Shading Flows and Scenel Bundles: A New Approach to Shape from Shading

Pierre Breton[1], Lee A. Iverson[1], Michael S. Langer[1], Steven W. Zucker[1,2]

[1] McGill University, Research Center for Intelligent Machine, 3480 rue Université, Montréal, Québec, Canada, H3A 2A7, e–mail: zucker@mcrcim.mcgill.edu
[2] Fellow, Canadian Institute for Advanced Research

Abstract. The classical approach to shape from shading problems is to find a numerical solution of the image irradiance partial differential equation. It is always assumed that the parameters of this equation (the light source direction and surface albedo) can be estimated in advance. For images which contain shadows and occluding contours, this decoupling of problems is artificial. We develop a new approach to solving these equations. It is based on modern differential geometry, and solves for light source, surface shape, and material changes concurrently. Local scene elements (scenels) are estimated from the shading flow field, and smoothness, material, and light source compatibility conditions resolve them into consistent scene descriptions. Shadows and related difficulties for the classical approach are discussed.

1 Introduction

The shape from shading problem is classical in vision; E. Mach (1866) was perhaps the first to formulate a formal relationship between image [1] and scene domains, and to capture their inter-relationships in a partial differential equation. Horn set the modern approach by focusing on the solution of such equations by classical and numerical techniques [2, 3, 4], and others have built upon it [5, 6]. Nevertheless, problems remain which are not naturally treated in the classical sense, especially those related to discontinuities and shadows. We present a new approach to the shape from shading problem motivated by modern notions of fibre bundles in differential geometry [9]. The global shape from shading problem is posed as a coupled collection of "local" problems, each of which attempts to find that local scene element (or scenel[3]) that captures the local photometry, and which are then coupled together to form global piecewise smooth solutions.

The paper is written in a discursive style to convey a sense of the "picture" behind our approach rather than the formal treatment. We begin with an overview of the classical formulation, then proceed to define the structure of a "scenel" and our new conceptualization.

1.1 The "Classical" Shape from Shading Problem

We take the classical setting in computer vision for shape from shading to be the following: a point light source at infinity uniformly illuminates a smooth matte surface of constant albedo whose image is formed by orthographic projection.

The matte surface is traditionally modeled with Lambert's reflectance function so the image irradiance equation is

$$I(x,y) = \rho \lambda \mathbf{L} \cdot \mathbf{N}(x,y)$$

---

[3] *cf.* Pixel, voxel, ... scenel.

where $I(x, y)$ is the intensity of an image point (x, y); ρ, the albedo of the surface, i.e. the fraction of the shining light which is reflected; λ, the illumination, i.e. the amount of shining light; $\mathbf{L}$, the light source direction; $\mathbf{N}(x, y)$, the normal at the surface point corresponding to an image point (x, y).

The literature cited above describes the various attempts to solve this (or a closely related) problem[4] from first principles. We emphasize, however, that, to make these approaches tractable, certain parameters are assumed known (e.g. typically ρ, λ and $\mathbf{L}$). Operationally this decouples problems; e.g., it decouples the shape from shading problem from light source estimation problems[10].

1.2 Piecewise Smooth Shape from Shading Problems

We submit that such decoupling, while appropriate for certain highly engineered situations, is not always necessary; moreover, it can make shading analysis impotent precisely when it should be useful. For example, a human observer confronted with a static, monocular view of a scene will succeed in obtaining some estimate of the shapes of the surfaces within it even when some of the classical setting's constraints are relaxed. The presence of a shadow, a diffuse light source, or even a patterned surface does not necessarily interfere with our ability to recover shape from shading. Thus the classical constraints can be relaxed in principle; but how far, and once relaxed by what mechanism can solutions be found? These are precisely the questions with which we shall be concerned.

We retain the basic assumption that smooth variation in intensity is entirely due to smooth variation in surface orientation; thus:

$$\nabla I(x, y) \; = \; \rho \lambda \mathbf{L} \cdot \nabla \mathbf{N}(x, y) \; . \tag{1}$$

But we diverge from the classical model in two ways. First, we develop an approach that handles light source and surface properties concurrently; neither problem must be solved "before" the other. Second, we allow for discontinuities. Geometric discontinuities (in curvature, orientation, and depth) are unavoidable and their projection into the image has widely recognized importance [11, 12, 13]. We will therefore assume that the scene is composed of piecewise smooth surfaces.[5]

Since we are requiring that all smooth variation in the image intensity arises from variation in surface normal, we also assume that the albedo is constant on each smooth surface region.

When several light sources (point or non-point sources) are present, one can consider instead a single equivalent light source under the condition that all light sources are visible from the lambertian surface patch. As long as this equivalent light source is constant over a surface patch, (1) will be valid.

Shadow boundaries are problematic since their effect is to change the direction and magnitude of the equivalent light source. These problems are discussed in Sect. 5. Another problem is the mutual illumination between bright surfaces which can be considered as nearby large light sources. However, (1) is only valid when light sources are far from the surface.

In summary, we attribute all smooth variation in intensity to smooth variation in orientation of surface elements, and thus to surface shape. Thus, in general, any shape from shading process should reconstruct shape information identically for a photograph

---

[4] Other reflectance functions have been studied, e.g. for glossy surfaces or the moon.

[5] Since the reflectance function depends on the existence of a differentiable surface normal, allowing surfaces that are nowhere smooth is clearly inappropriate.

(where the albedo varies continuously), a projected slide (where the illumination on the screen varies continuously) and the scene itself.

2 Shape from Shading as a Coupled Family of Local Problems: Outline of the Scenel Bundle Approach

The key idea underlying our approach is to consider the shape from shading problem as a coupled family of local problems. Each of these is a "micro"-version of the shape from shading problem in which a lighting and surface model interact to produce the locally observed shading structure. We call each of these different models a scene element, or *scenel*, and, since many different scenels may be consistent with the local image structure, utilize fibre bundles to provide a framework to couple them together.
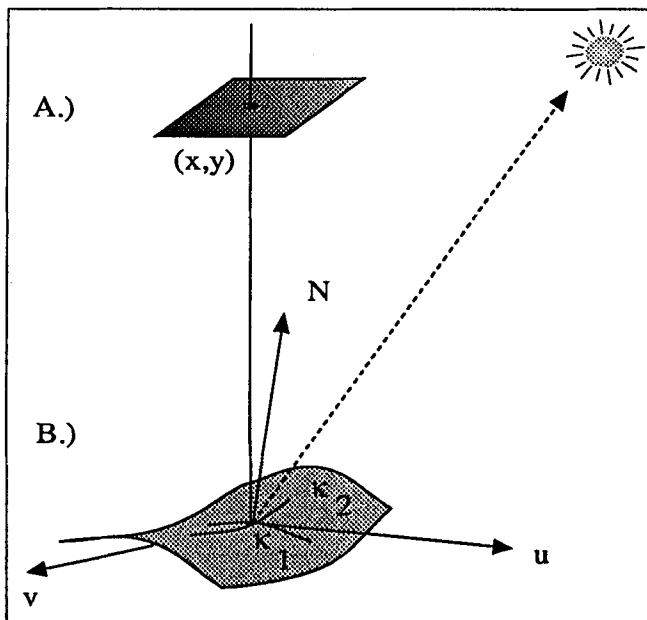
Fig. 1. Depiction of an abstract scene element, or scenel, corresponding to an image patch (A). The scenel (B) consists of a surface patch, described by its image coordinates, surface normal, and curvature. Its material properties (albedo) are also represented. Finally, a virtual light source completes the photometry.

The notion of FIBRE BUNDLES is fundamental to modern differential geometry [7]. A fibre bundle consists of a triple (E, π, M), where E is called the total space, π is a projection operator, and M is the base space. For each point $p \in M$, the subset $\pi^{-1}(p) \subset E$ is called the FIBRE over p. Denoting the fibre F, an example is the product bundle $(M \times F, \pi, M)$, which illustrates how the total space can be viewed as the base manifold crossed with the fibre space. While this construction is quite abstract, we use it

in the manner shown in Figs. 1 , 2 and 5. Essentially it allows us to consider the global shape from shading problem as a coupled family of local problems.

We take the image manifold as the base space, and consider the photometry for each point on it. Locally, in a neighbourhood around the point (x, y), the shading information can be described by a combination of local light source and surface property values (including curvature, albedo, etc.). Each of these defines a scenel (Fig. 1), and the space of all possible scenels defines the fibre over that point. Together, the collection of scenel fibres defines the scenel bundle (Fig. 2).
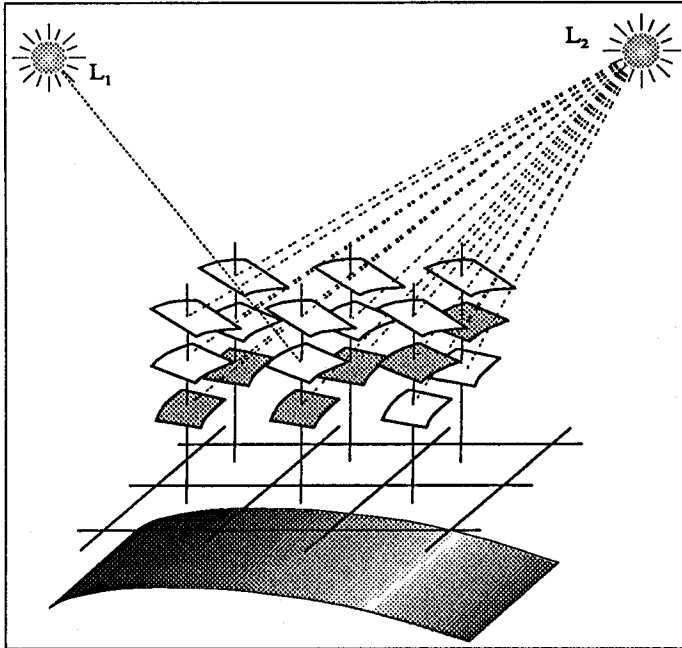

Fig. 2. Depiction of a Scenel Bundle over an image. At each point in the image there are many possible scene elements, or scenels. Each of these scenels is depicted along a fibre, or vertical space above each image coordinate. The union of scenel fibres over the entire image is called a scenel bundle. The shape from shading problem is formulated as determining sections through the scenel bundle. Such a section is depicted by the shaded scenels, and represents a horizontal slice across the bundle. Scenel participation in a horizontal section is governed by surface smoothness and material and light source constancy constraints.

We seek a solution of the shape from shading problem as connected sets of scenels in which neighbours are consistent; such a solution is called a CROSS SECTION through the scenel bundle. Formally, a cross section of a bundle (E, π, M) is a map $s : M \to E$ such that $\pi s = 1_M$. In other words, a cross section assigns a member of each fibre to each position in the manifold.

A sub-bundle of the scenel bundle is the TANGENT BUNDLE, in which the fibres consist of the tangent spaces at each point and sections correspond to vector fields; Sander and Zucker [8] previously used this bundle in their study of inferring principle direction fields

on surfaces.

We define what is meant by consistency shortly; first, we introduce the shading flow field as our initial data.

2.1 The Shading Flow Field as Initial Data

Observe that a sensitivity issue arises in the scenel framework; spatial quantization of the image induces a quantization of the scene domain. Analogously to the manner in which integer solutions are not always possible for algebraic equations, we begin with "quantized" initial data as well. In particular, we derive our initial estimates from *the shading flow field* instead of directly from the intensity image. This field is the first order differential structure of the intensity image expressed as the isoluminance direction and gradient magnitude (Fig. 3(a)); we supplement it with the intensity "edge" image (Fig. 3(b)). We suggest that dealing with uncertainties at the level of the shading flow field will expose more of the natural spatial consistency of the intensity variation, and will thus lead to more robust processing than the raw intensities. The shading flow field ideas are related to Kœnderink's isophotes [14].

Traditionally, the gradient of an image is computed by estimating directional derivatives by

$$\frac{\partial I}{\partial x} \approx I * G'_\sigma(x) G_\sigma(y) \ ,$$

$$\frac{\partial I}{\partial y} \approx I * G_\sigma(x) G'_\sigma(y) \ .$$

The gradient estimate follows immediately, and the isoluminance direction is simply perpendicular to the gradient. The one limitation of this approach is that (depending on the magnitude of σ) it always infers a smooth gradient field, even when the underlying image is non-continuous. We have investigated methods of obtaining stable, discontinuous shading flow fields using logical/linear operators [15, 16].

Our motivation for starting from the shading flow field is also biological. We take shading analysis to be an inherently geometric process, and hence handled within the same cortical systems that provide orientation selection and texture flow analysis. Shading flow is simply a natural extension.

2.2 Constraints between Local Scenels

The coupling between the local scenel problems dictates a consistency relationship over them, and derives from three principle considerations:

1. A SURFACE SMOOTHNESS CONSTRAINT, which states that the surface normal and curvatures must vary according to a Lipschitz condition between pairs of scenels which project to neighbouring points in the image domain. This notion is subtle to implement, because it involves comparison of normal vectors following parallel transport to the proper position (see [8]).
2. A SURFACE MATERIAL CONSTRAINT, which states that the surface material (albedo and reflectance) is constant between pairs of scenels which project to neighbouring points in the image domain.
3. A LIGHT SOURCE CONSTRAINT, which states that the virtual light source is constant for pairs of scenels which project to neighbouring points in the image domain.
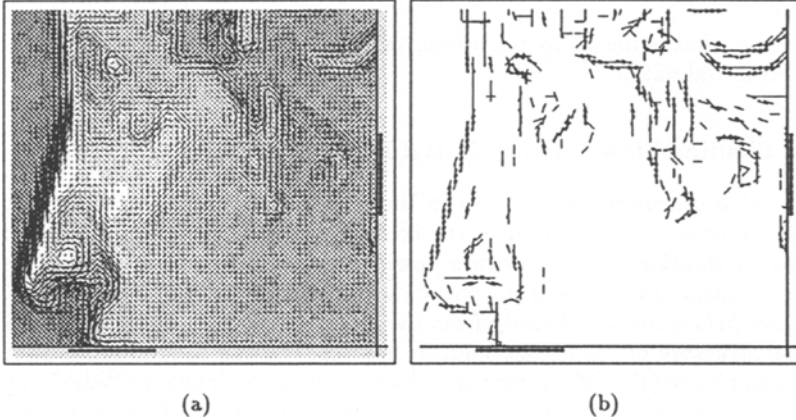
(a) (b)

Fig. 3. Typical shading flow (a) and edge (b) fields The left shading flow field depicts the first order differential structure of the image intensities, and the right is an edge map. Our shading analysis is based on these data, and not on the raw image intensities. This is to focus on the geometry of shape from shading analysis, and perhaps to capture something implicit in the biological approach.

Viewed globally, the solution we seek consists of sections in which a single (equivalent) light source illuminates a collection of surface patches with constant material properties but whose shape properties vary smoothly. The above constraints are embedded into a functional, and consistent sections through the scenel bundle are stationary points of this functional. More specifically, the constraints are expressed as compatibility relationships between pairs of neighbouring estimates within a relaxation labelling process. As background, we next sketch the framework of relaxation labelling.

2.3 Introduction to Relaxation Labelling

Relaxation labelling is an inference procedure for selecting labels attached to a graph according to optimal (symmetric) or variational (asymmetric) principles. Think of nodes in the graph as scenels, edges in the graph as links connecting "nearby" scenels with weights representing their compatibility. Formally, let $i \in \mathcal{I}$ denote the discrete scenel i, and let $l \in \mathcal{L}_i$ denote the set of labels for scenel i. The labels at each position are ordered according to the measure $p_i(l)$ such that $0 \le p_i(l) \le 1$ and $\sum_{l \in \mathcal{L}_i} p_i(l) = 1 \; \forall i$. (In biological terms, think of $p_i(l)$ as the firing rate for a neuron coding label l for scenel i.) Compatibility functions $r_{i,j}(l, l')$ are defined between label l at position i and label l' at position j such that increasingly positive values represent stronger compatibility. The abstract network structure is obtained from the support S that label l obtains from the labelling of it's neighbours $N(i)$; in symbols,

$$S_i(l) = \sum_{j \in N(i)} \sum_{l' \in \mathcal{L}_j} r_{i,j}(l, l') p_j(l') \; .$$

The final labelling is selected such that it maximizes the average local support

$$A(p) = \sum_{i \in I} S_i(l) p_i(l) = \sum_{i \in \mathcal{I}} \sum_{j \in N(i)} \sum_{l' \in \mathcal{L}_j} r_{i,j}(l, l') p_j(l') p_i(l) \ .$$

Such a labelling is said to be consistent [17]. We now simply remark that such "computational energy" forms have become common in neural networks, and observe that Hopfield [19] networks are a special case, as are polymatrix games, under certain conditions [20].

2.4 Overview of the Paper

The paper is organized as follows from this point on. We first formally define a scene element. We then show how to find those scene elements that are consistent with the shading flow field, and solve the forward problem of calculating the shading flow field expected for each scene element. The remainder of the paper is concerned with the inverse problems of inferring those scenels that are consistent with the shading flow field. The subtle interaction between expressing the scenel variables and the relaxation compatibilities is described, and scene element ambiguity addressed. An advantage of our technique is that, since both surface and lighting geometry are estimated, different types of shadow and illumination discontinuities can be handled; these are discussed in the Sect. 5.

3 Initial Estimation of Scene Attributes

We adopt a coarse coding of surface and light source attributes by quantizing the range of values of each attribute. Each scene element is defined by an assignment of a value to each of the scene attributes. The set of scene elements is viewed as a set of existence hypotheses of a surface patch of a fixed shape and orientation at a fixed image position, illuminated from a fixed direction with a fixed product of albedo and illumination.

3.1 The Scene Element

The attributes we consider are

1. IMAGE POSITION: The image pixels are themselves a set of discrete values of x and y position.
2. VIRTUAL ILLUMINANT DIRECTION: We take the light sources in the scene as a set of M distant point sources[6], $\{\lambda_{(i)} \mathbf{L}_{(i)} : 1 \leq i \leq M\}$ where $\lambda_{(i)}$ is the intensity and $\mathbf{L}_{(i)}$ is a unit vector. Let $V_i(x, y)$ be a binary "View" function such that $V_i(x, y) = 1$ if and only if light source $\mathbf{L}_{(i)}$ directly illuminates the surface element corresponding to pixel (x, y).
   We define the *virtual point source* at (x, y) by the two attributes λ and $\mathbf{L}$ such that

$$\lambda \, \mathbf{L}(x, y) \equiv \sum_i \lambda_{(i)} \, \mathbf{L}_{(i)} \, V_i(x, y) \ . \tag{2}$$

This virtual point source is constant for any neighbourhood in which all the V_i functions are constant. In such a neighbourhood, surface luminance satisfies

$$I(x, y) = \rho \lambda \mathbf{L} \cdot \mathbf{N}(x, y) \ .$$

---

[6] M would be quite large in the case of a diffuse source

The possible virtual light source directions map onto a unit sphere. We sample this sphere as uniformly as possible to get a discrete set of virtual illuminant directions. In viewer-centered coordinates, this unit vector is given as

$$\mathbf{L} = (L_x, L_y, L_z) .$$

3. MATERIAL PROPERTIES or the product $\rho\lambda$: We need only consider the product of the albedo and the illuminance (see (1)). Usually the imaging process normalizes to some maximum value, so we can assume a range between zero and one, and discretely sample this range.

4. SURFACE SHAPE DESCRIPTORS: The two principal curvatures (κ_1, κ_2) describe the shape up to rotation. Two angles slant σ and tilt τ are needed to describe the surface tangent plane orientation with respect to the viewer's coordinate frame. An additional angle ϕ is needed to describe the principal direction of the Darboux frame in the surface tangent plane.

(a) The two angles needed to orient the surface tangent plane in space describe the surface normal

$$\mathbf{N} = (N_x, N_y, N_z) = (\cos\tau\sin\sigma, \sin\tau\sin\sigma, \cos\sigma) .$$

The set of all such normals form a unit sphere. The surface of this sphere is sample as uniformly as possible to derive a discrete set of normals. Of these, only the ones in the hemisphere facing the viewer are used; the others are not visible; see Fig. 4.
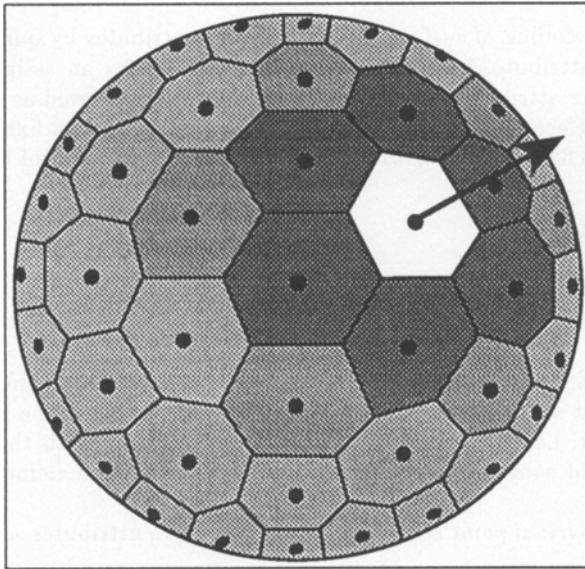

Fig. 4. The shape of surface patches is represented as a function of the principle curvatures mapped onto an abstract sphere through "curveness" and "shape index" measures (see text). Nearby positions on the sphere indicate smooth changes in either the shape or orientation of a scenel. This representation on the sphere facilitates the definition of scenel compatibilities later in the text.

(b) The two principal curvatures (κ_1, κ_2) are mapped into a curveness measure

$$c = \max(|\kappa_1|, |\kappa_2|)$$

and a shape index measure

$$s = \cos^{-1}\left(\frac{\kappa_1 + \kappa_2}{2c}\right) .$$

These are analogous to Kœnderink's curveness and shape index [21], with the choice of norm for the curveness and the spreading function for the shape index modified slightly.

The angles 2ϕ and s are, respectively, the longitude and latitude of a spherical coordinate system covering shape variation in the tangent plane. As we stated previously, the angle ϕ represents the principal direction, while s values 0 and π represent umbilic surfaces where principal directions are not defined. Any smooth curve on the $(2\phi, s)$ sphere represents a smooth deformation or rotation of the surface. We define a unit vector K as follows

$$\mathbf{K} = (K_1, K_2, K_3) = (\cos 2\phi \sin s, \sin 2\phi \sin s, \cos s) .$$

Thus by sampling the surface of the sphere uniformly we derive a discrete set of parameters which cover all variations of smooth, oriented shape in the tangent plane. Augmenting this with the curveness index provides a complete, discretely sampled shape descriptor.

Given the discrete sampling of the scene attributes as defined above, we derive a set of scenel labels

$$\mathcal{I} = \{x, y, \rho\lambda, \mathbf{L}, \mathbf{N}, c, \mathbf{K}\}$$

which represent all potential assignments of these scene attributes. Thus each i represents the hypothesis that the scene can be locally described by the scenel $(x_i, y_i, (\rho\lambda)_i, \mathbf{L}_i, \mathbf{N}_i, c_i, \mathbf{K}_i)$. To relate this to the *relaxation labelling* paradigm [17], we distribute a measure p_i over each scenel i representing confirmation of the hypothesis. The first step is to obtain an initial estimate for the confidence measure p_i from the shading flow field. The second is to extract locally consistent sets of hypotheses by relaxation labelling. The third step is to prune these sets by imposing appropriate boundary conditions. This is a large scale parallel computation, and we are currently implementing it on a massively parallel machine (MasPar MP-1).

3.2 The Expected Shading Flow Field

For each scenel i, we need an initial estimate for the associated weight p_i. We take this weight to reflect the match between the local properties of the shading flow field and the EXPECTED VALUES for the scene element i.

The EXPECTED SHADING FLOW FIELD is obtained by computing the light intensity gradient for a surface locally described by the scenel i. The paraboloid is an arbitrarily curved surface with the local parametric form

$$(u, v, w) \text{ where } w = -\frac{\kappa_1 u^2 + \kappa_2 v^2}{2} .$$

The u and v axes correspond to the two principal directions. The normal to the surface at a point (u, v) is given by:

$$\mathbf{N}(u,v) = \big(N_u(u,v), N_v(u,v), N_w(u,v)\big)$$

$$= \left(\frac{\kappa_1 u}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}}, \frac{\kappa_2 v}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}}, \frac{1}{(\kappa_1^2 u^2 + \kappa_2^2 v^2 + 1)^{\frac{1}{2}}} \right) .$$

Two rotations relate the viewer's coordinate frame to the paraboloid local frame. The first rotation takes care of the surface orientation and the second, of the principal directions. In matrix form

$$\mathbf{M} = \mathbf{M}_n \mathbf{M}_c$$

where

$$\mathbf{M}_n = \begin{pmatrix} \cos^2 \tau \cos \sigma + \sin^2 \tau & \sin \tau \cos \tau \cos \sigma - \sin \tau \cos \tau & \cos \tau \sin \sigma \\ \sin \tau \cos \tau \cos \sigma - \sin \tau \cos \tau & \sin^2 \tau \cos \sigma + \cos^2 \tau & \sin \tau \sin \sigma \\ -\cos \tau \sin \sigma & -\sin \tau \sin \sigma & \cos \sigma \end{pmatrix}$$

and

$$\mathbf{M}_c = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Since the inverse of the matrix $\mathbf{M}$ is simply its transpose, the light source expressed in the local surface coordinate is:

$$(L_u, L_v, L_w) = \mathbf{M}^t (L_x, L_y, L_z) .$$

Therefore the surface luminance can be computed.

$$I(u,v) = \rho \lambda \mathbf{L} \cdot \mathbf{N}(u,v) = \rho \lambda \big(L_u N_u(u,v) + L_v N_v(u,v) + L_w N_w(u,v) \big) .$$

Thus we can derive the local properties of the expected flow field. The orientation of the isoluminance line:

$$\theta_i = \tan^{-1} \left(\frac{L_u \kappa_1 M_{21} + L_v \kappa_2 M_{22}}{L_u \kappa_1 M_{11} + L_v \kappa_2 M_{12}} \right) .$$

The gradient magnitude:

$$|\nabla I(x,y)| = \rho \lambda \left| \frac{\partial(u,v)}{\partial(x,y)} (\kappa_1 L_u, \kappa_2 L_v) \right| .$$

The curvature of the isoluminance line:

$$\kappa_i = \frac{2\kappa_1^2 \kappa_2^2 L_w (L_u^2 + L_v^2) \det(M^{33}) - \kappa_1 \kappa_2 (\kappa_2 L_v \det(M^{32}) + \kappa_1 L_u \det(M^{31}))}{((M_{11} \kappa_2 L_v - M_{12} \kappa_1 L_u)^2 + (M_{21} \kappa_2 L_v - M_{22} \kappa_1 L_u)^2)^{\frac{3}{2}}} .$$

where

$$\det(M^{33}) = (M_{12} M_{21} - M_{11} M_{22}) ,$$
$$\det(M^{32}) = (M_{13} M_{21} - M_{23} M_{11}) ,$$
$$\det(M^{31}) = (M_{23} M_{12} - M_{13} M_{22}) .$$

The initial weight p_i is a decreasing function of some distance measure between the observed flow field in the neighbourhood of (x_i, y_i) and the expected flow field for scenel i.

$$p_i = \mathrm{G}(\theta_i - \theta_{obs.}) \cdot \mathrm{G}(|\nabla I|_i - |\nabla I|_{obs.}) \cdot \mathrm{G}(\kappa_i - \kappa_{obs.})$$

where $\theta_{obs.}$, $|\nabla I|_{obs.}$ and $\kappa_{obs.}$ are extracted from the initial flow field.

3.3 Scene Element Ambiguity

It should be clear that a local shading flow field will not select a unique scenel. In general, for an arbitrary shading flow field, several scene elements will be assigned a significant weight for each image position. Identical flow fields can be generated by surfaces of different shapes because of

- Intrinsic ambiguity. For example, the cases of convex (κ_1, κ_2), concave $(-\kappa_1, -\kappa_2)$, hyperbolic $(-\kappa_1, \kappa_2)$ or $(\kappa_1, -\kappa_2)$ surfaces facing the viewer and the light source all have the same intensity profile.
- Accidental correspondence between light source and surface orientations. For example, a surface with $\kappa_1 = \kappa_2$ will generate concentric circular isoluminance lines in the plane facing the light source. If the curvature is small, the isoluminance lines projected in the image plane form concentric ellipses. Such isoluminance lines can also be seen when an elliptic surface faces the viewer and the light source.

The example of elliptic isoluminance lines is particulary revealing. Such lines could be due to the shading of a spherical patch directly facing the light source but slanted away from the viewer; or it could be due to the shading of an elliptical (convex or concave) patch directly facing the light source and the viewer. Here, two phenomena are closely coupled: the formation of the isoluminance lines on the surface (related to the slant of the surface with respect to the light source) and the projection on the image plane (related to the slant of the surface with respect to the viewer).

However, these local ambiguities are not always accompanied by global ambiguities. Although, as Mach observed in 1866, "many curved surfaces may correspond to one light surface even if they are illuminated in the same manner" [1], these global ambiguities are often finite [18]. In the Sect. 4, we propose a relaxation labelling process to disambiguate the local surface geometry.

4 The relaxation labelling process

Recall that we are imposing three constraints on our surfaces: locally constant albedo, locally constant lighting conditions, and locally smooth geometry. For the relaxation labelling process, these translate into the following:

- A scenel j is *compatible* with the scenel i if they have the same constant $\rho\lambda$ and the same virtual illuminance direction L, and if scenel j's surface descriptors fall on scenel i's extrapolated surface at the corresponding relative position.
- A scenel j is *incompatible* with the scenel i if they have the same constant $\rho\lambda$ and the same virtual illuminance direction, and if there exists another scenel j', neighbouring scenel j along the fibre, that better fits the extrapolated surface from scenel i than scenel j. Observe that this incompatibility serves to localize information along each fibre.
- otherwise a scenel j is *unrelated* to the scenel i.

Using these guiding principles, we assign a value to the compatibility r_{ij} between two scenels i and j. This compatibility will be positive for compatible hypotheses, negative for incompatible hypotheses, and zero otherwise. In general, variation in r_{ij} is assumed to be smoothly varying between nearby points in the parameter space $\mathcal{I}$. The process is illustrated in Fig. 5.
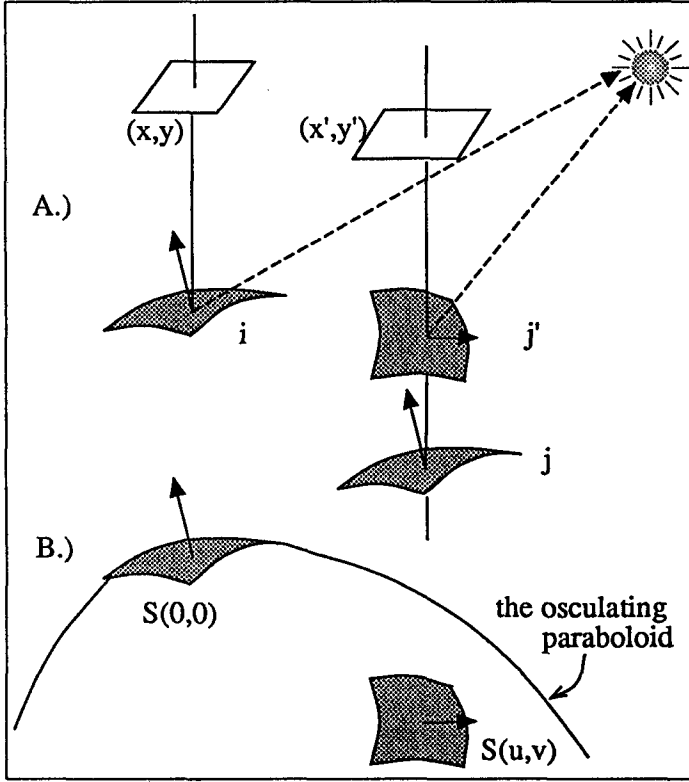
Fig. 5. Illustration of the compatibility relationship for scenel consistency. Two scenels are shown on the fibre at image location (x', y'), and are evaluated against the scenel (i) at (x, y). The surface represented in scenel$_{x,y}$ is modeled by the osculating paraboloid, and extended to (x', y'). It is now clear that one scenel (j') at (x', y') is consistent, because its surface patch lies on this paraboloid and light source and albedo agree. The other scenel (j) is inconsistent, because its surface does not match the extended paraboloid. Such osculating paraboloids are used to simulate the parallel transport of scenel$_{x',y'}$ onto scenel$_{x,y}$.

Consider the unique paraboloid $S_i(u, v)$ such that the neighbourhood of $S_i(0, 0)$ is described by the surface parameters of scenel i. The compatibility of a scenel j with i (r_{ij}) is then defined in terms of the relationship between scenel j and the point $S_i(u, v) \in S_i$ which is "closest to" scenel j. This operation is equivalent to the minimization of the distance measure

$$(x_j - x_i^*(u, v))^2 + (y_j - y_i^*(u, v))^2 + \left(\frac{\cos^{-1}(N_j \cdot N_i^*(u, v))}{\Delta \gamma_N} \right)^2$$

$$+ \left(\frac{\cos^{-1}(K_j \cdot K_i^*(u, v))}{\Delta \gamma_K} \right)^2 + \left(\frac{(c_j - c_i^*(u, v))}{\Delta c} \right)^2$$

over (u, v). Here, $\Delta \gamma_N, \Delta \gamma_K, \Delta c$ are the distance between neighbouring scenels for the given attribute and $(x_i^*, y_i^*, N_i^*, K_i^*, c_i^*)$ are the surface descriptors for the point $S_i(u, v)$.

So by physical consideration, the compatibility between scenel i and and scenel j can be expressed as

$$r_{ij} = \delta_{\rho\lambda_i, \rho\lambda_j} \cdot \delta_{\mathbf{L}_i, \mathbf{L}_j} \cdot G\left(\sqrt{(x_j - x_i^*)^2 + (y_j - y_i^*)^2}\right) \cdot -G''_{\Delta\gamma_N}\left(\cos^{-1}(\mathbf{N}_j - \mathbf{N}_i^*)\right)$$
$$\cdot -G''_{\Delta\gamma_K}\left(\cos^{-1}(\mathbf{K}_j - \mathbf{K}_i^*)\right) \cdot G_{\Delta c}\left(c_j - c_i^*\right)$$

where $G_\sigma(x)$ is the Gaussian, and $G''_\sigma(x)$ is its second derivative.

Notice that these values depend only on the relationship between scenel i and scenel j, which are fixed and constant throughout the computation. Therefore, these compatibilities can be calculated once and then stored in either a lookup table or as the weights in some sort of network.

5 A Case Study of Discontinuities

We have described a framework for computing sections of the scenel fibre bundle that are constrained to have a constant virtual light source, constant albedo, and smooth surface geometry. Any discontinuity in these attributes will demarcate a section boundary. Discontinuities can arise in many ways: the albedo can change suddenly along a smooth surface; a shadow can be cast across a smooth surface; the surface normal can change abruptly along a contour. In all three of these cases there will be a discontinuity in the image intensity. In this Sect., we address the question of how scenel discontinuities are manifested as discontinuities in the image. We restrict our attention to two types of image discontinuities : intensity discontinuities (edges) and shading flow discontinuities.

5.1 Illumination Discontinuities (Shadows)

Shadows are produced by variations in the virtual point source along a single surface patch (recall Sect.3.1 for the definition of λL). These variations are the direct result of a discontinuity in some $V_i(x, y)$. We say that there is a shadow boundary along this discontinuity. The image intensity cannot satisfy our model (1) in the neighbourhood of a shadow boundary since the image intensity on each side of the shadow boundary is due to a different virtual point source. We now briefly examine the two types of shadow boundary.

In general, an *attached shadow* boundary lies between two nearby pixels (x_0, y_0) and (x_1, y_1) when for some point source, $\mathbf{L}_{(j)}$, it is the case that $\mathbf{N}(x_0, y_0) \cdot \mathbf{L}_{(j)} > 0$ and $\mathbf{N}(x_1, y_1) \cdot \mathbf{L}_{(j)} < 0$.

The image intensity is continuous across the attached shadow boundary even though V_j is discontinuous since $\mathbf{N} \cdot \mathbf{L}_{(j)} = 0$ at the boundary. However, the shading flow is *typically* not continuous at the attached shadow boundary. The shading flow due to the source $\mathbf{L}_{(j)}$ will be parallel to the shadow boundary on the side where $V_j = 1$, but it will be zero on the side where $V_j = 0$. The shading flow due to the rest of the light sources will be smooth across the boundary but this shading flow will not typically be parallel to the boundary. Hence the sum of the two shading flows will typically be discontinuous at the boundary.

A *cast shadow* boundary is produced between two nearby points (x_0, y_0) and (x_1, y_1) when for some point source, $\mathbf{L}_{(j)}$, it is the case that both $\mathbf{N}(x_0, y_0) \cdot \mathbf{L}_{(j)} > 0$ and $\mathbf{N}(x_1, y_1) \cdot \mathbf{L}_{(j)} > 0$, and either $V_j(x_0, y_0) = 0$ or $V_j(x_1, y_1) = 0$ (but not both).
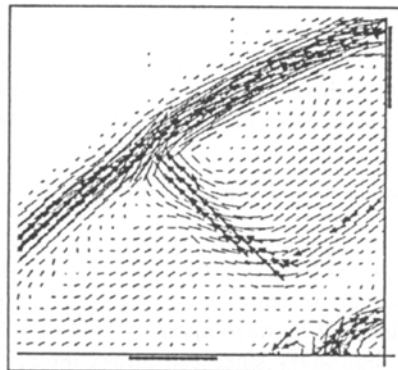
Examining (2), we note that for cast shadows, the discontinuity in V_j results in a discontinuity in image intensity, since the $\mathbf{N} \cdot \mathbf{L}_{(j)} > 0$. Furthermore, there is *typically* a
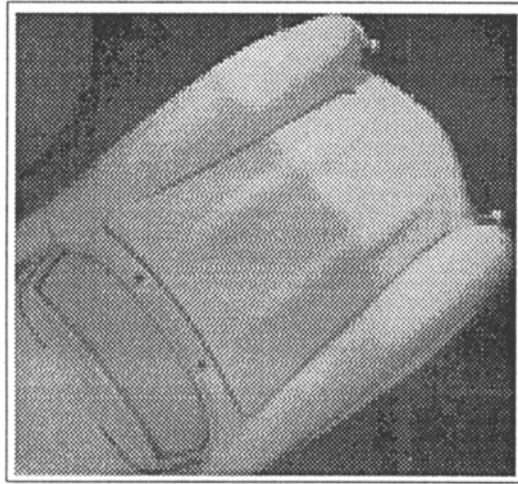
(a)
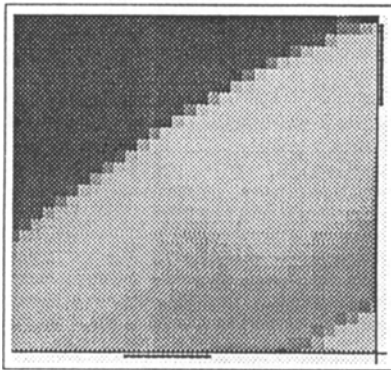
(b)

(c)

Fig. 6. An illustration of shadow boundaries and how they interact with flow fields. In (a) a shadow is cast across the hood of a car. In (b) we show a subimage of the cast shadow on the left fender, and in (c) we show the shading flow field (represented as a direction field with no arrowheads) and the intensity edges (represented as short arrows; the dark side of the edge is to the left of the arrow). Observe how the shading field remains continuous (in fact, virtually constant) across the intensity edge. This hold because the surface is cylindrical, and the shading flow field is parallel to the axis of the cylinder.

discontinuity in the shading flow since the virtual light source defined on the side of the boundary where $V_j = 0$ will usually produce a different shading flow across the boundary than is produced by $L_{(j)}$ on the side of the boundary where $V_j = 1$.

In the special case of parabolic (e.g. cylindrical) surfaces, the shading flow remains continuous across both cast and attached shadow boundaries because the flow is parallel to the axis of the cylinder. Note however that the attached shadow is necessarily parallel to the the shading flow field. This case is illustrated in Fig. 6.

To summarize, the image intensity in the neighbourhood of a pixel (x_0, y_0) can be modelled using a single virtual point source as long as there are neither attached nor cast shadow boundaries in that neighbourhood. Attached shadow boundaries produce continuous image intensities, but discontinuities in the shading flow. Cast shadows produce both intensity discontinuities and shading flow discontinuities.

5.2 Geometric Discontinuities

There are two different ways that the geometry of the scene can produce discontinuities in image. There can be a discontinuity in N along a continuous surface, or there can be a discontinuity in the surface itself when one surface occludes another. In the latter case, even if there is no discontinuity in the virtual light source direction there will still typically be a discontinuity in N which will usually result in both discontinuities in the image intensity and in the shading flow.

5.3 Material Discontinuities

If there is a discontinuity in the albedo along a smooth surface, then there will be a discontinuity in luminance across this material boundary. However, the shading flow will not vary across the boundary in the sense that the magnitude of the luminance gradient will change but the direction will not.

5.4 Summary of Discontinuities

In summary, shading flow discontinuities which are not accompanied by intensity discontinuities usually indicate attached shadows on a smooth surface. Intensity discontinuities which are not accompanied by shading flow discontinuities usually indicate material changes on a smooth surface. The presence of both types of image discontinuities indicates that either there is a cast shadow on a smooth surface, or that there is a geometric discontinuity.

6 Conclusions

We have proposed a new solution to the shape from shading problem based on notions from modern differential geometry. It differs from the classical approach in that light source and surface material consistency are solved for concurrently with shape properties, rather than independently. This has important implications for understanding light source and surface interactions, e.g., shadows, both cast and attached, and an example illustrating a cast shadow is included.

The approach is based on the notion of scenel, or unit scene element. This is defined to abstract the local photometry of a scene configuration, in which a single (virtual) light source illuminates a patch of surface. Since the image irradiance equation typically admits

many solutions, each patch of the image gives rise to a collection of scenels. These are organized into a fibre space at that point, and the collection of scenel fibres is called the scenel bundle. Algebraic and topological properties of the scenel bundle will be developed in a subsequent paper.

The solution of the shape from shading problem thus reduces to finding sections through the scenel bundle, and these sections are defined by material, light source, and surface shape consistency relationships. The framework thus provides a unification of these different aspects of photometry, and should be sufficently powerful to indicate the limitations of unification as well.

References

1. Ratliff, F.: **Mach Bands: Quantitative Studies on Neural Networks in the Retina**, Holden-Day, San Francisco (1965)
2. Horn, B.K.P.: "Obtaining Shape from Shading Information," P.H. Winston. Ed. in **The Psychology of Computer Vision**, McGraw-Hill, New York (1975)
3. Ikeuchi, K. and Horn, B.K.P.: "Numerical Shape from Shading and Occluding Boundaries," *Artificial Intelligence*, 17 (1981) 141–184
4. Horn, B.K.P. and Brooks, M.J.: "The Variational Approach to Shape from Shading," *Comput. Vision, Graph. Image Process.*, 33 (1986) 174–208
5. Nayar, S.K., Ikeuchi, K. and Kanade, T.: "Shape From Interreflections," *International Journal of Computer Vision*, 6 (1991) 173–195
6. Pentland, A.: "Linear Shape From Shading," *International Journal of Computer Vision*, 4 (1990) 153–162
7. Husemoller, D.: **Fibre Bundles**, Springer, New York (1966)
8. Sander, P., and Zucker, S.W.: "Inferring Surface Trace and Differential Structure from 3-D Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 9 (1990) 833–854
9. Spivak, M.: **A Comprehensive Introduction to Differential Geometry**, Publish or Perish, Berkeley (1979)
10. Pentland, A.: "Finding the Illuminant Direction," *J. Opt. Soc. Amer*, 72 (1982) 448–455
11. Kœnderink, J.J.: "What Does the Occluding Contour Tell us about Solid Shape?," *Perception*, 13 (1984) 321–330
12. Kœnderink, J.J. and van Doorn, A.J.: "The Shape of Smooth Objects and the Way Contours End," *Perception*, 11 (1982) 129–137
13. Biederman, I.: "Recognition-by-Components: A Theory of Human Image Understanding," *Psychological Review*, 94 (1987) 115–147
14. Kœnderink, J.J. and van Doorn, A.J.: "Photometric Invariants Related to Solid Shape," *Optica Acta*, 27 (1980) 981–996
15. Iverson, L. and Zucker, S.W.: "Logical/Linear Operators for Measuring Orientation and Curvature," *TR-CIM-90-06*, McGill University, Montreal, Canada (1990)
16. Zucker,S.W., Dobbins, A., and Iverson, L.: "Two Stages of Curve Detection Suggest Two Styles of Visual Computation," *Neural Computation*, 1 (1989) 68–81
17. Hummel, A.R. and Zucker, S.W.: "On the Foundations of Relaxation Labeling Processes," *IEEE Trans. Pattern Anal. Machine Intell.*, 5 (1983) 267–287
18. Oliensis, J.: "Shape from Shading as a Partially Well-Constrained Problem," *Comp. Vis. Graph. Im. Proc.*, 54 (1991) 163–183
19. Hopfield, J.J.: "Neurons with Graded Response Have Collective Computational Properties like Those of Two-State Neurons," *Proc. Natl. Acad. Sci. USA*, 81 (1984) 3088–3092
20. Miller, D.A. and Zucker, S.W.: "Efficient Simplex-like Methods for Equilibria of Nonsymmetric Analog Networks," *TR-CIM-91-3* McGill University, Montreal, Canada (1991); *Neural Computation*, (in press)
21. Kœnderink, J.J.: **Solid Shape**, MIT Press, Cambridge, Mass. (1990) p. 320