What Is the Set of Images of an Object
under All Possible Illumination Conditions?

PETER N. BELHUMEUR

Center for Comp. Vision and Control, Dept. of Electrical Engineering, Yale University, New Haven CT 06520

DAVID J. KRIEGMAN

Center for Comp. Vision and Control, Dept. of Electrical Engineering, Yale University, New Haven CT 06520

Abstract. *The appearance of an object depends on both the viewpoint from which it is observed and the light sources by which it is illuminated. If the appearance of two objects is never identical for any pose or lighting conditions, then – in theory – the objects can always be distinguished or recognized. The question arises: What is the set of images of an object under all lighting conditions and pose? In this paper, we consider only the set of images of an object under variable illumination, including multiple, extended light sources, shadows, and color. We prove that the set of n-pixel monochrome images of a convex object with a Lambertian reflectance function, illuminated by an arbitrary number of point light sources at infinity, forms a convex polyhedral cone in $\mathbb{R}^n$ and that the dimension of this* illumination cone *equals the number of distinct surface normals. Furthermore, the illumination cone can be constructed from as few as three images. In addition, the set of n-pixel images of an object of any shape and with a more general reflectance function, seen under all possible illumination conditions, still forms a convex cone in $\mathbb{R}^n$. These results immediately suggest certain approaches to object recognition. Throughout, we present results demonstrating the illumination cone representation.*

1. Introduction

One of the complications that has troubled computer vision recognition algorithms is the variability of an object's appearance from one image to the next. Wit1Gh slight changes in lighting conditions and viewpoint often come large changes in the object's appearance. To handle this variability methods usually take one of two approaches: either measure some property in the image of the object which is, if not invariant, at least insensitive to the variability in the imaging conditions, or model the object, or part of the object, in order to predict the variability.

Nearly all approaches to object recognition have handled the variability due to illumination by using the first approach; they have concentrated on edges in the images, i.e. the discontinuities in the image intensity. Because discontinuities in the albedo on the surface of the object or discontinuities in albedo across the boundary of the object generate edges in images, these edges tend to be insensitive to a range of illumination conditions [5].

Yet, edges do not contain all of the information useful for recognition. Furthermore, objects which are not simple polyhedra or are not composed of piecewise constant albedo patterns often produce inconsistent edge maps. The left half of Fig. 1 shows two images of a person with the same facial expression and photographed from the same viewpoint. The variabil-

ity in these two images due to differences in illumination is dramatic: not only does it lead to a change in contrast, but also to changes in the configuration of the shadows, i.e. certain regions are shadowed in the left image, but illuminated in the right, and vice versa. In fact, it has been observed that in face recognition, the variability in an image due to illumination is often greater than that due to a change in the person's identity [20]. The edge maps in the right half of Fig. 1 are produced from these images. Due to the variation in illumination, only a small fraction of the edges are common between images.

The reason most approaches have avoided using the rest of the intensity information is because its variability under changing illumination has been difficult to tame. Only recently have "appearance-based" approaches been developed in an effort to use intensity information to model or learn a representation that captures a large set of the possible images of an object under pose and/or illumination variation [21, 25, 26, 18, 31]. These methods have gone a long way in demonstrating the advantages of using much richer descriptions than simply sparse features like edges and corners for recognition.

Still, a drawback of these approaches is that in order to recognize an object seen from a particular pose and under a particular illumination, the object must have been previously seen under the same conditions. Yet, if one enumerates all possible poses and permutes these with all possible illumination conditions, things get out of hand quite quickly. This raises the question: Is there some underlying "generative" structure to the set of images of an object under varying illumination and pose such that to create the set, the object does not have to be viewed under all possible conditions?

In this paper, we address only part of this question, restricting our investigation to varying illumination. In particular, if an image with n pixels is treated as a point in $\mathbb{R}^n$, what is the set of all images of an object under varying illumination? Is this set an incredibly complex, but low-dimensional manifold in the image space? Or does the set have a simple, predictable structure? If the object is convex in shape and has a Lambertian reflectance function, can a finite number of images characterize this set? If so, how many images are needed?

The image formation process for a particular object can be viewed as a function of pose and lighting. Since an object's pose can be represented by a
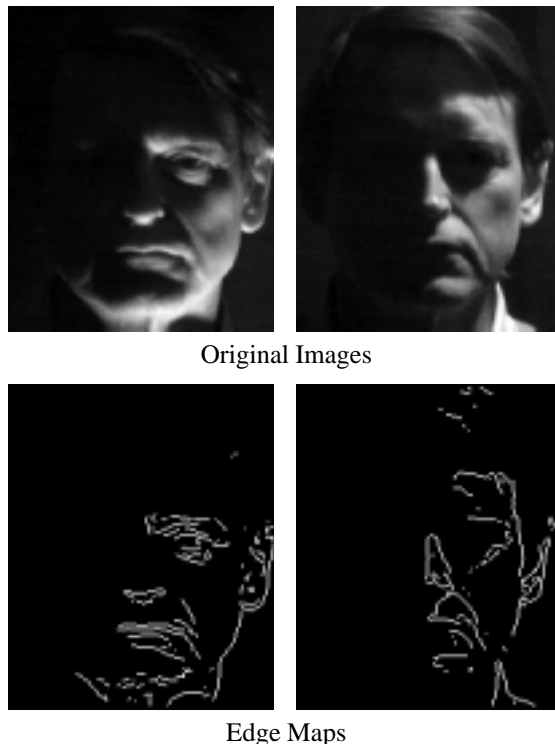

Original Images

Edge Maps

Fig. 1. **Effects of Variability in Illumination**: The left two images show the same face seen under different illumination conditions. The right two images show edge maps of the left two images. Even though the change in light source direction is less than $45°$, the change in the resulting image is dramatic.

point in $\mathbb{R}^3 \times SO(3)$ (a six dimensional manifold), the set of n-pixel images of an object under constant illumination, but over all possible poses, is at most six dimensional. Murase and Nayar take advantage of this structure when constructing appearance manifolds [21]. However, the variability due to illumination may be much larger as the set of possible lighting conditions is infinite dimensional.

Arbitrary illumination can be modeled as a scalar function on a four-dimensional manifold of light rays [19]. However, without limiting assumptions about the possible light sources, the bidirectional reflectance density functions, or object geometry, it is difficult to draw limiting conclusions about the set of images. For example, the image of a perfect mirror can be anything. Alternatively, if the light source is composed of a collection of independent lasers with one per pixel (which is admissible under [19]), then an arbitrary image of any object can be constructed by appropriately selecting the lasers' intensities.

Nonetheless, we will show that the set of images of an object with arbitrary reflectance functions seen un-

der arbitrary illumination conditions is a convex cone in $\mathbb{R}^n$ where n is the number of pixels in each image. Furthermore, if the object has a convex shape and a Lambertian reflectance function, the set of images under an arbitrary number of point light sources at infinity is a convex polyhedral cone in $\mathbb{R}^n$, which can often be determined from as few as three images. In addition, we will show that while the dimension of the illumination cone equals the number of distinct surface normals, the shape of the cone is "flat," i.e. the cone lies near a low dimensional linear subspace of the image space. Throughout the paper, empirical investigations are presented to complement the theoretical arguments. In particular, experimental results are provided which support the validity of the illumination cone representation and the associated propositions on the illumination cone's dimension and shape. Note that some results in this paper were originally presented in [3].

We are hopeful that the proposed illumination representation will prove useful for object recognition. While this paper provides no implementations of such, we believe the utility of the ideas presented in this paper will ultimately be determined through application. For problems such as face recognition, where pose is often fixed, we are building illumination cone representations for each individual in the database, thus allowing face recognition over extreme variation in illumination. For problems where pose is unknown, we envision marrying the illumination cone representation with a low-dimensional set of image coordinate transformations [32] or with the appearance manifold work of [21], thus allowing both illumination and pose variation. For problems in which occlusion and non-rigid motion cannot be discounted, we envision breaking the object down into sub regions and building "illumination subcones." These illumination subcones could then be glued together in a manner similar to the recent "body plans" work of [10].

2. The Illumination Cone

In this section, we develop the illumination cone representation. To start, we make two simplifying assumptions: first, we assume that the surfaces of objects have Lambertian reflectance functions; second, we assume that the shape of an object's surface is convex. While the majority of the propositions are based upon these two assumptions, we will relax them in Section 2.3

and show that the set of images is still a convex cone. In addition, the empirical investigations of Section 2.4 will demonstrate the validity of the illumination cone representation by presenting results on images of objects which have neither purely Lambertian reflectance functions, nor convex shapes.

2.1. Illumination Cones for Convex Lambertian Surfaces

To begin let us assume a Lambertian model for reflectance with a single point light source at infinity. Let $\mathbf{x}$ denote an image with n pixels. Let $B \in \mathbb{R}^{n \times 3}$ be a matrix where each row of B is the product of the albedo with the inward pointing unit normal for a point on the surface projecting to a particular pixel; here we effectively approximate a smooth surface by a faceted one and assume that the surface normals for the set of points projecting to the same image pixel are identical.

Let $\mathbf{s} \in \mathbb{R}^3$ be a column vector signifying the product of the light source strength with the unit vector for the light source direction. Thus, a convex object with surface normals and albedo given by B, seen under illumination $\mathbf{s}$, produces an image $\mathbf{x}$ given by the following equation

$$\mathbf{x} = \max(B\mathbf{s}, \mathbf{0}), \qquad (1)$$

where $\max(\cdot, \mathbf{0})$ zeros all negative components of the vector $B\mathbf{s}$ [14]. Note that the negative components of $B\mathbf{s}$ correspond to the shadowed surface points and are sometimes called *attached shadows* [28]. Also, note that we have assumed that the object's shape is convex at this point to avoid *cast shadows*, i.e. shadows that the object casts on itself.

If the object is illuminated by k point light sources at infinity, the image $\mathbf{x}$ is given by the superposition of images which would have been produced by the individual light source, i.e.

$$\mathbf{x} = \sum_{i=1}^{k} \max(B\mathbf{s}_i, \mathbf{0})$$

where $\mathbf{s}_i$ is a single light source. Note that extended light sources at infinity can be handled by allowing an infinite number of point light sources (i.e., the sum becomes an integral).

The product of B with all possible light source directions and strengths sweeps out a subspace in the n-dimensional image space [28, 13, 23]; we call the

subspace created by B the illumination subspace $\mathcal{L}$, where

$$\mathcal{L} = \{\mathbf{x} \mid \mathbf{x} = B\mathbf{s}, \forall \mathbf{s} \in \mathbb{R}^3\}.$$

Note that the dimension of $\mathcal{L}$ equals the rank of B. Since B is an $n \times 3$ matrix, $\mathcal{L}$ will in general be a 3-D subspace, and we will assume it to be so in the remainder of the paper. When the surface has fewer than three linearly independent surface normals, B does not have full rank. For example, in the case of a cylindrical object, both the rank of B and dimension of $\mathcal{L}$ are two. Likewise, in the case of a planar object, both the rank and dimension are one.

When a single light source is parallel with the camera's optical axis, all visible points on the surface are illuminated, and consequently, all pixels in the image have non-zero values. The set of images created by scaling the light source strength and moving the light source away from the direction of the camera's optical axis such that all pixels remain illuminated can be found as the relative interior of a set $\mathcal{L}_0$ defined by the intersection of $\mathcal{L}$ with the non-negative orthant of $\mathbb{R}^{n}$.[1]

Lemma 1. *The set of images $\mathcal{L}_0$ is a convex cone in* $\mathbb{R}^n$.

Proof: $\mathcal{L}_0 = \mathcal{L} \cap \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n$, with all components of $\mathbf{x} \geq 0\}$. Both $\mathcal{L}$ and the positive orthant are convex. For the definition of convexity and the definition of a cone, see [6, 27]. Because the intersection of two convex sets is convex, it follows that $\mathcal{L}_0$ is convex.

Because $\mathcal{L}$ is a linear subspace, if $\mathbf{x} \in \mathcal{L}$ then $\alpha\mathbf{x} \in \mathcal{L}$. And, if $\mathbf{x}$ has all components non-negative, then $\alpha\mathbf{x}$ has all components non-negative for every $\alpha \geq 0$. Therefore $\alpha\mathbf{x} \in \mathcal{L}_0$. So it follows that $\mathcal{L}_0$ is a cone. $\square$

As we move the light source direction further from the camera's optical axis, points on the object will fall into shadow. Naturally, which pixels are the image of shadowed or illuminated surface points depends on where we move the light source direction. If we move the light source all the way around to the back of the object so that the camera's optical axis and the light source are pointing in opposite directions, then all pixels are in shadow.

Let us now consider all possible light source directions, representing each direction by a point on the surface of the sphere; we call this sphere the *illumi-*nation sphere*. For a convex object, the set of light source directions for which a given pixel in the image is illuminated corresponds to an open hemisphere of the illumination sphere; the set of light source directions for which the pixel is shadowed corresponds to the other hemisphere of points. A great circle on the illumination sphere divides these sets.

Each of the n pixels in the image has a corresponding great circle on the illumination sphere. The collection of great circles carves up the surface of the illumination sphere into a collection of cells $\mathcal{S}_i$. See Figure 2. The collection of light source directions contained within a cell $\mathcal{S}_i$ on the illumination sphere produces a set of images, each with the same pixels in shadow and the same pixels illuminated; we say that these images have the same "shadowing configurations." Different cells produce different shadowing configurations.

We denote by $\mathcal{S}_0$ the cell on the illumination sphere containing the collection of light source directions which produce images with all pixels illuminated. Thus, the collection of light source directions from the interior and boundary of $\mathcal{S}_0$ produces the set of images $\mathcal{L}_0$. To determine the set of images produced by another cell on the illumination sphere, we need to return to the illumination subspace $\mathcal{L}$.

The illumination subspace $\mathcal{L}$ not only slices through the non-negative orthant of $\mathbb{R}^n$, but other orthants in $\mathbb{R}^n$ as well. Let $\mathcal{L}_i$ be the intersection of the illumination subspace $\mathcal{L}$ with an orthant i in $\mathbb{R}^n$ through which $\mathcal{L}$ passes. Certain components of $\mathbf{x} \in \mathcal{L}_i$ are always negative and others always greater than or equal to zero. Each $\mathcal{L}_i$ has a corresponding cell of light source directions $\mathcal{S}_i$ on the illumination sphere. Note that $\mathcal{L}$ does not slice through all of the 2^n orthants in $\mathbb{R}^n$, but at most $n(n-1) + 2$ orthants (see the proof of Proposition 1). Thus, there are at most $n(n-1) + 2$ sets $\mathcal{L}_i$, each with a corresponding cell $\mathcal{S}_i$ on the illumination sphere.

The set of images produced by the collection of light source directions from a cell $\mathcal{S}_i$ other than $\mathcal{S}_0$ can be found as a projection P_i of all points in a particular set $\mathcal{L}_i$. The projection P_i is such that it leaves the non-negative components of $\mathbf{x} \in \mathcal{L}_i$ untouched, while the negative components of $\mathbf{x}$ become zero. We denote the projected set by $P_i(\mathcal{L}_i)$.

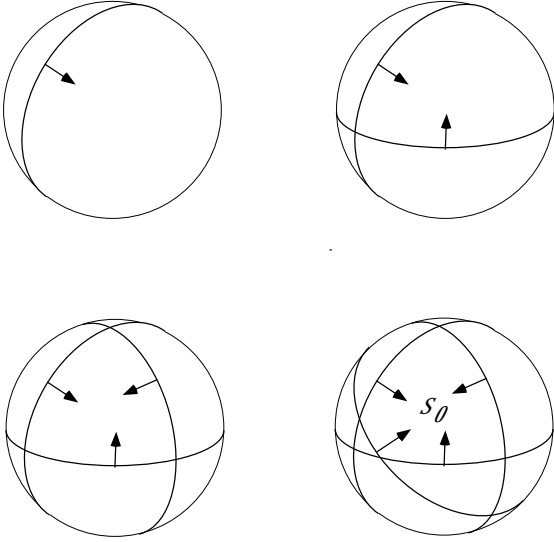**Lemma 2.** *The set of images $P_i(\mathcal{L}_i)$ is a convex cone in* $\mathbb{R}^n$.

Fig. 2. **The Illumination Sphere:** The set of all light source directions can be represented by points on the surface of a sphere; we call this sphere the *illumination sphere*. Great circles corresponding to individual pixels divide the illumination sphere into cells of different shadowing configurations. The arrows indicate the hemisphere of light directions for which the particular pixel is illuminated. The cell of light source directions which illuminate all pixels is denoted by $\mathcal{S}_0$. The light source directions within $\mathcal{S}_0$ produce $\mathcal{L}_0$ the set of images in which all pixels are illuminated. Each of the other cells produce the $\mathcal{L}_i$, $0 < i \leq n(n-1)+1$. The extreme rays of the cone are given by the images produced by light sources at the intersection of two circles.

Proof: By the same argument used in the proof of Lemma 1, $\mathcal{L}_i$ is a convex cone. Since the linear projection of a convex cone is itself a convex cone, $P_i(\mathcal{L}_i)$ is a convex cone.   □

Since $P_i(\mathcal{L}_i)$ is the projection of $\mathcal{L}_i$, it is at most three dimensional. Each $P_i(\mathcal{L}_i)$ is the set of all images such that certain pixels are illuminated, and the remaining pixels are shadowed. The dual relation between $P_i(\mathcal{L}_i)$ and $\mathcal{S}_i$ can be concisely written as $P_i(\mathcal{L}_i) = \{\alpha \max(B\mathbf{s}, \mathbf{0}) : \alpha \geq 0, \mathbf{s} \in \mathcal{S}_i\}$ and $\mathcal{S}_i = \{\mathbf{s} : |\mathbf{s}| = 1, \max(B\mathbf{s}, \mathbf{0}) \in P_i(\mathcal{L}_i)\}$. Let P_0 be the identity, so that $P_0(\mathcal{L}_0) = \mathcal{L}_0$ is the set of all images such that all pixels are illuminated. The number of possible *shadowing configurations* is the number of orthants in $\mathbb{R}^n$ through which the illumination subspace $\mathcal{L}$ passes, which in turn is the same as the number of sets $P_i(\mathcal{L}_i)$.

Proposition 1.   *The number of shadowing configurations is at most $m(m-1)+2$, where $m \leq n$ is the number of distinct surface normals.*

Proof: Each of the n pixels in the image has a corresponding great circle on the illumination sphere, but only m of the great circles are distinct. The collection of m distinct great circles carves up the surface of the illumination sphere into cells. Each cell on the illumination sphere corresponds to a particular set of images $P_i(\mathcal{L}_i)$. Thus, the problem of determining the number of shadowing configurations is the same as the problem of determining the number of cells on the illumination sphere. If every vertex on the illumination sphere is formed by the intersection of only two of the m distinct great circles (i.e., if no more than two surface normals are coplanar), then it can be shown by induction that the illumination sphere is divided into $m(m-1)+2$ cells. If a vertex is formed by the intersection of three or more great circles, there are fewer cells.   □

Thus, the set $\mathcal{U}$ of images of a convex Lambertian surface created by varying the direction and strength of a *single* point light source at infinity is given by the union of at most $n(n-1)+2$ convex cones, i.e.,

$$\mathcal{U} = \{\mathbf{x} \mid \mathbf{x} = \max(B\mathbf{s}, \mathbf{0}), \forall \mathbf{s} \in \mathbb{R}^3\}$$
$$= \bigcup_{i=0}^{n(n-1)+1} P_i(\mathcal{L}_i). \tag{2}$$

From this set, we can construct the set $\mathcal{C}$ of all possible images of a convex Lambertian surface created by varying the direction and strength of an *arbitrary number* of point light sources at infinity,

$$\mathcal{C} = \{\mathbf{x} : \mathbf{x} = \sum_{i=1}^{k} \max(B\mathbf{s}_i, \mathbf{0}), \forall \mathbf{s}_i \in \mathbb{R}^3, \forall k \in \mathbb{Z}^+\}$$

where $\mathbb{Z}^+$ is the set of positive integers.

Proposition 2.   *The set of images $\mathcal{C}$ is a convex cone in $\mathbb{R}^n$.*

Proof: The proof that $\mathcal{C}$ is a cone follows trivially from the definition of $\mathcal{C}$. To prove that $\mathcal{C}$ is convex, we appeal to a proposition for convex cones which states that a cone $\mathcal{C}$ is convex iff $\mathbf{x}_1 + \mathbf{x}_2 \in \mathcal{C}$ for any two points $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$ [6]. So the proof that $\mathcal{C}$ is convex also follows trivially from the above definition of $\mathcal{C}$.   □

We call $\mathcal{C}$ the *illumination cone*. Every object has its own illumination cone. Note that each point in the cone is an image of the object under a particular lighting configuration, and the entire cone is the set of images of the object under all possible configurations of point light sources at infinity.

Proposition 3.   *The illumination cone $\mathcal{C}$ of a convex Lambertian surface can be determined from as few as three images, each taken under a different, but unknown light source direction.*

Proof: The illumination cone $\mathcal{C}$ is completely determined by the illumination subspace $\mathcal{L}$. If the matrix of surface normals scaled by albedo B were known, then this would determine $\mathcal{L}$ uniquely, as $\mathcal{L} = \{\mathbf{x} \mid \mathbf{x} = B\mathbf{s}, \forall \mathbf{s} \in \mathbb{R}^3\}$. Yet, from images produced by differing, but unknown light source directions we can not determine B. To see this note that for any arbitrary invertible 3×3 linear transformation $A \in GL(3)$,

$$B\mathbf{s} = (BA)(A^{-1}\mathbf{s}) = B^*\mathbf{s}^*.$$

In other words, the same image is produced when the albedo and surface normals are transformed by A, while the light source is transformed by A^{-1}. Therefore, without knowledge of the light source directions, we can only recover B^* where $B^* = BA$, see [8, 13]. Nonetheless B^* is sufficient for determining the subspace $\mathcal{L}$: it is easy to show that $\mathcal{L} = \{\mathbf{x} \mid \mathbf{x} = B^*\mathbf{s}, \forall \mathbf{s} \in \mathbb{R}^3\} = \{\mathbf{x} \mid \mathbf{x} = B\mathbf{s}, \forall \mathbf{s} \in \mathbb{R}^3\}$, see [28]. □

Thus, for a convex object with Lambertian reflectance, we can determine its appearance under arbitrary illumination from as few as three images of the object – knowledge of the light source strength or direction is *not* needed, see also [28]. To determine the illumination cone $\mathcal{C}$, we simply need to determine the illumination subspace $\mathcal{L}$. In turn, we can choose any three images from the set $\mathcal{L}_0$, each taken under a different lighting direction, as its basis vectors. Naturally, if more images are available, they can be combined to find the best rank three approximation to $\mathcal{L}$ using singular value decomposition (SVD).

We should point out that for many convex surfaces the cone can be constructed from as few as three images; however, this is not always possible. If the object has surface normals covering the Gauss sphere, then there is only one light source direction – the viewing direction – such that the entire visible surface is illuminated. For any other light source direction, some portion of the surface is shadowed. To determine $\mathcal{L}$ each point on the surface of the object must be illuminated in at least three images; for this to be true over the entire visible surface, as many as five images may be required. See [15] for an algorithm for determining $\mathcal{L}$ from images with shadowed pixels.

What may not be immediately obvious is that any point within the cone $\mathcal{C}$ (including the boundary points) can be found as a convex combination of the rays (images) produced by light source directions lying at the $m(m-1)$ intersections of the great circles. Furthermore, because the cone is constructed from a finite number of extreme rays (images), the cone is polyhedral.

These propositions and observations suggest the following algorithm for constructing the illumination cone from three or more images:

Illumination Subspace Method: Gather images of the object under varying illumination without shadowing and use these images to estimate the three-dimensional illumination subspace $\mathcal{L}$. After normalizing the images to be of unit length, singular value decomposition (SVD) can be used to estimate the best orthogonal basis in a least squares sense. From the illumination subspace $\mathcal{L}$, the extreme rays defining the illumination cone $\mathcal{C}$ are then computed. Recall that an extreme ray is an image created by a light source direction lying at the intersection of two or more great circles. If there are m independent surface normals, there can be as many as $m(m-1)$ extreme rays (images). Let $\mathbf{b}_i$ and $\mathbf{b}_j$ be rows of B with $i \neq j$, the extreme rays are given by

$$\mathbf{x}_{ij} = \max(B\mathbf{s}_{ij}, \mathbf{0})$$

where

$$\mathbf{s}_{ij} = \mathbf{b}_i \times \mathbf{b}_j.$$

□

In Section 2.4, we use this method to experiment with images of real objects; we use a small number of images to build the illumination subspace $\mathcal{L}$ and then produce sample images from the illumination cone $\mathcal{C}$. To reduce space and computational requirements for applications using the cone, the images can be projected down to a low dimensional subspace; any image in

the projected cone can be found as convex combinations of the projected extreme rays. Note however, that some of the projected extreme rays are redundant since an extreme ray may project to the interior of the projected cone. As will be seen in the experiments of Section 3.4, the illumination cones of real objects do lie near a low dimensional subspace; thus dimensionality reduction by linear projection may be justified.

2.2. A Two-Dimensional Example

To illustrate the relationship between an object and its illumination cone, consider the simplified 2-D example in Fig. 3. An object composed of three facets is shown in Fig. 3.a. For facet i, the product of the albedo and surface normal is given by $\mathbf{b}_i \in \mathbb{R}^2$. In this 2-D world, the direction of a light source at infinity can be represented as a point on a circle.

Now, consider a camera observing the three facets from above such that the each facet projects to one pixel yielding an image $\mathbf{x} = (x_1, x_2, x_3)^t \in \mathbb{R}^3$. $\mathcal{L}$ is then a 2-D linear subspace of $\mathbb{R}^3$, and the set of images from a single light source such that all pixels are illuminated $\mathcal{L}_0 \in \mathcal{L}$ is the 2-D convex cone shown in Figure 3.b. The left edge of $\mathcal{L}_0$ where $x_3 = 0$ corresponds to the light source direction where Facet 3 just goes into shadow. Now, for a single light source, the set of images is formed by projecting $\mathcal{L}$ onto the positive orthant as shown in Figure 3.c. Note for example, that the 2-D cone $P_1(\mathcal{L}_1)$ corresponds to the set of images in which Facets 1 and 2 are illuminated while Facet 3 is in shadow, and the 1-D ray $P_3(\mathcal{L}_3)$ corresponds to the set of image with Facet 1 illuminated and Facets 2 and 3 shadowed. The union $\cup_{i=0}^{i=6} P_i(\mathcal{L}_i)$ defines the walls of the illumination cone $\mathcal{C}$, and the entire cone is formed by taking convex combinations of images on the walls.

As seen in Figure 3.d, the set of light source directions, represented here by a circle, can be partitioned into regions $\mathcal{S}_i$ such that all images produced by light sources within the region have the same shadowing configurations. That is, $\mathcal{S}_i = \{\mathbf{s} : |\mathbf{s}| = 1, \max(B\mathbf{s}, \mathbf{0}) \in P_i(\mathcal{L}_i)\}$. The corresponding partitioning of light source directions is shown in Figure 3.a.
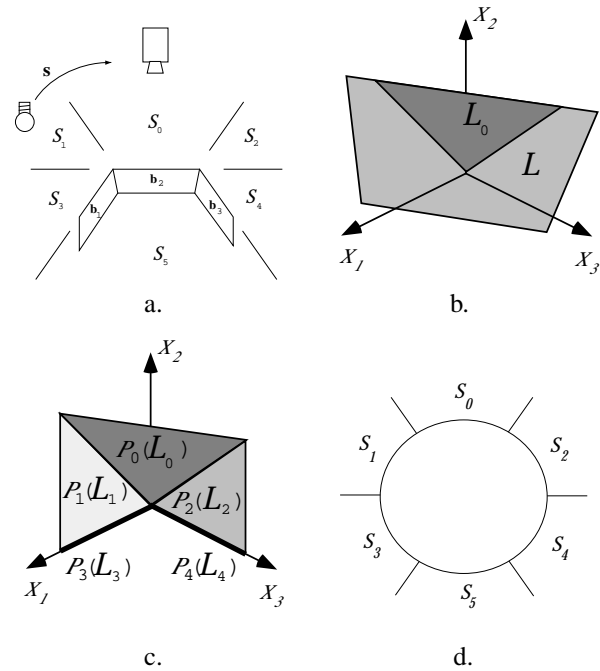

Fig. 3. **A 2-D Example:** a. A surface with three facets is observed from above and produces an image with pixels x_1, x_2 and x_3. b. The linear subspace $\mathcal{L}$ and its intersection with the positive quadrant $\mathcal{L}_0$. c. The "walls" of the cone $P_i(\mathcal{L}_i)$ corresponding to images formed by a single light source. The illumination cone $\mathcal{C}$ is formed by all convex combinations of images lying on the walls. d. The geometry of facets leads to a partitioning of the illumination circle.

2.3. Illumination Cones for Arbitrary Objects

In the previous subsection, we assumed that the objects were convex in shape and had Lambertian reflectance functions. Our central result was that the set of images of the object under all possible illumination conditions formed a convex cone in the image space and that this illumination cone can be constructed from as few as three images. Yet, most objects are nonconvex in shape and have reflectance functions which can be better approximated by more sophisticated physical [24, 29, 30] and phenomenological [17] models. The question again arises: What can we say about the set of images of an object with a nonconvex shape and a non-Lambertian reflectance function?

Nothing in the proof of Proposition 2 required assumptions about the shape of the object, the nature of the light sources, or the reflectance function for the object's surface. Consequently, we can state a more general proposition about the set of images of an object under varying illumination:

Proposition 4.   *The set of n-pixel images of any object, seen under all possible lighting conditions, is a convex cone in $\mathbb{R}^n$.*

Therefore, even for an arbitrary object (i.e., an object that does not have a convex shape or a Lambertian reflectance function), the set of images under all possible lighting conditions still forms convex cone in the image space. This result is in some sense trivial, arising from the superposition property of illumination: the image of an object produced by two light sources is simply the addition of the two images produced by the sources individually.

It is doubtful that the illumination cone for such objects can be constructed from as few as three images. This is not due to the nonconvexity of objects and the shadows they cast. The structure of objects with Lambertian reflectance, but nonconvex shapes, can be recovered up to a "generalized bas-relief" transformation from as few as three images [4]. From this, it is possible to determine the cast shadows exactly. Rather the difficulty is due to the fact that the reflectance function is unknown. To determine the reflectance function *exactly* could take an infinite number of images. However, the Illumination Subspace Method developed in Section 2.1 can be used to approximate the cone, as will be seen in the empirical investigation of Section 2.4. An alternative method for approximating the cone is presented below:

Sampling Method: Illuminate the object by a series of light source directions which evenly sample the illumination sphere. The resulting set of images are then used as the extreme rays of the approximate cone.   □

Note that this approximate cone is a subset of the true cone and so any image contained within the approximate cone is a valid image. The Sampling Method has its origins in the linear subspace method proposed by Hallinan [12]; yet, it differs in that the illumination cone restricts the images to be convex – not linear – combinations of the extreme rays. This method is a natural way of extending the appearance manifold method of Murase and Nayar to account for multiple light sources and shadowing [21].
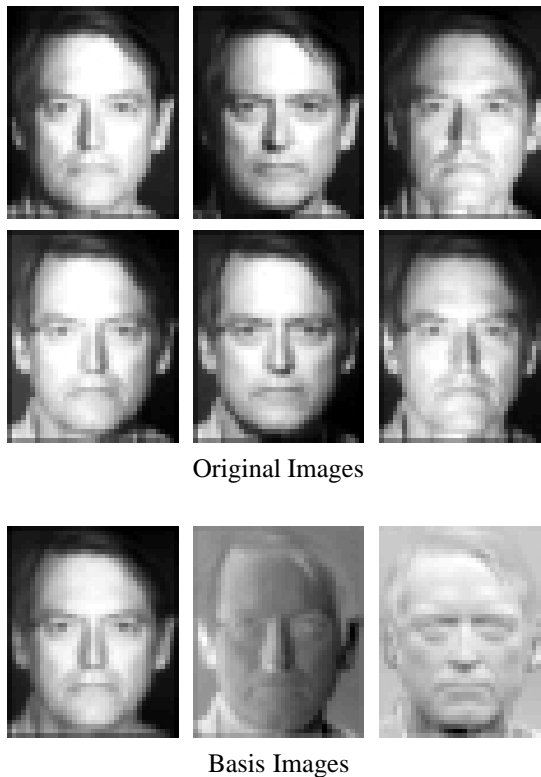

Original Images

Basis Images

Fig. 4.   **Illumination Subspace Method:** The top half of the figure shows all six of the original images used to construct the illumination subspace $\mathcal{L}$ of the face. The bottom half of the figure shows three basis images, lying in $\mathcal{L}_0$, that span the illumination subspace $\mathcal{L}$ for the face.

2.4. *Empirical Investigation: Building Illumination Cones*

To demonstrate the power of these concepts, we have used the Illumination Subspace Method to construct the illumination cone for two different scenes: a human face and a desktop still life. To construct the cone for the human face, we used images from the Harvard Face Database [12], a collection of images of faces seen under a range of lighting directions. For the purpose of this demonstration, we used the images of one person, taking six images with little shadowing and using singular value decomposition (SVD) to construct a 3-D basis for the illumination subspace $\mathcal{L}$. Note that this 3D linear subspace differs from the affine subspace constructed using the Karhunen-Loeve transform: the mean image is not subtracted before determining the basis vectors as in the Eigenpicture methods [18, 31].
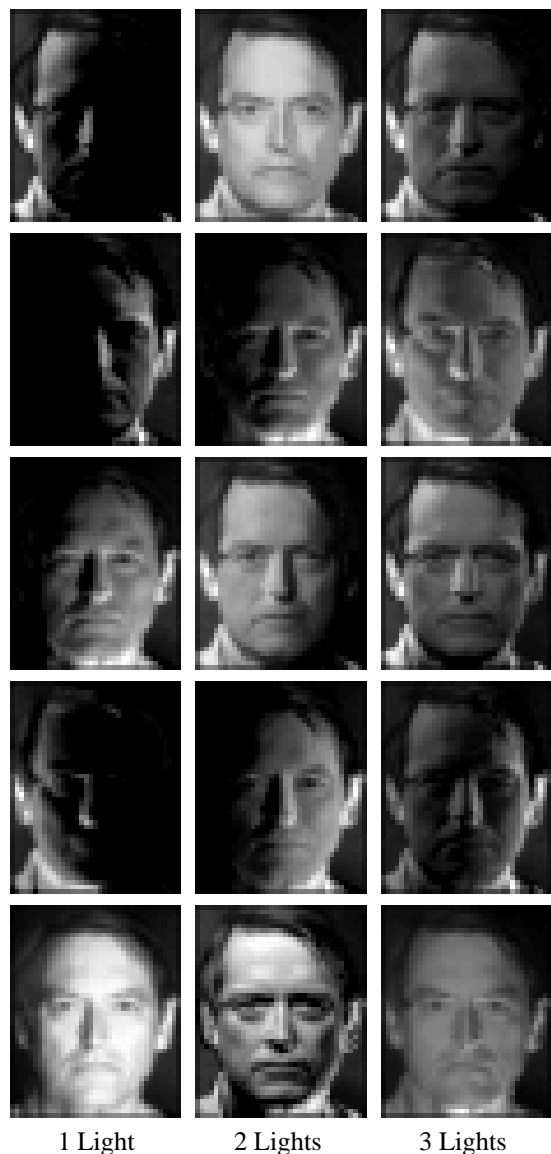
1 Light   2 Lights   3 Lights

Fig. 5. **Random Samples from the Illumination Cone of a Face:** Each of the three columns respectively comprises sample images from the illumination cone with one, two and three light sources.

The illumination subspace was then used to construct the illumination cone $\mathcal{C}$. We generated novel images of the face as if illuminated by one, two, or three point light sources by randomly sampling the illumination cone. Rather than constructing an explicit representation of the half-spaces bounding the illumination cone, we sampled $\mathcal{L}$, determined the corresponding orthant, and appropriately projected the im-

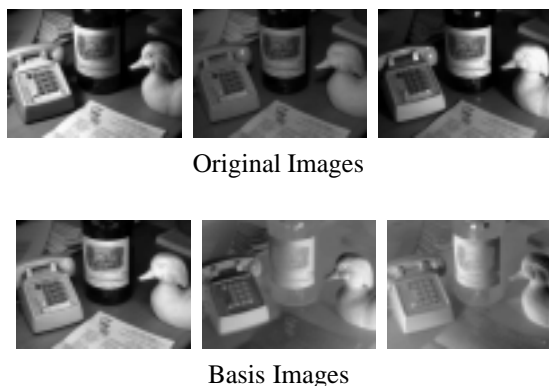Original Images

Basis Images

Fig. 6. **Illumination Subspace Method:** The top half of the figure shows three of the original nine images used to construct the illumination subspace $\mathcal{L}$ of the still life. The bottom half of the figure shows three basis images, lying in $\mathcal{L}_0$, that span the illumination subspace $\mathcal{L}$ for the still life.

age onto the illumination cone. Images constructed under multiple light sources simply correspond to the superposition of the images generated by each of the light sources.

The top half of Fig. 4 shows all six low resolution images of a person's face that were used to construct the basis of the linear subspace $\mathcal{L}$. The bottom half of Fig. 4 shows three basis images that span $\mathcal{L}$. The three columns of Fig. 5 respectively comprise sample images from the illumination cone for the face with one, two, or three light sources.

There are number of points to note about this experiment. There was almost no shadowing in the training images yet there are strong attached shadows in many of the sample images. They are particularly distinct in the images generated with a single light source. Notice for example the sharp shadow across the ridge of the nose in Column 1, Row 2 or the shadowing in Column 1, Row 4 where the light source is coming from behind the head. Notice also the depression under the cheekbones in Column 2, Row 5, and the cleft in the chin revealed in Column 1, Row 3. For the image in Column 3, Row 2, two of the light sources are on opposite sides while the third one is coming from below; notice that both ears and the bottom of the chin and nose are brightly illuminated while that rest of the face is darker.

To construct the cone for the desktop still life, we used our own collection of nine images with little shadowing. The top half of Fig. 6 shows three of these images. The bottom half of Fig. 6 shows the three ba-

Fig. 7. **Random Samples from the Illumination Cone of a Desktop Still Life:** Each of the three columns respectively comprises sample images from the illumination cone with one, two, and three light sources.

sis images that span $\mathcal{L}$. The three columns of Fig. 7 respectively comprise sample images from the illumination cone for the desktop still life with one, two, or three light sources.

The variability in illumination in these images is so extreme that the edge maps for these images would differ drastically. Notice in the image in Column 1, Row 4 that the shadow line on the bottle is distinct and that the left sides of the phone, duck, and bottle are brightly illuminated. Throughout the scene, notice that those points having comparable surface normals seem to be similarly illuminated. Furthermore, notice that all of the nearly horizontal surfaces in the bottom two images of the first column are in shadow since the light is coming from below. In the image with two light sources shown at the bottom of Column 2, the sources are located on opposite sides and behind the objects. This leads to a shadow line in the center of the bottle. The head of the wood duck shows a similar shadowing where the front and back are illuminated, but not the side.

3. Dimension and Shape of the Illumination Cone

In this section, we investigate the dimension of the illumination cone, and show that it is equal to the number of distinct surface normals. However, we conjecture that the shape of the cone is flat, with much of its volume concentrated near a low-dimensional subspace, and present empirical evidence to support this conjecture. Finally, we show that the cones of two objects with the same geometry, but with separate albedo patterns, differ by a diagonal linear transformation.

3.1. The Dimension of the Illumination Cone

Given that the set of images of an object under variation in illumination is a convex cone, it is natural to ask: What is the dimension of the cone in $\mathbb{R}^n$? By this we mean, what is the span of the vectors in the illumination cone $\mathcal{C}$? Why do we want to know the answer to this question? Because the complexity of the cone, may dictate the nature of the recognition algorithm. For example, if the illumination cones are 1-D, i.e., rays in the positive orthant of $\mathbb{R}^n$, then a recognition scheme based on normalized correlation could handle all of the variation due to illumination. However, in general the cones are not one dimensional unless the object is planar. To this end, we offer the following proposition.

Proposition 5. *The dimension of the illumination cone $\mathcal{C}$ is equal to the number of distinct surface normals.*

Proof: As for the proof of Proposition 1, we again represent each light source direction by a point on the surface of the illumination sphere. Each cell on the illumination sphere corresponds to the light source directions which produce a particular set of images $P_i(\mathcal{L}_i)$. For every image in a set $P_i(\mathcal{L}_i)$ certain pixels always equal zero, i.e., always in shadow. There exists a cell $\mathcal{S}_0$ on the illumination sphere corresponding to the light source directions which produce $\mathcal{L}_0$, the set of images in which all pixels are always illuminated. There exists a cell $\mathcal{S}_d$ corresponding to the light source directions which produce a set of images in which all pixels are always in shadow. Choose any point $\mathbf{s}_b \in \mathcal{S}_0$. The point $\mathbf{s}_d = -\mathbf{s}_b$ is antipodal to $\mathbf{s}_b$ and lies within $\mathcal{S}_d$. Draw any half-meridian connecting $\mathbf{s}_b$ and $\mathbf{s}_d$. Starting at $\mathbf{s}_b$, follow the path of the half-meridian; it crosses m distinct great circles, and

passes through m different cells before entering $\mathcal{S}_d$. Note the path of the half-meridian corresponds to a particular path of light source directions, starting from a light source direction producing an image in which all pixels are illuminated and ending at a light source direction producing an image in which all pixels are in shadow. Each time the half-meridian crosses a great circle, the pixel corresponding to the great circle becomes shadowed.

Take an image produced from any light source direction within the interior of each cell through which the meridian passes, including $\mathcal{S}_0$, but excluding $\mathcal{S}_d$. Arrange each of these m images as column vectors in an $n \times m$ matrix M. By elementary row operations, the matrix M can be converted to its echelon form M^*, and it is trivial to show that M^* has exactly m non-zero rows. Thus, the rank of M is m, and the dimension of $\mathcal{C}$ is at least m. Since there are only m distinct surface normals, the dimension of $\mathcal{C}$ cannot exceed m. Thus, the dimension of $\mathcal{C}$ equals m.
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Note that for images with n pixels, this proposition indicates that the dimension of the illumination cone is one for a planar object, is roughly $\sqrt{n}$ for a cylindrical object, and is n for a spherical object. But if the cone spans $\mathbb{R}^n$, what fraction of the positive orthant does it occupy? In Section 3.3, we investigate this question, conjecturing that the illumination cones for most objects occupy little volume in the image space.

3.2.   The Connection between Albedo and Cone Shape

If two objects are similar in geometry, but differ in their respective albedo patterns, then there is a simple linear relationship between their corresponding illumination cones. Here, we consider two Lambertian objects that have the same underlying geometry, but have differing albedo patterns (e.g., a Coke can and a Pepsi can). In this case, the product of albedo and surface normals for the two objects can be expressed as $B_1 = R_1 N$ and $B_2 = R_2 N$ where N is an $n \times 3$ matrix of surface normals and R_i is an $n \times n$ diagonal matrix whose diagonal elements are positive and represent the albedo. The following proposition relates the illumination cones of the two objects.

Proposition 6.   *If $\mathcal{C}_1$ is the illumination cone for an object defined by $B_1 = R_1 N$ and $\mathcal{C}_2$ is the illumina-*

tion cone for an object defined by $B_2 = R_2 N$, then

$$\mathcal{C}_1 = \{R_1 R_2^{-1}\mathbf{x} : \mathbf{x} \in \mathcal{C}_2\} \qquad and$$
$$\mathcal{C}_2 = \{R_2 R_1^{-1}\mathbf{x} : \mathbf{x} \in \mathcal{C}_1\}.$$

Proof: For every light source direction $\mathbf{s}$, the corresponding images are given by $\mathbf{x}_1 = \max(B_1\mathbf{s}, \mathbf{0}) = R_1 \max(N\mathbf{s}, \mathbf{0})$ and $\mathbf{x}_2 = \max(B_2\mathbf{s}, \mathbf{0}) = R_2 \max(N\mathbf{s}, \mathbf{0})$. Since R_1 and R_2 are diagonal with positive diagonal elements, they are invertible. Therefore, $\mathbf{x}_1 = R_1 R_2^{-1}\mathbf{x}_2$ and $\mathbf{x}_2 = R_2 R_1^{-1}\mathbf{x}_1$. $\square$

Thus, the cones for two objects with identical geometry but differing albedo patterns differ by a diagonal linear transformation. This fact can be applied when computing cones for objects observed by color cameras as noted in Section 4. Note that this proposition also holds when the objects are nonconvex; since the partitioning of the illumination sphere is determined by the objects' surface geometry, the set of shadowing configurations is identical for two objects with same shape. The intensities of the illuminated pixels are related by the transformations given in the proposition.

3.3.   Shape of the Illumination Cone

While we have shown that an illumination cone is a convex, polyhedral cone that can span n dimensions if there are n distinct surface normals, we have not said how big it is in practice. Note that having a cone span n dimensions does not mean that it covers $\mathbb{R}^n$, since a convex cone is defined only by convex combinations of its extreme rays. It is conceivable that an illumination cone could completely cover the positive orthant of $\mathbb{R}^n$. However, the existence of an object geometry that would produce this is unlikely. For such an object, it must be possible to choose n light source direction such that each of the n facets are illuminated independently.

On the other hand, if the illumination cones for objects are small and well separated, then recognition should be possible, even under extreme lighting conditions. We believe that the latter is true – that the cone has almost no volume in the image space. We offer the following conjecture:

Conjecture 1.   *The shape of the cone is "flat," i.e., most of its volume is concentrated near a low-dimensional subspace.*

While we have yet to prove this conjecture, the empirical investigations of [7, 12] and the one in the following section seem to support it.

3.4.   Empirical Investigation of the Shape of the Illumination Cones

To investigate Proposition 1 and Conjecture 1, we have gathered several images, taken under varying lighting conditions, of two objects: the corner of a cardboard box and a Wilson tennis ball. For both objects, we computed the illumination subspace using SVD from a set of images. Using the estimated illumination subspaces, we performed two experiments.

In the first experiment, we tried to confirm that the illumination spheres for both objects would appear as we would expect. For both the box and the tennis ball, we drew the great circles associated with each pixel on the illumination sphere, see Fig. 8. From Proposition 5, we would expect the illumination cone produced by the corner of the box to be three dimensional since the corner has only three faces. The illumination sphere should be partitioned into eight regions by three great circles, each meeting the other two orthogonally. This structure is evident in the figure. Yet, due to both image noise and the fact that the surface is not truly Lambertian, there is some small deviation of the great circles. Furthermore, the few pixels from the edge and corner of the box produce a few stray great circles. In contrast, the visible surface normals of the tennis ball should nearly cover half of the Gauss sphere and, therefore, the great circles should nearly cover the illumination sphere. Again, this structure is evident in the figure.

In the second experiment, we plotted the eigenvalues of the matrix of extreme rays for both the box and the tennis ball. The point of this experiment was to compare the size and "flatness" of both cones. As discussed in Section 2.1, an extreme ray $\mathbf{x}_{ij}$ is an image created by a light source direction $\mathbf{s}_{ij}$ lying at the intersection of two or more great circles on the illumination sphere. The matrix of extreme rays is simply the matrix whose columns are the vectorized images $\mathbf{x}_{ij}/|\mathbf{x}_{ij}|$. We then performed SVD on the matrix of extreme rays for the box corner and the matrix of extreme rays for the tennis ball. The corresponding eigenvalues are plotted in decreasing order in Fig. 9.

From this figure we make the following observations. First, in the plot of the box corner there is a
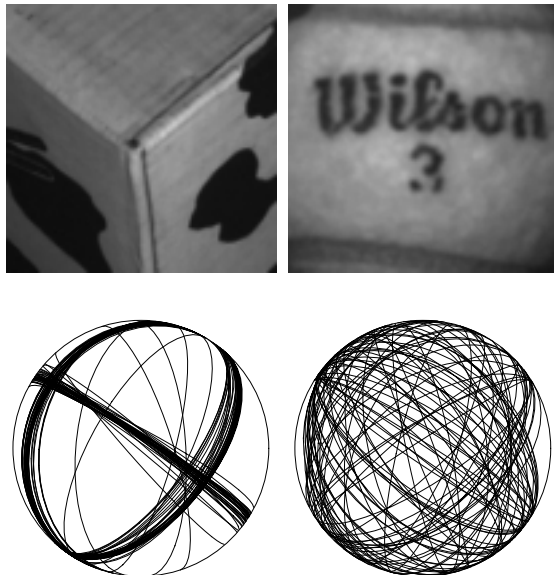

Fig. 8. **Examples of Illumination Spheres:** On the left, the figure shows an image of the corner of a cardboard box and its corresponding illumination sphere. Note that the illumination sphere is, for the most part, partitioned into eight regions by three great circles, each meeting the other two orthogonally. On the right, the figure shows an image of a Wilson tennis ball and its corresponding illumination sphere. Note that the great circles nearly cover the illumination sphere.

sharp drop-off after the third eigenvalue, indicating that most of the illumination cone is concentrated near a 3-D subspace of the image space. Second, the eigenvalues for the tennis ball do not drop-off as quickly as those for the box, indicating that the illumination cone for the tennis ball is larger than that for the box. And, third, the eigenvalues for both the box corner and the tennis ball diminish by at least two orders of magnitude within the first fifteen eigenvalues. Thus, in agreement with the above conjecture, the illumination cones appear to be concentrated near a low dimensional subspace.

We should point out that [7, 12] performed a related experiment on images created by *physically* moving the light source to evenly sample the illumination sphere. They too found that the set of images of an object under variable illumination lies near a low dimensional subspace. Our results using synthesized images from the cone seem to complement their findings.

4.   Color

Until now, we have neglected the spectral distribution of the light sources, the color of the surface, and the

Plot of Box Eigenvalues for Matrix of Extreme Rays

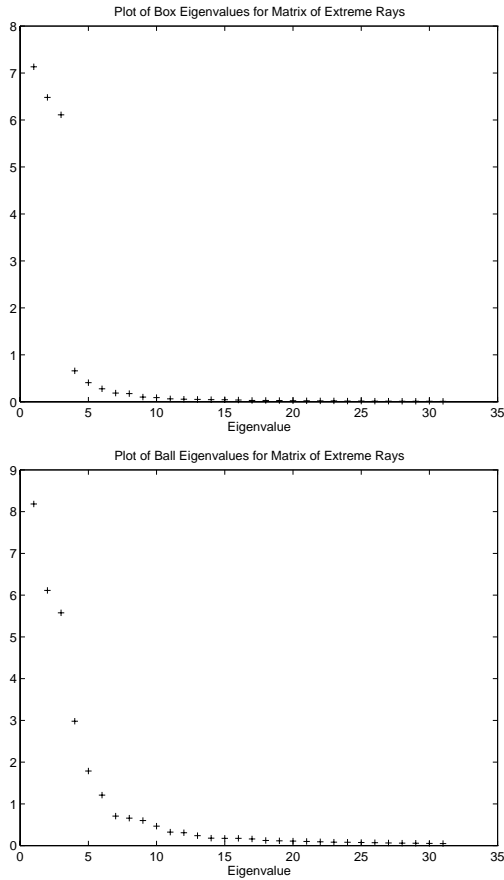Plot of Ball Eigenvalues for Matrix of Extreme Rays

Fig. 9. **Eigenvalues for the Matrix of Extreme Rays:** The figure shows a plot in decreasing order of the eigenvalues of the matrix of extreme rays for the illumination cone of the corner of a box and for the illumination cone of a tennis ball.

spectral response of the camera; here we extend the results of Section 2 to multi-spectral images.

Let λ denote the wavelength of light. Let $\rho_i(\lambda)$ denote the response for all elements of the ith color channel. Let $R(\lambda)$ be a diagonal matrix whose elements are the spectral reflectance functions of the facets, where the rows of $N \in \mathbb{R}^{n \times 3}$ are the surface normals of the facets. Finally, let $\tilde{s}(\lambda)$ and $\hat{s}$ be the power spectrum and direction of the light source, respectively. Then, ignoring attached shadows and the associated max operation, the n-pixel image $\mathbf{x}_i$ produced by color channel i of a convex Lambertian surface from a single colored light is [14, 16]

$$\mathbf{x}_i = \int \rho_i(\lambda)(R(\lambda)N)(\tilde{s}(\lambda)\hat{s})d\lambda. \qquad (3)$$

It is difficult to make limiting statements about the set of possible images of a colored object when $\rho(\lambda)$,

$R(\lambda)$ and $\tilde{s}(\lambda)$ are arbitrary. For example, if we consider a particular object with a spectral reflectance function $R(\lambda)$ and surface normals N, then without constraining assumptions on $\rho_i(\lambda)$ and $\tilde{s}(\lambda)$, any image x_i is obtainable. Consequently, we will consider two specific cases: cameras with narrow-band spectral response and light sources with identical spectral distributions.

4.1. Narrow-Band Cameras

Following [23], if the sensing elements in each color channel have narrow-band spectral response or can be made to appear narrow band [9], then $\rho_i(\lambda)$ can be approximated by a Dirac delta function about some wavelength λ_i, and Eq. 3 can be rewritten as

$$\begin{aligned} \mathbf{x}_i &= \rho(\lambda_i)(R(\lambda_i)N)(\tilde{s}(\lambda_i)\hat{s}) \\ &= \rho_i(R_iN)(\tilde{s}_i\hat{s}). \end{aligned} \qquad (4)$$

Note that ρ_i, R_i and N are constants for a given surface and camera whereas $\tilde{s}_i$ and $\hat{s}$ depend on properties of the light source. Eq. 4 can be expressed using the notation of Eq. 1 where $B = \rho_i R_i N$ and $\mathbf{s} = \tilde{s}_i\hat{s}$. The diagonal elements of $\rho_i R_i$ are the effective albedo of the facets for color channel i. For c narrow-band color channels, the color image $\mathbf{x} = [\mathbf{x}_1^t \mid \mathbf{x}_2^t \mid \cdots \mid \mathbf{x}_c^t]^t$ formed by stacking up the c images for each channel can be considered a point in $\mathbb{R}^{c\,n}$. Under a single light source, $\mathbf{x}$ is a function of $\hat{s}$ and $\tilde{s}_1 \cdots \tilde{s}_c$. Taken over all light source directions and spectral distributions, the set of images from a single light source without shadowing is a $c + 2$ dimensional manifold in $\mathbb{R}^{c\,n}$. It is easy to show that this manifold is embedded in a $3c$-dimensional linear subspace of $\mathbb{R}^{c\,n}$, and that any point (image) in the intersection of this linear subspace with the positive orthant of $\mathbb{R}^{c\,n}$ can be achieved by three colored light sources.

A basis for this $3c$-dimensional subspace can be constructed from three color images without shadowing. This is equivalent to independently constructing c three-dimensional linear subspaces in $\mathbb{R}^n$, one for each color channel. Note that $\rho_i R_i N$ spans subspace i. When attached shadows are considered, an illumination cone can be constructed in $\mathbb{R}^n$ for each color channel independently. The cones for each color channel are closely related since they arise from the same surface; effectively the albedo matrix R_i may be different for each color channel, but the surface normals N are the same. As demonstrated in Section 3.2, the

cones for two surfaces with same geometry, but different albedo patterns differ by a diagonal linear transformation. Now the set of all multi-spectral images of a convex Lambertian surface is a convex polyhedral cone in $\mathbb{R}^{cn}$ given by the Cartesian product of the c individual cones. Following Proposition 5, this color cone spans at most cm dimensions where m is the number of distinct surface normals.

4.2. *Light Sources with Identical Spectra*

Consider another imaging situation in which a color camera (c channels, not necessarily narrow-band) observes a scene where the number and location of the light sources are unknown, but the power spectral distributions of all light sources are identical (e.g., incandescent bulbs). Equation 3 can then be rewritten as

$$\mathbf{x}_i = \left(\int \rho_i(\lambda)\tilde{s}(\lambda)R(\lambda)d\lambda \right) N\hat{\mathbf{s}}. \qquad (5)$$

In this case, the integral is independent of the light source direction and scales with its intensity. If we define the intensity of the light source to be $\tilde{s} = \int \tilde{s}(\lambda)d\lambda$, then $\mathbf{s} = \tilde{s}\hat{\mathbf{s}}$ and $R_i = \frac{1}{s}\int \rho_i(\lambda)\tilde{s}(\lambda)R(\lambda)d\lambda$. Equation 5 can then be expressed as

$$\mathbf{x}_i = R_i N\mathbf{s}.$$

For c color channels, the color image $\mathbf{x} \in \mathbb{R}^{cn}$ is given by

$$\mathbf{x} = [R_1 \,|\, R_2 \,|\, \cdots \,|\, R_c]^t \, N\mathbf{s}.$$

Consequently, the set of images of the surface without shadowing is a three-dimensional linear subspace of $\mathbb{R}^{cn}$ since R_i and N are constants. Following Section 2, the set of all images with shadowing is a convex polyhedral cone that spans m dimensions of $\mathbb{R}^{cn}$. Thus, when the light sources have identical power spectra (even if the camera is not narrow-band), the set of all images is significantly smaller than considered above since the color measured at each pixel is independent of the light source direction.

5. Discussion

In this paper, we have shown that the set of images of a convex object with a Lambertian reflectance func-

tion, under all possible lighting conditions at infinity, is a convex, polyhedral cone. Furthermore, we have shown that this cone can be learned from three properly chosen images and that the dimension of the cone equals the number of distinct surface normals. We have shown that for objects with an arbitrary reflectance function and a nonconvex shape, the set of images is still a convex cone and that these results can be easily extended to color images.

Nevertheless, there remain a number of extensions and open issues which we discuss below. While we have focused this paper solely on *characterizing* the set of images under varying illumination, we believe the ideas presented within have natural applications to recognition algorithms.

5.1. *Interreflection*

A surface is not just illuminated by the light sources but also through interreflections from points on the surface itself [1, 11]. For a Lambertian surface, the image with interreflection $\mathbf{x}'$ is related to the image that would be formed without interreflection $\mathbf{x}$ by

$$\mathbf{x}' = (I - RK)^{-1}\mathbf{x}$$

where I is the identity matrix, R is a diagonal matrix whose diagonal elements denote the albedo of facet i, and K is known as the interreflection kernel [22]. When there is no shadowing, all images lie in a 3-D linear space that would be generated from Eq. 1 by a pseudo-surface whose normals and albedo B' are given by $B' = (I - RK)^{-1}B$ [22, 23]. From Proposition 4, the set of all possible images is still a cone. While B' can be learned from only three images, the set of shadowing configurations and the partitioning of the illumination sphere is generated from B, not B'. So, it remains an open question how the cone can be constructed from only three images.

5.2. *Effects of Change in Pose*

All of the previous analysis in the paper has dealt solely with variation in illumination. Yet, a change in the object's pose creates a change in the perceived image. If an object undergoes a rotation or translation, how does the illumination cone deform? The illumination cone of the object in the new pose is also convex, but almost certainly different from the illumi-

nation cone of the object in the old pose. Which raises the question: Is there a simple transformation, obtainable from a small number of images of the object seen from different views, which when applied to the illumination cone characterizes these changes? Alternatively, is it practical to simply sample the pose space constructing an illumination cone for each pose? Nayar and Murase have extended their appearance manifold representation to model illumination variation for each pose as a 3-D linear subspace [23]. However, their representation does not account for the complications produced by attached shadows.

5.3. Object Recognition

Ultimately, we intend to apply the illumination cone concept to recognition. In earlier face recognition work, we implemented a classification method based on the minimum distance of an image to a three dimensional linear subspace (i.e., $\mathcal{L}$) that was constructed for each person. Experimentally, recognition rates were perfect when there was only moderate variation in illumination with minor shadowing [2]. While the distance to illumination subspace algorithm performed impressively, the experiments also indicated that its performance deteriorates as the test images move further and further away from the illumination subspace $\mathcal{L}$. We believe that nearly perfect recognition rates could be achieved under extreme variation in illumination, by measuring distance to the illumination cone, rather than distance to the illumination subspace. Measuring the squared distance to a convex cone can be posed as a non-negative least-squares optimization problem.

It is important to stress that the illumination cones are convex. If they are non-intersecting, then the cones are linearly separable. That is, they can be separated by a $n - 1$ dimensional hyperplane in $\mathbb{R}^n$ passing through the origin. Furthermore since convex sets remain convex under linear projection, then for any projection direction lying in the separating hyperplane, the projected convex sets will also be linearly separable. For d different objects represented by d linearly separable convex cones, there always exists a linear projection of the image space to a $d - 1$ dimensional space such that all of the projected sets are again linearly separable. So an alternative to classification based on measuring distance to the cones in $\mathbb{R}^n$ is to find a much lower dimensional space in which to do classification. In our Fisherface method for recognizing faces under variable illumination and facial expression, projection directions were chosen to maximize separability of the object classes [2]; a similar approach can be taken here.

Notes

1. By orthant we mean the high-dimensional analogue to quadrant, i.e., the set $\{ \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n$, with certain components of $\mathbf{x} \geq 0$ and the remaining components of $\mathbf{x} < 0 \}$. By non-negative orthant we mean the set $\{ \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n$, with all components of $\mathbf{x} \geq 0 \}$.

References

1. R. Bajcsy, S. Lee, and A. Leonardis. Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation. *Int. J. Computer Vision*, 17(3):241–272, March 1996.

2. P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. In *European Conf. on Computer Vision*, pages 45–58, 1996.

3. P. Belhumeur and D. Kriegman. What is the set of images of an object under all possible lighting conditions. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 270–277, 1996.

4. P. Belhumeur, D. Kriegman, and A. Yuille. The bas-relief ambiguity. In *IEEE Proc. Conf. Computer Vision and Pattern Recognition*, 1997.

5. T. Binford. Generic surface interpretation: Observability model. In *Proc. of the 4th International Symposium on Robotics Research*, Santa Cruz, CA, August 1987.

6. M. Canon, C. Cullum Jr., and E. Polak. *Theory of Optimal Control and Mathematical Programming*. McGraw-Hill, New York, 1970.

7. R. Epstein, P. Hallinan, and A. Yuille. 5 ± 2 Eigenimages suffice: An empirical investigation of low-dimensional lighting models. Technical Report 94-11, Harvard University, 1994.

8. R. Epstein, A. Yuille, and P. N. Belhumeur. Learning and recognizing objects using illumination subspaces. In *Proc. of the Int. Workshop on Object Representation for Computer Vision*, 1996.

9. G. Finlayson, M. Drew, and B. Funt. Spectral sharpening: Sensor transformations for improved color constancy. *J. Opt. Soc. Am. A*, 11:1553–1563, 1994.

10. D. Forsyth and M. Fleck. Body plans. In *IEEE Proc. Conf. Computer Vision and Pattern Recognition*, 1997.

11. D. Forsyth and A. Zisserman. Reflections on shading. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 13(7):671–679, 1991.

12. P. Hallinan. A low-dimensional representation of human faces for arbitrary lighting conditions. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 995–999, 1994.

13. H. Hayakawa. Photometric stereo under a light-source with arbitrary motion. *JOSA-A*, 11(11):3079–3089, Nov. 1994.

14. B. Horn. *Computer Vision*. MIT Press, Cambridge, Mass., 1986.

15. D. Jacobs. Linear fitting with missing data: Applications to structure from motion and characterizing intensity images. In *IEEE Proc. Conf. Computer Vision and Pattern Recognition*, 1997.

16. G. Klinker, S. Shafer, and T. Kanade. Image segmentation and reflection analysis through color. *Int. J. Computer Vision*, 2(1):7–32, June 1988.

17. J. Koenderink and A. vanDoorn. Bidirectional reflection distribution function expressed in terms of surface scattering modes. In *European Conf. on Computer Vision*, pages II:28–39, 1996.

18. L. Sirovitch and M. Kirby. Low-dimensional procedure for the characterization of human faces. *J. Optical Soc. of America A*, 2:519–524, 1987.

19. M. Langer and S. Zucker. A ray-based computational model of light sources and illumination. In *Physics Based Modeling Workshop in Computer Vision*, page SESSION 4, 1995.

20. Y. Moses, Y. Adini, and S. Ullman. Face recognition: The problem of compensating for changes in illumination direction. In *European Conf. on Computer Vision*, pages 286–296, 1994.

21. H. Murase and S. Nayar. Visual learning and recognition of 3-D objects from appearence. *Int. J. Computer Vision*, 14(5–24), 1995.

22. S. Nayar, K. Ikeuchi, and T. Kanade. Shape from interreflections. *IJCV*, 6(3):173–195, August 1991.

23. S. Nayar and H. Murase. Dimensionality of illumination in appearance matching. *IEEE Conf. on Robotics and Automation*, 1996.

24. M. Oren and S. Nayar. Generalization of the Lambertian model and implications for machine vision. *Int. J. Computer Vision*, 14:227–251, 1996.

25. A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 84–91, 1994.

26. T. Poggio and K. Sung. Example-based learning for view-based human face detection. In *Proc. Image Understanding Workshop*, pages II:843–850, 1994.

27. R. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970.

28. A. Shashua. *Geometry and Photometry in 3D Visual Recognition*. PhD thesis, MIT, 1992.

29. H. Tagare and R. deFigueiredo. A framework for the construction of reflectance maps for machine vision. *Comp. Vision, Graphics, and Image Proces.*, 57(3):265–282, May 1993.

30. K. Torrance and E. Sparrow. Theory for off-specular reflection from roughened surfaces. *JOSA*, 57:1105–1114, 1967.

31. M. Turk and A. Pentland. Eigenfaces for recognition. *J. of Cognitive Neuroscience*, 3(1), 1991.

32. S. Ullman and R. Basri. Recognition by a linear combination of models. A.I. Memo 1152, MIT, Aug. 1989.