

A COOPERATIVE ALGORITHM FOR DETERMINING SURFACE ORIENTATION FROM A SINGLE VIEW

Robert J. Woodham
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts, 02139

INTRODUCTION

This paper addresses the problem of determining local surface orientation from the intensity information contained in a single monocular view. It represents another look at the problem of obtaining shape from shading information. This problem was first formulated by <Horn 1970> as the solution of first-order non-linear partial differential equations. It was subsequently reformulated by <Horn 1975> to take advantage of the geometric insight provided by the gradient space approach popularized by <Huffman 1971,1975> and <Mackworth 1973>.

The goal of this work is to understand how the observed intensity variation across surfaces of objects forces conclusions about the local topography of those surfaces. The problem is formulated as a problem in image analysis. This paper adopts the position that it is important to examine ways of squeezing out the last ounce of information from the intensity values recorded in an image before taking recourse to high-level knowledge.

In order to exploit intensity information, it is vital to understand how intensity values arise in an image. The paper begins by introducing gradient space and reflectance map techniques. Having developed these tools, the physical constraints imposed by the light source, object surface and viewer geometry are re-examined to recast the <Horn 1975> formulation as a cooperative computation which determines local surface orientation by matching points in an image to their corresponding points in gradient space. Finally, the method is illustrated using the simple example of a Lambertian sphere illuminated by a single distant light source.

GRADIENT SPACE AND THE REFLECTANCE MAP

In order to understand the correspondence between intensity data and local surface orientation, it is necessary to relate the geometry of the image forming process to the photometry of the object being imaged. The problem of determining local surface orientation from intensity can be characterized as a mapping:

$$T: I(u,v) \rightarrow LSO(u,v)$$

which assigns to each image intensity point $I(u,v)$ a local surface orientation $LSO(u,v)$. In a visual world consisting of opaque smooth objects immersed in a transparent medium, the mapping T is well defined since each image intensity point $I(u,v)$ arises from a unique object point (which, in turn, defines a unique surface orientation). What is not obvious, however, is whether T can be determined from intensity data alone. The difficulty, of course, is that T is not a local operator. In

formulating an algorithm for determining local surface orientation from image intensity, it will be important to keep track of how additional physical constraints are used in the computation.

Now, $LSO(u,v)$ is a two-valued function of u and v since two parameters are required to specify an arbitrary orientation in space. One might expect surface orientation to be difficult to represent explicitly. Gradient space is introduced as the appropriate formalism for reasoning about surface orientation.

Gradient space can be derived in several ways. For present purposes, it will be related directly to surface orientation. If the equation of a smooth surface is given by:

$$z = f(x,y)$$

then $(\partial f/\partial x, \partial f/\partial y, -1)$ defines a local (outward) surface normal. (Throughout this paper, smooth will be interpreted to mean $z = f(x,y)$ has continuous first and second partial derivatives.) Now, let:

$$p = \partial f/\partial x \\ q = \partial f/\partial y$$

so that the local (outward) surface normal becomes $(p,q,-1)$. The quantity (p,q) will be called the *gradient* and *gradient space* is defined to be the two-dimensional space of all such points (p,q) .

The fraction of light reflected by a surface in a given direction depends upon the optical properties of the object material, the surface micro-structure and the spatial and spectral distribution of the incident light. The important photometric observation underlying this work is the following: no matter how complex the distribution of incident illumination, for most surfaces, the fraction of the incident light reflected in a particular direction depends only on the local light source, object surface and viewer geometry. To make this observation more concrete, one can standardize the local representation of the light source, object surface and viewer geometry and tie this representation down to gradient coordinates p and q .

A surface photometric function $\phi(i,e,g)$ is typically defined in terms of the three angles i , e and g illustrated in figure 1. These angles are called, respectively, the incident, emergent and phase angle. Here, the emergent angle e will be referred to as the *view angle*.

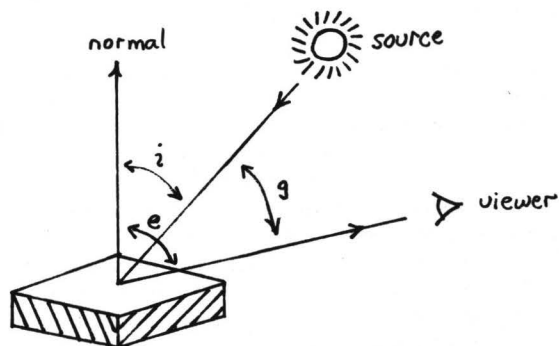


Figure 1

If both the viewing direction and the direction of incident illumination are known, then expressions for $\cos(i)$, $\cos(e)$ and $\cos(g)$ can be derived in terms of gradient space coordinates p and q . To simplify matters, consider an image forming system

that performs an orthographic projection. The important simplification inherent in the assumption of an orthographic projection is that the viewing direction, and hence the phase angle g , is constant for all image points (u,v) . To further simplify the mathematics, one can align the viewing direction with the negative Z-axis and assume a scaling in U and V that takes object point (x,y,z) to image point (u,v) where $u = x$ and $v = y$ (see figure 2). Under this projection, the use of separate image coordinates (u,v) is redundant. Henceforth, image coordinates (x,y) and object coordinates (x,y) will be referred to interchangeably. With this geometry, object space vector $(0,0,-1)$ points in the direction of the viewer. That is, the viewer is located at gradient space point $(0,0)$.

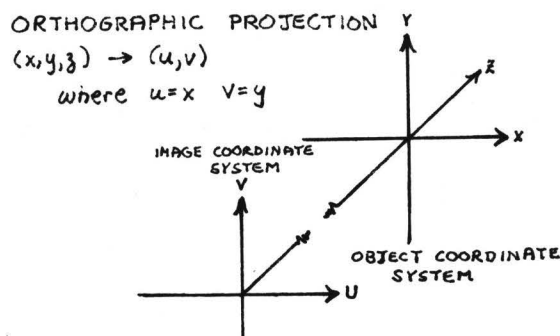


Figure 2

Now, for a given distribution of incident illumination, a given surface-viewer geometry and a given object material, the image intensity corresponding to a surface point with gradient (p,q) is unique. The intensities recorded at each (p,q) is called the *reflectance map* $R(p,q)$. Reflectance maps can be determined empirically, derived from phenomenological models of surface reflectivity or derived from analytic models of surface micro-structure. Once determined, however, the reflectance map is independent of the shape of the objects being viewed. It represents explicit knowledge of intensities that can be recorded from objects made of a given material and viewed under a particular light source and viewer geometry. The reflectance map accounts for the physical intensities that can be recorded under varying components of diffuse and specular reflection. It does not, of course, account for the human perception of those intensities. (eg. A local specular component can make an otherwise "dull" surface appear "glossy".)

Consider an image whose intensity values $I(x,y)$ have been normalized to equal the reflectance map values at the corresponding gradients (p,q) . Then, the following equation describes the image forming process:

$$I(x,y) = R(p,q)$$

This is the basic equation relating image intensity to the geometry of the image forming process. It is one equation in the two unknowns p and q . Thus, the problem of determining local surface orientation from intensity becomes the problem of finding the point in gradient space (p,q) corresponding to the image intensity point $I(x,y)$.

Now, the simplest case for incident illumination is that of a single distant point source. Choose such a source and place it so that object space vector $(p_s, q_s, -1)$ points in the direction of the source. That is, the source is located at gradient space

point (p_s, q_s) . Specifying a single distant point source is not a fundamental restriction on the development. Non-point sources can be modeled as the superposition of single point sources. The development does, however, assume equal illumination at all surface points. For non-convex surfaces, the reflectance map does not account for the fact that certain surface points can be shadowed with respect to one or more of the sources nor for the fact that certain surface points can receive additional illumination due to light reflected from other sections of surface (mutual illumination).

Recall that, object space vector $(p,q,-1)$ is an (outward) normal to the surface point (x,y,z) . That is, (p,q) is the gradient point corresponding to the surface point (x,y,z) . Then, using standard vector algebra, the expressions for $\cos(i)$, $\cos(e)$ and $\cos(g)$ become:

$$\begin{aligned}\cos(i) &= \frac{1 + pp_s + qq_s}{\sqrt{1 + p_s^2 + q_s^2} \sqrt{1 + p^2 + q^2}} \\ \cos(e) &= \frac{1}{\sqrt{1 + p^2 + q^2}} \\ \cos(g) &= \frac{1}{\sqrt{1 + p_s^2 + q_s^2}}\end{aligned}$$

Using the above expressions, it is clear that one can transform an arbitrary surface photometric function ϕ (i.e.g) into a reflectance map $R(p,q)$. By now, it should be clear how points in gradient space correspond to orientations in object space. However, it is possible to be more explicit about the correspondence between movement in gradient space and changes in local surface orientation.

Using elementary trigonometry, the expression for $\cos(e)$ can be rewritten as:

$$\tan(e) = \sqrt{p^2 + q^2}$$

This gives a first useful interpretation that the inclination of the surface with respect to the viewing direction varies monotonically with the distance from the origin in gradient space. Specifically, the distance from the origin is the tangent of the angle between the surface normal and the view vector. As the gradient (p,q) moves away from the origin, the inclination of the surface with respect to the viewer increases. As the gradient (p,q) moves toward the origin, the inclination of the surface with respect to the viewer decreases.

Now, the view angle e characterizes one of the two degrees of freedom associated with an arbitrary orientation in space. Note that the locus of points in object space having a constant view angle e defines a right circular cone oriented along the viewing direction. The angular position of each gradient (p,q) on the circle $p^2 + q^2 = \tan^2(e)$ defines the direction of steepest descent in image space along this cone. This gives a second useful interpretation that the angular position of a point (p,q) in gradient space corresponds to the direction of steepest descent in image space along the original surface. Rotating object space about the view vector induces an equal rotation in gradient space.