PROOF COPY [A-8466] 002301JOA

KGBR Viewpoint–Lighting Ambiguity

Alan Yuille, James M. Coughlan, and S. Konishi

Smith-Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, CA 94115

Received April 8, 2002; revised manuscript received July 23, 2002; accepted July 26, 2002

We address the visual ambiguities that arise in estimating object and scene structure from a set of images when the viewpoint and lighting are unknown. We obtain a novel viewpoint–lighting ambiguity called the KGBR that corresponds to a group of three-dimensional affine transformations on the object or scene geometry combined with transformations on the object or scene albedo. Our analysis assumes orthographic projection with an affine camera model. We include photometric cues, such as shadowing and shading, that we model using Lambertian reflectance functions with shadows (cast and attached) and multiple light sources (but no interreflections). We relate the KGBR to affine ambiguities in estimating shape and to the generalized bas-relief ambiguity. © 2003 Optical Society of America

OCIS codes: 150.0150, 100.0100.

PROOF COPY [A-8466] 002301JOA

1. INTRODUCTION

When is it possible to estimate object and scene structure from a series of images? (In the rest of this paper we will usually discuss only objects but with the understanding that all our results apply to scenes composed of many objects provided that we ignore intereflections.) This issue is of practical importance particularly in situations where the viewpoints and lighting conditions are unknown. It arises in visual tasks such as binocular stereo, structure from motion, structure from viewpoint, and reconstruction of a scene from a set of images. Phenomena such as the convex-versus-concave ambiguity and the bas-relief ambiguity (see Fig. 1) show that human observers often do not estimate the correct shape. Indeed, human observers seem able to estimate shape only up to an affine transform[1] (see Fig. 2).

This issue has been studied in detail in cases where the viewpoint changes but the shading and shadow cues are ignored. The studies are based on pointlike features, corresponding to the output of feature detectors, which are assumed to be independent of the lighting conditions. In a classic paper, Koenderink and van Doorn[2] showed how structure from motion could be obtained up to a three-dimensional affine ambiguity. Faugeras[3] then obtained a hierarchy of ambiguities depending on the knowledge of the viewpoints and the calibration of the cameras. The geometric ambiguities of scene reconstruction from multiple viewpoints are now well understood, and there are calibration strategies for disambiguating them.[4]

Ambiguities have also been found for the opposite case where shading and shadow cues are present but the viewpoint is fixed (e.g., photometric stereo). This topic has received less attention, but recent work by Belhumeur *et al.*[5] shows that there is a generalized bas-relief (GBR) ambiguity (see Fig. 3) in the estimation of shape from multiple images with unknown lighting and fixed viewpoint. This ambiguity corresponds to a group of transformations that act on both the geometry and the albedo of objects. The GBR ambiguity assumes that objects have Lambertian reflectance functions but allows for shadows (cast and attached) and multiple light sources (but no intereflections). It includes the convex versus concave ambiguity and the bas-relief ambiguity as special cases and is of practical importance for photometric stereo.[6] See also Ref. 7 for other work on photometric ambiguities.

In this paper we propose that these two strands of work should be combined by studying the ambiguities when both the viewpoint and the lighting are unknown. In particular, we were motivated by the need to understand how the GBR photometric ambiguity[5] related to the more standard geometric ambiguities.[4]

Our main technical result is the KGBR viewpoint and lighting ambiguity (see also Ref. 8). This viewpoint makes the same photometric assumptions used for the GBR.[5] It consists of an affine transformation on the geometry of objects, as occurs in pure geometric ambiguities,[2–4] in conjunction with a transformation on the albedo. We assume that the imaging is performed by orthographic projection that allows for two-dimensional affine warps in the image plane[9] or, alternatively, by affine cameras. By restricting the KGBR to a single viewpoint, we obtain a novel and intuitive derivation of the GBR ambiguity. More generally, we show that any KGBR transformation can be decomposed as a GBR transformation, a rotation, and a two-dimensional affine transformation.

We study the effect of the KGBR ambiguity on specific vision problems involving multiple images. Suppose that the object and lighting remain fixed but the viewpoint varies. Then all the linear ambiguities that arise in multiple-view geometry[4] persist despite the new photometric information. We cannot, for example, use photometric information to resolve geometric ambiguities in binocular stereopsis (without, of course, using non-Lambertian models). On the other hand, photometric cues may be able to resolve ambiguities in cases where the object moves and the viewpoint and the lighting are fixed, such as structure from motion.

In Section 2 we define the KGBR and prove that it preserves the shading and shadow properties on the surfaces of objects as the lighting changes. In Section 3 we con-

1084-7529/2003/010001-08$15.00 © 2003 Optical Society of America

PROOF COPY [A-8466] 002301JOA

PROOF COPY [A-8466] 002301JOA

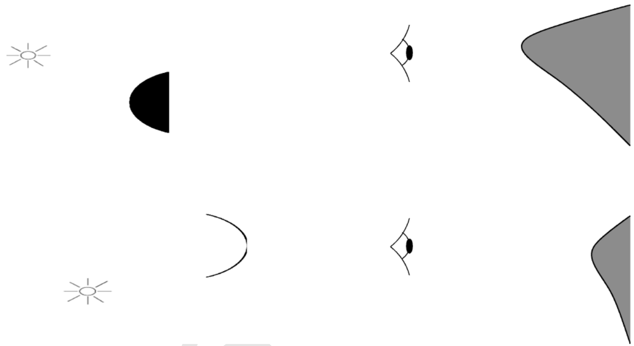

Fig. 1. Left panel: convex versus concave ambiguity. A convex object lit from above looks like a convex object lit from below. Right panel: the bas-relief ambiguity. The perception of shape is relatively insensitive to a linear scaling in the viewing direction.

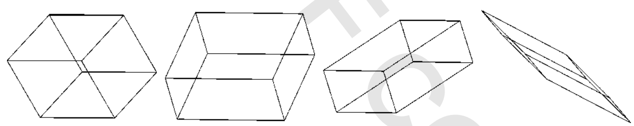

Fig. 2. Cube viewed from direction (0.51, 0.63, 0.58) (far-left panel) and the same cube undergoing affine transformations (remaining panels) seen from the same viewpoint.

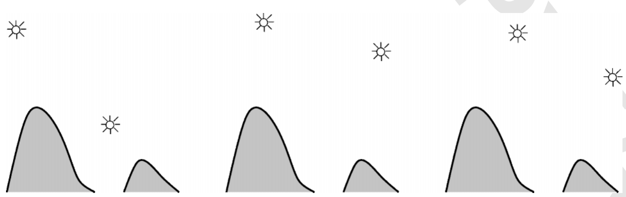

Fig. 3. If the lighting conditions are unknown, then it is impossible to distinguish between two objects related by a GBR transform.[5] For any image of the first object, under one illumination condition, we can always find a corresponding illumination condition that makes the second object appear identical (i.e., we can generate an identical image). We show two objects under three different, but corresponding, lighting conditions.

sider viewpoint projection and prove that there is a joint viewpoint–lighting ambiguity on the images of objects. In Section 4 we derive the GBR as a special case of the KGBR and give a general decomposition of the KGBR in terms of elementary transformations. In Section 5 we give examples of the KGBR. In Section 6 we describe the implications of the KGBR for structure from motion and structure from viewpoint.

2. KGBR ON SURFACES OF OBJECTS

In this section, we define the KGBR transform on the surface and albedo of objects together with the corresponding transform on the lighting. For objects with Lambertian reflectance functions, the image intensity at a fixed point on the surface is independent of the direction of the viewer. This feature enables us to study the shading and shadows as if they were painted onto the surface. In Section 3 we will see what this implies about the projection of the object to different viewpoints.

Let the object have surface position $\mathbf{r} = \mathbf{r}(u, v)$, surface normal $\mathbf{n}(u, v)$, and albedo $a(u, v)$ as functions of intrinsic coordinates (u, v) defined on the surface; see Fig. 4. We assume m light-source vectors $\{\mathbf{s}_1, \ldots, \mathbf{s}_m\}$.

Definition 1. A KGBR transform $\mathbf{K}$ transforms $\{\mathbf{r}(u, v), \mathbf{n}(u, v), a(u, v), \mathbf{s}_1, \ldots, \mathbf{s}_m, \mathbf{v}\}$ to $\{\hat{\mathbf{r}}(u, v), \hat{\mathbf{n}}(u, v), \hat{a}(u, v), \hat{\mathbf{s}}_1, \ldots, \hat{\mathbf{s}}_m, \hat{\mathbf{v}}\}$, where

$$\hat{\mathbf{r}}(u, v) = \mathbf{K}\mathbf{r}(u, v), \qquad \hat{\mathbf{n}}(u, v) = \frac{\mathbf{K}^{-1,T}\mathbf{n}(u, v)}{|\mathbf{K}^{-1,T}\mathbf{n}(u, v)|}, \tag{1}$$

$$\hat{a}(u, v) = a(u, v)|\mathbf{K}\sum_{j=1}^{M} \mathbf{S}_j||\mathbf{K}^{-1,T}\mathbf{n}(u, v)|, \tag{2}$$

$$\hat{\mathbf{s}}_i = \frac{1}{|\mathbf{K}\sum_{j=1}^{m} S_j|}\mathbf{K}\mathbf{s}_i \cdot i = 1, \ldots, m. \tag{3}$$

The matrix $\mathbf{K}$ of the KGBR can take any form. It gives an affine transformation on the three-dimensional surface that can involve squashing, skewing, or rotating the surface (or any combination of these operations). The form of $\hat{\mathbf{n}}(u, v)$ in Eq. (2) is derived directly from the transformation on the surface shape $\mathbf{r}(u, v)$ [recall that $\hat{\mathbf{n}}(u, v)$ must be orthogonal to the surface tangent vectors $\hat{\mathbf{r}}_u(u, v)$ and $\hat{\mathbf{r}}_v(u, v)$ that transform by the KGBR]. The normalization factor $|\mathbf{K}\Sigma_{j=1}^{m}\mathbf{s}_j|$ for transforming the light sources [see Eq. (2)] is chosen to ensure that the total lighting $\Sigma_{j=1}^{m}\hat{\mathbf{s}}_j$ has unit magnitude.

The definition of the KGBR transform is motivated by the following result.

Theorem 1. If two objects and their lighting are related by a KGBR, then their shading and shadows are preserved as functions of the surface coordinates (u, v). The shading is given by a Lambertian model $I(u, v) = \Sigma_{j=1}^{m} \max\{a(u, v)\mathbf{n}(u, v) \cdot \mathbf{s}_j, 0\}$ with cast shadows removed; see Fig. 5.

Proof. From Eqs. (1)–(3), we obtain

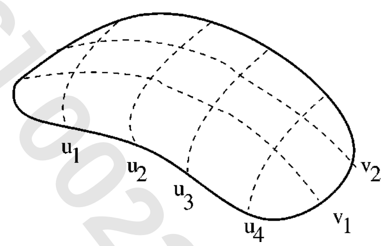

Fig. 4. We define intrinsic coordinates (u, v) on the surface of the object.

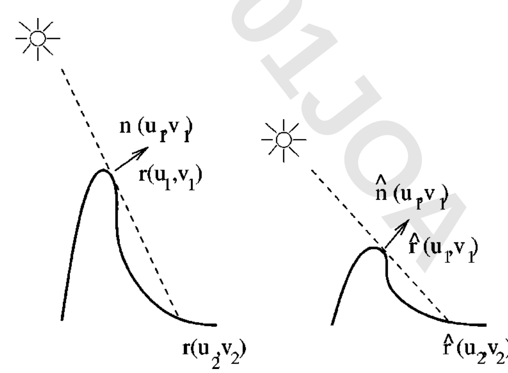

Fig. 5. The cast shadow boundaries, and hence the cast shadows, are preserved by the KGBR. Similar results were shown for the GBR.[5]

PROOF COPY [A-8466] 002301JOA

PROOF COPY [A-8466] 002301JOA

$$\max\{\hat{a}(u, v)\hat{\mathbf{n}}(u, v) \cdot \hat{\mathbf{s}}_j, 0\} = \max\{a(u, v)\mathbf{n}(u, v) \cdot \mathbf{s}_j, 0\}\ \forall j = 1, \ldots, m, \quad (4)$$

and so the shading (and attached shadows) are preserved.

The conditions for a point $\mathbf{r}(u_2, v_2)$ to be on the shadow boundary for light source $\mathbf{s}_j$ cast by $\mathbf{r}(u_1, v_1)$ (with normal $\mathbf{n}(u_1, v_1)$) are that $\mathbf{n}(u_1, v_1) \cdot \mathbf{s}_j = 0$ and $\mathbf{r}(u_2, v_2) - \mathbf{r}(u_1, v_1) \propto \mathbf{s}_j$; see Fig. 5.[5] These conditions are preserved by the KGBR; see Eqs. (1)–(3). Hence if a point (u, v) is in shadow on surface $\mathbf{r}(u, v)$ from light source $\mathbf{s}_j$, then it is also in shadow on surface $\hat{\mathbf{r}}(u, v)$ from light source $\hat{\mathbf{s}}_j$.

There are variations of the KGBR for which Theorem 1 will also hold. For example, the formulation in Definition 1 has been chosen to ensure that the magnitude of the light source $\Sigma_{j=1}^{m}\mathbf{s}_j$ is preserved by the KGBR. This choice, however, has the undesirable property that the transformation on the albedo is dependent on the light sources. An alternative is to transform $\mathbf{s}_i \mapsto \mathbf{K}\mathbf{s}_i/F(\mathbf{K})$ $i = 1, \ldots, m$ and $a(u, v) \mapsto a(u, v)F(\mathbf{K})|\mathbf{K}^{-1,T}\mathbf{n}(u, v)|$, where $F(\mathbf{K})$ is a function of the KGBR $\mathbf{K}$. One possibility is to set $F(\mathbf{K}) = |\det \mathbf{K}|$, which improves on the choice $F(\mathbf{K}) = \det \mathbf{K}$ described in previous work.[8] The modulus sign is necessary for handling cases where $\det \mathbf{K}$ is negative; such cases which occur, for example, in the convex-versus-concave ambiguity.

3. KGBR VIEWPOINT–LIGHTING AMBIGUITY

Section 2 has established that shading and shadow properties at points (u, v) on a surface are preserved by a KGBR transformation. We now determine what happens as we project the surface onto an image plane.

Our projection model assumes an affine camera or orthographic projection with a two-dimensional affine transform on the image coordinates. These projection models are equivalent because of the two-dimensional affine transform, but they lead to slightly different mathematical formulations. They can be derived as approximations to perspective projection[4] which results from (i) assuming that the camera parameters are only approximately known and (ii) modifying the orthographic projection equations to allow for perspective effects (by making assumptions about the shape of the viewed object and its position relative to the camera).

Definition 2. A projection is specified by three vectors $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{v}$ that we call camera parameters. For the orthographic camera these vectors are constrained to be orthogonal unit vectors. For the affine camera we constrain only $\mathbf{c}_1$, $\mathbf{c}_2$ to be orthogonal to $\mathbf{v}$. So in both cases, we have $\mathbf{v} \propto \mathbf{c}_1 \times \mathbf{c}_2$; see Fig. 6. A point (u, v) on the surface $\mathbf{r}(u, v)$ is projected to a point $[p_1(u, v), p_2(u, v)]$ in the image plane by

$$p_1(u, v) = \mathbf{c}_1 \cdot \mathbf{r}(u, v), \quad p_2(u, v) = \mathbf{c}_2 \cdot \mathbf{r}(u, v). \quad (5)$$

As stated above, the use of the Lambertian reflectance function means that the image intensity of a point on the surface is independent of the viewpoint. Therefore the intensity at an image point $[p_1(u, v), p_2(u, v)]$ is equal to the intensity $I(u, v)$ of the corresponding point (u, v) on the surface of the object. Theorem 1 has already shown that shading and shadows are preserved under a KGBR at each surface point (u, v). Hence the remaining analysis is concerned only with the geometry of projection. In particular, we can rederive existing results for pointlike features[2–4] by replacing the surface with a discrete set of N points $\{\mathbf{r}_i : i = 1, \ldots, N\}$ and setting the images to be $I(p_1, p_2) = \Sigma_{i=1}^{N}\delta(p_1 - \mathbf{c}_1 \cdot \mathbf{r}_i)\delta(p_2 - \mathbf{c}_2 \cdot \mathbf{r}_i)$, where $\delta(x)$ is the Dirac delta function.

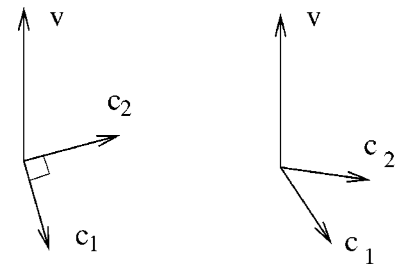

Fig. 6. For the orthographic camera (left) the vectors $\mathbf{v}$, $\mathbf{c}_1$, $\mathbf{c}_2$ are orthogonal unit vectors. For the affine camera (right) $\mathbf{c}_1$, $\mathbf{c}_2$ are constrained to be orthogonal only to the unit vector $\mathbf{v}$.

We now define how the camera parameters change under a KGBR.

Definition 3. The camera parameters of the affine camera transform by:

$$\hat{\mathbf{c}}_1 = \mathbf{K}^{-1,T}\mathbf{c}_1, \quad \hat{\mathbf{c}}_2 = \mathbf{K}^{-1,T}\mathbf{c}_2, \quad \hat{\mathbf{v}} = \frac{\mathbf{K}\mathbf{v}}{|\mathbf{K}\mathbf{v}|}. \quad (6)$$

The camera parameters of the orthographic camera transform as

$$\hat{\mathbf{c}}_1 = A_{11}\mathbf{K}^{-1,T}\mathbf{c}_1 + A_{12}\mathbf{K}^{-1,T}\mathbf{c}_2,$$

$$\hat{\mathbf{c}}_2 = A_{21}\mathbf{K}^{-1,T}\mathbf{c}_1 + A_{22}\mathbf{K}^{-1,T}\mathbf{c}_2, \quad \hat{\mathbf{v}} = \frac{\mathbf{K}\mathbf{v}}{|\mathbf{K}\mathbf{v}|}, \quad (7)$$

where the parameters A_{11}, A_{12}, A_{21}, A_{22} are functions of $\mathbf{K}$ (their exact form is given in Theorem 5) chosen to ensure that $\hat{\mathbf{c}}_1$, $\hat{\mathbf{c}}_2$, $\hat{\mathbf{v}}$ are orthogonal unit vectors.

These transformations ensure that the viewpoint direction vector $\hat{\mathbf{v}}$ is orthogonal to the projections vectors $\hat{\mathbf{c}}_1$, $\hat{\mathbf{c}}_2$. For the orthographic camera, the transformation in Eq. (7) must, of course, correspond to a rotation. Theorem 5 shows how this rotation is a function of the KGBR transformation $\mathbf{K}$.

We now show that there is a viewpoint–lighting ambiguity; see Fig. 7. Our results are given in Theorems 2 and 3 for the affine camera and for orthographic projection, respectively.

Theorem 2. If two objects' geometry and albedo are related by a KGBR transform $\mathbf{K}$, then for any viewpoint and illumination condition of one object there exists a viewpoint, affine camera setting, and illumination condition for the second object such that the images of the two objects are identical.

Proof. We use Definitions 1, 2, and 3 to specify the KGBR transformations on the objects and the affine cameras. It follows from projection equations (5) that points labeled (u, v) on the two objects project to the same point $[\mathbf{c}_1 \cdot \mathbf{r}(u, v), \mathbf{c}_2 \cdot \mathbf{r}(u, v)] = [\hat{\mathbf{c}}_1 \cdot \hat{\mathbf{r}}(u, v), \hat{\mathbf{c}}_2 \cdot \hat{\mathbf{r}}(u, v)]$

PROOF COPY [A-8466] 002301JOA

PROOF COPY [A-8466] 002301JOA

in the two images. By Theorem 1 it follows that the images of the two objects are identical.

We next extend our result to orthographic projection (it is also straightforward to deal with scaled orthographic by relaxing the conditions on A_{11}, $A_{12}.A_{21}$, A_{22}).

Theorem 3. If two objects are related by a KGBR transform $\mathbf{K}$, then for any viewpoint and illumination condition of one object there exists a viewpoint and illumination condition for the second object such that the images of the two objects are identical up to the affine transformation

$$\mathbf{A}_2 = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

given by camera transformation equation (7).

Proof. Use Definitions 1, 2, and 3 to specify the KGBR transformations on the objects and the orthographic cameras. It follows from projection equations (5) that points labeled (u, v) on the two objects project to points in the images related by the two-dimensional affine transformation $\mathbf{A}_2$:

$$\begin{pmatrix} \hat{\mathbf{c}}_1 \cdot \hat{\mathbf{r}}(u, v) \\ \hat{\mathbf{c}}_2 \cdot \hat{\mathbf{r}}(u, v) \end{pmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} \mathbf{c}_1 \cdot \mathbf{r}(u, v) \\ \mathbf{c}_2 \cdot \mathbf{r}(u, v) \end{pmatrix}. \quad (8)$$

It follows from Theorem 1 that the images of the two objects are identical up to affine warping.

We note that Werman and Weinshall[9] loosened the notion of image similarity by proposing an affine-invariant measure on the set of images. Two images $I(\mathbf{p})$ and $\hat{I}(\mathbf{p})$ are identical up to a two-dimensional affine transformation $\mathbf{A}_2$ provided that $\hat{I}(\hat{\mathbf{p}}) = I(\mathbf{p})$, where $\hat{\mathbf{p}} = \mathbf{A}_2\mathbf{p}$.

Theorems 2 and 3 summarize the KGBR lighting–viewpoint ambiguity. In the following sections we will explore the ambiguity and, in particular, determine the relationship between the KGBR transform $\mathbf{K}$ and the two-dimensional affine warp $\mathbf{A}_2$.

We have avoided the special case where $\mathbf{K}$ is not invertible. In this degenerate case, Eq. (1) implies that one object is planar. If the second object is also planar then the theorems above are easy to prove. If the second object is nonplanar, however, then the results no longer hold. This is because all views of the planar object are equivalent to within an affine transformation and therefore correspond only to a single view of the nonplanar object (the front-on view). Thus for almost all views of the nonplanar object there is no corresponding view of the planar object. This applies both for the affine camera and for orthographic/scaled-orthographic projections.

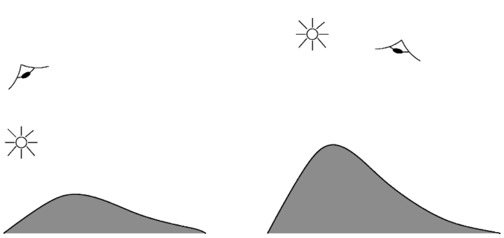

Fig. 7. Joint viewpoint–lighting ambiguity. If two objects are related by a KGBR, then for any view of one there is a corresponding view of the other that is identical (after adjusting the setting of the affine camera) or identical up to a two-dimensional affine warp (for orthographic projection). The lighting is also transformed by the corresponding KGBR.

Finally, an attractive conceptual picture of the KGBR lets the viewer and the light source be at finite distance from the object. In this case, we let $\mathbf{v}$ and $\mathbf{s}$ be vectors representing the viewer position and the light-source position, respectively. They can transform by $\mathbf{s} \mapsto \mathbf{Ks}$ and $\mathbf{v} \mapsto \mathbf{Kv}$, exactly like surface points $\mathbf{r}(u, v)$ on the object. In other words, the KGBR simply corresponds to an affine transformation on the space. But because the light source is at finite distance from the object, we cannot use the Lambertian model except as an approximation (because there will be an inverse-square-law fall-off of intensity with distance from the light source). Therefore the theorems will hold in the limit only as the light source position tends to infinity. [For this transformation the albedo will transform as $a(u, v) \mapsto a(u, v)|\mathbf{K}^{-1,T}\mathbf{n}(u, v)|$].

4. DECOMPOSITION OF THE KGBR AND ITS RELATIONS TO THE GBR

This section investigates the structure of the KGBR. We first describe the GBR ambiguity and rederive it as a special case of the KGBR. Then we proceed to give a general decomposition of the KGBR transform in terms of a GBR transform, a three-dimensional rotation, and a two-dimensional affine transform. This decomposition explains the form of the two-dimensional affine transformation $\mathbf{A}_2$ required for transforming the orthographic camera images; see Definition 3 and Theorem 3.

The GBR was originally derived by assuming a fixed viewpoint with orthographic projection.[5] Objects are represented by their albedoes $a(x, y)$ and surface normals $\mathbf{n}(x, y)$ by use of a coordinate system x, y in the image plane. The imaging model is Lambertian, so that $I(x, y) = \Sigma_j \max\{a(x, y)\mathbf{n}(x, y) \cdot \mathbf{s}_j, 0\}$, and cast shadows are removed. To derive the GBR, observe that the image is invariant to the transformations $a(x, y)\mathbf{n}(x, y) \mapsto \mathbf{G}a(x, y)\mathbf{n}(x, y)$ and $\mathbf{s}_j \mapsto \mathbf{G}^{-1,T}\mathbf{s}_j \,\forall j$, where $\mathbf{G}$ is any invertible matrix. Belhumeur *et al.*[5] prove that the surface integrability condition (i.e., the requirement that the surface normals be consistent with a real surface) restricts $\mathbf{G}$ to be of form

$$\begin{bmatrix} 1 & 0 & -\mu/\lambda \\ 0 & 1 & -\nu/\lambda \\ 0 & 0 & 1/\lambda \end{bmatrix}$$

for constants μ, ν, λ. Moreover, if the surface is represented by $z = f(x, y)$, then the transformation on the surface normals corresponds to a transformation on the surface $f(x, y) \mapsto \lambda f(x, y) + \mu x + \nu y$. Therefore a GBR transform $\mathbf{G}$ on the surface normals corresponds to a transformation $\mathbf{G}^{-1,T}$ on the surface points.

We now rederive the GBR from the KGBR by requiring that the images of objects be identical when viewed from a specific viewpoint. This requirement means that the GBR transforms are only a subgroup of the KGBR transforms. It can be argued that our result is more intuitive than the original proof, briefly summarized in the previous paragraph, because it bypasses the surface-integrability constraint.

PROOF COPY [A-8466] 002301JOA

PROOF COPY [A-8466] 002301JOA

Theorem 4. If two objects are related by a KGBR $\mathbf{K}$ and there exists a special viewpoint $\mathbf{v}^*$ such that the images of the two objects are identical, then $\mathbf{K}$ must be of form $\mathbf{G}^{-1,T}$, where $\mathbf{G}$ is a GBR.

Proof. Let the two objects be $\mathbf{r}(u, v)$ and $\hat{\mathbf{r}}(u, v) = \mathbf{K}\mathbf{r}(u, v)$. If there exists a special viewpoint (with orthographic projection) such that the images are identical, then we can find vectors $\mathbf{c}_1^*$, $\mathbf{c}_2^*$ such that

$$\mathbf{c}_1^* \cdot \mathbf{r}(u, v) = \mathbf{c}_1^* \cdot \mathbf{K}\mathbf{r}(u, v), \forall u, v,$$
$$\mathbf{c}_2^* \cdot \mathbf{r}(u, v) = \mathbf{c}_2^* \cdot \mathbf{K}\mathbf{r}(u, v), \forall u, v. \tag{9}$$

This implies that $\mathbf{K}^T$ has two unit eigenvectors (unless the surface is a plane). Hence by a suitable choice of coordinate system we can express $\mathbf{K}$ and $\mathbf{K}^{-1,T}$ as

$$\mathbf{K} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{bmatrix}, \quad \mathbf{K}^{-1,T} = \begin{bmatrix} 1 & 0 & -\mu/\lambda \\ 0 & 1 & -\nu/\lambda \\ 0 & 0 & 1/\lambda \end{bmatrix}, \tag{10}$$

where λ, μ, ν are constants. This shows that $\mathbf{K}^{-1,T}$ is of GBR form.[5] Conversely, if $\mathbf{K}^T$ has two unit eigenvalues, then we can define the projections to be their corresponding eigenvectors.

We now go further by showing how to decompose any KGBR in terms of a GBR, a rotation, and a two-dimensional affine transformation. The proof assumes orthographic cameras with a two-dimensional transform, but the result is a property of matrices (i.e., linear algebra) and so is independent of the choice of camera.

Theorem 5. Any KGBR transform $\mathbf{K}$ can be decomposed as $\mathbf{K} = \mathbf{\Phi}\mathbf{G}^{-1,T}\mathbf{A}_3$, where $\mathbf{\Phi}$ is a rotation, $\mathbf{A}_3$ is given by

$$\mathbf{A}_3 = \begin{bmatrix} \mathbf{A}_2 & 0 \\ 0 & 1 \end{bmatrix},$$

and

$$\mathbf{A}_2 = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

is the two-dimensional affine warp that $\mathbf{K}$ induces on the orthographic projection. If the images are identical under orthographic projection (i.e., $\mathbf{A}_3 = \mathbf{I}$), then $\mathbf{K} = \mathbf{\Phi}\mathbf{G}^{T,-1}$.

Proof. Equation (7) of Definition 3 specifies how an orthographic camera transforms under a KGBR transformation $\mathbf{K}$. This transformation can be reexpressed as

$$\mathbf{K}^{-1,T}\mathbf{c}_1 = B_{11}\hat{\mathbf{c}}_1 + B_{12}\hat{\mathbf{c}}_2, \qquad \mathbf{K}^{-1,T}\mathbf{c}_2 = B_{21}\hat{\mathbf{c}}_1 + B_{22}\hat{\mathbf{c}}_2, \tag{11}$$

where the matrix

$$\mathbf{B}_2 = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

is the inverse of

$$\mathbf{A}_2 = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

But this transformation must also correspond to a rotation $\mathbf{\Phi}$ of the coordinate axes (because the vectors $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{v}$ must remain orthonormal unit vectors) and hence can be written as

$$\hat{\mathbf{c}}_1 = \mathbf{\Phi}\mathbf{c}_1, \qquad \hat{\mathbf{c}}_2 = \mathbf{\Phi}\mathbf{c}_2, \qquad \hat{\mathbf{v}} = \mathbf{\Phi}\mathbf{v}. \tag{12}$$

By combining Eqs. (11) and (12), we obtain

$$\mathbf{K}^{-1,T}\mathbf{c}_1 = \mathbf{\Phi}(B_{11}\mathbf{c}_1 + B_{12}\mathbf{c}_2),$$
$$\mathbf{K}^{-1,T}\mathbf{c}_2 = \mathbf{\Phi}(B_{21}\mathbf{c}_1 + B_{22}\mathbf{c}_2). \tag{13}$$

Without loss of generality, let $\mathbf{c}_1 = (1, 0, 0)$ and $\mathbf{c}_2 = (0, 1, 0)$. Then Eqs. (13) determine the first two columns of the matrix $\mathbf{K}^{-1,T}$ in terms of $\mathbf{B}_2$ and $\mathbf{\Phi}$ but places no constraints on the third column. We can therefore express

$$\mathbf{K}^{-1,T} = \mathbf{\Phi}\begin{bmatrix} B_{11} & B_{21} & \alpha \\ B_{12} & B_{22} & \beta \\ 0 & 0 & \gamma \end{bmatrix} = \mathbf{\Phi}\mathbf{G}\mathbf{B}_3, \tag{14}$$

where α, β, γ are constants and

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & \beta \\ 0 & 0 & \gamma \end{bmatrix}$$

is a GBR transform[5] and

$$\mathbf{B}_3 = \begin{bmatrix} B_{11} & B_{12} & 0 \\ B_{21} & B_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The first result follows from Eq. (14) by taking the inverse. If the images are identical under orthographic projection, then $\mathbf{A}_3 = \mathbf{I}$, by definition, and the second result follows directly.

Theorem 5 shows how to decompose a KGBR transform into its components. Observe that if images of a KGBR transform are identical under orthographic projection (i.e., $\mathbf{A}_3 = \mathbf{I}$), then the KGBR reduces to a GBR and a rotation. It is the two-dimensional affine transformation $\mathbf{A}_2$ (and hence $\mathbf{A}_3$) that is necessary for $\mathbf{K}$ to be a full three-dimensional affine transformation.

It is possible to take $\mathbf{K}$ and decompose it into its parts. For example, the following theorem shows how to determine $\mathbf{A}_2$ from $\mathbf{K}$. More precisely, the theorem distinguishes between those properties of $\mathbf{A}_2$ that are dependent on the choice of coordinate systems in the two viewing planes and those properties that are invariant to this choice. These invariant properties are then related to $\mathbf{K}$.

Theorem 6. The affine transformation $\mathbf{A}_2$ changes as $\mathbf{A}_2 \mapsto \mathbf{\Psi}\mathbf{A}_2\mathbf{\Phi}^T$, where $\mathbf{\Psi}$ and $\mathbf{\Phi}$ are two-dimensional rotation matrices corresponding to changing the coordinate systems in the two viewing planes. The properties of $\mathbf{A}_2$ that are invariant to these rotations are specified in terms of $\mathbf{K}$ by $\det \mathbf{A}_2 = \det \mathbf{K}/|\mathbf{K}\mathbf{v}|$ and Trace $\{\mathbf{A}_2\mathbf{A}_2^T\} = \text{Trace}\{\mathbf{K}\mathbf{K}^T\} - |\mathbf{K}^T\mathbf{K}\mathbf{v}|^2/|\mathbf{K}\mathbf{v}|^2$.

Proof. First observe that we have the freedom to change the axes $\mathbf{c}_1$, $\mathbf{c}_2$ by a rotation $\mathbf{\Phi}$ in the first viewing plane and similarly rotate $\hat{\mathbf{c}}_1$, $\hat{\mathbf{c}}_2$ by a similar rotation $\mathbf{\Psi}$ in the second viewing plane. Hence, using Eqs. (7), we

PROOF COPY [A-8466] 002301JOA

PROOF COPY [A-8466] 002301JOA

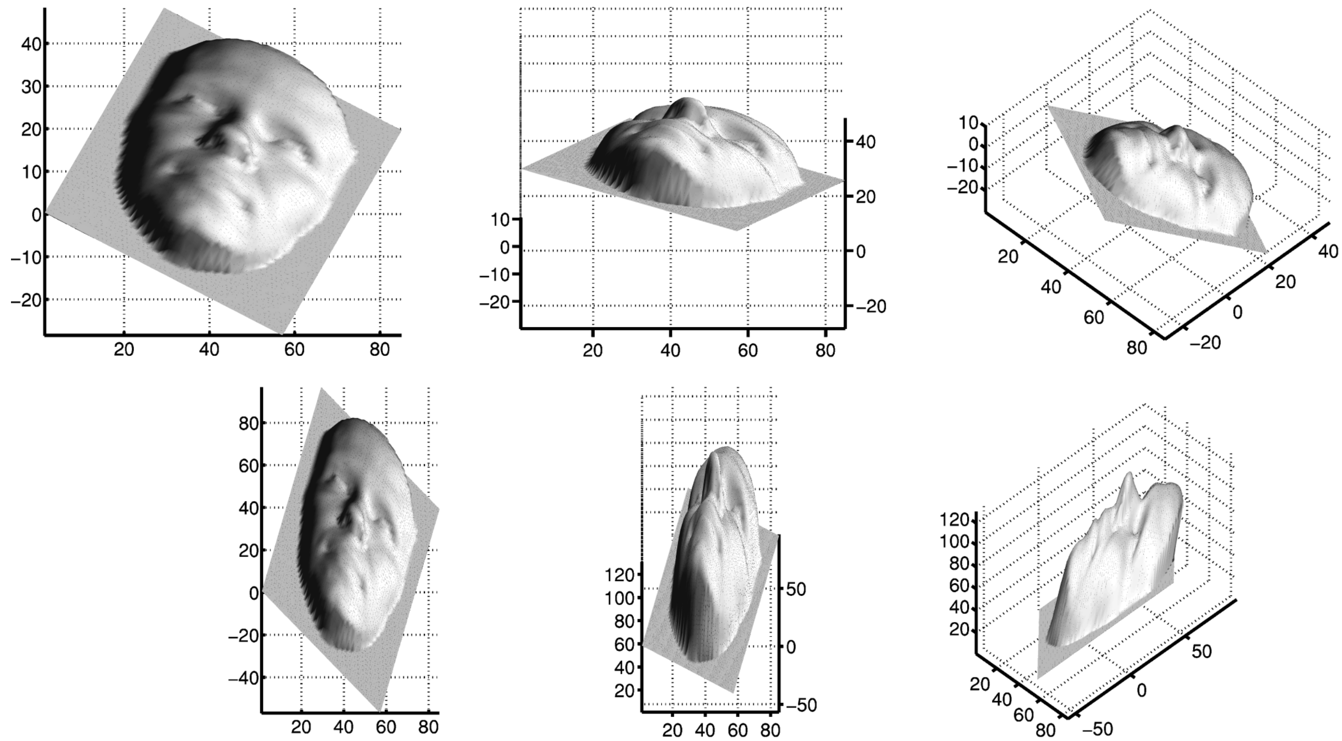

Fig. 8. Top row, original object; bottom row, object after a KGBR (see text). Left panels, the objects look the same when viewed from direction (1, 0, 0) up to an affine warp on the images. But they look very different when viewed from direction $(1/\sqrt{2}, 0, 1/\sqrt{2})$ (center panels) and direction $(1/2, 1/2, 1/\sqrt{2})$ (right panel).

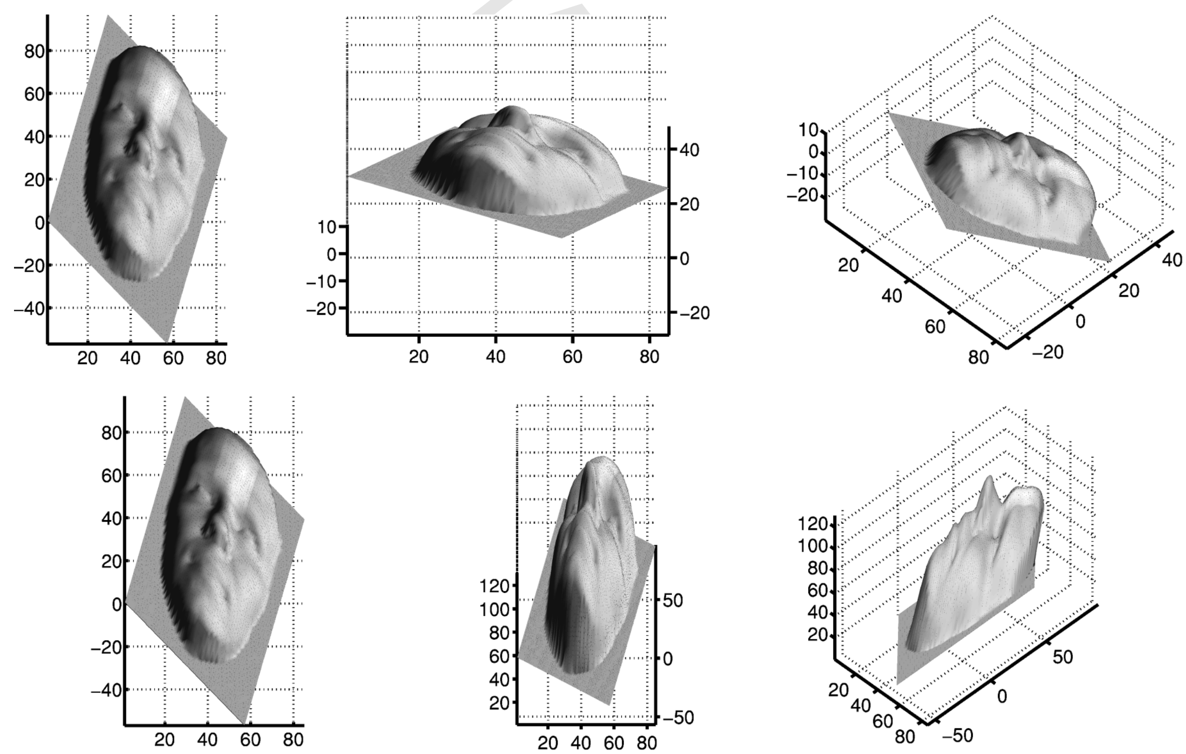

Fig. 9. Top row, original object (with albedo); bottom row, object after a KGBR (see text). Left panels, the objects look the same when viewed from direction (1, 0, 0) up to an affine warp on the images. But they look very different when viewed from direction $(1/\sqrt{2}, 0, 1/\sqrt{2})$ (center panels), and direction $(1/2, 1/2, 1/\sqrt{2})$ (right panel).

have the freedom to send $\mathbf{A}_2 \mapsto \mathbf{\Psi}\mathbf{A}_2\mathbf{\Phi}^T$. Next we choose the coordinate system so that $\mathbf{v} = (0, 0, 1)$ and use the decomposition of $\mathbf{K}$ given by Theorem 5. It follows that $\det \mathbf{K} = \det \mathbf{G}^{-1,T} \times \det \mathbf{A}_3 = \lambda \det \mathbf{A}_2$ and $|\mathbf{K}\mathbf{v}| = \lambda$, and hence $\det \mathbf{A}_2 = \det \mathbf{K}/|\mathbf{K}\mathbf{v}|$. Finally, multiplying Eqs. (7) by $\mathbf{K}^T$ and taking the dot product of each yields Trace

PROOF COPY [A-8466] 002301JOA

PROOF COPY [A-8466] 002301JOA

$\{\mathbf{A}_2\mathbf{A}_2^T\} = |\mathbf{K}^T\hat{\mathbf{c}}_1|^2 + |\mathbf{K}^T\hat{\mathbf{c}}_2|^2$. This can be expressed as Trace $\{\mathbf{A}_2\mathbf{A}_2^T\} = \text{Trace}\{\mathbf{K}\mathbf{K}^T\} - |\mathbf{K}^T\hat{\mathbf{v}}|^2$ [because $\hat{\mathbf{c}}_1$, $\hat{\mathbf{c}}_2$, $\hat{\mathbf{v}}$ are orthogonal unit vectors, see Eq. (7)]. The result follows by setting $\hat{\mathbf{v}} = \mathbf{K}\mathbf{v}/|\mathbf{K}\mathbf{v}|$.

5. EXAMPLES OF THE KGBR

We now show a few simple example of the KGBR.

The object is a face on a planar background rotated by euler angles (0.2, 0.5, 0.3). The albedo is initially set to be constant. We obtain a second object by applying a KGBR composed of an identity rotation, a GBR $\mathbf{G}$ chosen so that

$$\mathbf{G}^{-1,T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1.2 & 1.3 & 4 \end{bmatrix},$$

and a two-dimensional warp

$$\mathbf{A}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

The two objects look identical from viewpoint (0, 0, 1) but are clearly very different when viewed from other directions; see Fig. 8.

We also modify the first object by changing its albedo to be $a(x, y) = 1 + 0.5\exp(-0.01\{(x - 52)^2 + (y - 52)^2\})$, where $x = 1,.., 64$ and $y = 1,..., 64$. Observe that from extreme views, (see right-hand panels of Fig. 9), the albedo starts being interpreted as the geometry of the surface.

6. IMPORTANT CASES

There are three variables factors: the lighting, the observer, and the object. Suppose we move one and fix the others. This gives (I) photometric stereo when we move the lighting, (II) structure from viewpoint when we move the viewer, and (III) structure from motion when we move the object. For photometric stereo we know that the KGBR ambiguity reduces to the GBR. We now study what happens to the ambiguity for structure from viewpoint and structure from motion. If photometric effects are ignored, then these two shape cues have the same ambiguities. We will ignore special cases such as, for example, when the objects are planar and/or the KGBR is a pure rotation. So our results will apply to generic objects and KGBR transformations

Following recent work[4] we use affine transformations to model changes of viewpoint and motion. This reduces to the standard definitions if the transformation is restricted to being a rotation. Otherwise, it corresponds to a rotation, or viewpoint change, combined with an affine distortion on the object. This affine distortion can be considered to arise from failure to calibrate the camera(s).[4]

Our result for structure from viewpoint shows that there are ambiguities if we allow for affine changes of viewpoint, but these ambiguities disappear if we restrict ourselves to pure rotations.

Theorem 7. Let O and $\hat{O}$ be two objects related by a KGBR transform $\mathbf{K}$. Then for any affine change of viewpoint $\mathbf{M}$ of O there exists a corresponding affine change of viewpoint $\mathbf{N} = \mathbf{K}^{-1,T}\mathbf{M}\mathbf{K}^T$ for $\hat{O}$ such that the images are identical up to a constant two-dimensional affine transformation. If $\mathbf{M}$ and $\mathbf{N}$ are both rotation matrices, then they must both commute with $\mathbf{K}\mathbf{K}^T$ and hence must be the identity matrix unless $\mathbf{K}$ is nongeneric.

Proof. Let the object O have shape $\mathbf{r}(u, v)$, surface normals $\mathbf{n}(u, v)$, albedo $a(u, v)$, and light source $\mathbf{s}$. Let the first viewpoint be specified by $\{\mathbf{c}_1, \mathbf{c}_2\}$ and the second by $\{\mathbf{M}\mathbf{c}_1, \mathbf{M}\mathbf{c}_2\}$, where $\mathbf{M}$ is an affine transform. It is straightforward to calculate the projections of the surface before, (x^1, y^1), and after, (x^2, y^2), the change of viewpoint:

$$x^1(u, v) = \mathbf{c}_1^T\mathbf{r}(u, v), \qquad y^1(u, v) = \mathbf{c}_2^T\mathbf{r}(u, v),$$
$$I^1(u, v) = \max\{a(u, v)\mathbf{n}^T(u, v)\mathbf{s}, 0\},$$
$$x^2(u, v) = \mathbf{c}_1^T\mathbf{M}^T\mathbf{r}(u, v),$$
$$y^2(u, v) = \mathbf{c}_2^T\mathbf{M}^T\mathbf{r}(u, v),$$
$$I^1(u, v) = \max\{a(u, v)\mathbf{n}^T(u, v)\mathbf{s}, 0\}. \tag{15}$$

Now let object $\hat{O}$ be related to O by a KGBR transformation $\mathbf{K}$; see Eqs. (1)–(3). Let the initial viewpoint be given by $\{\hat{\mathbf{c}}_1, \hat{\mathbf{c}}_2\}$ and the second by $\{\mathbf{N}\hat{\mathbf{c}}_1, \mathbf{N}\hat{\mathbf{c}}_2\}$, where $\mathbf{N}$ is an affine transform. The projections before $\hat{x}^1, \hat{y}^1$ and after $\hat{x}^2, \hat{y}^2$ change of viewpoint are

$$\hat{x}^1(u, v) = \hat{\mathbf{c}}_1^T\mathbf{K}\mathbf{r}(u, v), \quad \hat{y}^1(u, v) = \hat{c}_2^T\mathbf{K}\mathbf{r}(u, v),$$
$$\hat{I}^1(u, v) = \max\{a(u, v)\mathbf{n}^T(u, v)\mathrm{s}, 0\},$$
$$\hat{x}^2(u, v) = \hat{\mathbf{c}}_1^T\mathbf{N}^T\mathbf{K}\mathbf{r}(u, v),$$
$$\hat{y}^2(u, v) = \hat{\mathbf{c}}_2^T\mathbf{N}^T\mathbf{K}\mathbf{r}(u, v),$$
$$\hat{I}^1(u, v) = \max\{a(u, v)\mathbf{n}^T(u, v)\mathrm{s}, 0\}. \tag{16}$$

From Theorem 2 we have $\hat{\mathbf{c}}_1 = \mathbf{K}^{-1,T}\mathbf{c}_1$ and $\hat{\mathbf{c}}_2 = \mathbf{K}^{-1,T}\mathbf{c}_2$. Combining these with Eqs. (15) and (16) implies that $\mathbf{K}^{-1}\mathbf{N}^T\mathbf{K} = \mathbf{M}^T$ or, equivalently, that $\mathbf{N} = \mathbf{K}^{-1,T}\mathbf{M}\mathbf{K}^T$. If $\mathbf{M}$ is a rotation matrix then $\mathbf{M}\mathbf{M}^T = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, which implies that $\mathbf{K}\mathbf{K}^T = \mathbf{N}^T\mathbf{K}\mathbf{K}^T\mathbf{N}$. If $\mathbf{N}$ is also a rotation matrix then $\mathbf{N}^T = \mathbf{N}^{-1}$, and so $\mathbf{N}$ commutes with $\mathbf{K}\mathbf{K}^T$. But it is impossible for a rotation matrix to commute with a symmetric matrix, like $\mathbf{K}\mathbf{K}^T$, unless one or both are degenerate.[10] A similar argument shows that $\mathbf{M}$ must commute with $\mathbf{K}\mathbf{K}^T$. We conclude that it is impossible for both $\mathbf{M}$ and $\mathbf{N}$ to be rotation matrices except for nongeneric KGBRs.

We now consider the case of structure from motion. Here we find that the photometric cues can disambiguate the geometric ambiguities (except for nongeneric special cases).

Theorem 8. Let O and $\hat{O}$ be specified as for Theorem 7. If we transform O by an affine transformation $\mathbf{M}$, then there is no affine transformation on $\hat{O}$ that yields the same image (unless O is a degenerate surface or the KGBR is a pure rotation).

Proof. Transforming O by an affine transformation $\mathbf{M}$ induces a map on the surface positions and the surface normals by $\mathbf{r} \mapsto \mathbf{M}\mathbf{r}$ and $\mathbf{n} \mapsto \mathbf{M}^{-1,T}\mathbf{n}/|\mathbf{M}^{-1,T}\mathbf{n}|$. The projec-

PROOF COPY [A-8466] 002301JOA

PROOF COPY [A-8466] 002301JOA

tion (x^1, y^1) before the affine transformation is given by Eqs. (15) (top two lines), and the projection (x^2, y^2) after is given by

$$x^2(u, v) = \mathbf{c}_1^T \mathbf{M}\mathbf{r}(u, v), \qquad y^2(u, v) = \mathbf{c}_2^T \mathbf{M}\mathbf{r}(u, v),$$

$$I^2(u, v) = \max\left\{\frac{a(u, v)\mathbf{n}^T(u, v)\mathbf{M}^{-1}\mathbf{s}}{|\mathbf{M}^{-1,T}\mathbf{n}(u, v)|}, 0\right\}. \tag{17}$$

Suppose that object $\hat{O}$ is transformed by an affine transformation $\mathbf{N}$. This is equivalent to a KGBR transform $\mathbf{NK}$ acting on the geometry of O but a KGBR transform $\mathbf{K}$ on the albedo of O. The projection $(\hat{x}^1, \hat{y}^1)$ before the transformation is given by Eqs. (16) and the projection $(\hat{x}^2, \hat{y}^2)$ after by

$$\hat{x}^2(u, v) = \hat{\mathbf{c}}_1^T \mathbf{NK}\mathbf{r}(u, v), \qquad \hat{y}^2(u, v) = \hat{\mathbf{c}}_2^T \mathbf{NK}\mathbf{r}(u, v),$$

$$\hat{I}^2(u, v) = \max\left\{\frac{a(u, v)|\mathbf{K}^{-1,T}\mathbf{n}(u, v)|\mathbf{n}^T(u, v)\mathbf{K}^{-1}\mathbf{N}^{-1}\mathbf{K}\mathbf{s}}{|\mathbf{N}^{-1,T}\mathbf{K}^{-1,T}\mathbf{n}(u, v)|}, 0\right\}. \tag{18}$$

Comparing Eqs. (17) and (18), we find that corresponding points (u, v) on the two objects are projected to the same positions in the two viewpoints provided that $\mathbf{K}^{-1}\mathbf{NK} = \mathbf{M}$. But their intensities will not correspond unless $|\mathbf{N}^{-1,T}\mathbf{K}^{-1,T}\mathbf{n}(u, v)| = |\mathbf{K}^{-1,T}\mathbf{n}(u, v)| \times |\mathbf{M}^{-1,T}\mathbf{n}(u, v)|$. This is equivalent to $|\mathbf{K}^{-1,T}\mathbf{M}^{-1,T}\mathbf{n}(u, v)| = |\mathbf{K}^{-1,T}\mathbf{n}(u, v)| \times |\mathbf{M}^{-1,T}\mathbf{n}(u, v)|$. This can be solved directly provided that $\mathbf{K}$ is a rotation matrix [since both sides reduce to $|\mathbf{M}^{-1,T}\mathbf{n}(u, v)|$]. Otherwise no solutions exist (except for nongeneric objects). Hence the ambiguity occurs only if $\mathbf{K}$ is a rotation matrix.

7. CONCLUSION

The goal of this paper was to propose the study of joint viewpoint–lighting ambiguities. Another motivation was to resolve the differences between the affine ambiguities found when ignoring photometric effects[2–4] and the generalized bas-relief (GBR) photometric ambiguity from fixed viewpoint.[5]

In particular, we derived the KGBR ambiguity by restricting the reflectance functions to be Lambertian (allowing for multiple light sources and shadows but no interreflections). The KGBR consists of a standard affine transformation on a surface shape in conjunction with a transformation on the surface albedo. We derived the GBR as a special case of the KGBR, which yielded a new intuitive proof of the GBR ambiguity. We obtained a general decomposition of the KGBR in terms of a rotation, a GBR, and a two-dimensional affine transformation. Then we illustrated the KGBR and discussed its implications for shape from shading and structure from viewpoint.

We hope that this work will motivate further studies of joint viewpoint–lighting ambiguities.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Health (National Eye Institute) by grant number RO1-EY 12691-01. We thank Dan Kersten for helpful discussions and two reviewers for useful feedback.

Please contact A. L. Yuille at yuille@stat.ucla.edu or phone, 310-267-5383. He is currently at Department of Psychology, 7461 Franz Hall, University of California, Los Angeles, Los Angeles, California 90095-1563.

REFERENCES

1. F. Phillips, J. T. Todd, J. J. Koenderink, and A. M. L. Kappers, “Perceptual localization of surface position,” J. Exp. Psychol. Hum. Percep. Perform. **23**, 1481–1492 (1997).
2. J. J. Koenderink and A. Van Doorn, “Affine structure from motion,” J. Opt. Soc. Am. A **8**, 377–385 (1991).
3. O. Faugeras, “Stratification of three-dimensional vision: projective, affine, and metric representations,” J. Opt. Soc. Am. A **12**, 465–484 (1995).
4. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision* (Cambridge U. Press, Cambridge, UK, 2000).
5. P. N. Belhumeur, D. Kriegman, and A. L. Yuille, “The generalized bas relief ambiguity,” Int. J. Comput. Vision **35**, 33–44 (1999).
6. A. L. Yuille, D. Snow, R. Epstein, and P. Belhumeur, “Determining generative models for objects under varying illumination: shape and albedo from multiple images using SVD and integrability,” Int. J. Comput. Vision **35**, 203–222 (1999).
7. R. Rosenholtz and J. Koenderink, “Affine structure and photometry,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (IEEE Computer Society Press, Los Alamitos, Calif., 1996), pp. **XX–XX**.
8. A. L. Yuille, J. M. Coughlan, and S. Konishi, “The KGBR viewpoint–lighting ambiguity and its resolution by generic constraints,” in *Proceedings of the International Conference of Computer Vision* (IEEE Computer Society Press, Los Alamitos, Calif. 2001), pp. 376–382.
9. M. Werman and D. Weinshall, “Similarity and affine invariant distance between point sets,” IEEE Trans. Pattern Anal. Mach. Intell. **17**, 810–814 (1995).
10. G. Strang, *Introduction to Applied Mathematics* (Wellesley-Cambridge Press, Wellesley, Mass., 1986).

PROOF COPY [A-8466] 002301JOA

PROOF COPY [A-8466] 002301JOA