradibus Luminus, Colorum et Umbrae", Eberhard Klett, 1760.

[29] Lamdan, Y., J.T. Schwartz and H.J. Wolfson, 1990, "Affine Invariant Model-Based Object Recognition," *IEEE Trans. Robotics and Automation*, **6**: 578–589.

[30] C. H. Lee and A. Rosenfeld, "Improved Methods of Estimating Shape from Shading Using the Light Source Coordinate System", *Artificial Intelligence*, No.26: 125–143, 1985.

[31] Y. Moses. *Face recognition: generalization to novel images*, Ph.D. Thesis, Weizmann Institute of Science, 1993.

[32] Y. Moses, S. Ullman. "Generalization to Novel Views: Universal, Class-based, and Model-based Processing", *Int. J. of Comp. Vis.*, Vol.29, No.3: 233-253, 1998.

[33] Y. Moses, S. Ullman. "Limitations of Non Model-Based Recognition Schemes", *Sec. Eur. Conf. on Comp. Vis.*: 820-828, 1992.

[34] J. Mundy, A. Zisserman (eds.). *Geometric Invariance in Computer Vision*, MIT Press, Cambridge, 1992.

[35] H. Murase, S. Nayar. Visual learning and recognition of 3D objects from appearance. *Int. J. of Comp. Vis.* , Vol.14, No.1: 5–25, 1995.

[36] S. Nayar, R. Bolle. "Reflectance Based Object Recognition", *Int. J. of Comp. Vis.*, Vol.17, No.3: 219-240, 1996.

[37] R. Onn, F. Bruckstein. "Integrability Disambiguates Surface Recovery in Two-Image Photometric Stereo", *Int. J. of Comp. Vis.*, Vol.5, No.1: 105–113, 1990.

[38] A. P. Pentland, "Finding the Illuminant Direction", J. Opt. Soc. Am. **72**: 448-455, 1982.

[39] R. Rao, and D. Ballard, "Object Indexing Using an Iconic Sparse Distributed Memory," *IEEE Int. Conf. on Comp. Vis.* :24-31, 1995.

[40] D. Ruderman, "The statistics of natural images", *Network: Computation in Neural Systems* **5** (1994), 517-548.

[41] D. Ruderman and W. Bialek, "Statistics of Natural Images: Scaling in the Woods", *Phyical Review Letters* **Vol 73**, No. 6 (1994), 814-817.

[42] C. Schmid, R. Mohr. "Local Grayvalue Invariants for Image Retrieval", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.19, No.5:530-535, 1997.

[43] A. Shashua. "On Photometric Issues in 3D Visual Recognition from a Single 2D Image", *Int. J. of Comp. Vis.*, Vol.21, No.1-2: 99–122, 1997.

[44] M. Turk, A. Pentland. "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, Vol.3, No.1: 71-96, 1991.

[45] S. Ullman, R. Basri. "Recognition by Linear Combinations of Models", *IEEE Trans. PAMI*, Vol.13, No.10: 992-1007, 1991.

[46] D. Weinshall, "The Shape and the Direction of Illumination from Shading on Occluding Contours", *MIT Artificial Intelligence Memo* 1264, 1990.

[47] L. Wolff, E. Angelopoulou. *Eur. Conf. on Comp. Vis.*: 247–258, 1994.

[48] L. Wolff, J. Fan. "Segmentation of Surface Curvature with a Photometric Invariant", *J. Opt. Soc. Am. A*, Vol.11, No.11: 3090-3100, Nov. 1994.

[49] Y. Yang and A. Yuille, "Source from Shading", *Proc. of the Conf. on Comp. Vision and Pat. Recog.*: 534–539, 1991.

[50] Y. Zhang and Y. Yang, "Illuminant Direction Determination for Multiple Light Sources", *Proc. of the Conf. on Comp. Vision and Pat. Recog.*: 269–276, 2000.

[51] M. Zerroug, and R. Nevatia, (1996). Three-dimensional descriptions based on the analysis of the invariant and quasi-invariant properties of some curved-axis generalized cylinders. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 18, 3,* 237-966.

[52] Q. Zheng and R. Chellappa, "Estimation of Illumination, Albedo, and Shape from Shading", *IEEE TRans. Pattern Anal. Mach. Intell.* **13**: 680–702, 1991.