AAAI -87

Shadow Stereo -- Locating Object Boundaries Using Shadows

William B. Thompson
Computer Science Department
University of Minnesota
Minneapolis, MN 55455

Michael T. Checky
Computer Science Department
University of Minnesota
Minneapolis, MN 55455

William F. Kaemmerer
Systems Development Division
Honeywell Inc.
Golden Valley, MN 55427

Abstract

Shadows are a useful source of information about object structure. Shadows cast under oblique lighting often indicate the location of the silhouette of an object. This paper describes a method for reliably detecting shadow edges corresponding to object edges. It is able to distinguish between detected edges due to shadows and those due to surface markings. The basis of the technique is to observe the differences in shadows due to changes in the direction of illumination. Analysis is further aided by a simple stereo technique that does not require a solution to the general correspondence problem. Both the multi-light source and multi-camera methods can be implemented in an extremely efficient manner.

I. Introduction.

This paper outlines a method for finding part boundaries using an approach combining structured lighting and stereo techniques. The method uses multiple light sources and multiple cameras to determine object boundaries based on detected shadows in the images. It is effective at distinguishing dark portions of an image due to shadows from those due to surface markings. The combined approach allows for significant simplifications in each component technique. The structured light required consists of collimated illumination from a small number of fixed light sources, rather than a more complex requirement that patterned illumination be projected onto or scanned over the part. Likewise, since the multi-camera portion of the analysis is used only to determine whether a surface point is on or off of the ground plane, the correspondence problem involved is simplier than that inherent in most other stereo techniques.

The basis of the method is that obliquely lit objects cast shadows in such a way that shadow boundaries in an image are often coincident with object boundaries. The method deals in a direct way with two critical problems in analyzing shadows: 1) simple thresholding is insufficient to accurately recognize shadows in situations with significant variations in surface reflectance, and 2) many shadows are physically detached from the objects generating them. The use of multiple light sources allows a simple and computationally efficient filtering of dark portions of the image, leaving only those regions corresponding to true, attached shadows. The use of multiple views is able to further filter shadow edges, eliminating those due to internal structure of the object and leaving only the object silhouette.

In its current form, the method is designed to find the boundaries of isolated parts on a flat supporting surface. The method is subject to some limitations on the geometry of the parts. However, a strength of the technique is the ability to cope with supporting surfaces which are not visually distinct from the parts. The system is appropriate for the control of automated pick and place operations off a typical conveyor belt or pallet, with an almost arbitrary surface pattern and coloring.

The use of shadows has received surprisingly little attention within the computer vision community, and as a result many of the vision methods which have been developed work effectively only in situations in which the illumination is highly diffuse. Methods which do deal directly with shadows tend to be computationally complex. In addition, most assume that shadows can be easily detected by simple thresholding operations. In fact, thresholding is often ineffective due to variations in surface reflectance, secondary reflections, and other related effects. [Waltz, 1975] demonstrated that shadow information can add constraints that simplify the analysis of simple blocks world problems. [Huertas and Nevatia, 1983] and [Hambrick and Loew, 1985] infer the shapes of objects casting shadows based on an analysis of shadow shape and knowledge of the direction of illumination. [Shafer and Kanade, 1983] use shadow shape to infer surface orientation. [Kender and Smith, 1986] introduces the idea of using moving shadows (generated by moving light sources) to aid in determining surface shape.

II. Types of Apparent Shadow Boundaries.

We are interested in inferring the location of the silhouette of an object from the locations of apparent shadow boundaries in the image. To do this, we need to classify each apparent shadow boundary as arising either from an "actual" object boundary, or from some other factor in the scene. We refer to object boundaries making up the silhouette as *exterior boundaries*. Non-convex objects may generate shadows due to surface structure unrelated to the object's silhouette. The object structures generating such shadows will be

This work was supported by National Science Foundation Grant DCR-8500899.

William F. Kaemmerer was with the Control Data Corporation Image Systems Technology Center during his participation in this work.

referred to as *interior boundaries*.

Shadow regions identified in an image (for example, by thresholding) can be *true* shadows, correlated to actual shadows in the scene, or *false* shadows resulting, for example, from surface markings. Each point on the boundary of a true shadow region can be either *attached* or *detached*. An attached shadow boundary point is coincident (in the scene) with the object generating the shadow. Detached shadow boundary points correspond either to the "far" side of the shadow, or to parts of "cast" shadows associated with non-convex objects. Attached shadow boundary points always are associated with a point on objects. Detached shadow boundary points may lie on either the supporting ground place, or an object in the scene (possibly the same as the object casting the shadow.)

Any automatic procedure identifying shadows from raw intensity data will also find false shadow regions. These are regions of dark intensity in the image which arise from *surface markings*, that is, spatial patterns in the colors on a surface in the scene, rather than from differences in the illumination falling on the surface. Such markings can occur on either objects or on the ground plane.

Finally, shadow boundaries can be generated by occlusions. One type of such a boundary arises from the occlusion of a portion of a shadow region by another illuminated surface nearer to the image plane. The resulting shadow boundary is false in the sense that it does not reflect the shape of the true shadow in the scene; nevertheless, it provides useful information about the location of an external boundary of an object in the image. Another type arises from a self-occlusion, which occurs when a shadowed surface in the scene curves out of view in front of an illuminated background.

In summary, a consideration of the ways shadow regions can appear in an image leads to seven types of apparent shadow boundaries:

1) attached shadow boundaries
2) detached shadow boundaries on ground plane
3) detached shadow boundaries on an object
4) surface markings on ground plane
5) surface markings on object
6) occlusion of shadow by object
7) self-occlusion of shadowed surface

Of these seven, types 1, 6, and 7 provide information on the location of external boundaries of objects in the scene. A goal of the following analysis is to develop a technique for identifying these types of shadow boundaries among all the shadow boundaries apparent in an image.

III. Assumptions.

A number of simplifying assumptions are made to facilitate the analysis:

1) Only primary illumination effects are considered; secondary illuminations from reflections are assumed to be faint enough to be ignored.

2) A supporting ground plane at a known location is assumed to exist.

3) All objects are assumed to be resting directly on the ground plane.

4) The surface curvature of the objects at the points corresponding to their exterior boundaries in the image is assumed to be high, relative to moderate changes in the angle of illumination of the scene. In particular, moderate changes in the angle of illumination should not significantly affect the location of attached shadow boundaries in the image.

5) It is assumed that it is sufficient to find the location of boundaries defined as the image of silhouettes produced by a projection along the direction of illumination of the scene, rather than by a projection from the point of view. In many applications, the relative positions of light sources, cameras, and the angle of the object surface at exterior boundaries are such that the difference between projections are within the tolerance of any actions based on the visual analysis.

One further assumption will be introduced during the course of the analysis to lead to the final solution:

6) Detected shadow regions are assumed to be a manifestation of a single type of underlying cause. For example, it is presumed that detected regions are not combinations of real shadows and surface markings. (See section VII.)

IV. Leading and Trailing Edges.

A useful first step in classifying the apparent shadow boundaries in an image is to identify *leading* versus *trailing* shadow edges [Hambrick and Loew, 1985], as follows: Define a *projected illumination direction* by projecting the direction of illumination onto the image plane. Define a shadow leading edge as a transition from light to dark while moving in the projected illumination direction. A trailing edge is a transition from dark to light, while moving in the same direction. Leading edge-trailing edge pairs are identified by associating with each leading edge point the first trailing edge point found by moving in the projected illumination direction.

Each leading edge and each trailing edge in the image is part of an apparent shadow boundary, and hence must be one of the seven types described in the last section. As a result, seven types of leading edges, and seven types of trailing edges may be conceptually identified, for a total of 49 cases. These cases are shown in Table 1. Ideally, one would like to have an image analysis technique which allows all of the unique cases among the 49 to be distinguished. However, for the current purpose it is sufficient to have a technique which will distinguish cases in which leading edges provide information about boundaries from among the other cases:

1 – Exterior boundary of back lit object.

8 – Exterior boundary (casting "typical" shadow).

13 – Shadow cast by an overhanging part of an object, partly occluded by another overhanging part of an object.

15 – Interior boundary (surface protrusion of object casting shadow onto the same object) *or* shadow cast by an object onto another nearby object.

20 – Shadow cast by an overhanging part of an object onto another object, partly occluded by an overhanging part of a third object (possible with highly oblique lighting).

22 – Shadow cast onto a surface marking on ground.

27 – Surface mark on ground partly occluded by (unmarked) object.

29 – Shadow cast onto a surface marking on an object.

34 – Surface mark on object partly occluded by an overhanging part of an object.

36 – Shadow cast by an object partly occluded by an overhanging part of an object.

41 – Shadow of one object viewed through hole in an overhanging object, or gap between overhanging parts of objects.

43 – Self-shadowed surface curving out of view before light background.

47 – Surface mark on self-shadowed object that curves out of view.

V. Multiple Light Source -- Multiple Camera Technique.

Two sources of information are used to assist in interpreting the shadow boundaries in an image.

Multiple light sources --

A set of images is taken from the same camera position but using different illumination directions. Illumination is varied in such a way that the direction of the projection of the illumination direction vector onto the ground plane remains constant. The image of detached shadow boundaries will move when the illumination direction changes, because the location of the detached boundary is a trigonometric function of illumination angle. Attached, occluded, and surface marking boundaries will not move due to small changes in illumination. (There is an exception for attached boundaries due to low curvature surfaces -- but note assumption 5, above.) It is useful to have several sets of illumination sources, each set individually satisfying the requirement for a common projected direction. In this way, all exterior boundaries of objects in the scene cast shadows in at least some of the images, and it is possible to deal with object boundaries parallel to the projected direction of illumination.

Multiple cameras--

Imaging the scene using multiple cameras allows stereo techniques to be used to determine whether boundaries are on

Trailing Edge:	Leading Edge: *attached*	*detached on ground*	*detached on object*	*surf. mark on ground*	*surf. mark on object*	*occlusion of region*	*self-occlusion*
attached	1	2	3	4	5	6	7
detached on ground	8	9	10	11	12	13	14
detached on object	15	16	17	18	19	20	21
surf. mark on ground	22	23	24	25	26	27	28
surf. mark on object	29	30	31	32	33	34	35
occlusion of region	36	37	38	39	40	41	42
self-occlusion	43	44	45	46	47	48	49

Table 1: Situations underlying combinations of leading edge - trailing edge apparent shadow boundary types.

or off of the ground plane. The analysis is simple, because stereo triangulation is required only at shadow boundaries and because only an on versus off ground plane determination is required, not an actual depth measurement. *On-off stereo* solves this problem without the need for a solution to the general correspondence problem. The technique is based on the fact that different views of the ground plane will be distorted in systematic ways due to the camera projection functions. Knowing the camera models, it is possible to determine the view of the ground plane seen from one camera given the view in the other. In fact it is not necessary to know the camera models. A target such as a checkerboard can be placed on the ground plane and then conventional image warping techniques can be used to determine the transformation between two different views. (This involves a "correspondence problem," but one of a very simple sort.) The warping will be the same for any patterns on the ground plane. Given two views of a shadow edge lying on the ground plane, the location of the edge in one view will be the same as the transformed location of the same edge in the second view. If the edge is at a height different from the ground plane, however, the warping transformation will not accurately predict the change between different views. To identify edges not on the ground plane, we need only warp one image to correspond to the viewing point of the other, and then look for edges that "move."

VI. LS–LM/CS–CM Designation.

Given the change in the illumination direction from image-to-image, and the determination of on/off ground plane from the stereo views, each apparent shadow boundary in an image of the scene may be labeled as moving or stationary with respect to the lighting and camera position changes. Shadows which move with changes in illumination may be labeled LM; those which do not may be labeled LS. Shadows which move (relative to their predicted location if they were at the level of the ground plane) with a change in camera viewpoint may be labeled CM; those which do not may be labeled CS.

Each apparent shadow boundary thus may be given one of four labels based on information from the images: LS–CS, LS–CM, LM–CS, or LM–CM. When these labels are applied to leading-trailing edge pairs, sixteen possible cases result. These are the sixteen cases which may be distinguished using information available in the image. The goal of the current analysis is attained if the 49 different situations (from Table 1) can be mapped onto these 16 cases in a manner such that the shadows arising from exterior object boundaries are shown to be distinguishable from all other shadows, using the information available in an image.

The desired mapping can be constructed by assigning an LS–LM/CS–CM label to each of the seven types of apparent shadow boundaries, based on a consideration of their behavior under changes in lighting and camera positions. Attached shadow boundaries do not move with lighting changes, but they move with changes in camera position, since actual three-dimensional structure to which they are attached must be above the ground plane (LS–CM). Detached shadow boundaries do move with changes in illumination direction. Those on the ground plane appear stationary from different camera viewpoints (LM–CS), and those falling onto objects appear to move (LM–CM). The boundaries of surface markings are stationary regardless of illumination direction, and move with camera viewpoint depending on whether they are on the ground plane (LS–CS) or an object (LS–CM).

VII. Leading Edge Interpretation and Filtering.

Given this labeling of the types of apparent shadow boundaries, the mapping of the 49 situations to the 16 cases involves a straight-forward "collapsing" of the rows and columns of Table 1, as shown in Table 2. (Cells in Table 1 which are not physically realizable are omitted.)

Trailing Edge:	Leading Edge: LS–CS	LS–CM	LM–CS	LM–CM
LS–CS	25	*22*, 26, *27*	23	24
LS–CM	32, 39	*1*, *29*, 33, *34*, *36*, 40, *41*, *43* *47*, 48	30, 37	31, 38, 45
LM–CS	11	*8*, 12, *13*	9	10
LM–CM	18	*15*, 19, *20*	16	17

Table 2: Mapping of Shadow-Producing Situations Onto Labeling Possibilities for Leading-Trailing Edge Pairs.

Numbers in cells correspond to the situations in Table 1. Italicized numbers are those situations for which the leading edge of the pair can provide information on the location of an exterior boundary.

In Table 2, the situation numbers which are italicized are those for which the leading edge of the apparent shadow boundary pair is indicative of the location of an actual exterior object boundary in the scene. Inspection of this table shows that there is no image information condition (i.e., joint labeling of leading and trailing edges as LS–LM/CS–CM types) which uniquely selects the situations providing the location of exterior boundaries. However, the additional simplifying assumption that shadow regions in the image are due to a single manifestation rules out the more "esoteric" shadow-producing situations, (situation numbers 22, 29, and 47.) The assumption that shadows are due to a single manifestation is not as restrictive as it might first appear. In fact, we do not require that the whole shadow satisfy this constraint, but only that each leading/trailing edge pair be due to a single cause. Table 3 shows the mapping when these situations are eliminated. Under this assumption, there are three joint labeling conditions which may interpreted as providing information (from the leading edge) concerning the location of an exterior boundary in the scene: LS–CM/LS–CS pairs, LS–CM/LM-CS pairs, and LS–CM/LM–CM pairs. Situation 15 can arise due to two distinct causes and thus is associated with an intrinsic ambiguity. We adopt a conservative approach by ignoring all LS–CM/LM–CM pairs.

Trailing Edge:	Leading Edge: LS–CS	LS–CM	LM–CS	LM–CM
LS–CS	25	**27**		
LS–CM	39	*1*, 33, *34*, *36*, 40, *41*, *43*, 48	37	38
LM–CS		**8**, **13**	9	
LM–CM		**15***, **20**		17

Table 3: Table 2 After Deleting Situations In Which The Shadow Region Has Multiple Causes.

Boldface numbers (**27**, **8**, **13**, **15**, and **20**) correspond to cases which can be unambiguously discriminated.

The result of the analysis is the finding that two types of leading-trailing edge pair labelings can be assumed to provide information on the location of exterior boundaries (LS–CM/LS–CS and LS–CM/LM–CS pairs). Operationally, this may be implemented by considering all leading shadow edges which are labeled LS–CM to be candidates for exterior boundary locations, and checking the labeling of their corresponding trailing edges. Those with trailing edges labeled LS–CS or LM–CS are retained, and all other leading edges are discarded.

VIII. Examples and Discussion.

Figures 1 and 2 show a simple example. We want the outline of the large box. Simple edge detection alone is obviously not appropriate. Thresholding can be used to find the shadows, but the problem is complicated by surface markings on the supporting surface, and a smaller structure on top of the box that both casts a shadow and has a surface marking. Figures 1a and 1b show two views of the scene differing in illumination position. Figures 1a and 2 show two views of the scene differing in camera position. Figure 3 shows potential shadows detected in figure 1a by thresholding (marked in white). Note that the shadow corresponding to the silhouette of the box is found, along with a variety of interior shadows and surface makings on both the box and ground plane. The thresholded results for the other images are similar. Figure 4 shows the leading and trailing edges of the regions indicated in figure 3. Leading edges are marked by a thin white line, trailing edges by a thicker white line. Figure 5a shows the

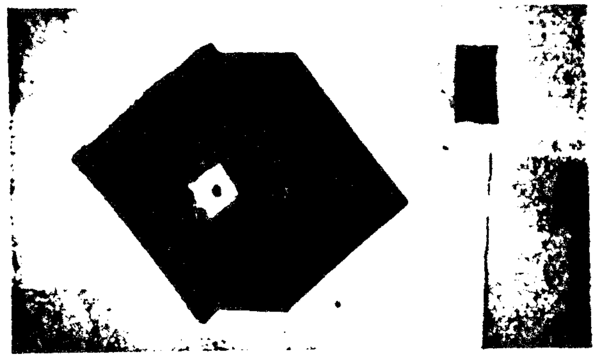

Figure 1a: View of box and shadows.

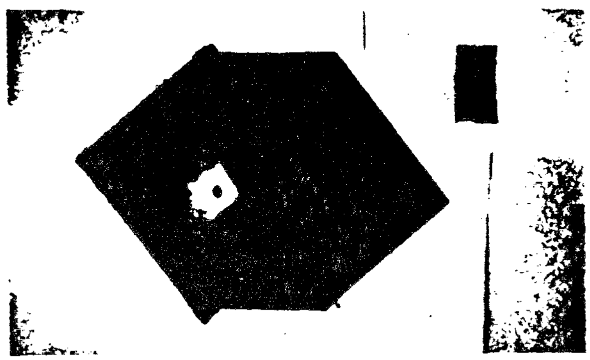

Figure 1b: Same as 1a, but with different illumination direction.

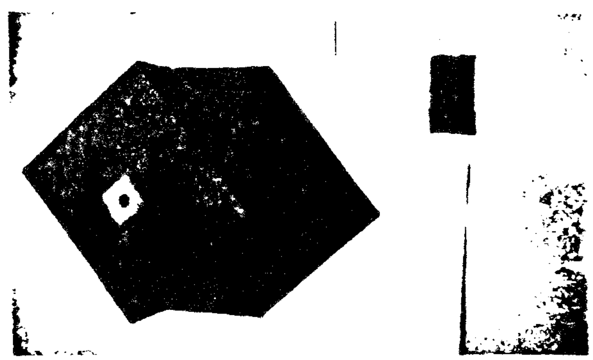

Figure 2: Same as figure 1a, but from different camera.

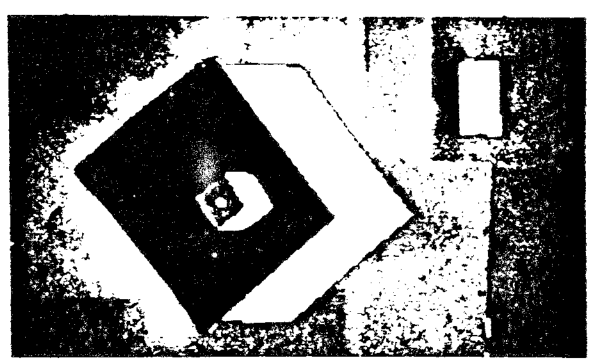

Figure 3: Thresholded regions in figure 1a.

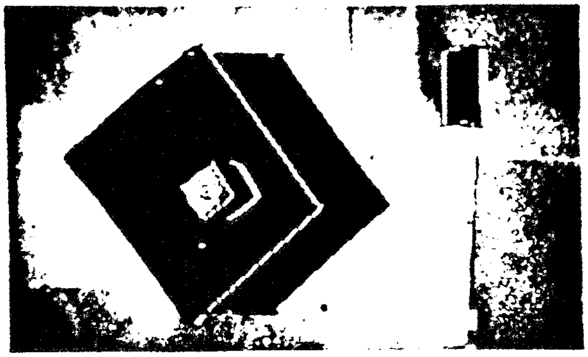

Figure 4: Leading and trailing edges in figure 3.

edges that are stationary or moving with respect to moving illumination. Stationary edges are marked by a thin white line, moving edges by a thicker white line. Figure 5b shows the edges that are stationary or moving with respect to alternate view points (the coding is the same as for figure 5a).

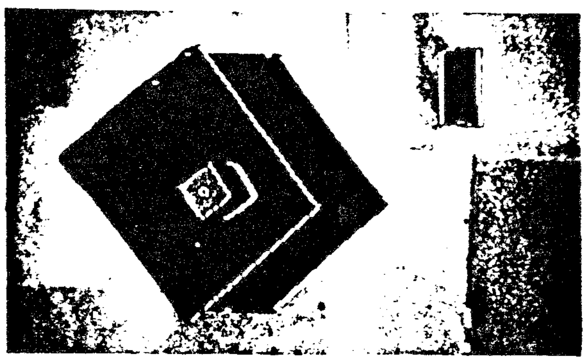

Figure 5a: Stationary/moving edges with respect to light source.

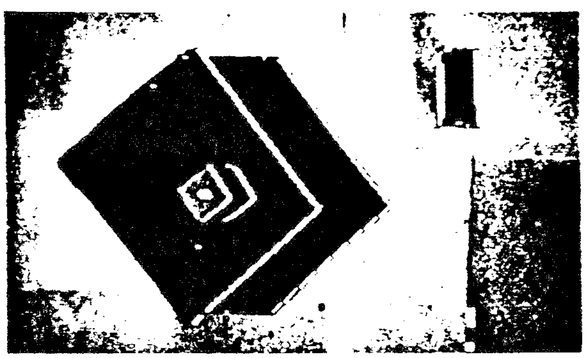

Figure 5b: Stationary/moving edges with respect to alternate cameras.

Figure 6 shows the results of filtering out all edges save those corresponding to external object boundaries.

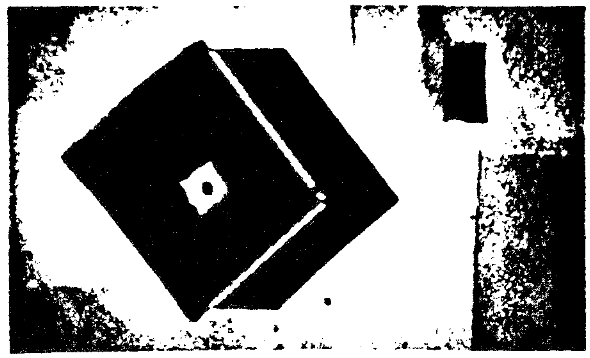

Figure 6: Detected object boundary.

Shadow stereo provides a computationally simple method for determining which dark regions of an image correspond to shadows providing information about the shape of object silhouettes. Its most important feature is the ability to easily distinguish between shadows associated with object boundaries, detached shadows and interior shadows not associated with object boundaries, and dark surface markings. The method is most appropriate for tasks involving isolated objects lying on flat (or at least smooth) supporting surfaces. This is true for many aerial reconnaissance and industrial vision problems. The method is much less useful in other situations.

References

[Hambrick and Loew, 1985] L.N. Hambrick and M.H. Loew, "Entry-exit method of shadow boundary segmentation," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, June 1985.

[Huertas and Nevatia, 1983] A. Huertas and R. Nevatia, "Detection of buildings in aerial images using shape and shadows," *Proc. 8th Int. Joint Conf. on Artificial Intelligence*, pp. 1099-951, August 1983.

[Kender and Smith, 1986] J.R. Kender and E.M. Smith, "Shape from darkness: Deriving surface information from dynamic shadows," *Proc. National Conference on Artificial Intelligence*, pp. 664-667, 1986.

[Shafer and Kanade, 1983] S.A. Shafer and T. Kanade, "Using shadows in finding surface orientations," *Computer Vision, Graphics, and Image Processing*, pp. 145-176, 1983.

[Waltz, 1975] D. Waltz, "Understanding line drawings of scenes with shadows," in *The Psychology of Computer Vision*, P.H. Winston, ed., New York: McGraw-Hill, 1975.