MASSACHUSETTS INSTITUTE OF TECHNOLOGY

ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo 490 August 1978

DETERMINING SHAPE AND REFLECTANCE USING MULTIPLE IMAGES

Berthold K. P. Horn, Robert J. Woodham & William M. Silver

Abstract. Distributions of surface orientation and reflectance factor on the surface of an object can be determined from scene radiances observed by a fixed sensor under varying lighting conditions. Such techniques have potential application to the automatic inspection of industrial parts, the determination of the attitude of a rigid body in space and the analysis of images returned from planetary explorers. A comparison is made of this method with techniques based on images obtained from different viewpoints with fixed lighting.

This report describes research done 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 Office of Naval Research under contract N00014-77-C-0389.

INTRODUCTION

Reflected radiance depends on incident radiance, surface orientation and the properties of the surface material. An imaging device produces image irradiances proportional to scene radiances. Consequently the gray levels, quantized measurements of the image irradiance values, contain a great deal of information about the scene being imaged. This information is present in the raw gray levels in a coded form however. To learn more about the objects being viewed, it is necessary to carefully analyze the gray levels in order to infer the surface properties of the objects. This is not easy since a single gray level value depends on multiple factors. To constrain possible solutions to this problem it is helpful to have available more than one image taken from the same viewpoint but with varied lighting conditions. At each point in the image one then has more than one constraint on the possible surface orientation and surface reflectance factor.

Such reconstructions of a scene being viewed depend on a detailed understanding of the imaging process. In addition, the imaging instrument must be of high caliber, so that the gray levels produced can be dependably related to scene radiances. Fortunately our understanding of image formation and the physics of light reflection has advanced sufficiently, and the quality of imaging devices is now high enough, to make this endeavour feasible. The details of a number of particular cases have been developed. Much remains to be done, particularly testing of these ideas on images of practical importance. Further analytic development will also be required to define the exact conditions for which unambiguous solutions are possible and to construct algorithms that can find these solutions efficiently.

It will perhaps be useful to list here some cases that have been explored so far.

(1) Determining the shape from a single image

Understanding of the human visual system, and the construction of artificial vision systems, require an appreciation of shading, the variation of reflected light intensity with surface orientation. The scene radiance seen by the image sensor is the product of the scene irradiance, and the reflectance. The reflectance in turn is a function of the direction of the surface normal relative to the direction to the light source and the direction to the image sensor. The reflectance function discussed here can be related to the bidirectional reflectance-distribution function (BRDF) defined by the National Bureau of Standards [1]. It depends on the material of which the surface is composed as well as its microstructure.

Let $R(p, q)$ be the scene radiance of a surface element with gradient (p, q), where p and q are the first partial derivatives of the surface elevation with respect to coordinate axes parallel to the image axes. We have called R(p, q) the "reflectance map" [2, 3]. It gives scene radiance as a function of surface gradient and can be determined from the BRDF for the surface material and the distribution of incident radiance. If the distribution of lightsources is known, and the object has a uniform surface covering, its shape can be calculated from the image irradiances in a single image [2, 4]. If I(x, y) is the image irradiance at the image point (x, y), then the relevant equation is

$$I(x, y) = R(p, q). \tag{1}$$

Until recently, the calculation of shape from shading could be performed only by a rather tedious method involving the direct solution of the above nonlinear first-order partial differential equation, using something like the method of characteristic strip expansion. Progress has been made in the development of an iterative, local method based on relaxation or "cooperative computation" [3, 5]. It is interesting, by the way, that a special case applicable to the material in the maria of the moon had been solved earlier by Rindfleisch [6]. This in turn was inspired by work on the special case of surfaces with very shallow angles between the local tangent plane and the incident light [7]. The method of Rindfleisch, incidentally, was based on an assumed reflectance function that violates basic physical laws [8].

(2) Determining shape from two images obtained with different lighting.

If two images taken from the same point of view are available, then two constraints apply at every image point. The possible directions of the local surface normal form a two parameter family. It seems reasonable then that enough information might be available to find the surface normal by local computations. This method, which has been called photometric stereo [5, 9, 10], produces information about the shape of an object in the form of a distribution of surface normals. This kind of representation for shape has been favored recently in the machine vision community because of several advantageous properties [2, 11]. As an object rotates, for example, the surface normals undergo a rather simple transformation, while distances to the surface change in less symmetrical fashion. Here one obtains a pair of equations,

$$I_1(x, y) = R_1(p, q) \tag{2}$$
$$I_2(x, y) = R_2(p, q) \tag{3}$$

The calculation of surface normal from the two image irradiances at each point is local and can be implemented by means of a simple two dimensional look-up table. This table in essence is the inverse of a table containing the two image irradiances indexed on the two components of the surface normal. The calculation is local, rapid and not subject to possible error propagation as is the serial method for single images mentioned above. There may be some ambiguity that cannot be resolved locally, since more than one surface normal orientation may produce the same pair of scene radiances. This difficulty may be resolved by an iterative "relaxation" computation in which compatible solutions for neighbouring points are reinforced while conflicting assignments are eliminated.

It is possible to produce inconsistent surface normals as a result of noise in the scene radiance measurements or as a result of depth discontinuities (perhaps even due to discontinuities in gradient). This possibility arises because the surface normal is perpendicular to a local tangent plane, which in turn is defined by the partial derivatives of distance to the object's surface with respect to two orthogonal directions parallel to the coordinate axes of the image plane. From the representation in terms of surface normals one can thus calculate the gradient [2] of the surface. For a smooth surface the partial derivatives along orthogonal directions are not independent -- the second partial derivative with respect to the two directions taken in one order must equal the second partial derivative obtained with the reverse order.

Little has been done to explore methods that use violations of this condition to segment the image into regions belonging to different objects, or to iteratively refine the solution using the redundant information available in areas where there are no depth discontinuities.

(3) Determining shape and reflectance using two images.

In many real situations the objects do not have uniform surface properties. In this case the methods described above will be misled, much as a human is by carefully applied facial make-up. It is thus of interest to explore methods that can solve for surface reflectance factor and shape simultaneously. There is not enough information in a single image to accomplish this task in general. Under appropriate circumstances it is not possible to distinguish a photographic print from the scene portrayed in the photograph. The same image irradiances are produced in a monocular image sensor in the one case by a flat surface with varying reflectance and in the other case by three-dimensional shapes that may have uniform surface reflectance properties.

Two images do provide enough information however. Indeed, it was apparent in the previous section that two images contain more information than is required to solve for shape alone. If the surface reflectance factor or "albedo" is $\rho(x, y)$ then one obtains two equations,

$$I_1(x, y) = \rho(x, y) * R_1(p, q) \qquad (4)$$
$$I_2(x, y) = \rho(x, y) * R_2(p, q) \qquad (5)$$

The variable surface reflectance factor can be removed by dividing the two images (provided the reflectance factor or "albedo" appears as a multiplicative factor). If we let I_{12} be the ratio of I_1 and I_2, and R_{12} be the ratio of R_1 and R_2, then

$$I_{12}(x, y) = R_{12}(p, q) \tag{6}$$

To this peculiar, new "image" one can then apply the shape from shading methods mentioned in section (1) above for the determination of shape from a single image. Finally, one can recover the reflectance factor by dividing one of the real images by a synthetic image obtained using the shape so determined [12].

$$\rho(x, y) = I_1(x, y)/S_1(x, y) \tag{7}$$

Here S_1 is a synthetic image computed using the reflectance map $R_1(p, q)$. In practice, to assure numerical accuracy, it would be better to perform this operation on the sum of the two images (using the sum of the two reflectance maps). The spatial distribution of surface gradient obtained in this fashion, as well as the spatial distribution of reflectance factor, are examples of "intrinsic images" [13], distributions of underlying information that can be extracted from the raw image irradiances. The method described here can be criticized on the same grounds as the basic shape from shading method in that it requires numerical solution of a non-linear first-order partial differential equation. Also, more complicated methods are required if the varying reflectance is caused by variations in surface cover that cannot be modelled simply as multiplication by a varying reflectance factor.

(4) Determining shape and reflectance using three images.

When three images taken under different lighting conditions are available, enough constraint is available to solve for the surface normal and reflectance factor locally. In fact, three pieces of information are available at every point and three variables are needed to pin down the two components of local surface orientation and the reflectance factor. If the reflectance function has a particularly simple form an analytic solution for the surface normal and the reflectance factor is possible [10]. If this is not the case, one may use a simple algorithm based on a three-dimensional look-up table indexed by the three gray levels found in the images at corresponding image points. Here the three equations of interest are,

$$I_1(x, y) = \rho(x, y) * R_1(p, q) \tag{8}$$
$$I_2(x, y) = \rho(x, y) * R_2(p, q) \tag{9}$$
$$I_3(x, y) = \rho(x, y) * R_3(p, q) \tag{10}$$

A better method again uses ratios of images to remove the variable surface reflectance factor. If for example two new "images" are generated by dividing two pairs of the original three images, these can be used to determine the shape as described above in section (2). If $\mathbf{I}_{12}$ is the ratio of $\mathbf{I}_1$ and $\mathbf{I}_2$ and so on, then the relevant equations are,

$$I_{12}(x, y) = R_{12}(p, q) \tag{11}$$
$$I_{23}(x, y) = R_{23}(p, q) \tag{12}$$

A two dimensional look-up table is appropriate for this computation. The surface reflectance factor can once again be determined after the surface shape has been found, by division of one of the real images with a synthetic image created using the shape information. For reasons of numerical accuracy it is actually more appropriate to use other combinations of two images than their ratio -- it may be convenient to work with the ratio of their difference and their sum for example.

$$[I_2 - I_1]/[I_2 + I_1] = [R_2 - R_1]/[R_2 + R_1] \tag{13}$$

Many details remain to be worked out. It appears for example that the positions of the light sources must be choosen carefully so as to insure that the three measurements of image intensity are in fact independent [10]. When the directions from the objects being viewed to the light sources all lie in a plane no new information may be obtained from the third image.

Comparison with ordinary stereo methods

Shading is not usually the first depth cue that comes to mind. Stereo, based on two images taken from different points of view, with the same lighting, will usually be mentioned first. This is curious, since shading in many cases is the more important cue, particularly for printed material containing pictures of smooth, rounded objects such as people. Nevertheless, stereo is a method of great practical importance, since topographic maps and digital terrain models [14] are derived from aerial photography by this technique.

Elevations can be found by a human operator who fuses corresponding image details from the two pictures [15, 16], or by means of "automated" methods, usually based on correlation [17, 18, 19, 20]. While automated methods still leave much to be desired in terms of robustness in the face of highly variable terrain and terrain cover (or smooth areas without surface markings), many ideas from the machine vision community point to more sophisticated techniques for the solution of the correspondence problem [21, 22, 23, 24]. At the cost of greater computational effort, these promise to produce excellent results. While much remains to be done, it seems likely that really automated stereo machines will eventually be commonplace.

Methods using radiometric measurements in one or more image from the same viewpoint to determine surface shape and reflectance factor can be considered complementary to methods based on the identification of corresponding points in two images taken from different viewpoints:

(1) Stereo allows the accurate determinations of distances to objects. Radiometric methods on the other hand are best when the surface gradient is to be found.

(2) Stereo works well on rough surfaces with discontinuities of surface orientation. Radiometric methods work best on smooth surfaces with few discontinuities.

(3) Stereo does well with "painted" or textured surfaces with varying surface reflectance. Radiometric methods are best when applied to surfaces with uniform surface properties.

Methods based on analysis of radiometric measurements do have some unique advantages however:

(1) There is no difficulty identifying points in the two images that correspond to the same object point since the images are taken from the same point of view. This is the major computational task in the analysis of stereo information.

(2) Under appropriate circumstances, the surface reflectance factor can be determined because the effect of surface orientation on scene radiance can be removed. Stereo provides no such capability.

(3) Describing the shape of an object in terms of its surface normal is preferable in a number of situations to description in terms of elevation about some reference plane. This is the natural form of output from the radiometric methods.

References

[1] Nicodemus, F. E., Richmond, J. C. & Hsia, J. J. (1977) "Geometrical Considerations and Nomenclature for Reflectance," NBS Monograph 160, National Bureau of Standards, Washington, D. C.

[2] Horn, B. K. P. (1977) "Understanding Image Intensities," Artificial Intelligence, Vol. 8, No. 11, pp 201-231.

[3] Woodham, R. J. (1977) "A Cooperative Algorithm for Determining Surface Orientation from a Single View," Proc. 5th Int. Joint Conf. on Artif. Intell., M. I. T., Cambridge, Mass, August 1977, pp 635-641.

[4] Horn, B. K. P. (1975) "Determining Shape from Shading," Chapter 4, in Winston, P. H. (Ed) The Psychology of Computer Vision, McGraw-Hill, N. Y.

[5] Woodham, R. J. (1978) "Reflectance Map Techniques for Analyzing Surface Defects in Metal Castings," TR-457, A. I. Laboratory, M. I. T., Cambridge, Mass.

[6] Rindfleisch, T. (1966) "Photometric Method for Lunar Topography," Photogrammetric Engineering, March 1966.

[7] Diggelen, J. van (1951) "A Photometric Investigation of the Slopes and Heights of the Ranges of Hills in the Maria of the Moon," Bulletin of the Astronomical Institute of the Netherlands, July 1951.

[8] Minnaert, M. (1941) "The Reciprocity Principle in Lunar Photometry," Astrophysical Journal, Vol. 93, pp 403-410.

[9] Woodham, R. J. (1978) "Photometric Stereo: A Reflectance Map Technique for Determining Surface Orientation from Image Intensity", Proceedings of SPIE's 22nd Annual Technical Symposium, Vol. 155, August 1978.

[10] Woodham, R. J. (1978) "Photometric Stereo," A. I. Memo 479, M. I. T., Cambridge, Mass. June 1978.

[11] Marr, D. & Nishihara, H. K. (1977) "Representation and Recognition of the Spatial Organziation of Three-Dimensional Shapes," Proc. Roy. Soc. Lond. B. 200 pp 269-294.

[12] Horn, B. K. P. & Bachman, B. L. (1978) "Using Synthetic Images to Register Real Images with Surface Models," C. A. C. M., Nov 1978.

[13] Barrow, H. G. & Tenenbaum, J. M. (1978) "Recovering Intrinsic Scene Characteristics from Images," in Hanson, A. & Riseman, E. (Eds) Computer Vision Systems, Academic Press, N. Y.

[14] Proc. of the Digital Terrain Models (DTM) Symposium, (1978), American Society of Photogrammetry, St. Louis, Missouri, May 9-11, 1978.

[15] Wolf, P. R. (1974) Elements of Photogrammetry, McGraw-Hill, N. Y.

[16] Liebes, S. Jr. & Swartz, A. A. (1977) "Viking 1975 Mars Lander Interactive Computerized Video Stereophotogrammetry," Journal of Geophysical Research, Vol. 82, No. 28, pp 4421-4429.

[17] "B8 Stereomat Automated Plotter," Wild Heerbrugg & Raytheon (1965), unpublished company brochure.

[18] "AS-11B-X Automated Stereo Mapper," (1976) RADC-TR-76-100 Rome Air Development Center, Griffis Air Force Base, N. Y.

[19] Kelly, R., McConnell, P. & Mildenberger, S. (1977) "The Gestalt Photomapping System," Photogrammetric Engineering and Remote Sensing, Vol. 42, No. 11.

[20] Brunson, E. B. (1977) "Operational use of the Gestalt Photo Mapper II," paper prepared for American Society of Photogrammetry Semi-Annual Meeting, Little Rock Arkansas.

[21] Gennery, D. B. (1977) "A Stereo Vision System for an Autonomous Vehicle," Proc. 5th Int. Joint Conf. Artif. Intell., August 1977, M.I.T., Cambridge, Mass.

[22] Arnold, R. D. (1978) "Local Context in Matching Edges for Stereo Vision," Proc. DARPA Image Understanding Workshop, Baumann, L. (ed), Science Applications, Inc., Arlington, VA. May 1978.

[23] Marr, D. & Poggio, T. (1976) "Cooperative Computation of Stereo Disparity," Science, Vol. 194, pp 283-287.

[24] Marr, D. & Poggio, T. (1977), "A Theory of Human Stereo Vision," A. I. Memo 451, M. I. T., Cambridge, Mass. November 1977.