

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/103983>

Please be advised that this information was generated on 2019-06-01 and may be subject to change.

# Extraction of three-dimensional shape from optic flow: a geometric approach

T. M. H. Dijkstra, P. R. Snoeren, and C. C. A. M. Gielen

Laboratory of Medical Physics and Biophysics, University of Nijmegen, P.O. Box 9101, NL-6500 HB, Nijmegen, The Netherlands

Received November 16, 1993; revised manuscript received February 22, 1994; accepted February 24, 1994

We show how a scale-invariant measure of three-dimensional shape can be derived from the velocity field generated by a rigid curved surface patch under perspective projection. We use invariance under rotation of the image plane [the Lie group  $SO(2)$ ] to decompose the second-order velocity field in differential invariants. From a combination of these invariants we construct an approximation of the absolute value of Koenderink's shape index [Image Vis. Comput. **10**, 557 (1992)]. We show that the effect of these approximations on the shape index is small, especially under parallel projection. Furthermore, we provide an explanation for the psychophysical finding that elliptical shapes are more readily detected than parabolic or hyperbolic shapes. From the invariants we can also derive approximations of the principal directions, the curvedness, the slant, and the tilt.

## 1. INTRODUCTION

Almost two decades have passed since original paper by Koenderink and van Doorn<sup>1</sup> on the relation between the geometry of a surface and the induced optic flow was published. In that paper they relate the geometrical properties of rigid objects to invariants of the velocity field generated by a moving object. They were quite successful for the first-order object properties (e.g., slant), but for the second-order properties (e.g., curvature) they could calculate only the sign of the Gaussian curvature. The research initiated by Koenderink and van Doorn took a more algebraic turn in, e.g., Refs. 2 and 3. In a recent paper Koenderink and van Doorn<sup>4</sup> focused on the second-order structure. They were able to describe the structure fully but only in terms of a somewhat unusual quantity, viz., the projected indicatrix of Dupin. In this paper we go back to a geometric approach based on differential invariants, as in Ref. 1, and extend it to second-order properties.

Central to our approach to the extraction of structure from motion is the use of geometry. Geometrical properties are those properties that are invariant when the coordinate system is changed. More specifically, we study the invariance of the velocity field under the group of rotations of the plane [the Lie group  $SO(2)$ ]. Because the velocity field is a vector field, its invariants are not necessarily scalars but are usually vectorlike quantities. Paradoxically, this means that the invariants of the velocity field can change when the coordinate system is rotated. A definition of the invariance of vector fields is beyond the scope of this paper, but in Section 3 we indicate how the invariants of a vector field are calculated. For a definition of the invariance of nonscalar fields, the reader is referred to Ref. 5 and the references therein. From the point of view of machine vision, rotational invariance allows one to do the calculations independently of the orientation of the camera. Moreover, one can view rotational invariance as a convenient computational tool: through expression of the relations among unknowns and observables in invariants, the equations become simpler.

Group theory is often used for this purpose in mathematics and in image processing. From the point of view of human vision, the use of the group of rotations of the plane (corresponding to torsional eye movements or torsional movements of the object) is not so easily defended. It would be more natural to study the invariants of the velocity field under the full rotation group in three dimensions,  $SO(3)$ . Since this is computationally much more complex and since  $SO(2)$  is a subgroup of  $SO(3)$ , one can view the current approach as a first step. Group theory was introduced into psychology by Hoffman,<sup>6</sup> but it has attracted few adherents.<sup>7</sup> Recently it has been applied successfully in neurophysiology.<sup>8</sup> In the rest of this paper we use invariance to mean rotational invariance. The only exception is Section 2, in which we introduce the shape index,<sup>9</sup> which is a scale-invariant descriptor of shape.

Smooth rigid objects have many differential geometric properties, which can be classified according to order. The zeroth-order property is the distance of the object from the observer. First-order properties can be described by the slant and the tilt. Note that the distance, the slant, and the tilt of an object will in general change when the object moves relative to the observer. The second-order properties can be described by, e.g., the two principal curvatures and the direction of maximal normal curvature. Our main focus will be on these properties. Note that the principal curvatures are the lowest-order properties, which do not change when the object moves relative to the observer; i.e., they are intrinsic to the object. The third-order properties can be described by the gradients of the two principal curvatures. We will not deal with these, although the method presented here could conceivably be used to deduce the third-order properties from invariants of the velocity field, too.

An important point to keep in mind is that metric information cannot be obtained from the velocity field, because the velocity field is derived from a projection. This means that one cannot obtain the complete three-dimensional (3D) translation velocity of the observer relative to the object. Neither can one obtain the distance

to the object nor its curvatures. Of course, one can construct combinations of these properties that do not depend on metric information, e.g., the ratio of the principal curvatures or the ratio of the 3D velocity and the distance. Since the principal curvatures are the lowest-order properties that are intrinsic, the ratio of the principal curvatures (or any function thereof) is the only intrinsic property of the object that can be derived from the velocity field (up to second order).

Our calculations are based on a number of assumptions to make the problem manageable. We assume the existence of a dense smooth vector field on a planar camera, which is generated by perspective projection of a smooth rigid surface patch. The assumption of the existence of a dense vector field is probably not necessary, as algorithms have been proposed to extract first-order differential invariants directly from the spatio-temporal luminance pattern.<sup>10,11</sup> These algorithms could be extended to include the second order, as well. The assumption of a planar camera (equipped with the natural metric) is accurate for machine vision but probably not for human vision when wide-field stimuli are employed. The assumption of a rigid smooth surface patch is realistic for rigid surfaces away from their contour. Finally, perspective projection is the correct way of proceeding. Some of the approximations that we use below are not necessary under parallel (orthographic) projection. The use of perspective projection will allow us to give an expression for the error that we make in these approximations.

Finally, we point out the usefulness of the current approach to the problem of binocular vision. One can view the disparity field as the equivalent of the velocity field and use the same relations that are introduced below. The differences between stereo and passive motion are the possible use of nonvisual information in stereo (e.g., vergence angle of the eyes) and the fact that the eyes are usually horizontally aligned, making our principle of rotational invariance less relevant for stereo vision. Nevertheless, we are able to offer an explanation for some results of shape-from-stereo experiments.

This paper is structured as follows. In Section 2 we review some of the descriptors that have been proposed to characterize shape. We introduce the shape index<sup>9</sup> as the descriptor that comes closest to an intuitive idea of shape. In Section 3 we introduce the notion of an invariant under the group of rotations in the plane, and we decompose the second-order velocity field into invariants. In Section 4 we calculate the second-order invariants of the velocity field generated by a moving curved surface. We show that we can construct an approximation of the shape index from a combination of these invariants. In Section 5 we show that the approximations made in the construction of the shape index lead to small deviations for a field of view of 6 deg. In the Section 7 we show that the predictions from our theory are consistent with some recent psychophysical results. Finally, we compare our model with two other recent models: the spin variation theory<sup>12</sup> and the Dupin indicatrix theory.<sup>4</sup>

## 2. DESCRIPTION OF SHAPE MEASURES

In this section we introduce the shape measures to be used in what follows: the shape index and the curvedness.<sup>9</sup>

Then we introduce some notation to describe a surface patch up to second order.

Many descriptors of shape for the purpose of vision and image processing have been used, but the descriptors that have been used most often are the principal curvatures  $\kappa_{\max}$  and  $\kappa_{\min}$  and the Gaussian and mean curvatures  $K$  and  $H$ .<sup>13</sup> The principal curvatures are the maximum and minimum of the normal curvature  $\kappa_n$ . The normal curvature is obtained as the curvature of the curve that one gets when one cuts the surface with a plane through the normal of the surface. The relation between the normal curvature in direction  $\mathbf{e}_\alpha$  and the principal curvatures is given by Euler's formula<sup>14</sup>:

$$\kappa_n(\alpha) = \kappa_{\max} \cos^2(\alpha - \alpha_0) + \kappa_{\min} \sin^2(\alpha - \alpha_0), \quad (1)$$

with  $\alpha_0 \in [0, \pi)$  denoting the direction of maximal curvature. The direction of minimal curvature is always orthogonal to the direction of maximal curvature.

An important thing to note about these descriptors for curvature is that they are scale dependent; e.g., making a sphere twice as large will change the values of all these descriptors. This is not consistent with our intuition that spheres of different radii have the same shape. Koenderink and van Doorn proposed the following more intuitive descriptors<sup>9</sup>:

$$S = 2/\pi \arctan\left(\frac{\kappa_{\max} + \kappa_{\min}}{\kappa_{\max} - \kappa_{\min}}\right), \quad (2)$$

$$C = \left(\frac{\kappa_{\max}^2 + \kappa_{\min}^2}{2}\right)^{1/2}, \quad (3)$$

with the shape index  $S$  carrying all scale-independent information about shape and the curvedness  $C$  carrying all scale-dependent information. Algebraically one can view the shape index and the curvedness as scaled polar coordinates in the  $\kappa_{\max}, \kappa_{\min}$  half-plane (see Fig. 1). We take  $\kappa_{\max} \geq \kappa_{\min}$  because interchanging  $\kappa_{\max}$  and  $\kappa_{\min}$  has the same effect as interchanging  $x$  and  $y$  [see Eq. (1) with  $\alpha_0 = 0$ ] and thus changes only the orientation of the surface, not its shape.

We parameterize a smooth surface patch, using the range function  $Z(X, Y)$  with  $(X, Y, Z) \in \mathbf{R}^3$ . We write the surface patch up to second order in a Taylor series as follows:

$$Z(X, Y) = Z_0 + Z_X X + Z_Y Y + 1/2 Z_{XX} X^2 + Z_{XY} XY + 1/2 Z_{YY} Y^2, \quad (4)$$

where  $Z_0$  is the distance to the patch, the pair  $(Z_X, Z_Y)^T \equiv \nabla Z$  is the range gradient, and the remaining three parameters  $Z_{XX}$ ,  $Z_{XY}$ , and  $Z_{YY}$  denote the second-order derivatives of the range function. The range gradient is related to the attitude of the surface patch by  $\nabla Z = \tan \sigma (\cos \tau, \sin \tau)^T$ , with  $\sigma \in [0, \pi/2]$  denoting the slant and  $\tau \in [0, 2\pi)$  denoting the tilt. The slant equals the angle between the normal of the surface and the line of sight (we take the direction of the normal always toward the viewer). For a frontoparallel plane we have  $\sigma = 0$ ; for a plane viewed edge on we have  $\sigma = \pi/2$ . The tilt equals the angle between the projection of the normal onto the image plane and the  $X$  axis.

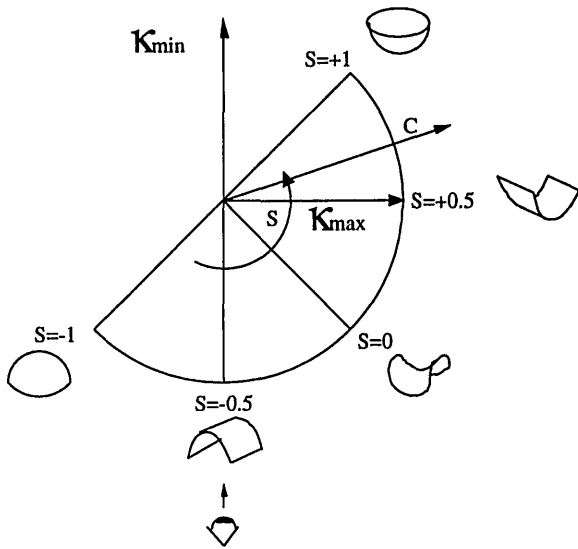


Fig. 1. Objects in  $\kappa_{\max}, \kappa_{\min}$  space. By a change of coordinate system in this space, objects can also be characterized by an angular coordinate only; the shape index  $S$ ; and a radial coordinate, the curvedness  $C$ . The shapes drawn to illustrate various values of  $S$  are to be viewed from below.

The relation between normal curvature and the second-order derivatives of the range function depends also on the first-order derivatives. Taking a fiducial direction  $\hat{e}_\alpha$  in the tangent plane (which has angle  $\alpha$  with the  $X$  axis), one can show that (Ref. 14, p. 221)

$$\kappa_n(\alpha) = \frac{1}{(1 + Z_X^2 + Z_Y^2)^{1/2}} \frac{Z_{XX} \cos^2 \alpha + 2Z_{XY} \cos \alpha \sin \alpha + Z_{YY} \sin^2 \alpha}{(1 + Z_X^2) \cos^2 \alpha + 2Z_X Z_Y \cos \alpha \sin \alpha + (1 + Z_Y^2) \sin^2 \alpha}.$$

Expressing the first-order derivatives in terms of slant and tilt, we find that

$$\kappa_n(\alpha) \frac{1 + \tan^2 \sigma \cos^2(\alpha - \tau)}{\cos \sigma} = Z_{XX} \cos^2 \alpha + 2Z_{XY} \cos \alpha \sin \alpha + Z_{YY} \sin^2 \alpha.$$

This expression can be simplified by introduction of the angles  $\zeta$  and  $\mu$ , defined by<sup>4</sup>

$$\begin{aligned} \cos \zeta &\equiv \cos \sigma [1 + \tan^2 \sigma \cos^2(\alpha - \tau)]^{1/2}, \\ \cos \mu &\equiv 1/[1 + \tan^2 \sigma \cos^2(\alpha - \tau)]^{1/2}. \end{aligned}$$

These angles describe the geometry of the normal of the surface patch relative to the plane (denoted  $\Gamma$ ) determined by the fiducial direction  $\hat{e}_\alpha$  and the line of sight  $\hat{e}_z$  (Fig. 2). The angle between the normal and the plane  $\Gamma$  is  $\zeta$ . Projecting the normal onto  $\Gamma$ , we obtain  $\mu$  as the angle between this projection and  $\hat{e}_z$ . Using these angles, we find that

$$\lambda \equiv \kappa_n / (\cos \zeta \cos^3 \mu) = Z_{XX} \cos^2 \alpha + 2Z_{XY} \cos \alpha \sin \alpha + Z_{YY} \sin^2 \alpha, \quad (5)$$

where we have introduced  $\lambda$  as a new curvature measure. In contrast to  $\kappa_n$ , which depends only on curvature,  $\lambda$  depends also on the attitude of the tangent plane and on

the fiducial direction. Later in the paper we show that it is impossible to obtain  $\kappa_n$  directly from the velocity field but that  $\lambda$  can be obtained directly. Note that Eq. (5) has the same structure as Euler's formula [Eq. (1)]. One can easily show that for

$$\tan(2\alpha_0) = 2Z_{XY} / (Z_{XX} - Z_{YY}) \quad (6)$$

one obtains maximal and minimal  $\lambda$ .  $\alpha_0$  gives the direction of maximal  $\lambda$ . Maximal and minimal  $\lambda$  are given by

$$\begin{aligned} \lambda_{(\max, \min)} &= 1/2(Z_{XX} + Z_{YY}) \\ &\pm 1/2(Z_{XX}^2 + 4Z_{XY}^2 + Z_{YY}^2 - 2Z_{XX}Z_{YY})^{1/2}. \quad (7) \end{aligned}$$

When the tangent plane is frontoparallel, we have  $\sigma = 0$ . In that case  $\lambda_{\max}$  and  $\lambda_{\min}$  in Eq. (7) become equal to  $\kappa_{\max}$  and  $\kappa_{\min}$ . Also, when the slant is not too large,  $\cos \zeta$  and  $\cos \mu$  will be close to 1, and thus  $\lambda_{\max}$  and  $\lambda_{\min}$  will not deviate much from  $\kappa_{\max}$  and  $\kappa_{\min}$  (see Section 5).

### 3. INVARIANT DECOMPOSITION OF THE SECOND-ORDER VELOCITY FIELD

In this section we decompose the second-order velocity field into differential invariants. The complete decomposition of the zeroth- and first-order velocity field has already been reported<sup>1</sup> and leads to four differential invariants: translation of order zero and divergence, rotation, and deformation of order one. For reference we give the expressions of the first-order differential invariants

in terms of derivatives of the velocity field. We denote the divergence by  $\nabla \cdot \mathbf{v}$ , the rotation by  $\nabla \times \mathbf{v}$ , and the deformation by  $\nabla \circ \mathbf{v}$ . We have<sup>1</sup>

$$\nabla \cdot \mathbf{v} = v_x^x + v_y^y, \quad (8)$$

$$\nabla \times \mathbf{v} = -v_y^x + v_x^y, \quad (9)$$

$$\nabla \circ \mathbf{v} = \begin{pmatrix} v_x^x - v_y^y \\ v_y^x + v_x^y \end{pmatrix}, \quad (10)$$

with, e.g.,  $v_y^x$  denoting the first-order derivative in the  $y$  direction of the  $x$  component of the velocity field. It

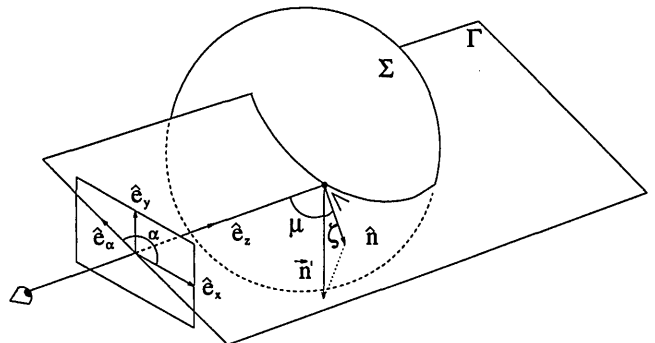


Fig. 2. Geometry of the normal  $\mathbf{n}$  to the surface patch  $\Sigma$  relative to the fiducial plane  $\Gamma$ .  $\Gamma$  is the plane through the fiducial direction  $\hat{e}_\alpha$  and the direction of looking  $\hat{e}_z$ .  $\zeta$  denotes the angle between  $\mathbf{n}$  and  $\Gamma$ .  $\mu$  denotes the angle between the projection ( $\mathbf{n}'$ ) of  $\mathbf{n}$  on  $\Gamma$  and  $\hat{e}_z$ .

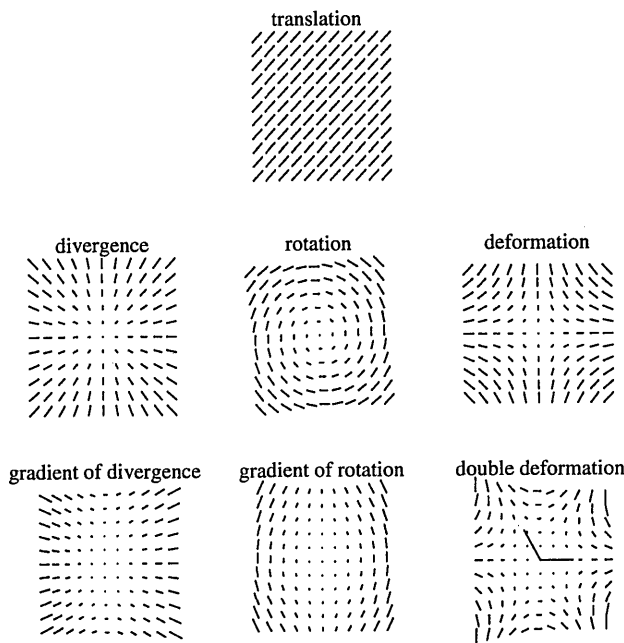


Fig. 3. Flow fields generated by pure differential invariants. Top row: the zeroth-order invariant, translation in direction (1, 1). Middle row: the first-order invariants, divergence, rotation, and the deformation in direction (1, 0). Bottom row, the second-order invariants, all in direction (1, 0). In the flow field of the double deformation we have drawn two lines of equal length, with an angle  $2\pi/3$  between them.

will be an easy exercise to show with the methods introduced below that these are actually invariant. In Fig. 3 we have plotted some examples of the vector field generated by these invariants. Discussing these will give us a chance to introduce the concept of the weight of an invariant.<sup>5</sup> The weight captures the symmetry properties of the invariant under rotation:  $2\pi$  divided by the weight is the smallest angle over which one has to rotate the invariant so that it becomes equal to itself again (except for weight zero). The translation has weight 1, meaning that a rotation over  $2\pi$  maps the vector field onto itself. The divergence and rotation have weight zero, meaning that every rotation maps the vector field onto itself. Thus they are scalars. The deformation has weight 2, meaning that a rotation over  $\pi$  maps the vector field onto itself (see Fig. 3).

The importance of invariance under a rotation stems from the fact that we do not have a preferred direction in the image plane. This is a natural approach because the shape measures listed in Section 2 were also constructed so as to be independent of the choice of coordinate system. A nice bonus of calculating invariants is that invariants have weights, which correspond directly to geometrical properties of the observables. Only linear combinations of invariants of the same weight are invariants. It is important to realize that invariance does not mean that an observable does not change its value when the coordinate system is rotated. One can already observe such a change from the zeroth-order invariant: rotating the coordinate system over  $\theta$  rotates the translation over  $-\theta$ . The observables that do not change value when the coordinate system is rotated are called scalar invariants in this context. Examples are the divergence and the rotation (they have weight zero). The construction of invari-

ants proceeds as follows: we take a set of observables measured relative to some coordinate system. We then rotate the coordinate system over an infinitesimal angle. The observables relative to the new coordinate system will be a linear transformation of the observables relative to the old coordinate system. Denote this transformation  $T$ . We now call our set of observables invariant when  $T$  is a (complex) diagonal matrix. Our notion of invariance is perhaps best illustrated by an example of a quantity that is not invariant (and thus variant). Examples of variants are the spatial derivatives in a Cartesian coordinate system. Because we will come back to the following concept later in the paper, we take the spin variation in the direction of the  $x$  axis  $SV(0) = v_{xx}^y$  as an example of a variant. Rotating  $v_{xx}^y$  over an arbitrary angle  $\theta$ , we find that

$$\begin{aligned}
 SV(\theta) = & \cos^3 \theta (2v_{xy}^x + v_{xx}^y - v_{yy}^y) \\
 & + \sin^3 \theta (-2v_{xy}^y + v_{xx}^x - v_{yy}^x) \\
 & + \cos \theta (-2v_{xy}^x + v_{yy}^y) + \sin \theta (2v_{xy}^y + v_{xx}^x).
 \end{aligned} \tag{11}$$

Clearly, when we rotate the coordinate system over an arbitrary angle,  $v_{xx}^y$  transforms in a complicated way. The transformed  $v_{xx}^y$  depends not only on the original  $v_{xx}^y$  but also on all other second-order spatial derivatives. Thus the transformation  $T$  is not a complex diagonal matrix. The situation is different for the invariants: a transformed invariant depends only on the original invariant, not on the others.

The image plane is two dimensional and has the nice property that rotations commute; i.e., the result of two rotations is independent of the order in which they are performed. Because of this, one can show that all invariants are necessarily complex numbers,<sup>5</sup> with the exception of the invariants of weight zero, which form a pair of real numbers. Because the group of rotations in the plane [SO(2)] is also easily denoted by a complex number of unit length, we construct the invariants by using a complex number notation. We denote the velocity field  $v(z, \bar{z})$ , with  $z = x + iy$  and the complex conjugate  $\bar{z} = x - iy$ . Now  $v$  is a mapping from the complex plane to the complex plane. We write the second-order development of the velocity field as

$$v(z, \bar{z}) = 1/2 v_{zz} z^2 + v_{z\bar{z}} z \bar{z} + 1/2 v_{\bar{z}\bar{z}} \bar{z}^2,$$

with  $v_{zz}$  the velocity field differentiated twice with respect to  $z$  and with similar notation for the other derivatives. Now  $v_{zz}$ ,  $v_{z\bar{z}}$ , and  $v_{\bar{z}\bar{z}}$  are observables of our velocity field, which we measure relative to some coordinate system. To make these observables independent of our choice of coordinate system, we rotate the system with an arbitrary angle and see what happens. As we see below, it turns out that  $v_{zz}$ ,  $v_{z\bar{z}}$ , and  $v_{\bar{z}\bar{z}}$  are already invariant. This actually means that we have been cheating by starting with a favorable representation from the outset. We denote an arbitrary rotation  $U(\theta) = \exp(in\theta)$  with  $U(\theta) \in \text{SO}(2)$ ,  $\theta \in S^1$ , and an arbitrary integer  $n$  as the weight. Rotating  $v(z, \bar{z})$  over  $\theta$ , we find that

$$U(\theta)v[U^{-1}(\theta)z, U^{-1}(\theta)\bar{z}] = \exp(in\theta)\{1/2 v_{zz}[\exp(-in\theta)z]^2 + v_{z\bar{z}}[\exp(-in\theta)z][\exp(in\theta)\bar{z}] + 1/2 v_{\bar{z}\bar{z}}[\exp(in\theta)\bar{z}]^2\}.$$

This results in

$$v(z, \bar{z}) = 1/2 \exp(-in\theta)v_{zz}z^2 + \exp(in\theta)v_{z\bar{z}}z\bar{z} + 1/2 \exp(3in\theta)v_{\bar{z}\bar{z}}\bar{z}^2.$$

Thus we have three differential invariants of second order:  $v_{zz}$  is an invariant of weight  $-1$ ,  $v_{z\bar{z}}$  is an invariant of weight  $1$ , and  $v_{\bar{z}\bar{z}}$  is an invariant of weight  $3$ . The sign of the weight has to do with symmetry under reflection in the  $x$  axis and need not concern us here. One can change the sign of an invariant by complex conjugation. The differential invariants are not unique: linear combinations of invariants with the same weight are also invariant.

By rewriting the differential invariants in real coordinates, we can get a better picture of them. As above, we use superscripts to denote the components of the velocity field  $v$ ; e.g.,  $v^x$  denotes the  $x$  component. We use subscripts to denote spatial derivatives; e.g.,  $v_{yy}^x$  denotes the  $x$  component twice differentiated in the  $y$  direction. The reader can easily verify the following relations:

$$2(\bar{v}_{zz} + v_{z\bar{z}}) = \begin{pmatrix} v_{xx}^x + v_{yy}^y \\ v_{xy}^x + v_{yx}^y \end{pmatrix} = \nabla(\nabla \cdot \mathbf{v}), \quad (12)$$

$$2i(\bar{v}_{zz} - v_{z\bar{z}}) = \begin{pmatrix} -v_{xy}^x + v_{yx}^y \\ -v_{yy}^x + v_{xx}^y \end{pmatrix} = \nabla(\nabla \times \mathbf{v}), \quad (13)$$

$$4v_{\bar{z}\bar{z}} = \begin{pmatrix} v_{xx}^x - v_{yy}^y - 2v_{xy}^y \\ 2v_{xy}^x + v_{xx}^y - v_{yy}^x \end{pmatrix} = \nabla \circ \nabla \circ \mathbf{v}. \quad (14)$$

Here we have introduced the three second-order differential invariants of the velocity field. The gradient of the divergence  $\nabla(\nabla \cdot \mathbf{v})$  and the gradient of the rotation  $\nabla(\nabla \times \mathbf{v})$  have already been introduced in Ref. 1. The double deformation  $\nabla \circ \nabla \circ \mathbf{v}$  is the remaining differential invariant.

In Fig. 3 (bottom row, left-hand panel) we have plotted the vector field of a pure gradient of the divergence. In Section 4 we show this invariant to depend on both the second- and the lower-order terms of the geometry of the surface. Although the dependence on the curvature is the most important, we can use the first-order term to gain an intuitive idea of the vector field generated by this invariant [see Eq. (16) below]: imagine a frontoparallel plane rotating around an axis in the plane. Then part of the plane comes toward the observer (leading to positive divergence) and part goes away from the observer (leading to negative divergence). For the flow pattern of Fig. 3 (bottom row, left-hand panel) the rotation would be around a vertical axis. Note that the vector field thus generated leads not only to a pure gradient of divergence but also to a nonzero gradient of rotation (see below). Thus the divergence changes with position, and we have a gradient. The gradient of the divergence has weight  $1$ ; thus it is a normal vector: only rotation over  $2\pi$  transforms it into itself.

The vector field of a pure gradient of the rotation is plotted in Fig. 3 (bottom row, middle panel). Just as for

the gradient of the divergence, this invariant depends on both the second- and lower-order terms of the geometry of the surface, with the second-order term being the important one. Again, we can get an idea of the vector field generated by this invariant by considering a rotating plane: in the part moving toward the observer the velocity field will not only be oriented away from the axis of rotation but will also be slightly curved inward. On the other side of the axis of rotation the velocity field will be slightly curved outward. Thus the direction of rotation is different at opposite sides of the point of fixation, leading to a gradient in the rotation. This is harder to imagine than for the gradient of divergence, because the noncurvature-dependent terms are three times as small for the gradient of rotation [see Eq. (17) below]. The gradient of the rotation has weight  $1$ ; thus it is a normal vector.

In Fig. 3 (bottom row, right-hand panel) we have plotted the vector field of a pure double deformation. In Section 4 we show that this invariant depends only on the shape of the object: it is zero when the object is not curved. The double deformation has weight  $3$ , and a rotation over  $2\pi/3$  transforms it into itself. As this is difficult to see directly, we have drawn two additional lines under an angle of  $2\pi/3$  in the plot of the double deformation.

The tools that we have introduced above are very powerful: it is easy to see that the invariants of third order have weights  $-2$ ,  $0$ ,  $2$ , and  $4$ . One should be able to identify the invariants of weights  $-2$  and  $2$  with the deformation of the gradients of rotation and divergence, respectively. The invariant of weight  $4$  could be identified with the triple deformation. Finally, the pair of scalar invariants of weight zero could be identified with the divergence of the gradient of the divergence or rotation.

## 4. CONSTRUCTION OF THE SHAPE INDEX

In this section we calculate the velocity field generated by a moving rigid surface patch under perspective projection, and we calculate the invariants of the velocity field. From these invariants we obtain an approximation to the shape index and to some other properties related to shape.

### Velocity Field and the Second-Order Invariants

The expression of the velocity field on a planar camera has been derived in many studies, e.g. Refs. 2 and 3. We consider a moving observer viewing a stationary object. We locate the origin of our coordinate system at the vertex of perspective projection and the positive  $Z$  axis along the line of sight. Representing the patch by  $Z(X, Y)$  and subjecting the observer to a translation  $\mathbf{V} = (V^X, V^Y, V^Z)^T$  and rotation  $\Omega = (\Omega^X, \Omega^Y, \Omega^Z)^T$ , we find for the velocity field on a planar camera at unit focal distance from the point of projection that

$$\begin{pmatrix} v^x \\ v^y \end{pmatrix} = \begin{pmatrix} V^Z/Zx - V^X/Z + \Omega^X xy - \Omega^Y(1+x^2) + \Omega^Z y \\ V^Z/Zy - V^Y/Z + \Omega^X(1+y^2) - \Omega^Y xy - \Omega^Z x \end{pmatrix}, \quad (15)$$

with  $x = X/Z$ ,  $y = Y/Z$  Cartesian coordinates on the

camera. Here, to keep the results general, we have introduced the 3D translation  $\mathbf{V}$  and the 3D rotation  $\Omega$  as independent. In Section 5 below we assume fixation. Now, using the definition of the surface patch in Eq. (4), we approximate  $1/Z$  in camera coordinates<sup>15</sup>:

$$1/Z(x, y) = (1/Z_0)(1 - Z_X x - Z_Y y) - 1/2 Z_{XX} x^2 - Z_{XY} xy - 1/2 Z_{YY} y^2 + O^3(x, y).$$

Substituting this relation into Eq. (15) and then using Eqs. (12)–(14), we get for the second-order differential invariants

$$\nabla(\nabla \cdot \mathbf{v}) = \begin{bmatrix} Z_{XX} & Z_{XY} \\ Z_{XY} & Z_{YY} \end{bmatrix} \mathbf{V}_{\parallel} + 3J\Omega_{\parallel} - 3V_{\perp}/Z_0 \nabla Z, \quad (16)$$

$$\nabla(\nabla \times \mathbf{v}) = - \begin{bmatrix} Z_{XX} & Z_{XY} \\ Z_{XY} & Z_{YY} \end{bmatrix} J\mathbf{V}_{\parallel} - \Omega_{\parallel} - V_{\perp}/Z_0 J\nabla Z, \quad (17)$$

$$\nabla \circ \nabla \circ \mathbf{v} = \begin{bmatrix} Z_{XX} - Z_{YY} & -2Z_{XY} \\ 2Z_{XY} & Z_{XX} - Z_{YY} \end{bmatrix} \mathbf{V}_{\parallel}, \quad (18)$$

where we have used the notation  $\mathbf{V}_{\parallel} \equiv (V^X, V^Y)^T$  for translation parallel to the camera and  $V_{\perp}$  for translation orthogonal to the camera and identical notation for  $\Omega$ . The matrix

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

denotes a rotation over  $\pi/2$ . The first two expressions have already been derived,<sup>1</sup> albeit in a spherical coordinate system. Note that the gradients of divergence and rotation depend both on curvature of the patch and on lower-order terms and that the double deformation depends only on curvature.

### Shape Index

The double deformation is related in an interesting way to the curvature of the surface patch. To see this we calculate the length of the double deformation and the angle between double deformation and  $\mathbf{V}_{\parallel}$ :

$$|\nabla \circ \nabla \circ \mathbf{v}| = (Z_{XX}^2 + 4Z_{XY}^2 + Z_{YY}^2 - 2Z_{XX}Z_{YY})^{1/2} |\mathbf{V}_{\parallel}|, \quad (19)$$

$$\tan \angle(\mathbf{V}_{\parallel}, \nabla \circ \nabla \circ \mathbf{v}) = \frac{-2Z_{XY}}{Z_{XX} - Z_{YY}}. \quad (20)$$

These relations are easily understood from Eqs. (6) and (7). We have

$$|\nabla \circ \nabla \circ \mathbf{v}| = |\lambda_{\max} - \lambda_{\min}| |\mathbf{V}_{\parallel}|, \quad (21)$$

$$\angle(\mathbf{V}_{\parallel}, \nabla \circ \nabla \circ \mathbf{v}) = -2\alpha_0. \quad (22)$$

Geometrically the operation of the double deformation on  $\mathbf{V}_{\parallel}$  can be viewed as a rotation over  $-2\alpha_0$  followed by a scaling with size  $\lambda_{\max} - \lambda_{\min}$ . For slants that are not too large,  $\cos \zeta$  and  $\cos \mu$  are close to 1, and it follows from Eq. (5) that  $\lambda_{\max}$  and  $\lambda_{\min}$  are close to  $\kappa_{\max}$  and  $\kappa_{\min}$  (see Section 5 below). Looking back at the definition of the shape index [Eq. (2)], we can see that we also need the sum of the two principal curvatures. This can be found

from a linear combination of the gradients of divergence and rotation. We define<sup>1</sup>

$$\beta \equiv \nabla(\nabla \cdot \mathbf{v}) + J\nabla(\nabla \times \mathbf{v}) = \mathbf{v}_{xx} + \mathbf{v}_{yy}.$$

This is actually the differential invariant  $v_{zz}$  used above. We get

$$\beta = \begin{bmatrix} Z_{XX} + Z_{YY} & 0 \\ 0 & Z_{XX} + Z_{YY} \end{bmatrix} \mathbf{V}_{\parallel} + 2J\Omega_{\parallel} - 2V_{\perp}/Z_0 \nabla Z. \quad (23)$$

Neglecting the last two terms for the moment (these terms are zero under parallel projection; see below), we have

$$|\beta| = |\lambda_{\max} + \lambda_{\min}| |\mathbf{V}_{\parallel}|.$$

Geometrically the operation of  $\beta$  on  $\mathbf{V}_{\parallel}$  can be viewed as a scaling with size  $\lambda_{\max} + \lambda_{\min}$ . Assuming that  $\lambda_{\max}$  and  $\lambda_{\min}$  are close to  $\kappa_{\max}$  and  $\kappa_{\min}$ , we can extract the absolute value of the shape index by

$$|S_e| = 2/\pi \arctan\left(\frac{|\beta|}{|\nabla \circ \nabla \circ \mathbf{v}|}\right). \quad (24)$$

The sign of the shape index can be obtained from the relation between  $\beta$  and  $\mathbf{V}_{\parallel}$ . Because  $\beta$  is always parallel to  $\mathbf{V}_{\parallel}$ , we give  $S_e$  a positive sign when  $\beta$  and  $\mathbf{V}_{\parallel}$  point in the same direction, and we give  $S_e$  a negative sign when the angle between  $\beta$  and  $\mathbf{V}_{\parallel}$  is  $\pi$ . So all we need is an estimate of the direction of  $\mathbf{V}_{\parallel}$ . Unfortunately, this cannot be obtained from second-order optic flow when we neglect the noncurvature-dependent terms, as can be seen geometrically by introduction of the sister of  $\beta$  (called  $\alpha$ ), which we define by

$$\alpha \equiv M[\nabla(\nabla \cdot \mathbf{v}) - J\nabla(\nabla \times \mathbf{v})],$$

with  $M$  a reflection in the  $x$  axis ( $\alpha$  is the differential invariant  $v_{zz}$  used above). We neglect the noncurvature-dependent terms, just as we did for  $\beta$ , but it should be noted that they are twice as large for  $\alpha$  as for  $\beta$ . It is easy to show that the operation of  $\alpha$  on  $\mathbf{V}_{\parallel}$  can be viewed as a rotation over  $2\alpha_0$  followed by a scaling with size  $\lambda_{\max} - \lambda_{\min}$ . Thus one could also construct an estimate of the shape index from  $\alpha$  and  $\beta$  at the expense of some extra approximations. More interestingly, we find the axial direction of  $\mathbf{V}_{\parallel}$  as the bisector of  $\alpha$  and  $\nabla \circ \nabla \circ \mathbf{v}$ . Unfortunately, we do not know which direction to take on the bisector, leaving the sign of  $\mathbf{V}_{\parallel}$  undetermined. For now, we assume that we obtained the direction of  $\mathbf{V}_{\parallel}$  by some other means, but we will come back to this issue below. We obtain a signed estimate of the shape index from

$$S_e = \text{sign}(\beta \cdot \mathbf{V}_{\parallel}) 2/\pi \arctan\left(\frac{|\beta|}{|\nabla \circ \nabla \circ \mathbf{v}|}\right). \quad (25)$$

It turns out to be possible to derive an approximation to the shape index directly from the velocity field, without the need to calculate any other parameter, not even the absolute value of the 3D translation velocity. It is important to realize that  $S_e$  is defined only for curved surfaces and does not signal the presence of any significant curvature. Thus the shape index is not relevant for tasks

in which curved surfaces have to be discriminated from planar ones.

**Direction of Maximal Curvature and the Curvedness**

In the derivation above we have also found another shape characteristic: the direction of maximal curvature  $\alpha_0$  [see Eq. (6)]. Because  $\beta$  and  $\mathbf{V}_{\parallel}$  are parallel, we obtain an estimate of  $\alpha_0$  by

$$2\alpha_e = \angle(\nabla \circ \nabla \circ \mathbf{v}, \text{sign}(\beta \cdot \mathbf{V}_{\parallel})\beta).$$

Thus we can also find an approximation to the direction of maximal curvature directly from the velocity field. Note that when we do not know the direction of  $\mathbf{V}_{\parallel}$ , we can only obtain the orientations of the principal curvatures, but we do not know which orientation is the direction of maximal curvature and which of minimal curvature. Below we show that this ambiguity is connected with the sign ambiguity of the shape index.

Furthermore, it is easy to derive the following estimate for the velocity scaled curvedness, denoted by  $\Gamma_e$ :

$$\Gamma_e \equiv C_e |\mathbf{V}_{\parallel}| = 1/2(|\nabla \circ \nabla \circ \mathbf{v}|^2 + |\beta|^2)^{1/2}.$$

As we have already stated in Section 1, it is impossible to obtain metric information from the velocity field. Here we obtain an approximation to the curvedness scaled with the absolute value of the 3D velocity parallel to the camera.

**Slant and Tilt**

In order to find relations for the attitude, we need the first-order differential invariants. We have from Eqs. (15) and (8)–(10)

$$\nabla \cdot \mathbf{v} = 1/Z_0(\nabla Z \cdot \mathbf{V}_{\parallel} + 2V_{\perp}), \tag{26}$$

$$\nabla \times \mathbf{v} = 1/Z_0(\nabla Z \cdot J\mathbf{V}_{\parallel}) + 2\Omega_{\perp}. \tag{27}$$

$$\nabla \circ \mathbf{v} = 1/Z_0 \begin{bmatrix} Z_X & -Z_Y \\ Z_Y & Z_X \end{bmatrix} \mathbf{V}_{\parallel}. \tag{28}$$

Just as for the double deformation, it is interesting to look at the length of the deformation and its angle with  $\mathbf{V}_{\parallel}$ :

$$|\nabla \circ \mathbf{v}| = 1/Z_0(Z_X^2 + Z_Y^2)^{1/2} |\mathbf{V}_{\parallel}| = 1/Z_0 \tan \sigma |\mathbf{V}_{\parallel}|, \\ \angle(\nabla \circ \mathbf{v}, \mathbf{V}_{\parallel}) = \angle(\nabla Z, \mathbf{e}_x) = \tau. \tag{29}$$

The first of these equations gives the velocity scaled slant. We obtain an estimate of the tilt from

$$\tau_e = \angle(\nabla \circ \mathbf{v}, \text{sign}(\beta \cdot \mathbf{V}_{\parallel})\beta).$$

Note that when we do not know the direction of  $\mathbf{V}_{\parallel}$  we can obtain only the axis to which  $\tau$  is restricted: both  $\tau$  and  $\pi + \tau$  lead to the same estimate of the tilt. Below we show that this ambiguity is connected with the sign ambiguity of the shape index.

**Velocity Field under Parallel Projection**

In our derivation of a linear estimate for the shape index we had to make the assumption that the noncurvature-dependent terms in  $\beta$  are small, that  $\lambda_{\max}$  and  $\lambda_{\min}$  are

close to  $\kappa_{\max}$  and  $\kappa_{\min}$ , and that we can obtain the direction of  $\mathbf{V}_{\parallel}$ . We also indicated an ambiguity in the shape index, in the orientation of the principal curvatures, and in the tilt when the direction of  $\mathbf{V}_{\parallel}$  is unknown. We will show that the assumption that the noncurvature-dependent terms in  $\beta$  are zero is identical to the use of parallel (orthographic) projection in the derivation of the velocity field. From the velocity field under parallel projection it is easy to understand the ambiguities.

The expression of the velocity field of a rigid moving second-order surface patch under parallel projection is derived in Ref. 4. Because it is important to our argument, we give a summary of this derivation. Representing the patch by  $Z(X, Y)$  and subjecting the observer to a translation  $\mathbf{V} = (V^X, V^Y, V^Z)^T$  and rotation  $\Omega = (\Omega^X, \Omega^Y, \Omega^Z)^T$ , we find for the velocity field on a planar camera under parallel projection

$$\begin{pmatrix} v^x \\ v^y \end{pmatrix} = \begin{pmatrix} -V^X - \Omega^Y Z + \Omega^Z y \\ -V^Y + \Omega^X Z - \Omega^Z x \end{pmatrix}.$$

Note that the roles of  $\mathbf{V}$  and  $\Omega$  are reversed when we compare this velocity field with the velocity field that we obtained under perspective projection: under parallel projection  $\Omega$  generates information about the structure of the object, and under perspective projection  $\mathbf{V}$  generates this information. Introducing polar coordinates  $r, \alpha$  in the image plane, we can write this succinctly as

$$\mathbf{v}(r, \alpha) = -\mathbf{V}_{\parallel} - \Omega_{\perp} J\mathbf{e}_{\alpha} r + J\Omega_{\parallel} Z.$$

Now, substituting the Taylor development of  $Z$  [Eq. (4)] and the expressions for the attitude and principal curvatures [Eq. (5)] and rewriting them in polar coordinates, we find that

$$\mathbf{v}(r, \alpha) = -\mathbf{V}_{\parallel} + J\Omega_{\parallel} Z_0 + [-\Omega_{\perp} J\mathbf{e}_{\alpha} + (\nabla Z, \mathbf{e}_{\alpha}) J\Omega_{\parallel}] r \\ + \kappa_n(\alpha) \frac{1 + \tan^2 \sigma \cos^2(\alpha - \tau)}{\cos \sigma} J\Omega_{\parallel} r^2/2.$$

It is now easy to calculate the second-order invariants, as in Eqs. (16)–(18). The second-order invariants have only the curvature-dependent terms but are otherwise similar to those in Eqs. (16)–(18). Further, one can easily show that the zeroth order disappears when the observer fixates and that the first- and second-order velocity fields  $\mathbf{v}(r, \alpha)$  do not change when we make the following substitutions:

$$\begin{aligned} \Omega_{\parallel} &\rightarrow -\Omega_{\parallel}, \\ \kappa_n(\alpha) &\rightarrow -\kappa_n(\alpha), \\ \tau &\rightarrow \pi + \tau. \end{aligned}$$

The change of the tilt by  $\pi$  is equivalent to a change of the sign of  $\nabla Z$ . Because of Euler's formula [Eq. (1)], the substitution  $\kappa_n(\alpha) \rightarrow -\kappa_n(\alpha)$  is equivalent to

$$\begin{aligned} \kappa_{\max} &\rightarrow -\kappa_{\min}, \\ \kappa_{\min} &\rightarrow -\kappa_{\max}, \\ \alpha_0 &\rightarrow \pi/2 + \alpha_0. \end{aligned}$$

Thus the velocity field of a surface patch under parallel projection is ambiguous in exactly the same way as



we found above when we neglected the noncurvature-dependent terms in  $\beta$  and  $\alpha$ . This should come as no great surprise, as we have shown that the second-order invariants under parallel projection contain only curvature-dependent terms. The ambiguities are easily imagined one by one: the concave/convex ambiguity is easily understood for rotating spheres with a frontoparallel tangent plane, the interchange of the directions of maximal and minimal curvature is easily understood for a rotating symmetric saddle with a frontoparallel tangent plane, and the tilt ambiguity is easily understood with a rotating plane.

As we showed above, the ambiguities can be lifted once we know the direction of  $\mathbf{V}_{\parallel}$ . A detailed discussion of how this direction can be obtained is beyond the scope of this paper. There are two possibilities regarding how the information about the direction of  $\mathbf{V}_{\parallel}$  could be obtained. The first is the use of extraretinal signals, such as knowledge that one's left eye is at the left of one's right eye in stereo vision or knowledge of movement direction in active vision. The second possibility is to use a third view, as described in Ref. 16.

## 5. SIMULATIONS

In deriving the various estimates related to shape, we made some approximations. In this section we show the effects of these approximations to be generally small for a field of view of 6 deg. We do not give analytic expressions for the effect of the approximations, because they tend to get unwieldy, but we rely on computer simulations instead.

In this section we assume fixation and zero torsion; i.e., we take

$$\Omega_{\parallel} = -J\mathbf{V}_{\parallel}/Z_0, \quad (30)$$

$$\Omega_{\perp} = 0. \quad (31)$$

We take quadratic surface patches. It should be noted that these patches generally do not have a constant shape index: except for the parabolic (cylinderlike) patches, shape index varies with position. This is not a big problem, though, as long as the field of view is small. To make a comparison with experimental results, we took most of the parameters from Ref. 17: a distance of 2.5 m, slant zero, a curvedness of  $5 \text{ m}^{-1}$ , a direction of maximal curvature in the  $x$  direction, and a velocity of 1 m/s in the  $x$  direction. We took a square field of view of  $6 \text{ deg} \times 6 \text{ deg}$ , whereas in the experiment the field of view was circular. When we take a nonzero slant, we always take zero tilt (and thus rotate the tangent plane around a horizontal axis). We use this set of parameters in all simulations reported below, unless otherwise noted. We took a square grid of  $5 \times 5$  in the image plane and calculated the velocity in each grid point. The results did not depend on the number of grid points, except when the curvature was high. Therefore we took an  $11 \times 11$  grid for the curve for curvedness  $30 \text{ m}^{-1}$  in Fig. 4. We fitted the velocity field with a polynomial up to second order, using a linear least-squares algorithm. From the fitted parameters we calculated the differential invariants and from there the estimates of the shape index  $S_e$ , the velocity scaled curvature  $\Gamma_e$ , and the angle between direction of maximal curvature and the  $x$  axis,  $\alpha_e$ . In the experiment of

van Damme and van de Grind,<sup>17</sup> the subject was shown many views of the object and was actively moving. Thus we can assume that the subject knows the direction of movement, and we take the correct direction of  $\mathbf{V}_{\parallel}$  in our simulations.

The vector  $\beta$  enters into all our shape measures. Therefore it is important to show the effect of the noncurvature-dependent terms in  $\beta$ . Substituting Eq. (30) into Eq. (23), we find that

$$\beta = (Z_{XX} + Z_{YY} + 2/Z_0)\mathbf{V}_{\parallel} - 2(V_{\perp}/Z_0)\nabla Z. \quad (32)$$

There are two terms in this equation that we previously neglected: a zeroth-order term  $2/Z_0$ , which we assumed to be small relative to  $Z_{XX} + Z_{YY}$ , and a first-order term  $2(V_{\perp}/Z_0)\nabla Z$ , which we assumed to be small relative to  $(Z_{XX} + Z_{YY})\mathbf{V}_{\parallel}$ . Note that the zeroth-order term changes only the length of  $\beta$ , never its direction. The first-order term has a more complicated influence on  $\beta$ , depending on the tilt (which gives the direction of  $\nabla Z$ ) and  $\mathbf{V}_{\parallel}$ . When  $\nabla Z$  and  $\mathbf{V}_{\parallel}$  are parallel, the first-order term changes only the length, not the direction, of  $\beta$ . When  $\nabla Z$  and  $\mathbf{V}_{\parallel}$  are orthogonal, the first-order term influences not so much the length as the direction of  $\beta$ .

### Shape Index

First, we show the effect of neglecting the zeroth-order term in Eq. (32) on the estimate of the shape index  $S_e$ . In Fig. 4 we have plotted the bias in shape index (difference between the estimated shape index and the shape index in the fixation point) for several values of the curvedness. Note that we would have obtained the same curves for different values of  $Z_0$ : keeping the curvedness constant at  $5 \text{ m}^{-1}$  and  $Z_0$  equal to 15, 7.5, 5, and 2.5 m results in the same curves. The bias is zero at the extremes of the shape-index scale and increases toward the middle. One can show that the effect that the bias is zero for the spherical shapes is caused by the vanishing of the double deformation at the extremes of the shape-index scale. The bias is smaller for higher values of the curvedness. Incidentally, one can also see the problems that one would have for a series development of the bias: the lowest

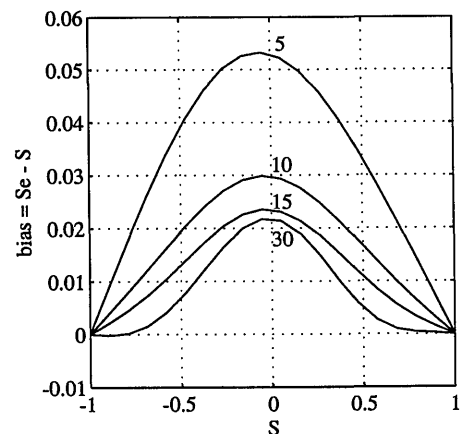


Fig. 4. Bias in shape index (difference between the estimated shape index and the shape index at the fixation point) as a function of shape index at the fixation point. The different curves are for different values of the curvedness: 5, 10, 15, and  $30 \text{ m}^{-1}$ .

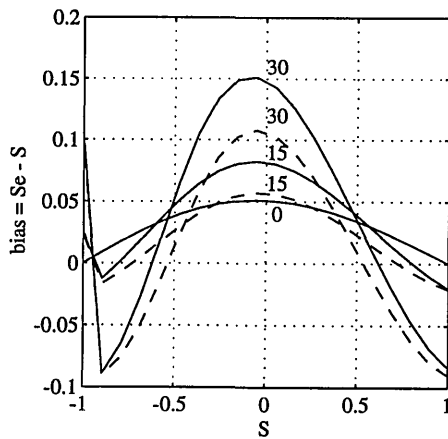


Fig. 5. Bias in shape index as a function of shape index at the fixation point. The different curves are for different combinations of the slant and  $V_{\perp}$ . The two solid curves labeled 30 and 15 are for slant 30 and 15 deg with  $V_{\perp} = 1$ ; the remaining solid curve is for slant zero ( $V_{\perp}$  does not matter). The dashed curves are for slant 30 and 15 deg with  $V_{\perp} = -1$ . The other parameters are curvedness  $5 \text{ m}^{-1}$ , tilt zero, distance 2.5 m, and  $\mathbf{V}_{\parallel} = (1, 0) \text{ m/s}$ .

curve (for the highest value of curvedness) dips below zero for  $S$  around  $-0.9$ , indicating that one would have to go to fourth order in this series development (the effect is even stronger for higher values of curvedness). The effect that the bias is slightly larger for cylindrical and hyperbolic shapes than for elliptic shapes is reported experimentally in Refs. 17 and 18. Increasing the curvature decreases the bias, a weak effect that was also reported. Finally, note that the bias is slightly larger for concave than for convex surfaces.<sup>17</sup>

Second, we show the effect of neglecting the first-order term in Eq. (32) on the estimate of the shape index  $S_e$ . In Fig. 5 we have plotted the bias in shape index for different values of the slant and of the sign of  $V_{\perp}$ . It should be noted that the bias is caused not only by the first-order term but also by the zeroth-order term and by the difference between  $\lambda$  and  $\kappa$ . In fact, as we show in the next paragraph, the latter effect is larger than the effect of the first-order term. The curve for zero slant is the same as the upper curve in Fig. 4. We see that when  $V_{\perp} > 0$ , i.e., when the observer is moving toward the surface, the bias is smaller than when  $V_{\perp} < 0$ . This is the case because the first-order terms have different signs when  $V_{\perp} > 0$ . In the figure we have plotted the best- and worst-case situations: when  $\nabla Z$  and  $\mathbf{V}_{\parallel}$  are not parallel we find intermediate values of the bias. The sharp discontinuity in the bias around shape index  $-0.9$  is discussed in the next paragraph.

A third approximation that we made in deriving  $S_e$  was to neglect the slant; i.e., we assumed  $\lambda_{\max}$  and  $\lambda_{\min}$  to be close to  $\kappa_{\max}$  and  $\kappa_{\min}$ . In Fig. 6 we have plotted the bias in shape index for different values of the slant. It should be noted that the bias is caused not only by the difference between  $\lambda$  and  $\kappa$  but also by the zeroth-order term. We observe the effect of slant to be very small for the cylindrical shapes. This can be understood from the fact that one of the two principal curvatures is zero. Thus one of the  $\lambda$ 's is zero, and the ratio in the argument of arctan in the definition of the shape index [Eq. (2)] equals 1. We observe that the slant increases

the bias relative to the bias of the zeroth order for the hyperbolic shapes and decreases the bias relative to the bias of the zeroth order for the spherical shapes. The sharp discontinuity in the bias around shape index  $-0.9$  is caused by a change of the sign of the double deformation. Also, the bias cannot be very negative there, because the shape index cannot become smaller than  $-1$ . Still, the bias does not exceed 0.1 around shape index  $-0.9$ . Even a slant of 30 deg does not have much influence on the shape index, a result reported in Ref. 19.

To illustrate the stability of our estimate of the shape index against noise, we added 10% multiplicative Gaussian white noise to the velocity field (independent for the  $x$  and  $y$  directions). We ran 250 simulations for each value of the shape index and calculated the mean and standard deviation of  $S_e$ . In Fig. 7 we have plotted the difference between the mean of  $S_e$  and the shape index in the fixation point. Because we used the term bias for the deterministic difference, we denote this difference as the statistical bias. In Fig. 7 we have also plotted the bias in shape index from the noiseless simulation. Except for  $S = \pm 1$ , the statistical bias does not deviate from

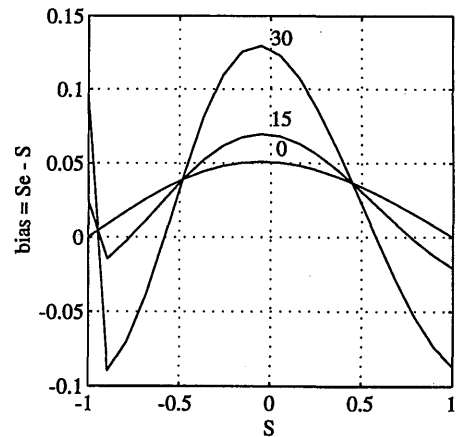


Fig. 6. Bias in shape index as a function of shape index at the fixation point. The different curves are for different values of the slant: 0, 15, and 30 deg. The other parameters are curvedness  $5 \text{ m}^{-1}$ , tilt zero, distance 2.5 m, and  $\mathbf{V} = (1, 0, 0) \text{ m/s}$ .

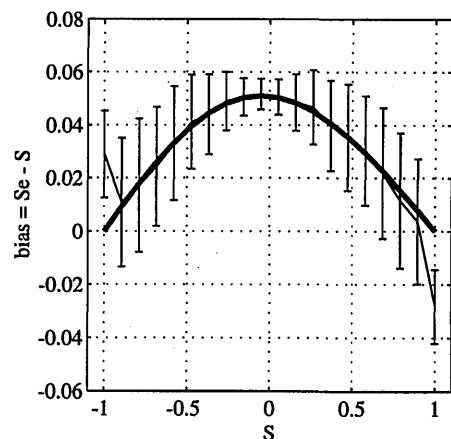


Fig. 7. Bias in shape index as a function of shape index at the fixation point. The thick curve gives the bias without noise; the thin curve gives the mean bias of 250 simulations with 10% multiplicative Gaussian white noise. The error bars denote the standard deviation. The other parameters are curvedness  $5 \text{ m}^{-1}$ , slant zero, distance 2.5 m, and  $\mathbf{V} = (1, 0, 0) \text{ m/s}$ .

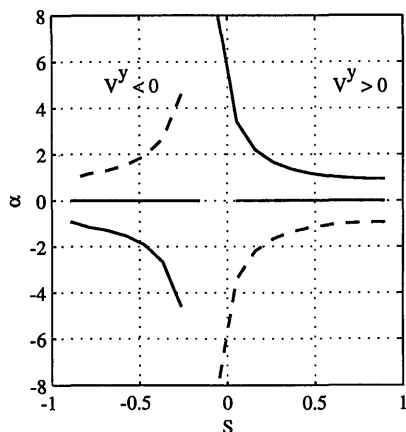


Fig. 8. Bias in direction of maximal curvature relative to the  $x$  axis in degrees as a function of shape index at the fixation point. The middle curve is for slant zero, the dashed discontinuous curve for slant 30 deg and  $\mathbf{V} = (0, -1, 1)$  m/s, and the solid discontinuous curve for slant 30 deg and  $\mathbf{V} = (0, 1, 1)$  m/s. The other parameters are curvedness  $5 \text{ m}^{-1}$ , tilt zero, and distance 2.5 m.

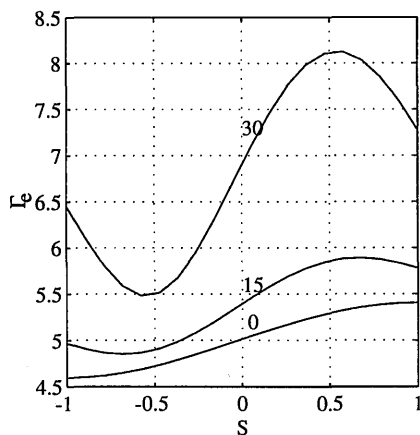


Fig. 9. Velocity scaled curvature  $\Gamma_e$  as a function of the shape index at the fixation point. The correct value is  $5 \text{ s}^{-1}$ . The different curves are for different values of the slant: 0, 15, and 30 deg. The other parameters are curvedness  $5 \text{ m}^{-1}$ , tilt zero, distance 2.5 m, and  $\mathbf{V} = (1, 0, 0)$  m/s.

the bias, showing our estimate to be statistically unbiased. The deviation at  $S = \pm 1$  is caused by the arctan function in the shape index, which is very nonlinear close to  $S = \pm 1$ . The standard deviation seems to be larger for parabolic surfaces than for elliptic or hyperbolic ones. We performed some more simulations with noise (multiplicative and additive) that showed our estimate of the shape index to be highly resistant to noise. Variations of slant, the direction of  $\mathbf{V}_{\parallel}$ , and the curvedness had negligible influence beyond the deterministic bias. Only when we added a large movement orthogonal to the camera ( $V_{\perp} \neq 0$ ) in combination with multiplicative noise did the estimates become very noisy, because the velocities are high in the periphery, and so is the noise.

#### Direction of Maximal Curvature and the Curvedness

In Fig. 8 we have plotted the bias in the direction of maximal curvature (difference between  $\alpha_e$  and  $\alpha_0$  in the fixation point) for different values of the slant and velocity in the  $y$  direction (and for  $V_{\perp} > 0$ ). We have shown above

that  $\alpha_e$  can be calculated from the angle between  $\beta$  and the double deformation. When either of these becomes very small, the angle cannot be determined accurately. Therefore we did not calculate  $\alpha_e$  when the smaller of  $\beta$  and the double deformation was less than 10% of the larger of the two in length. This happens around  $S = \pm 1, 0$  which is why the curves shown are discontinuous at these values. For  $S = \pm 1$  this is not a problem, as  $\alpha_0$  is not defined. That the direction of maximal curvature is hard to determine for hyperbolic shapes is an interesting prediction from our theory. The horizontal curve is for slant zero, the dashed discontinuous curve is for slant 30 deg and  $V^Y < 0$ , and the solid discontinuous curve is for slant 30 deg and  $V^Y > 0$ . In discussing the effects of neglecting the zeroth- and first-order terms in  $\beta$ , we have already noted that only the first-order term can change the direction of  $\beta$  when  $\nabla Z$  and  $\mathbf{V}_{\parallel}$  are not parallel. We have chosen  $\nabla Z$  and  $\mathbf{V}_{\parallel}$  orthogonal (the worst case), and still the difference does not exceed 8 deg.

In Fig. 9 we have plotted the estimate of velocity scaled curvedness  $\Gamma_e$  for different values of the slant. The value that we used for the simulation was  $5 \text{ s}^{-1}$ . The figure shows  $\Gamma_e$  to be quite sensitive for slant. For higher values of the slant,  $\Gamma_e$  leads to a considerable overestimation of the true value. Further,  $\Gamma_e$  increases with  $S$  for slant zero. One can easily show this to be caused by the zeroth-order term in  $\beta$ .

## 6. DISCUSSION

We have found that we can obtain an approximation of the shape index, a scale-independent descriptor of 3D shape, directly from the velocity field generated by a curved patch moving rigidly relative to a planar camera. We have tried to strike a middle ground between human perception and machine vision. With our results we are able to explain some of the outcomes of psychophysical experiments (see below). We think that our model is relevant for machine vision for those cases in which one does not know the movement of the camera through 3D space, e.g., for imaging techniques with hand-held cameras.

#### Comparison with Psychophysical Findings

Shape index and curvedness have been used in a number of psychophysical experiments in which both optic flow<sup>17,20</sup> and stereo vision<sup>18,19</sup> were used. In these experiments it was established that human observers can use shape index and curvedness quite independently of each other. In a detection experiment<sup>17,18</sup> in which subjects were asked to classify a surface patch of unknown shape index in one of eight shape-index categories, it was found that subjects could more readily classify the elliptic shapes. Performance was lower for the cylindrical and hyperbolic shapes. Performance increased only slightly with increasing curvedness. Furthermore, it was remarked<sup>17</sup> that the velocity of the subject did not seem to have a significant influence. In another experiment<sup>19</sup> subjects were shown two patches, and their task was to detect which one was the reference patch. By use of this paradigm and stereo vision, it was established that the discrimination of shape index is independent of the slant of the surface patch for slants up

to 30 deg. All these results are in qualitative agreement with the results of the simulations in Section 5.

**Comparison with the Approach of Koenderink and van Doorn**

As we stated in Section 1, our approach is close to that of Koenderink and van Doorn.<sup>1,4</sup> We improved on their 1975 paper by introducing the remaining invariant of the second-order velocity field: the double deformation. We showed this invariant to be proportional to the difference between the maximal and minimal curvatures. From this we were able to construct an approximation of the shape index directly from the velocity field. In their original approach they could calculate only the sign of the Gaussian curvature, which would be the equivalent of calculating whether the absolute value of the shape index is smaller or larger than 0.5.

Our approach is somewhat different from the approach of Ref. 4. First, Koenderink and van Doorn employed parallel projection, whereas we used both perspective and parallel projection, which allowed us to estimate the difference. For parallel projection we found the same expressions for the second-order differential invariants as we did for perspective projection, except for the gradients of divergence and rotation, which retained only the curvature-dependent term. A second difference is that the approach of Koenderink and van Doorn results in the projected indicatrix of Dupin, scaled by the slant and the 3D velocity, whereas some of our shape measures (the shape index and the orientation of the principal curvatures) are relatively independent of both slant and 3D velocity.

**Comparison with Spin Variation Theory**

A recent theory on the perception of curvature from optic flow is based on the concept of spin variation,<sup>12,21</sup> i.e., the bending of lines in the image. The spin variation is not a geometrical object like the differential invariants, but it is defined independently of a coordinate system. Therefore it can be expressed in differential invariants as follows. The spin variation in the direction of the *x* axis is given by

$$SV(0) = v_{xx}^y.$$

By rotating this over an arbitrary angle  $\theta$  we obtain the spin variation function [see Eq. (11)]. Because we know that the spin variation depends only on second-order spatial derivatives of the range function, we construct differential invariants that have this property, too. The double deformation already has this property: it depends only on second-order spatial derivatives of the range function. We define

$$\delta \equiv \nabla(\nabla \cdot \mathbf{v}) + 3J\nabla(\nabla \times \mathbf{v})$$

as a second differential invariant dependent only on second-order derivatives of the range function. Using Eq. (11), we can easily show that

$$SV(\theta) = -1/4 \nabla \circ \nabla \circ \mathbf{v} J \begin{pmatrix} \cos 3\theta \\ \sin 3\theta \end{pmatrix} - 1/4 \delta J \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}. \tag{33}$$

This shows again nicely the invariance of  $\nabla \circ \nabla \circ \mathbf{v}$  (weight 3) and  $\delta$  (weight 1).

Spin variation theory has been compared with the findings of psychophysical experiments in which human subjects had to discriminate cylinders and planes. A central prediction of spin variation theory was the asymmetry in detection thresholds, depending on the axis of the cylinder relative to the direction of movement of the observer. Movement parallel to the cylinder axis leads to lower detection thresholds than movement orthogonal to the cylinder axis. This effect has indeed been observed both in studies employing optic flow<sup>21,22</sup> and in studies employing stereo vision.<sup>23</sup>

From the expression of the spin variation in differential invariants [Eq. (33)], it is easy to see that this asymmetry for cylinders must be due to  $\delta$ , because the double deformation is symmetric [take  $Z_{XY} = 0$  in Eq. (18)]. The vector field  $\delta$  is indeed asymmetric, as can be seen by substitution of Eqs. (16) and (17) in the definition of  $\delta$ . We obtain

$$\delta = \begin{bmatrix} Z_{XX} + 3Z_{YY} & -2Z_{XY} \\ -2Z_{XY} & 3Z_{XX} + Z_{YY} \end{bmatrix} \mathbf{V}_{\parallel}.$$

By taking  $Z_{XY} = 0$  and  $\mathbf{V}_{\parallel} = (1, 0)$  we find that  $\delta$  is three times larger for a horizontal cylinder ( $Z_{XX} = 0$ ) than for a vertical cylinder ( $Z_{YY} = 0$ ). The curvedness  $\Gamma_e$  is constructed from the symmetric vector fields  $\beta$  and  $\nabla \circ \nabla \circ \mathbf{v}$ . Thus our curvedness measure cannot be relevant for this particular aspect of human perception. Precisely because of its symmetry, however, it might be interesting for machine vision.

In summary, the spin variation theory is not so different from the current approach in its mathematics. The difference is more in the emphasis. Whereas spin variation theory stresses more the computational aspects and the detection of curved, versus planar surfaces, our emphasis is more on geometry and on the detection of shape as given by the shape index. Beyond that, our formulation solves a few of the problems that occur in the spin variation theory: in our formulation the shape index is independent of 3D velocity and relatively independent of slant, whereas the spin variation function depends on both.

**Comparison with Discrete Algorithms**

We have employed the velocity field as an input to our calculations; i.e., we assumed a small disparity between two views of a rigid curved surface patch. The discrete algorithms of structure from motion do not make the assumptions of small motion and of a smooth surface patch. Two recent studies in this field are reported in Refs. 24 and 25. For several reasons it is hard to compare our results directly with the results of these studies. First, the discrete algorithms start with the calculation of the velocity and rotation of the camera and, given these, calculate the 3D position of the points. They do not calculate an explicit structure of the environment, although one could do this in an extra step by fitting a model to the 3D points. This contrasts with our approach, in which we calculate the structure of the environment directly. Second, the discrete algorithms employ a two-step approach: in the first step a linear algorithm is used to estimate the velocity and rotation of the camera. Because this estimate is statistically biased, this estimate is used as

starting point for an iteration, which minimizes the discrepancy between fitted image points and observed image points. In contrast, our approach is linear in the parameters. Third, the discrete algorithms are exact, whereas our results are approximate. Notwithstanding these differences, there is an interesting point in the study by Weng *et al.* (Fig. 16 of Ref. 25), where some simulations were done for small image motion. Using parameters comparable with theirs, we found the statistical bias in the estimates to be small, generally of the order of a few percent. This contrasts with their estimate of the bias in the velocity, which can be of the order of 100%, for small image motion.

### Consequences of the Theory

From our approach we are able to derive new hypotheses that can be tested in psychophysical experiments. Most of them are related to the approximations that we used to calculate our estimates of shape index, orientation of the principal curvatures, velocity scaled curvedness, and tilt. These approximations are reasonable for many real-life situations but can cause a breakdown of shape perception in experimental situations. In particular, the effects of the zeroth- and first-order terms in  $\beta$  lead to testable predictions:

1. *Ambiguities.* We showed that our way of calculating shape amounts to the use of parallel projection. This introduces ambiguities in  $V_{\parallel}$ , the shape index, the direction of maximal curvature, and the tilt. In contrast, the curvedness and slant can be extracted without ambiguity. In particular, this means that the difference between a planar and a curved surface can be extracted unambiguously from a two-frame motion sequence. This is relevant, since most psychophysical studies searching for an effect of sequence length have used the curvedness in their tasks (see, e.g., Ref. 26 and the references therein). For a two-frame motion sequence we would predict ambiguities in  $V_{\parallel}$ , the shape index, the direction of maximal curvature, and the tilt but not in the curvedness and slant. These ambiguities could be lifted by the use of longer sequences.

2. *Shape index.* In the calculation of the absolute value of the shape index we neglected the noncurvature-dependent terms in  $\beta$ . The first-order term is especially interesting because it would predict a slant-dependent influence of movement orthogonal to the camera ( $V_{\perp}$ ) on the shape index. More specifically, we would predict no influence of  $V_{\perp}$  for zero slant and an increasing influence with increasing slant, depending on the sign of  $V_{\perp}$ . For  $V_{\perp} > 0$  and for  $\nabla Z$  and  $V_{\parallel}$  parallel, we would predict a decreasing bias with increasing slant because the zeroth- and first-order terms have different signs. For  $V_{\perp} < 0$  we would predict the opposite effect (see Fig. 5).

3. *Orientation of the principal curvatures.* The orientation of the principal curvatures could be recovered from the angle between  $\beta$  and the double deformation. When either of these is small, the angle is hard to determine. The double deformation is small for spheres, which is not a big problem, as the orientation of the principal curvatures is not defined for spheres (all directions have the same curvature). The prediction that the orientation of the principal curvatures is also hard to determine for hyperbolic surfaces where  $\beta$  is small and where the two cur-

vatures are of opposite sign is a surprising consequence of our theory.

Since the zeroth-order term does not change the direction of  $\beta$  but changed only its length, we would predict the estimate of the orientation of the principal curvatures to be independent of 3D velocity, curvedness, and distance, just as for the shape index we predict a slant-dependent influence of movement orthogonal to the camera ( $V_{\perp}$ ). In this case we expect this effect to be strongest when  $\nabla Z$  and  $V_{\parallel}$  are orthogonal (see Fig. 8).

4. *Velocity scaled curvedness.* We already discussed that the velocity scaled curvedness is a symmetrical quantity and thus cannot be used to explain the asymmetry found in the detection of curvature of cylinders. Still, it could be used by human observers for the detection of symmetrical surfaces. An interesting prediction of the theory is that the velocity scaled curvedness would be overestimated for slanted objects and that this overestimation would depend on shape index (see Fig. 9).

5. *Tilt.* The tilt of the tangent plane could be recovered from the angle between  $\beta$  and  $\nabla \circ v$ . We would predict the pattern of dependencies to be the same as for the orientation of the principal curvatures.

### ACKNOWLEDGMENTS

This research was supported by the Foundation for Biophysics and by European Strategic Programme for Research and Development in Information Technology Basic Research Project 6615. We thank Peter Werkhoven for stimulating criticism.

### REFERENCES

1. J. J. Koenderink and A. J. van Doorn, "Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer," *Opt. Acta* **22**, 773–791 (1975).
2. H. C. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image," *Proc. R. Soc. London Ser. B* **208**, 385–387 (1980).
3. A. M. Waxman, B. Kamgar-Parsi, and M. Subbarao, "Closed-form solutions to image flow equations for 3D structure and motion," *Int. J. Comput. Vision* **1**, 239–258 (1987).
4. J. J. Koenderink and A. J. van Doorn, "Second-order optic flow," *J. Opt. Soc. Am. A* **9**, 530–538 (1992).
5. K. Kanatani, *Group-Theoretical Methods in Image Understanding* (Springer-Verlag, Berlin, 1990).
6. W. C. Hoffman, "The Lie algebra of visual perception," *J. Math. Psychol.* **3**, 65–98 (1966).
7. P. C. Dodwell, "The Lie transformation group model of visual perception," *Percept. Psychophys.* **34**, 1–16 (1983).
8. J. L. Gallant, J. Braun, and D. C. Van Essen, "Selectivity for polar, hyperbolic and cartesian gratings in macaque visual cortex," *Science* **259**, 100–103 (1993).
9. J. J. Koenderink and A. J. van Doorn, "Surface shape and curvature scales," *Image Vis. Comput.* **10**, 557–565 (1992).
10. P. Werkhoven and J. J. Koenderink, "Extraction of motion parallax structure in the visual system I & II," *Biol. Cybern.* **63**, 185–199 (1990).
11. T. M. H. Dijkstra, E. Argante, and C. C. A. M. Gielen, "Motion parallax from catastrophes in scale-space," in *Proceedings of the International Conference on Artificial Neural Networks, Amsterdam, The Netherlands*, S. Gielen and B. Kappen, eds. (Springer-Verlag, London, 1993), pp. 237–240.
12. J. Droulez and V. Cornilleau-Pérès, "Visual perception of surface curvature. The spin variation and its physiological consequences," *Biol. Cybern.* **62**, 211–224 (1990).
13. P. J. Besl and R. C. Jain, "Invariant surface characteristics for 3D object recognition and range images," *Comput. Vis. Graphics Image Process.* **33**, 33–80 (1986).

14. B. O'Neill, *Elementary Differential Geometry* (Academic, New York, 1966).
15. M. Subbarao, "Interpretation of image flow: rigid curved surfaces in motion," *Int. J. Comput. Vision* **2**, 77–96 (1988).
16. J. J. Koenderink and A. J. van Doorn, "Affine structure from motion," *J. Opt. Soc. Am. A* **8**, 377–385 (1991).
17. W. J. M. van Damme and W. A. van de Grind, "Active vision and the identification of three-dimensional shape," *Vision Res.* **33**, 1581–1587 (1993).
18. S. C. de Vries, A. M. L. Kappers, and J. J. Koenderink, "Shape from stereo: a systematic approach using quadratic surfaces," *Percept. Psychophys.* **53**, 71–80 (1993).
19. S. C. de Vries, A. M. L. Kappers, and J. J. Koenderink, "Influence of surface attitude and curvature scaling on discrimination of binocularly presented curved surfaces," *Vision Res.* (to be published).
20. W. J. M. van Damme, F. H. Oosterhoff, and W. A. van de Grind, "Discrimination of 3D shape and 3D curvature in active vision," *Percept. Psychophys.* **55**, 340–349 (1994).
21. V. Cornilleau-Pérès and J. Droulez, "Visual perception of surface curvature: psychophysics of curvature detection induced by motion parallax," *Percept. Psychophys.* **46**, 351–364 (1989).
22. J. F. Norman and J. S. Lappin, "The detection of surface curvatures defined by optical motion," *Percept. Psychophys.* **51**, 386–396 (1992).
23. B. Rogers and R. Cagenello, "Disparity curvature and the perception of three-dimensional surfaces," *Nature (London)* **339**, 135–137 (1989).
24. K. Kanatani, "Unbiased estimation and statistical analysis of 3-D rigid motion from two views," *IEEE Trans. Pattern Anal. Mach. Intell.* **15**, 37–50 (1993).
25. J. Weng, N. Ahuja, and T. S. Huang, "Optimal motion and structure estimation," *IEEE Trans. Pattern Anal. Mach. Intell.* **15**, 864–884 (1993).
26. J. T. Todd and J. F. Norman, "The visual perception of smoothly curved surfaces from minimal apparent motion sequences," *Percept. Psychophys.* **50**, 509–523 (1991).