

RE-EXAMINING PHYSICAL CONSTRAINTS

One can formulate the problem of determining the point in gradient space (p,q) corresponding to the image intensity point $I(x,y)$ analytically as has been done in <Horn 1970,1975>. This work will not be reviewed here. Instead, the formal nature of the computation will be put aside in order to first re-examine its basis in the physical world.

Two constraints of importance can be identified:

1. A given point on a physical surface has a unique orientation in space.
2. Matter is cohesive. It is separated into objects. The surfaces of objects are generally smooth compared with their distance from the viewer.

The reader will note that these are essentially the same two constraints <Marr and Poggio 1976> use as a basis for their computation of stereo disparity. As in their paper, the goal is to translate the above two physical constraints into rules for how points in an image can be matched to points in gradient space.

In their most general form, these rules can be expressed as:

1. **UNIQUENESS:** Each image point may be assigned to at most one location in gradient space
2. **CONTINUITY:** Surfaces vary smoothly almost everywhere. Only a small fraction of the area of an image is composed of boundaries that correspond to discontinuities of surface.

The task ahead is to demonstrate that these rules can be explicitly embedded in a computation. The result is an algorithm which attempts to achieve a global correspondence between image points and points in gradient space via local, interactive constraints.

It has become fashionable to call such algorithms "cooperative" after similar phenomena in physics. Perhaps the reader will be more comfortable if the algorithm is presented as a "relaxation scheme". Nonetheless, what is important is to get the flavor of how local constraints can propagate back and forth to globally constrain possible matches between image points and points in gradient space.

SPECIFYING LOCAL CONSTRAINT

The basic equation

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

is one equation in the two unknowns p and q . With this equation alone, the gradient corresponding to a particular image point is constrained to lie on a one parameter (family of) contour(s) in gradient space. The goal is to apply further constraint in order to assign a unique location in gradient space to each image point.

The essential physical constraint to be exploited is the assumption that, compared to the viewing distance, surfaces vary smoothly almost everywhere. This surface smoothness assumption is translated into monotonicity rules on changes to view angle and changes to direction of steepest descent permitted between (closely spaced) image points.

One can illustrate how physical constraint adds additional constraint to the possible gradient space solutions to the basic equation $I(x,y) = R(p,q)$. Suppose, two (closely spaced) image points (x_1,y_1) and (x_2,y_2) are hypothesized to correspond to object points on the same section of smooth surface. Further, suppose that the view angle increases in going from (x_1,y_1) to (x_2,y_2) and that the direction of steepest descent decreases in going from (x_1,y_1) to (x_2,y_2) . Let (p_1,q_1) and (p_2,q_2) be the gradient locations corresponding to (x_1,y_1) and (x_2,y_2) .

Suppose, further, that the basic equation $I(x,y) = R(p,q)$ constrains (p_1,q_1) and (p_2,q_2) to lie on the contours C_1 and C_2 respectively as shown in figure 3.

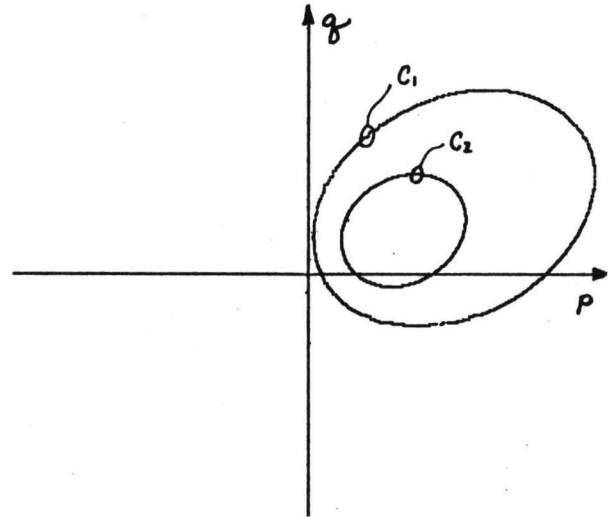


Figure 3

Figure 4 shows, superimposed on the two contours of figure 3, the gradient space circle corresponding to the maximum view angle interpretation of (p_2,q_2) and the gradient space line corresponding to the minimum direction of steepest descent interpretation of (p_2,q_2) . Since the view angle increases in going from (x_1,y_1) to (x_2,y_2) , the contour of permissible (p_1,q_1) can be restricted to include only those gradient points on C_1 lying on or within the circle of the maximum view angle interpretation of (p_2,q_2) . Similarly, since direction of steepest descent decreases in going from (x_1,y_1) to (x_2,y_2) , the contour of permissible (p_1,q_1) can be restricted to include only those gradient points on or above the line of the minimum direction of steepest descent interpretation of (p_2,q_2) . Thus, without any additional constraint on (p_2,q_2) , the assumed monotonicity relations between (x_1,y_1) and (x_2,y_2) have been applied to the topography of the reflectance map to constrain the possible interpretation of (p_1,q_1) to include only those points of C_1 indicated by the solid line of figure 4.

Let (r,θ) denote the polar representation of the gradient space point (p,q) . That is:

$$r = \sqrt{p^2 + q^2}$$

$$\theta = \text{atan}(q/p)$$

Generalizing from the above illustration, the following two rules are stated:

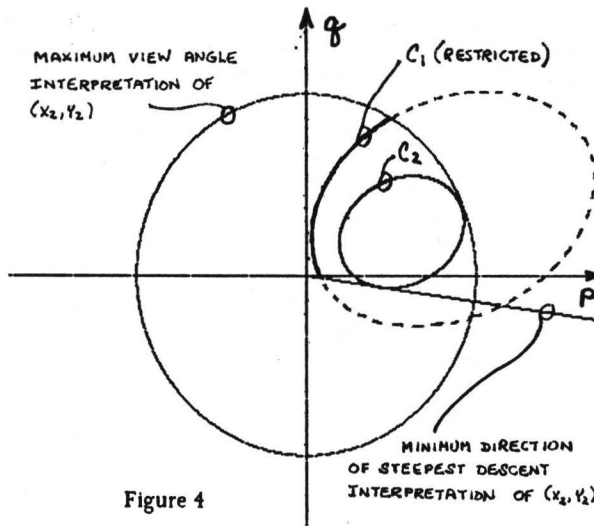


Figure 4

RULE I: CHANGES IN VIEW ANGLE

Let $I_1, I_2, \dots, I_n$ be a set of (closely spaced) image points hypothesized to correspond to object points on the same section of smooth surface that are monotonically non-decreasing in view angle. Let $C_1, C_2, \dots, C_n$ be the corresponding set of gradient contours determined from the reflectance map. Then, each contour C_i can be further constrained such that, for each $(r, \theta) \in C_i$, $i = 2, 3, \dots, n-1$

$$\min \{r \mid (r, \theta) \in C_{i-1}\} \leq r \leq \max \{r \mid (r, \theta) \in C_{i+1}\}$$

Similarly, if $I_1, I_2, \dots, I_n$ is hypothesized to correspond to a set of object points on the same section of smooth surface that are monotonically non-increasing in view angle, then each contour C_i can be further constrained such that, for each $(r, \theta) \in C_i$, $i = 2, 3, \dots, n-1$

$$\min \{r \mid (r, \theta) \in C_{i+1}\} \leq r \leq \max \{r \mid (r, \theta) \in C_{i-1}\}$$

RULE II: CHANGES IN DIRECTION OF STEEPEST DESCENT

Let $I_1, I_2, \dots, I_n$ be a set of (closely spaced) image points hypothesized to correspond to object points on the same section of smooth surface that are monotonically non-decreasing in direction of steepest descent. As above, let $C_1, C_2, \dots, C_n$ be the corresponding set of gradient contours. Then, each contour C_i can be further constrained such that, for each $(r, \theta) \in C_i$, $i = 2, 3, \dots, n-1$

$$\min \{\theta \mid (r, \theta) \in C_{i-1}\} \leq \theta \leq \max \{\theta \mid (r, \theta) \in C_{i+1}\}$$

Similarly, if $I_1, I_2, \dots, I_n$ is hypothesized to correspond to a set of object points on the same section of smooth surface that are monotonically non-increasing in direction of steepest descent, then each contour C_i can be further constrained such that, for each $(r, \theta) \in C_i$, $i = 2, 3, \dots, n-1$

$$\min \{\theta \mid (r, \theta) \in C_{i+1}\} \leq \theta \leq \max \{\theta \mid (r, \theta) \in C_{i-1}\}$$

HYPOTHESIZING MONOTONICITY RELATIONS

It is now time to turn to the question of how to hypothesize monotonicity relations between selected image points. To begin with, consider the worst possible approach. For some small value of n , one might explore all possible orderings, with respect to both view angle and direction of steepest descent, of selected (closely spaced) image points $I_1, I_2, \dots, I_n$. The hope would be that only a small fraction of those orderings would have

admissible interpretations (ie. interpretations that included at least one gradient point for each image point). The constraints imposed by each local interpretation would propagate to neighboring sets of selected image points to provide further mutual constraint. Again, the hope would be that propagation of local constraint would converge to a correct global interpretation while "incorrect" propagations would (quickly) die out.

Consider a second possible approach. Suppose a particular surface interpretation is forced onto the data. Such an interpretation would provide a framework to (partially) order selected (closely spaced) image points with respect to changes in both view angle and direction of steepest descent. Instead of allowing all possible orderings to compete, this second approach pursues a particular interpretation. Again, the hope would be that propagation of local constraint would converge to a single global interpretation that represents a simple distortion of the particular interpretation being forced. (Here, simple distortion implies any surface that preserves the assumed monotonicity relations concerning changes to view angle and changes to direction of steepest descent.)

The method actually implemented corresponds to this second approach. The program has a small set of interpretations it is willing to pursue. Some are quite rigid, others are quite flexible. In the next section, a specific example is presented. For now, a brief analysis is given to show of how convexity can be used to hypothesize monotonicity relations between (closely spaced) image points.

By partially differentiating the basic equation $I(x, y) = R(p, q)$ with respect to X and Y we obtain two equations which can be written as the single matrix equation:

$$\begin{bmatrix} I_x \\ I_y \end{bmatrix} = \begin{bmatrix} p_x & q_x \\ p_y & q_y \end{bmatrix} \begin{bmatrix} R_p \\ R_q \end{bmatrix}$$

(subscripts denote partial differentiation). Similarly, the two first-order equations $dp = p_x dx + p_y dy$ and $dq = q_x dx + q_y dy$ can be written as the single matrix equation:

$$\begin{bmatrix} dp \\ dq \end{bmatrix} = \begin{bmatrix} p_x & p_y \\ q_x & q_y \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix}$$

Recalling the original definitions of $p = \partial f / \partial x$ and $q = \partial f / \partial y$, observe that $p_y = q_x$ (for smooth surfaces, the order of differentiation can be interchanged). Thus, one can define a matrix H by:

$$H = \begin{bmatrix} p_x & q_x \\ p_y & q_y \end{bmatrix} = \begin{bmatrix} p_x & p_y \\ q_x & q_y \end{bmatrix}$$

H is the standard Hessian matrix (which captures the notion of surface curvature). The first observation to be made is that smoothness guarantees that H is symmetric. Thus, it is reasonable to ask under what conditions is H positive semidefinite (A real symmetric $n \times n$ matrix A is positive semidefinite if and only if $x^T A x \geq 0$, for all vectors x in R^n). Now, H is positive semidefinite if and only if $z = f(x, y)$ is convex. This result is used to show how convexity adds constraint. Suppose $z = f(x, y)$ is convex. Thus, H is positive semidefinite. Multiplying the two matrix equations on the left by (non-zero) $[R_p \ R_q]$ and (non-zero) $[dx \ dy]$ respectively, gives the two inequalities:

$$I_x R_p + I_y R_q \geq 0 \\ dp dx + dq dy \geq 0$$

This first inequality can be viewed as an additional *a priori* constraint on the contour in gradient space of possible solutions to the basic equation $I(x,y) = R(p,q)$. The tangent vector $[R_p, R_q]$ to the contour of constant reflectance at any point (p,q) hypothesized to be a solution to the basic equation $I(x,y) = R(p,q)$ must have a non-negative component in the direction of the tangent vector $[I_x, I_y]$ to the contour of constant intensity at (x,y) .

The second inequality can be viewed as an additional constraint on the possible movement $[dp, dq]$ in gradient space corresponding to a movement $[dx, dy]$ in the image. As above, the vector $[dp, dq]$ must have a non-negative component in the direction $[dx, dy]$. Thus, by choosing $[dx, dy]$ appropriately, it is possible to guarantee either the sign of the change to the view angle or the sign of the change to the direction of steepest descent.

ACHIEVING GLOBAL CONSTRAINT

Regardless of what mechanism is used to hypothesize (local) monotonicity relations between points in image space, it is still necessary to embed that mechanism in a computation to achieve global constraint. The implementation approach taken here is admittedly ad hoc. The program selects nine points in a simple 3×3 square pattern as its basic set of (closely spaced) image points. This pattern serves as the set $I_1, I_2, \dots, I_n$ for applying the local constraint criteria (Rules I and II). First, however, the set $I_1, I_2, \dots, I_n$ is passed to the chosen hypothesizing routines to be (partially) ordered with respect to view angle and direction of steepest descent. The reflectance map $R(p,q)$ is then used to determine the initial contour of possible gradient space locations for each point I_i . Rules I and II are iteratively applied to these contours until no further mutual constraint is provided.

The above describes the basic application of local constraint to each 3×3 template. The selection of successive 3×3 square patterns is allowed to overlap. Thus, each image point I_i will eventually belong to nine templates. Each time a particular image point I_i is further constrained by the application of local constraint to a template of which it is a member, each of its eight other templates is marked for reconsideration. Before moving on to a previously unconsidered template local constraint is applied recursively to each marked template, with additional marking added as required, until no marked templates remain to be reconsidered. Each time an image point I_i is considered, any additional constraint on the gradient space contour of possible solutions to the basic equation $I(x,y) = R(p,q)$ propagates through this local filtering mechanism to all other image points under consideration.

The next issue to arise is the question of how to terminate the growth of templates. Currently, the program terminates only when some image point under consideration has no admissible gradient space interpretation, in which case the "forced" interpretation is deemed to have failed, or when all boundary points to existing templates have view angle greater than some preassigned value, in which case the surface is deemed to be

too oblique for further expansion. Clearly, these termination conditions are not adequate for a general surface analysis routine. The question of termination also raises deeper issues. How important are boundaries to surface interpretation? How can a complex surface be segmented into simpler (convex/concave) subsections? How does one handle the inherent indentation/protrusion ambiguity in trying to piece together surfaces? These issues must be dealt with. But, for now, let us consider a specific example in order to illustrate the method in operation.

AN EXAMPLE

Consider the simple example of a "Lambertian" sphere illuminated by a single distant light source. The intensity space to gradient space correspondence will be derived analytically and then it will be used as a basis for judging the performance of the computation.

The first task is to generate a reflectance map. To do this, a reflectivity function must be specified. Now, the term Lambertian refers to a phenomenological model of a perfect diffuse reflector such that the surface appears equally bright from all viewing directions. For such a surface, the reflectivity depends only on the foreshortening effect of the varying angle of incidence. In particular, the reflectance function $\phi(i.e.g)$ is given by:

$$\phi(i.e.g) = \cos(i)$$

This reflectivity function is transformed into a reflectance map by simply recalling the expression derived earlier for $\cos(i)$. The reflectance map is thus given by:

$$R(p,q) = \frac{1 + p p_s + q q_s}{\sqrt{1 + p_s^2 + q_s^2} \sqrt{1 + p^2 + q^2}}$$

The second task is to determine the local surface orientation of each image point. In the example, this result is particularly easy to obtain. Let the object sphere be centered at the origin and have radius r . Thus, the equation of the sphere is:

$$x^2 + y^2 + z^2 = r^2$$

Elementary calculus will verify that the vector (x,y,z) defines an (outward) normal at each point (x,y,z) on the surface. The appropriate gradient is obtained by rewriting this normal as $(-x/z, -y/z, 1)$. Now, for each (x,y) , there are actually two possible z values to be considered. Note, however, that in this example the hemi-sphere actually in view corresponds to negative values of z (recall figure 2).

The parameters of the image forming system can be factored out by assuming that the image intensity has been normalized to correspond directly to reflectivity. In this case, the intensity recorded at each image point (x,y) is equal to the value of the reflectance map at the corresponding (p,q) . Thus, the equation for image intensity is:

$$I(x,y) = R(-x/z, -y/z) \text{ where } z = -\sqrt{r^2 - x^2 - y^2}$$

Finally, the mechanism by which monotonicity relations were hypothesized for this example must be specified. One additional assumption was used. The algorithm assumes that it knows at least one image point corresponding to an object point oriented directly facing the viewer. Now, let such a point define a *pseudo-origin* in image space. Since the view angle is

assumed to be zero at the pseudo-origin, the only possible interpretation is that, in a particular direction, the view angle is locally non-decreasing with increasing image distance from the pseudo-origin. In general, one can not hope to assert any local monotonicity relation on direction of steepest descent based on angular position about the pseudo-origin. If, however, the surface is known to be convex, the direction of steepest descent is (locally) non-decreasing with increasing angular position about the pseudo-origin. For the example, the set of image points $I_1, I_2, \dots, I_n$ was ordered in view angle according to their distance in image space from the pseudo-origin and ordered in direction of steepest descent according to their angular position about the pseudo-origin. Applied together, these hypothesis rules are equivalent to the strong assumption that the surface in question is a convex solid of revolution.

For the example, $p_s = 0.7$, $q_s = 0.3$ and $R = 60$. A 128×128 test image was considered. The algorithm was applied to this image using 3×3 square templates sampled at an image spacing of 5 points (in both X and Y). The pseudo-origin was defined as $x = 0$ and $y = 0$. The results are first presented as a pair of figures. Figure 5 shows the reflectance map $R(p,q) = \cos(i)$ drawn as a series of contours (spaced 0.1 units apart). Superimposed are crosses marking the known gradient points (p,q) corresponding to the image points (x,y) sampled from the image. All gradient points to the left of the contour $R(p,q) = 0.0$ correspond to surface points oriented more than 90° away from the direction of incident illumination and hence to the self-shadowed region of the sphere. Figure 6 shows the restricted subsection of contour determined for each sampled image point. Again, crosses are superimposed to illustrate how well the algorithm has performed. The crosses mark the correct gradient points, determined analytically, while the corresponding subsection of contour marks how well the algorithm has isolated those points.

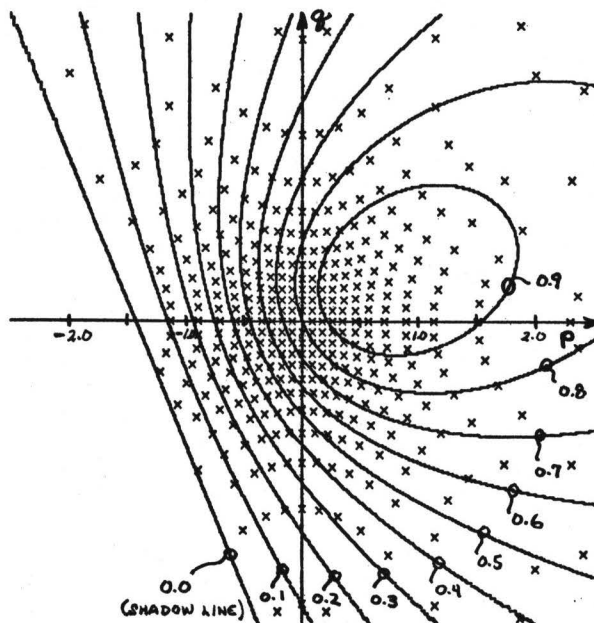


Figure 5

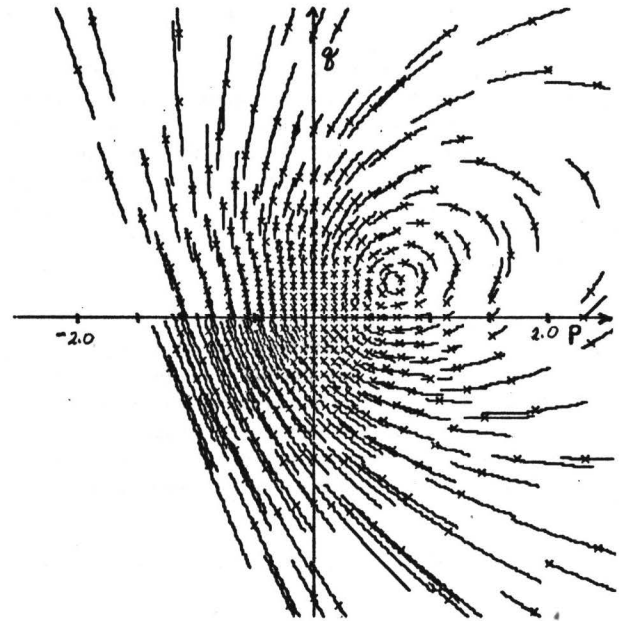


Figure 6

DISCUSSION

The ultimate criteria for judging the performance of this method is, of course, to discover whether it can be used to solve interesting problems. For now, however, the discussion is restricted to some simple statistics and some qualitative observations.

How well has the algorithm performed on this example? To answer this question, the freedom remaining at each image point I_i can be examined. Since each point in gradient space defines an orientation, this freedom can be characterized by the angular spread remaining in the corresponding gradient contour C_i . This freedom has been measured in two ways. First, define the angular spread in local surface orientation at image point I_i by:

$$\max \{ \theta \mid \cos(\theta) = \frac{(p_1, q_1, -1) \cdot (p_2, q_2, -1)}{\|(p_1, q_1, -1)\| \|(p_2, q_2, -1)\|} \mid (p_1, q_1), (p_2, q_2) \in C_i \}$$

Second, define the angular spread in view angle e at image point I_i :

$$\max \{ (\theta_1 - \theta_2) \mid \text{where } j = 1, 2; (p_j, q_j) \in C_i; \tan(\theta_j) = \sqrt{p_j^2 + q_j^2} \}$$

(View angle is useful because it determines the degree of surface foreshortening at each image point. Knowing the view angle, one can calculate the area of surface equivalent to a given area of image.) To tabulate the results, the sample points were split into two classes. First, consider all sample points within 45° view angle (ie. lying within gradient space circle $p^2 + q^2 = 1$). Second, consider all sample points within 60° view angle (ie. lying within gradient space circle $p^2 + q^2 = 3$). Table 1 summarizes the results:

	Points Within 45° View Angle	Points Within 60° View Angle
<u>Local Surface Orientation:</u>		
Mean Angular Spread	8.9°	9.9°
Standard Deviation	6.1°	7.0°
Best Case	0.5°	0.5°
Worst Case	36.3°	41.7°
<u>View Angle:</u>		
Mean Angular Spread	4.7°	5.0°
Standard Deviation	2.4°	2.5°
Best Case	0.5°	0.5°
Worst Case	10.9°	11.9°

Table 1

Note that these measures refer to total angular freedom. If a choice algorithm is adopted which selects the "correct" answer to be at the midpoint of the angular spread, this choice is guaranteed to be no more than half the angular spread in error. Thus, the upper right portion of table 1 can be interpreted as follows:

On the average, by sampling at an image spacing of 5 points in X and Y, the algorithm was able to position image points, corresponding to surface points less than 60° in view angle, to within 5° of their true orientation in space. The standard deviation of this measure over all such points was 3.5° while the worst case point was located to within 21° of its true orientation in space.

The performance of the algorithm depends critically on two factors: the ability to hypothesize monotonicity relations between selected image points and the topography of the reflectance map. The discussion of convexity gives some indication of how one can use *a priori* assumptions about surface geometry to specify monotonicity relations. In any event, given a particular hypothesis mechanism, the algorithm generates a constrained interpretation consistent with that hypothesis mechanism (or demonstrates that no such consistent interpretation is possible).

The topography of the reflectance map is determined by two factors: the local photometry of the surface being viewed and the light source, object surface and viewer geometry. In general, one would not expect to have much control over the local photometry of the surfaces being viewed. In principle, however, the user is free to vary the light source and viewer geometry to achieve optimal results with the algorithm. This approach has proven particularly useful in extending these techniques to the problem of inspecting for surface defects in metal castings (where it is feasible to construct inspection stations with independently controllable light sources <Woodham 1977>).

CONCLUSION

The notions of local surface orientation, gradient space and reflectance maps represent important conceptual tools for image analysis. The beauty of the gradient space formulation is that

it allows one to formulate physical constraints on the object surface as simple geometric constraints on the gradient space contour of possible solutions to the basic image forming equation $I(x,y) = R(p,q)$. The beauty of the "cooperative" algorithm is that it provides a simple mechanism for propagating these geometric constraints.

ACKNOWLEDGEMENTS

I wish to thank Berthold K. P. Horn, Mark Lavin, Tomas Lozano-Perez, David Marr and Patrick Winston for help with the ideas presented in this paper. Anna R. Bruss provided valuable proof-reading assistance. This research was conducted at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defence under Office of Naval Research contract number N00014-75-C-0643.

REFERENCES

- Horn, Berthold K. P. (1970), "Shape From Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View", MAC TR-79, M.I.T. Project MAC, November 1970.
(See also: Horn, Berthold K. P., "Obtaining Shape from Shading Information", in *The Psychology of Computer Vision*, P. H. Winston (editor), McGraw-Hill, 1975, pp 115-155.)
- Horn, Berthold K. P. (1975), "Image Intensity Understanding", AI-MEMO 335, AI Laboratory, MIT, August 1975.
(See also: Horn, Berthold K. P., "Understanding Image Intensities", in *Artificial Intelligence*, Vol 8, 1977 <in publication>)
- Huffman, David A. (1971), "Impossible Objects as Nonsense Sentences", in *Machine Intelligence 6*, R. Meltzer and D. Michie (editors), Edinburgh University Press, pp 295-323.
- Huffman, David A. (1975), "Curvature and Creases: A Primer on Paper", *Proceedings of Conference on Computer Graphics, Pattern Recognition and Data Structures*, May 1975, pp 360-370.
- Mackworth, Alan K. (1973), "Interpreting Pictures of Polyhedral Scenes", in *Artificial Intelligence*, Vol 4, 1973 pp 121-137.
- Marr, David & Poggio, Tomaso (1976), "Cooperative Computation of Stereo Disparity", AI-MEMO 364, AI Laboratory, MIT, June 1976.
- Woodham, Robert J. (1977), "Reflectance Map Techniques for Analyzing Surface Defects in Metal Castings", Ph.D. Thesis, AI Laboratory, M.I.T., <in preparation>