November 1994

Combining Color and Geometry for the Active, Visual Recognition of Shadows

Gareth Funka-Lea
University of Pennsylvania

Ruzena Bajcsy
University of Pennsylvania

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-94-62.

Combining Color and Geometry for the Active, Visual Recognition of Shadows

Abstract

In computer vision for object recognition or navigation, shadows are a frequent occurrence. However, shadows are difficult to recognize because they cannot be infallibly recognized until a scene's geometry and lighting are known. We present a number of cues which *together* strongly suggest the identification of a shadow and which can be examined without a high computational cost. The techniques developed are: a color model for shadows and a color image segmentation method that recovers single material surfaces as single image regions irregardless of whether the surface is partially in shadow; a method to recover the penumbra and umbra of shadows; and, a method for determining whether some object could be obstructing a light source. These cues either depend on or their reliability improves with the examination of some well understood shadows in a scene. Our observer is equipped with an extendable probe for casting its own shadows. These actively obtained shadows allow the observer to experimentally determine the number, location, and rough extent of the light sources in the scene. The system has been tested against a variety of indoor and outdoor environments.

Keywords

integration of cues, active vision, color, shadows

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-94-62.

Combining color and geometry for the active, visual recognition of shadows

MS-CIS-94-62
GRASP LAB 384

Gareth Funka-Lea
Ruzena Bajcsy

University of Pennsylvania
School of Engineering and Applied Science
Computer and Information Science Department
Philadelphia, PA 19104-6389

December 1994

Combining color and geometry for the active, visual recognition of shadows

Gareth Funka-Lea and Ruzena Bajcsy
GRASP Lab., Dept. of Computer and Information Science
University of Pennsylvania, Philadelphia, PA 19104
U. S. A.
lea@grip.cis.upenn.edu bajcsy@central.cis.upenn.edu

November 21, 1994

Keywords: Integration of cues, active vision, color, shadows

ABSTRACT

In computer vision for object recognition or navigation, shadows are a frequent occurrence. However, shadows are difficult to recognize because they cannot be infallibly recognized until a scene's geometry and lighting are known. We present a number of cues which *together* strongly suggest the identification of a shadow and which can be examined without a high computational cost. The techniques developed are: a color model for shadows and a color image segmentation method that recovers single material surfaces as single image regions irregardless of whether the surface is partially in shadow; a method to recover the penumbra and umbra of shadows; and, a method for determining whether some object could be obstructing a light source. These cues either depend on or their reliability improves with the examination of some well understood shadows in a scene. Our observer is equipped with an extendable probe for casting its own shadows. These actively obtained shadows allow the observer to experimentally determine the number, location, and rough extent of the light sources in the scene. The system has been tested against a variety of indoor and outdoor environments.

1 Introduction

Many existing computer vision modules assume that shadows in an image have been accounted for prior to their application. For instance, object recognition systems assume that one can recover intrinsic properties of an object despite irradiance changes such as shadows. Similarly, shape from shading algorithms assume that surface radiance does not include shadow effects. In addition, it has long been recognized that identifying shadows in an image constrains the geometric interpretation of the pictured scene [Waltz, 1975; Shafer, 1985a]. However, despite the acknowledged importance of identify shadows for image understanding, attempts to date have been overly simplistic.

We begin with an examination of the nature of shadows and then present our technique for their detection.

1.1 The Nature of Shadows

Shadows result from the obstruction of light from a source of illumination. As such, shadows have two components: one spectral and one geometric.

The spectral nature of a shadow derives from the characteristics of the light illuminating the shadow as compared to the additional light that would illuminate the same area if there was no obstruction. Hence, shadows reveal themselves as a spectral change in radiance due to a change in the local irradiance.

The geometry of a shadow is determined by the nature of the illumination obstruction and the relative position of objects in a scene. The nature of the illumination obstruction is important in determining the shape of the shadow. The other component of the shape of a shadow is the geometry of the surface or surfaces on which a shadow is cast. The relative position of objects in a scene determines where a shadow is cast.

Shadows can be classified as either *cast shadows* or *self shadows* (also called *attached shadows*). For a cast shadow, there is always free space between a point on the shadow and the point at which the light from a light source is obstructed for that point. For a self shadow, there is no free space between where the shadow is and where the obstruction is. Note that an object can cast a cast shadow on itself if the object has a concavity. We are interested in detecting cast shadows.

To further examine the nature of shadows, consider that the light from a light source may be only partially obstructed. In fact, for any non-point light source, the outer portion of a shadow results from the partial obstruction of the light source. This is the *penumbra* of the shadow. The *umbra* of a shadow is the part of the shadow where the light source is completely obstructed.

This work deals with shadows at a particular scale. We assume that part of the shape of a shadow is visible. Generally, the more of a shadow's shape that is visible, the better the system will be able to recognize it as a shadow.

1.2 Shadow Cues

Recognizing shadows in a scene is, in general, a difficult problem. Shadows can only be confidently recognized once the scene geometry, materials, and a characterization of the flux of light throughout the scene is known. With knowledge of these factors and the determination that light from a source of illumination has actually been obstructed, one can conclude that a shadow is present. However, this is more knowledge than one can expect an observer to have when recognizing shadows in a scene. Although we cannot hope to distinguish shadows from material and geometric changes with certainty in an environment that conspires against us, there are a number of cues that suggest the presence of a shadow in natural and indoor scenes:

1. The intensity, hue, and saturation changes due to shadows tend to be predictable. A material in shadow appears darker than the same material not in shadow for the same viewing conditions.

2. Shadows are only possible if there is an object obstructing light from a light source.

3. The shape of a shadow is the projection of a silhouette of the object obstructing light emitted from a light source. For an extended light source (not a point light source), the projection is unlikely to be a perspective projection, hence shadow penumbra.

4. For an extended light source, shadows can be expected to have a penumbra and umbra structure.

5. Shadow boundaries tend to change direction with changes in the geometry of the surfaces on which they are cast.

6. Shadows remain stationary relative to the surfaces on which they are cast for a fixed scene geometry.

7. Shadows of objects moving relative to a stationary light source have a corresponding motion.

8. Surface markings (or texture) tend to continue across a shadow boundary and vice versa under general viewing conditions.

Some of these cues have been used in the explicit or implicit recognition of shadows. For instance, within the field of aerial image understanding, shadow detection is used as a cue for detecting objects that rise above the ground, such as buildings or clouds [Nagao *et al.*, 1979; Huertas and Nevatia, 1988; Irvin and McKeown, 1988; Liow and Pavlidis, 1990; Wang *et al.*, 1991]. Generally, all dark regions of an image, as determined by a threshold, which have an "object" adjacent to them in the direction of the light source are considered shadows. [Gershon *et al.*, 1986] looks at just a color indicator for shadows, but uses a more sophisticated model then a single threshold applied to brightness. The color model we use is similar, but of a more general format, than that of [Gershon *et al.*, 1986]. [Jiang and Ward, 1992] use a detected penumbra and umbra structure in gray-scale images as a cue for shadows along with relative position besides an "object" and the light source. However, their techniques require strong restrictions on the layout of objects within the scene. [Witkin, 1982] uses texture continuation across a change in brightness as a cue to shadows.

This body of this paper sketches out a system to recognize shadows. The details are presented elsewhere [Funka-Lea, 1994]. We make use of cues 1, 2, and 4, as numbered above. (The use of cue 6 is presented in [Funka-Lea, 1994].) This work represents the most general system for detecting shadows to date and is consequently able to work on scenes of greater complexity than have been previously addressed. The conclusion examines why we use the various shadow cues in the way that we do and looks at an alternative way of organizing a shadow recognition system.

2 Shadows and Color

We, first, examine a scene with a single, extended, strong source source of illumination D. Let $D(\lambda)$ be the amount of energy emitted at each wavelength λ by D as measured at a given material surface. For the moment, we assume that $D(\lambda)$ does not vary with location across a given surface (hence that there is no shading). In addition to the light $D(\lambda)$, the light $A(\lambda)$ that has been reflected or scattered in the environment also strikes the surface. $A(\lambda)$ is the ambient light of the scene for the surface. Ambient light in this model is defined to be independent of position or direction in the environment. The total illumination striking a surface partially in shadow is

$$\alpha(p)D(\lambda) + A(\lambda),$$

where $\alpha(p) \in [0 \ldots 1]$ indicates that the light source D will be only partially obstructed at some locations on the surface.

Locations are measured relative to image locations (pixels). It is well known that the irradiance of a camera is proportional to the radiance of an imaged surface [Horn, 1986]. We ignore the constant of proportionality and use an image location as a measure of surface radiance. The identity of the surface in question is uniquely determined by the inverse imaging transformation for non-transparent surfaces. Consequently, we can define the term $\alpha(p)$ as follows:

$$\alpha(p) = \begin{cases} 1 & \text{if the back-projection of } p \text{ is not part of the shadow.} \\ x \in (0 \ldots 1) & \text{if the back-projection of } p \text{ is in the penumbra, where} \\ & \quad x \text{ is the percentage of total light reaching the surface.} \\ 0 & \text{if the back-projection of } p \text{ is part of an umbra.} \end{cases}$$

We will not consider the case of a point light source. The justification being that point light sources are rare outside of scientific laboratories. (The sun, despite its tremendous distance, still has a resolvable extent in the sky.)

Let $S(\lambda)$ be the surface reflectance (albedo). For the moment, assume that $S(\lambda)$ is independent of position, viewing direction, and direction of illumination; and hence, that there are no specularities (high-lights) or material variations. Let $Q_j(\lambda)$ be the weighting function of the observer's camera system for the jth filter ($j \in [1, \ldots, m]$). Let $L_i(\lambda)$ for $i \in [1, \ldots, n]$ be direct source of illumination whose light is not obstructed by the shadow casting object. Then, the camera measurement from the surface directly lit and in shadow for one filter is

$$I_j(p) = \int_{\Lambda} (\alpha(p)D(\lambda) + E(\lambda))\; S(\lambda)\, Q_j(\lambda)\, d\lambda \tag{1}$$

where $E(\lambda, p) = A(\lambda) + \sum_{i=1}^{n} L_i(\lambda)$, and Λ is the range of visible light.

Let $\mathbf{D}$ be the vector whose elements are

$$\int_{\Lambda} D(\lambda)\, S(\lambda)\, Q_j(\lambda)\, d\lambda$$

and let $\mathbf{E}$ be the vector whose elements are

$$\int_{\Lambda} E(\lambda)\, S(\lambda)\, Q_j(\lambda)\, d\lambda,$$

$j \in [1, \ldots, m]$. It follows that the image of the surface directly lit and in shadow is

$$\mathbf{I}(p) = \alpha(p)\mathbf{D} + \mathbf{E}. \tag{2}$$

An important part this work involves looking at color values plotted in a color space. The axes of the space will be a set of basis functions such as (red, green, blue) or $Q_j(\lambda)$, ($j \in [1, \ldots, m]$). Histogramming color images has become a common technique in computer vision (see, for example, [Shafer, 1985b; Klinker *et al.*, 1990; Lee, 1991; Novak and Shafer, 1992]). The histogram in a color space of an image of a surface directly lit and in shadow (i.e. dropping the dependence on position p) will be

$$\mathbf{I} = \alpha\mathbf{D} + \mathbf{E}. \tag{3}$$

Equation 3 is the parametric form of a line with parameter α. The end-point of the line at $\alpha = 0$ corresponds to the shadow umbra. The end-point at $\alpha = 1$ corresponds to the surface directly lit. The open interval of the line ($0 < \alpha < 1$) corresponds to the shadow penumbra.

Our images were taken with three color filters ($m = 3$ in Equation 1): red, green, and blue. Some images were converted to the form $(S_0, S_1, S_2) = (\mathit{brightness}, \sin\lambda, \cos\lambda)$, where λ ranges over the visible wavelengths [Lee, 1991]. Equations 2 and 3 hold in both color spaces.

In general, the description of the spectral appearance of a surface under shadow will also depend on the following factors.

1. **Shading:** Whether or not the reflection variation due to shadow is confounded with that due to shading.

2. **Multiple compact light sources:** Whether there are any compact light source in addition to the obstructed light. This affects the nature of the radiance within the shadow and whether or not multiple shadows are cast.

3. **Material changes:** Whether or not a shadow is cast on a single material, multiple materials with distinct constant albedos, or on materials with smoothly varying albedos.

4. **Specular materials:** Whether or not there can be specularities (high-lights) on the same surface material where a shadow is cast. All real objects are to some degree specular.

5. **Discontiguous shadows:** Whether or not a single shadow is visible in discontiguous parts of an image due to obstruction or the shadow being cast on discontiguous surfaces in the scene.

6. **Transparency of obstructing object:** Whether or not a shadow making object transmits light.

Unfortunately, from knowledge of total radiance it is not easy to tease out knowledge of shadow or any of the other individual factors above. Consequently, we restrict the scene conditions in order to have a manageable problem. We restrict the system to scenes with a singly colored light source and piece-wise constant albedos. In addition, we will assume that specularities and inter-reflections are local phenomenon which can be recognized independently of the shadows in the scene. In addition, the system will be attempting to detect cast shadows of opaque objects whose brightness structure is visible in a contiguous region of the image.

The restrictions placed on the problem limit the spectral model of shadows to that of Equation 2. We will refer to the above conditions where Equation 2 is applicable as the *Linear Color Cluster Assumption for Penumbrae.*

We have not restricted the problem to the case of the *Gray World Assumption* [Rubin and Richards, 1988] where the ambient and direct light source must have the same hue and saturation. The spectral characteristics of the ambient light can be quite different from those of the direct light and our analysis still applies.

There are a number of common scenarios where the lighting conditions match those to which the problem is restricted. Any sunny day in an environment without strongly specular materials qualifies. Similarly, an indoor environment lit by a single type of illumination, such as overhead fluorescent or incandescent lights (but not both together). In the indoor environment, the restriction that albedos be piece-wise constant is more likely to be met because of the nature of manufactured materials. In the experimental results, we will show that under the Linear Color Cluster Assumption for Penumbrae, the scenarios addressed still pose a challenging, practical problem.

2.1 Color Image Segmentation

The goal of our color image segmentation is to segment an image into regions such that, if shadows are present, a uniformly colored surface material directly lit and in shadow will be recovered as a single region under the Linear Color Cluster Assumption for Penumbrae.

Our color image segmentation produces a labeling of an image into multiple color-model regions and areas where no appropriate model was found to apply. In order to find regions of an image that could represent the same surface material lit and in shadow, we present a scheme in which an image is segmented into regions of line-like or uniform color clusters. The regions modeled by line-line color clusters will be shadow candidate regions.

Under the Linear Color Cluster Assumption for Penumbrae, line-like color clusters can still originate from physical phenomenon other than shadows. Shading, inter-reflections, highlights, or material changes may also produce line-like color clusters [Shafer, 1985b; Lee, 1991]. Multiple cues are necessary before a shadow can be recognized with any confidence – color alone is not sufficient.

Three ideas underlie the color image segmentation strategy. First, the realization that segmentation should be the search for the best description of an image in terms of model primitives [Leonardis *et al.*, 1990]. Second, the use of line-like color clusters as a model primitive in order to recover shadow candidate regions. Third, to dove-tail the image segmentation and model recovery between color-space and image-space in order to take into account aspects of both.

The algorithm is organized into iterations of finding seed regions based on color-space or image-space and then growing each seed to find regions of the image that are well described by a model. The models are either uniform linear functions in 3D color space. The region growing is accomplished through a region growing and pruning algorithm based on [Leonardis *et al.*, 1990]. The region growing and pruning algorithm in outline is:

1. WHILE change DO LOOP
2. Grow each region based on extrapolating the region model and the use of a tolerance criteria.
3. Update each region model to fit the new and old data.
4. Consider a higher order model.
5. Prune away entire regions based on the overlap, size, model goodness of fit, and model order.
6. END LOOP

Region growing is initiated by finding image regions with a strong color signal. These are the seed regions for the region growing. Seed regions are found based either on compact, strong peaks found in a 2D color histogram of the image, or based on uniformly colored image areas within a grid of 7x7 pixel squares which is over laid on the image. The specific 2D color spaces used for histogramming were $(\frac{\sin\lambda}{intensity}, \frac{\cos\lambda}{intensity})$ where λ ranges over the visible wavelengths [Lee, 1991] and the 2D color space $(\frac{green}{red}, \frac{blue}{red})$. We believe that any 2D color space that tends to de-emphasize intensity is a suitable choice for the initial histogramming. The process first attempts to find seed regions through color histogramming because these regions are often large. The algorithm switches to the grid method when compact strong peaks with low color variance can no longer be found in the histogram.

The details of our color segmentation algorithm have already been described in [Funka-Lea and Bajcsy, 1993]. Sample results will be presented in Section 4.

2.2 Brightness Structure of Shadows

Image regions produced by the color image segmentation that have a clearly linear (versus point) color model, may potentially contain shadows as well as a directly lit surface.

Therefore, all regions with a linear model become shadow candidate regions. Within each shadow candidate region the system examines the spatial layout of the brightness variation in order to determine if the brightness structure is consistent with that of a shadow.

In a scene with an extended light source, shadows will have a two part structure consisting of a penumbra and an umbra. The simplest brightness structure for shadows results from shadows cast by convex objects illuminated by a convex light source. In such cases, shadows cast onto contiguous locations in space will have a penumbra darkening towards an umbra, curve, or central point. The case of a point or curve apply when there is no umbra, only a partial obstruction of the light source.

Shadows cast by objects with holes onto non-contiguous surfaces can have a very complex brightness pattern. However, in our work we have found it useful to consider each umbra or local minimum in the brightness and its surrounding neighborhood where the brightness is increasing to be an individual shadow. Consequently, a single concave object may cast multiple but contiguous shadows on a single planar surface.

2.2.1 Recovering Umbrae

From our color model, we know that if an image region with a line-like color cluster model contains the umbra of a shadow then the umbra will be the darkest part of the region. In addition, the umbra will be uniformly dark. Consequently, our system begins its analysis of the brightness structure within a segmented image region by determining if any uniformly dark regions exist. Because of noise in the image, one cannot expect a shadow umbra to have the absolute minimum brightness value within a segmented image region. Instead uniformly dark regions are found using a modification of the algorithm used for color image segmentation. In this case, the models are restricted to being uniform (zeroth order). In addition, the method of finding seed regions is different.

The identification of uniformly dark seed regions within a color segmented region is initiated with a brightness threshold. We have two strategies for setting this threshold. The first strategy relies on the empirically based heuristic that the histogram of brightness for a region containing shadow often has a bimodal distribution where the darker mode originates with the shadow and the brighter mode originates with the directly lit part of the surface. (See Figure 6.) In this case, the umbra brightness threshold is set to the darker mode. Our criteria for determining if a distribution is bimodal is listed later. The second, and fall back strategy for determining a brightness threshold, uses a cut off at the lower 10% of brightness within the segmented region.

Once the brightness threshold has been determined, seed regions are then found, if possible, within the image region values below the threshold by the grid method discussed above. The color segmentation algorithm is then run on the entire shadow candidate region in order to recover uniformly dark regions from these uniformly dark seed regions. Because the system runs a region growing algorithm, regions can grow into values that are above the brightness threshold used to initiate the process if the image values are still within the tolerance for growing a uniform surface patch. Consequently, the algorithm is much more flexible than those proposals to find shadows based rigidly on a threshold (such as [Nagao *et al.*, 1979; Huertas and Nevatia, 1988; Irvin and McKeown, 1988; Liow and Pavlidis, 1990; Scanlan *et al.*, 1990; Wang *et al.*, 1991]). The brightness threshold in our approach serves as a starting point for the recovery of shadows rather than as a judgment criteria. In addition, the brightness threshold our system uses is derived from image values under a reflectance based shadow model (the linear color clusters).

2.2.2 Recovering Penumbrae

Once umbrae have been found, their corresponding penumbrae must be recovered and those shadows that consist only of penumbra must be found. Finding a penumbra is the the problem of finding an area of locally concave brightness values within the segmented image regions.

We again use image region growing as our principle tool. For each umbra, the system iteratively grow outward from the umbra as long as the brightness values are increasing. There are two types of criteria for terminating the penumbra growing: local and regional.

- Local criteria:
  - The local brightness does not increase (less than 1% brighter).
  - The local brightness is above the penumbra brightness threshold (see below for how this is determined).
- Regional criteria:
  - The average increase in brightness of the region is small (less than 10% of the previous iteration).

– The increase in size of the region is small (less than 5% of the second iteration).

Penumbra growing is terminated by the the regional brightness criteria. The local criteria determine how many, if any, new pixels are added to the penumbra at each iteration of growing. Note that the local brightness increase criteria is very weak so as to account for noise in the brightness.

The penumbra brightness threshold is set based on similar criteria to those used for the umbra seed threshold. If the histogram of brightness has a bimodal distribution, take the lower end of the brighter distribution (i.e. the minimum brightness value found in the brighter peak). If the histogram of brightness does not have a clear brightness distribution, use the upper 10% cut-off as a threshold. The penumbra brightness threshold is intended to handle those cases where there is gradually increasing shading across an entire surface that coincides with the brightening due to a shadow penumbra.

Penumbra without umbra are found in basically the same manner as shadows with both penumbra and umbra. The system finds dark patches within a segmented region and then try to grow a penumbra outward from them.

3 Shadows and Geometry

In order to insure that an image region corresponds to a shadow in a scene, an observer must determine that a light source is being obstructed in a manner consistent with a shadow visible at that position in an image. In order to determine this, the location of any light sources within the scene must be determined. A light source is defined to be anything that can cause shadows to be cast; consequently, we use detected cast shadows to locate light sources.

3.1 Locating a Light Source

It has long been recognized that a point to point correspondence in an image between a shadow and the terminator along a shadow making object indicates the direction of the obstructed light source as projected onto the image plane. [Shafer, 1985a] examined the ability to locate a light source in the case where a polyhedra casts a known shadow in a single image. The results are summarized in Table 1. The conclusion is that a point light source cannot be fully located from a known shadow and shadow casting object from

	The Point Light Source		
	Infinitely far	Finite distance	Distinguish finite from infinite
Orthography	2 parameters - (slant, tilt); 1 recoverable	3 parameters - (X, Y, Z); 2 recoverable	Converging versus parallel illumination rays
Perspective	2 parameters - (slant, tilt); 2 recoverable	3 parameters - (X, Y, Z); 2 recoverable	No method

Table 1: Shafer's four cases for locating a point light source based on the correspondence between the image of a shadow and the image of the object casting the shadow. Two cases are examined: whether or not the point light is infinitely far away from the observer. The rightmost column indicates if the observer can determine which case holds. The two middle columns indicate the number of parameters needed to define the location of the light source and how many can be recovered from the image. [Shafer, 1985a]

a single view. However, if the observer moves, this, in general, provides an independent constraint sufficient to locate a point light source.

We, however, are interested in locating an extended light source under measurement uncertainty in uncontrolled environments. The solution is to bound the location of the light source using the shadow of a known shading casting object.

In order to have a shadow whose validity is assured for locating a light source, we allow the observer to cast its own shadows with a probe that can be extended into the environment. Any visible shadows cast by the probe can be easily detected because they are new to the environment. We assume the environment is otherwise static for the time it takes to extend the probe.

The penumbra and umbra of a shadow provide different information about the location of the light source. This can be seen for the case of shadows in a 2D world in Figure 1. Pictured are *illumination rays* from the outer limit of the umbra or penumbra, tangential to the object casting the shadow and tangential to the light source. If the extent of the shadow and the location of the object casting the shadow was known, a tight bound on the location of the light source could be recovered from these illumination rays.

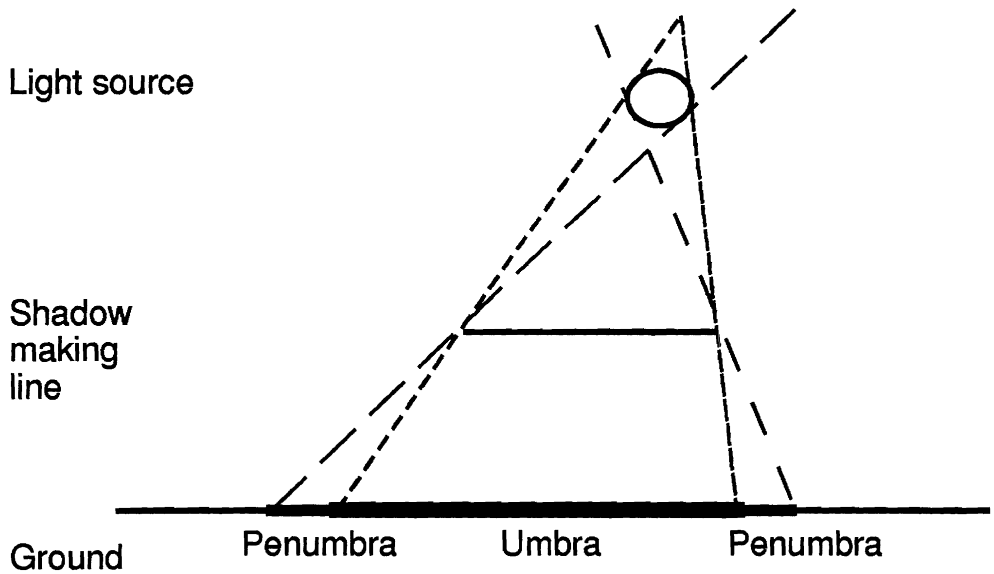

Figure 1: Umbra and penumbra of a 2D shadow. The lines with long dashes are the penumbra illumination rays and the lines with short dashes are the umbra illumination rays.

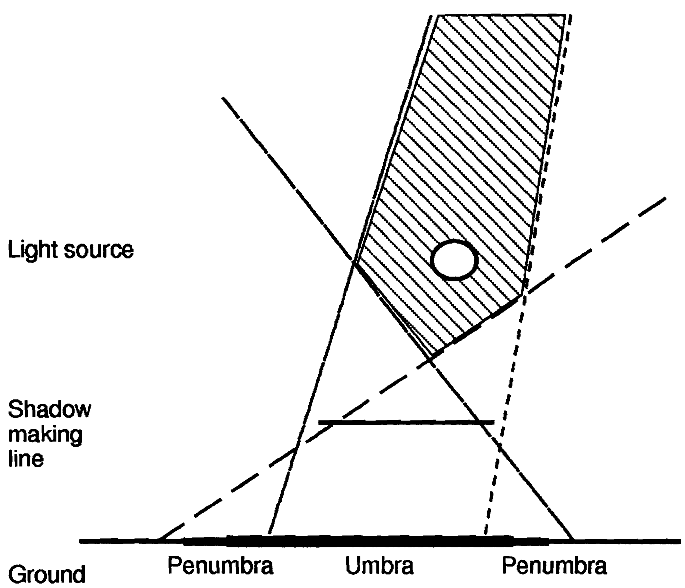

Figure 2: Adjusting the location of the shadow boundaries to take into account data errors while still maintaining a bound on the location of the light source. The light source is necessarily located within the shaded region. The short dashes are the *umbra bounds* and the long dashes are the *penumbra bounds* on the light source location.

However, the outer limits of the umbra and penumbra of a shadow are not practically recoverable because they represent an infinitesimal change in the illumination conditions and a camera has limited spatial and intensity resolution. Figure 2 demonstrates that by under-estimating the size of a shadow's umbra while over-estimating the size of the shadow casting object one is still guaranteed of having the location of the light source bounded by the umbra illumination rays. Similarly, over-estimating the size of a shadow's penumbra while under-estimating the size of the object also still guarantees that the penumbra illumination rays bound the location of the light source.

For the image of a shadow in the 3D world, it is important to under-estimate the size of the umbra and over-estimate the size of the penumbra perpendicular to the direction of the light source projected onto the image plane. This is the direction from the image of the shadow towards the image of the probe. If the probe is not visible, its projection onto the image plane can be calculated since the observing agent knows where the shadow probe is. These bounds on the location of the light source will be referred to as the *image bounds* on the location of the light source.

We can also find bounds on the location of the light source in depth relative to the image plane based on the known location of the probe. We refer to these as the *depth bounds* on the location of the light source.

The following definition of the depth bounds on the location of a light source refers to Figure 3. Figure 3 shows a 2D slice through the focal plane of the camera, the center of mass of the light source, and the center of mass of the probe. Let L be a line through the center of mass of the image of the probe's shadow and the center of mass of the projection of the probe onto the image plane. The intersection of each of the planes $\mathbf{F_u}, \mathbf{N_u}, \mathbf{F_p}$, and $\mathbf{N_p}$ with the image plane will be perpendicular to L. Planes $\mathbf{F_u}$ and $\mathbf{N_u}$ will go through a the point on the probe closest to the probe's shadow. Planes $\mathbf{F_p}$ and $\mathbf{N_p}$ will go through a the point on the probe farthest from the probe's shadow.

- $\mathbf{F_u}$ will be defined to go through a point u_{max} on line L that lies on the boundary of the shadow umbra at the maximum distance from the probe.
- $\mathbf{F_p}$ will be defined to go through a point s_{max} on line L that lies on the boundary of the shadow penumbra at the maximum distance from the probe.
- $\mathbf{N_u}$ will be parallel to the line through the focal point of the camera and point u_{max}.

Figure 3: The depth bounds on the location of the light source in 2D based on both the shadow umbra and penumbra. The light source must lie in the shaded region defined by the lines $\mathbf{F_u}, \mathbf{N_p}$, and $\mathbf{F_p}$. See the text for an explanation.

- $\mathbf{N_p}$ will be parallel to the line through the focal point of the camera and point s_{max}.

As can be seen from Figure 3, $\mathbf{F_u}$ always provides a weaker bounds on the location of a light source than $\mathbf{F_p}$; so, $\mathbf{F_u}$ is not used. Together $\mathbf{F_p}, \mathbf{N_u}$, and $\mathbf{N_p}$ make up the depth bounds on the location of the light source.

The intersection of the volumes enclosed by the image bounds and the depth bounds define the area where the light source is located. As the observer moves and does more experiments with the shadow probe, new bounds on the location of the light source can be found and these can be intersected with the previous bounds. In order to combine the results from multiple experiments it is important that the observer know its relative motion. Errors in the estimate of the relative motion are compensated for by further loosening the bounds on the location of the light source for each experiment before combining the results from multiple experiments.

3.2 Is the Light Source Obstructed?

Once the light source has been localized within a scene, the observer can examine the area between a shadow candidate and the light source to determine if a there is an object that could be obstructing the light and, hence, casting the shadow in question. However, the observer does not know the actual 3D location of its shadow candidates, nor do we intend for the system to determine their full 3D location. Instead, we examine what can be achieved by observing a single view or multiple individual views of the shadow candidates.

Consider a shadow candidate region R. The projection onto the image plane of the shadow making object would lie between R and the bounds of the light source projected onto the image plane. The bounds on the image plane between R and the projection of the light source bounds onto the image plane we call the *shadow casting bounds.* (We assume perspective projection under a pin-hole camera model.) The system calculates the actual shadow casting bounds as the convex hull of the projection of the light source bounds onto the image plane and R.

If all the image plane area in the shadow casting bounds can be discounted as being an obstruction to the light, then the observer knows that R cannot be a shadow. If some area A within the shadow casting bounds cannot be discounted, then that area may be either casting shadow R or that area may be obstructing the observer's view of the actual object casting shadow R. The observer can further test the consistency of the hypothesis that A

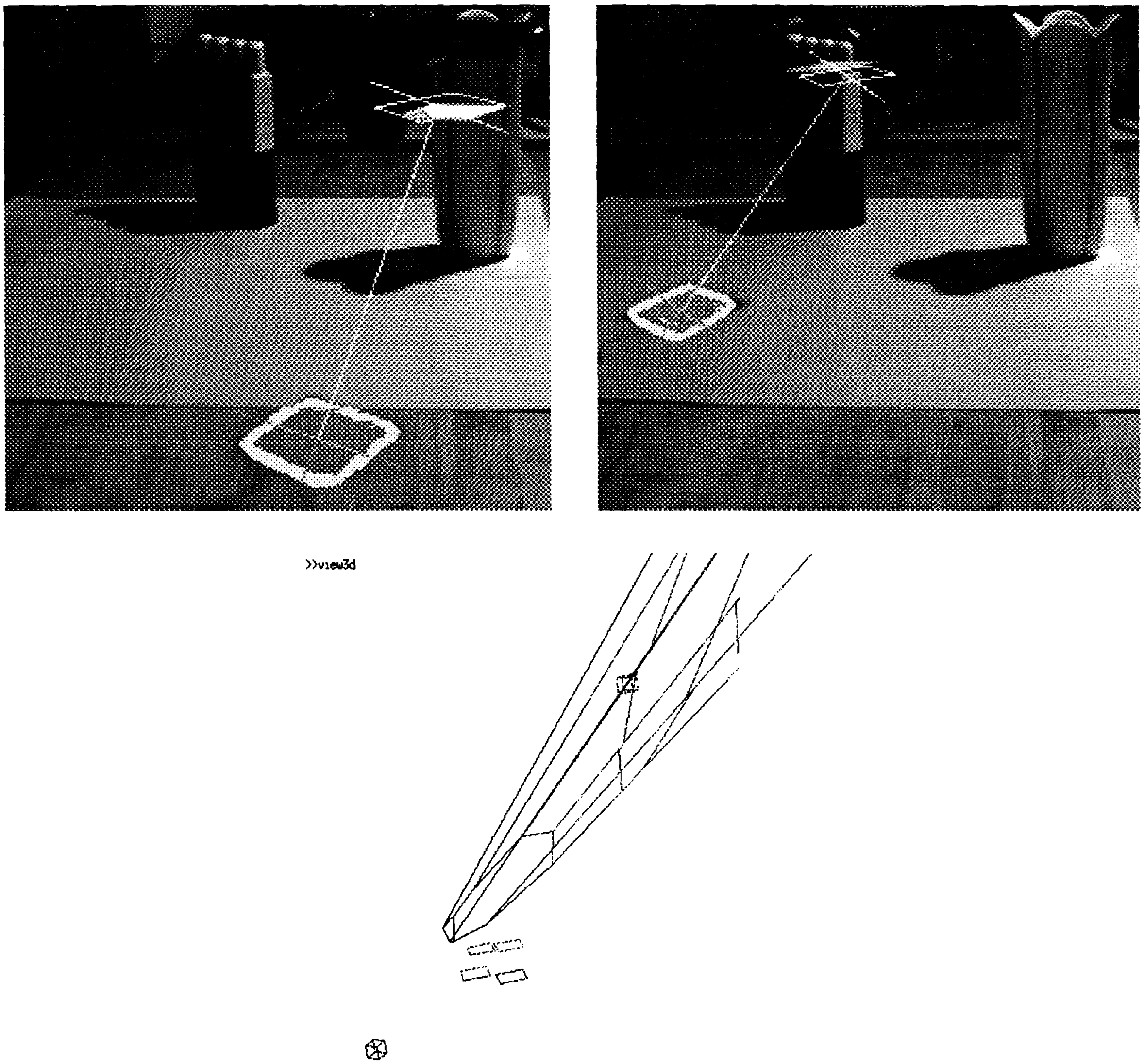

Figure 4: Locating a light source based on actively cast shadows of a probe. The probe consists of a flat plate held in a gripper of a robot arm. Top: two of the four positions in which the probe was placed. The expected position of the probe is projected onto the image plane and superimposed on the image. The location of the umbra of the probe's shadow is indicated by dark gray and the location of the penumbra by light gray. Also shown are the line L and the over-estimated extent of the penumbra and probe. Bottom: Line drawing representation of the recovered geometry. The actual light source location is indicated by the small cube inside the large wire mesh. The large wire mesh is the calculated bounds on the light source location. The four squares beneath the light source indicate the four positions the probe was placed in. The cube at the lower left indicates the viewer position.

is casting shadow R or obscuring the view of the object casting the shadow by examining the penumbra bounds on A. If part of R does not lie within the penumbra bounds of A, then that part of R could not be a shadow of A. The penumbra bounds of a A is the area between the penumbra illumination rays on the far side of A from the light source calculated assuming that the light source is as big as the entire light source bounds. Figure 9 demonstrates the use of this technique.

If the light source is not visible in the same image as R, we assume that the observer can pan in the known direction of the light source in order to acquire a sequence of images in which the area between R and the light source is visible. We leave it to a different module to decide if this effort is warranted in relation to the other constraints and goals of the observer. If multiple light source exist in a scene, then the observer must examine the shadow casting bounds between R and each of the light sources.

The primary difficulty of this test of shadow consistency is to decide that parts of the scene are not casting a particular shadow. For a single view, a significant part of the scene may be visible between a shadow candidate region R and the light source. However, ours is a mobile observer, and as the it moves, it can observe the area between R and the light source from multiple views. Many objects that appear visible between R and the light source for a particular view will not lie between the two for a different view, assuming a general trajectory for the observer. This idea is based on parallax and hence it works best if R and the light source can be viewed from very different angles. In addition, there is the problem of tracking objects or finding the correspondence between segmented image regions across multiple image frames. We do not address this problem here, but instead assume it can be done.

For a single view, additional knowledge is necessary if the observer is to dismiss parts of the visible scene from being a shadow casting object. Certain parts of the world can often be discounted from casting any shadows, for instance – a level ground plane or the sky.

4 Results

The first image segmentation and shadow recovery result is presented in Figure 5. The original image contains surfaces with a variety of albedos and surface characteristics. There are four wooden blocks and one plastic block. The plastic block is green and has four pegs on top of it and is at the center top of the image. The wooden blocks at the top left and top

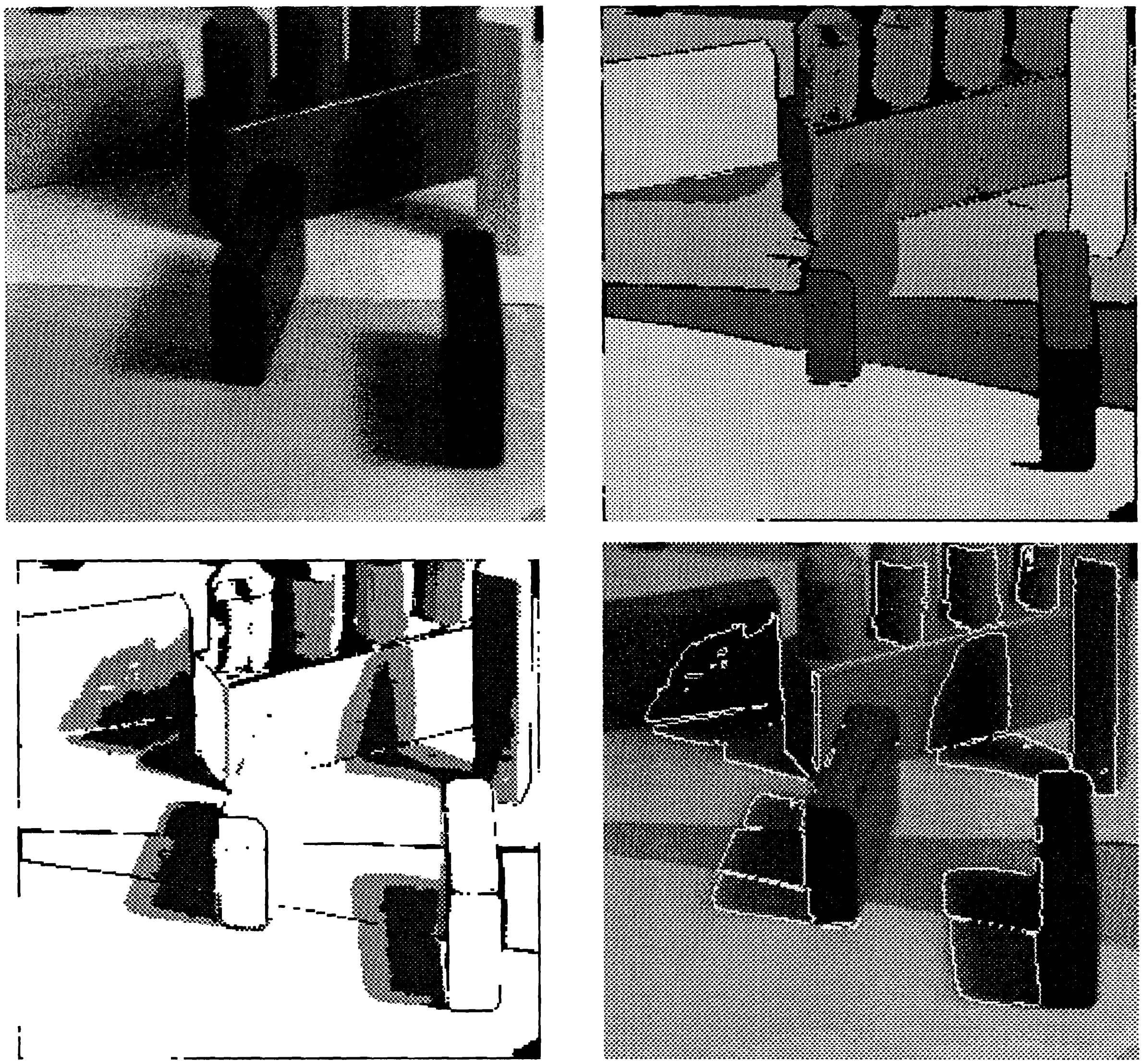

Figure 5: Top left: a color image of a block world. Top right: the color image segmentation results. Different shades of gray indicate the different regions. Black indicates that no region was recovered. Bottom left: white indicates no shadow, dark gray indicated umbra, light gray indicates penumbra. Bottom right: the outer bounds of the shadow overlaid on the original image.

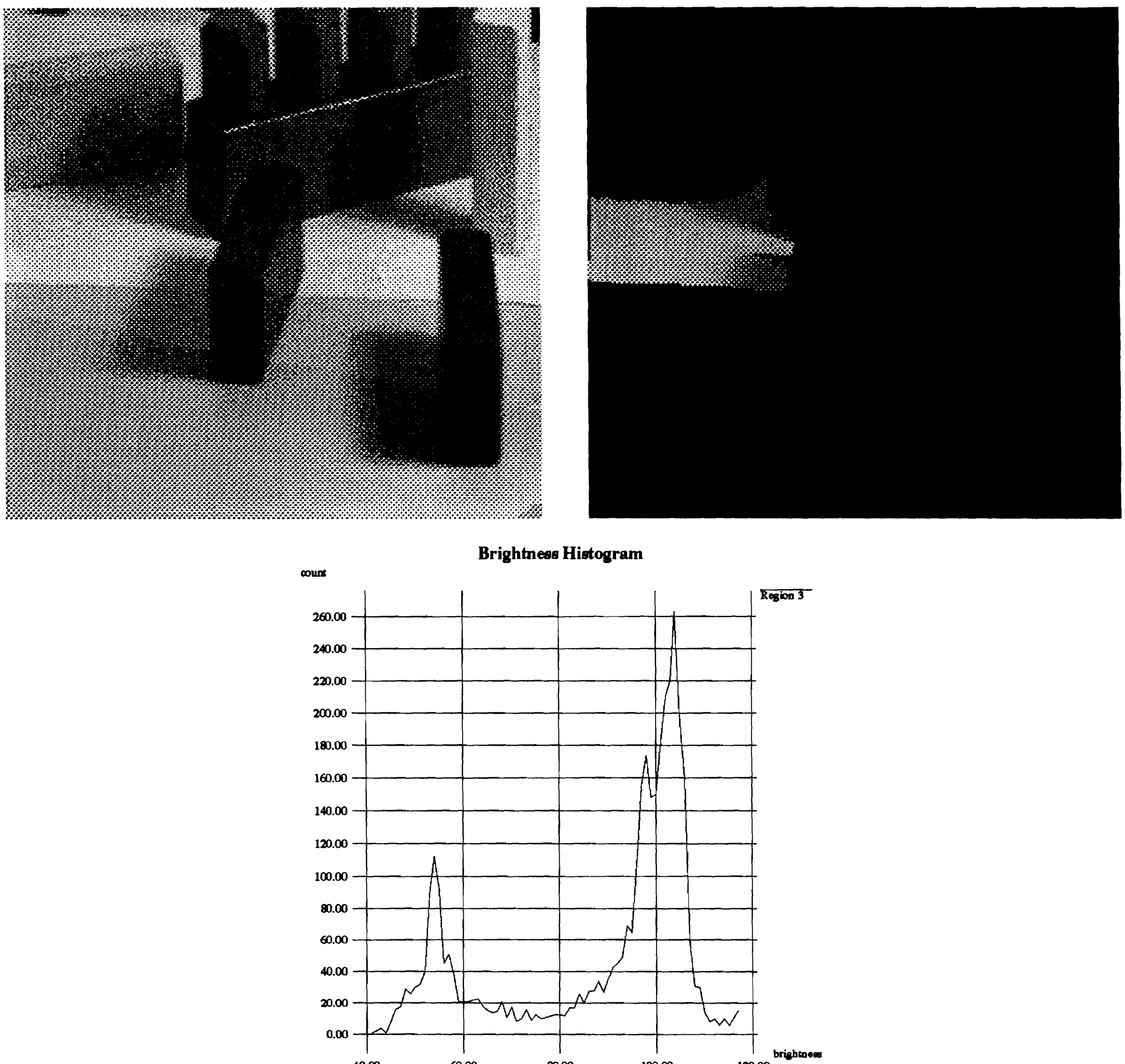

Figure 6: Brightness histogram for a region with a linear model recovered during color image segmentation. Top left: the original color image. Top right: the image data corresponding to a region recovered with a linear model. Bottom: the brightness histogram for the region.

right of the image are both red. The wooden block at the center of the image is green and the wooden block at the lower right is dark blue. The paper on the ground at the bottom of the image is white, the paper in the middle of the image is brown, and the paper at the top of the image is pink. The scene is illuminated by an incandescent bulb to the right of the field of view. Additional incandescent bulbs were used to illuminate the walls in which this scene was set in order to provide sufficient ambient light to illuminate the shadows for the CCD camera.

The final results of the segmentation shown at the top right of Figure 5 generally match the natural boundaries of material changes with three types of exceptions. First, self-shadow or shading changes that occur with a change in surface gradient are separated from directly lit surfaces of the same material on the two wooden blocks in the foreground and on the plastic block at the back. The change in illumination conditions for these cases is abrupt rather than gradual as with the penumbra of a shadow. Second, the specularities on top of the pegs on the plastic block become separate regions. Finally, the strong reflections off the front of the red block at the top left of the image produce a strong secondary illumination in the shadow on the paper in front of the block. The algorithm segments this region of increased red reflectance as a separate region. Note that each of these cases breaks the Linear Color Cluster Assumption for Penumbrae we made as a condition for the success of the segmentation algorithm. Techniques exist for successfully recognizing specularities and inter-reflections [Lee, 1991] that are complementary to our effort with shadows.

Small parts of image have no region assigned to them by the segmentation. These areas are best represented by a non-smooth model or a higher than first order model. Because the final iterations of finding seed regions begin with seed regions of a fixed size, we are unlikely to find regions smaller than the grid seed regions (7x7).

There are a number of things to note about the shadow recovery results at the bottom of Figure 5. Despite the fact that the umbra and penumbra boundaries are found independently for each segmented image region, there is strong agreement on the shadow boundaries that span multiple materials and image regions. The self-shadowing on the two blocks in the foreground and at the left end of the plastic block has not been recognized as shadow because these regions were segmented from the same surface material directly lit. However, for the case of the tall, narrow block at the far right, the self shadow is detected because the entire block was recovered as a single image region. The umbra and penumbra structure assigned to this block is somewhat counter-intuitive. One might expect the entire area to be labeled as umbra. However, there are strong reflections off the pink paper onto the block that brighten its base without changing the block's red color. This inter-reflection brightening is mistakenly interpreted as the brightening of a shadow penumbra. Also note, the shading on the pegs of the plastic block darken towards the point of self-shadowing and hence mimic the brightness structure of an umbra and penumbra. Distinguishing between such shading and a shadow can only be done based on the geometry of the scene. Overall, the cast shadows that exist in the scene are successfully located.

The original image in Figure 7 shows an unmarked road running through a stand of trees. The bottom of a tree trunk is visible at each of the top corners of the image. A green garbage can sits on the green grass to the left of the road. The asphalt of the road is cracked. Shadow are cast on the road and on the grass. Four image region are recovered from segmentation: one for the road, one for the garbage can, and one for the grass on each side of the road. Although the garbage can and the grass are both green they are recovered as separate regions. However, the tree trunks, which are very dark, are merged with the grass into single regions. The growing tolerance set by the segmentation algorithm when recovering the road region was very tight and consequently small variations on the road or the cracks in the road are not recovered as part of the road. In the final panel of Figure 7 the successful recovery of the shadows on the road is shown. The darkest gray indicated umbra. The medium gray indicates the seed used for growing shadows without umbra and light gray indicates penumbra. Because the shadow of the leaves of a tree has such a complex brightness pattern, a large number of shadows are recovered. This is a consequence of how we define what constitutes an individual shadow.

Figure 8 shows an image of a book, a ball, and blocks made of three different materials arranged on a wooden table. The blocks from left to right are made of Styrofoam, wood, and plastic. The ball is made of plastic. All the objects, except the table, are blue. Beyond the end of table can be seen a blue curtain. The objects are lit from above and in the distance so that we see the self shadowed side of the ball and blocks. A CCD camera mounted on robot arm was used to take this image. This image is one in a series.

The scene was arranged in an attempt to defeat the color image segmentation algorithm. It was assumed prior to the actual processing that the algorithm would have difficulty distinguishing between the different adjacent blue objects. As can be seen in Figure 8, the book cover, foam block and plastic ball have all been recovered as a single image region. In addition, the plastic block at right and part of the curtain are also represented by a single

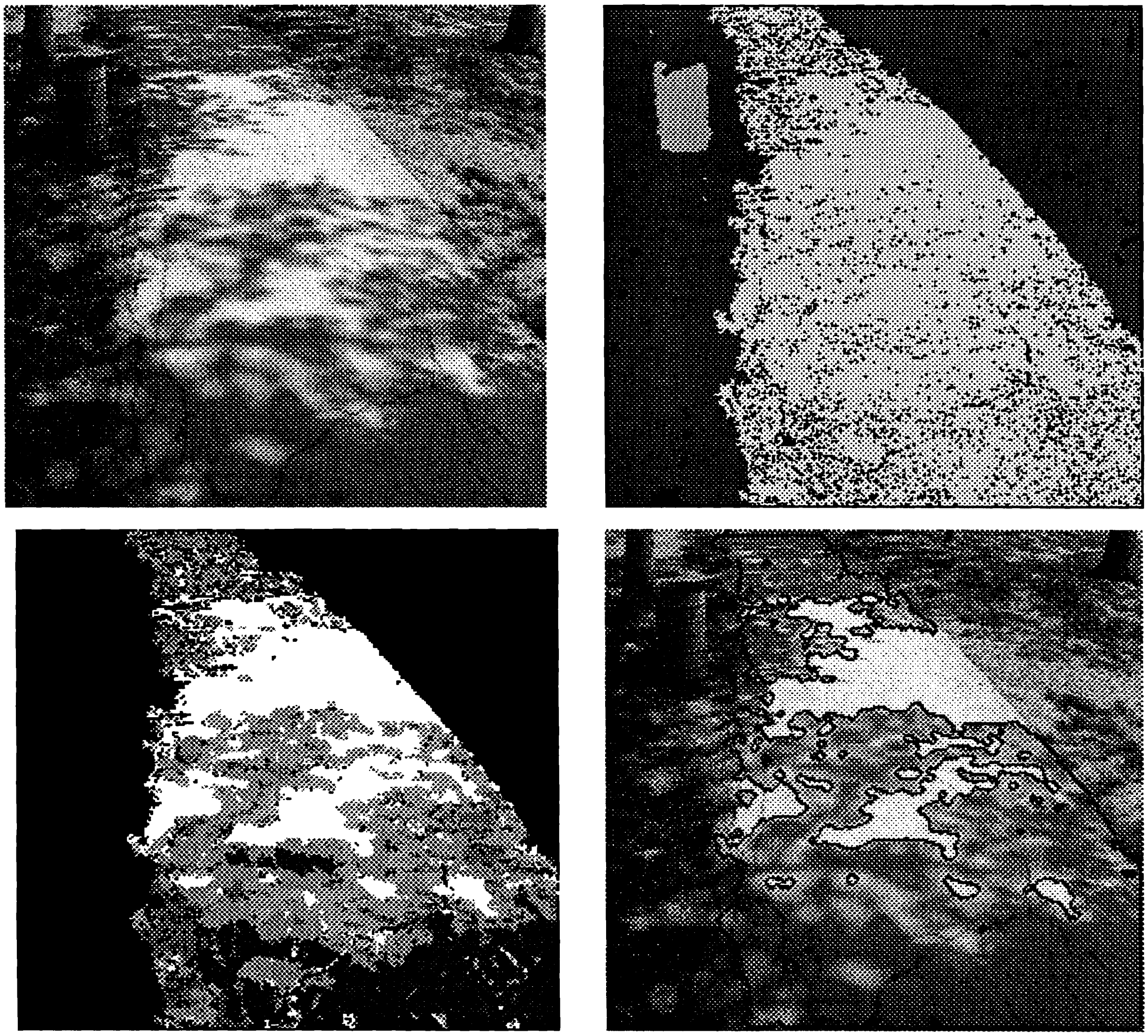

Figure 7: Top left: a color image taken from the CMU Navlab. Top right: the color image segmentation results. Bottom left: white indicates no shadow, dark gray indicated umbra, light gray indicates penumbra. Bottom right: the outer bounds of the shadow overlaid on the original image.

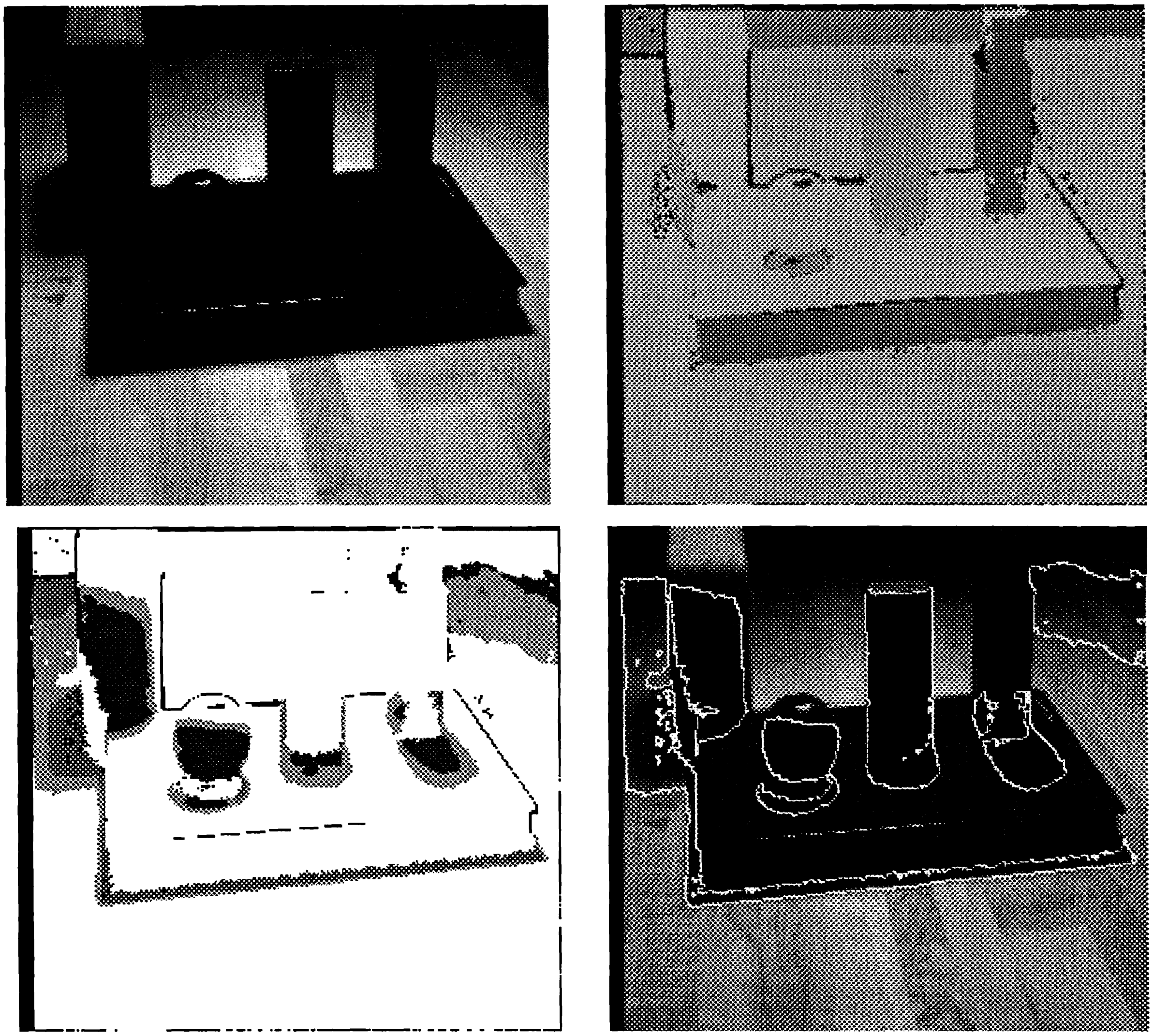

Figure 8: Top left: a color image of a block world. All the objects are dark blue. Top right: the color image segmentation results. Bottom left: white indicates no shadow, dark gray indicated umbra, light gray indicates penumbra. Bottom right: the outer bounds of the shadow overlaid on the original image.

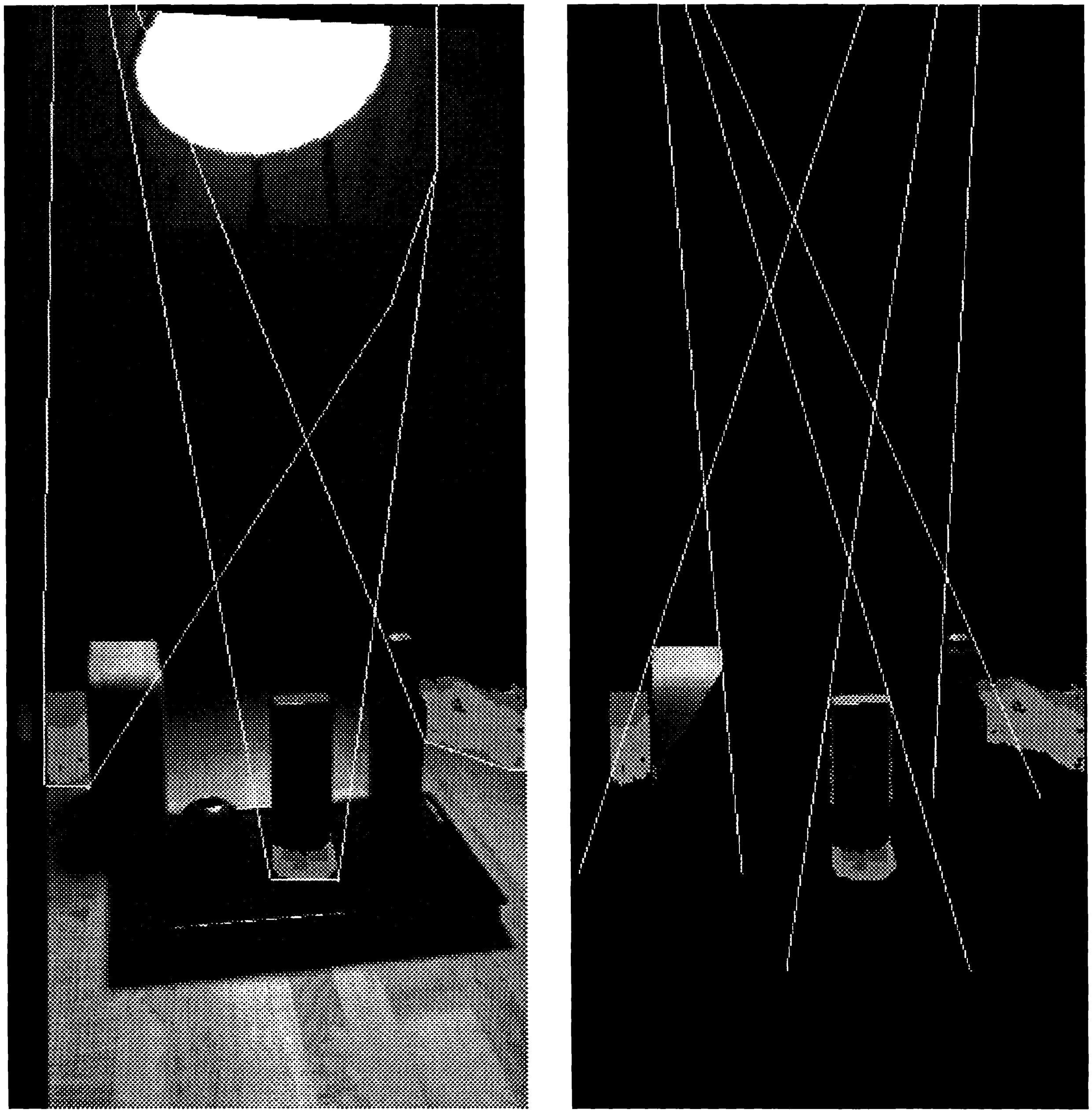

Figure 9: Distinguishing shadows from shading. Left: a mosaic of images panning from the table top in the direction of the light source. The light source is an incandescent bulb in a parabolic reflector. The white lines indicate the convex hull of light source bounds as projected onto the image plane and the shadow candidate regions. Right: the penumbra bounds for the light source and the possible shadow making objects. The parts of these shadow candidates outside of the penumbra bounds could not be a shadow because there is no object in the scene that could cast them.

image region. However, the table top, the wooden block, and two of the three parts of the curtain are successfully recovered.

The cast shadows that actually exist in the scene are successfully recognized. However, shading on the table top has also been mistaken for shadow. The light source in this scene is strongly directional meaning that parts of the scene away from the direction of the light source receive little illumination. The light is directed just behind the objects in the center of the scene. Away from this point there is progressively less illumination.

Figure 9 shows how, once the observer has found bounds on the location of the light source, shading and shadows can be distinguished. We assume that the table top and the curtain at the back of the scene are known not to cast any shadows on the table top. The shadow casting bounds around the light source and three shadow candidates are shown. The shadow candidate at center is the real shadow of a block. The block is clearly within the shadow casting bounds. The other two shadow candidates for which bounds are shown are not shadows, but rather artifacts of the directional lighting. For each of these shadow candidates only a small part of a block is within the shadow casting bounds. This implies that either that small part of the block is casting the shadow in question or there is an object behind that part of the block casting the shadow. We present the penumbra bounds for each of the blocks within the shadow casting bounds in right half of Figure 9. To a large extant, the shading that was mistaken for shadow has been recognized as not being shadow.

Additional results (indoors and out) and a quantitative evaluation of the indoor block-world examples compared with a manual segmentation is provided in [Funka-Lea, 1994].

5 Discussion and Conclusion

The primary goal and contribution of this work has been in the use of both color and geometry to recognize shadows. The achievement of the system has been to successfully recognize shadows in scenes of greater complexity than has previously been examined.

For our particular implementation of a system to recognize shadows, the image based sensory data was separated and analyzed in the following order:

1. color and relative image position (color image segmentation),
2. brightness and relative image position (image brightness pattern),

3. 3D scene geometry with the probe (shadow geometry).

Alternative orderings and arrangements of cues are possible. We will first look at the motivation behind the order we used and then discuss when other orderings would be appropriate.

Our system is primarily a bottom up strategy. Work on the system began by asking "What can a naive observer in an unknown environment achieve in recognizing shadows?" This lead to the examination of color images first. Grayscale images are inadequate for a bottom up recognition of shadows because in a grayscale image the difference between a single surface lit and in shadow is the same type of difference as between any two different materials - a change in brightness. However, for color images under the constraints of the Linear Color Cluster Assumption for Penumbrae, shadows produce a line-like color cluster in the color histogram rather than an arbitrary distribution. Once an image is segmented based on color, the brightness pattern within the segmented regions can be examined for local concavities, which is the brightness structure of shadows. Next, we looked at what additional information would be useful and how higher level information can be practically obtained. We enabled our observer to cast shadows to locate the light source rather than rely on a priori knowledge or high-level scene assumptions such as [Huertas and Nevatia, 1988; Jiang and Ward, 1992]. Because we recognize the prevalence of extended light sources and the difficulty of determining exact scene locations based on image positions, we bound the location of the light source rather than attempting to determine the light source's exact location (as is presented in [Shafer, 1985a]).

In scenes with complex illumination – a number of equally strong light sources with very different spectral characteristics – it will be necessary to have a system that could begin the recognition of shadows from the scene geometry. Consider the case of an observer who has identified the locations of a light source and an object in the foreground. The observer could then check to see if there was a shadow cast by this object. The area of the scene to be examined could be bounded in a way similar to how our system bounds the location of a light source. Within a small part of the scene, it might be possible to find illumination conditions that meet the Linear Color Cluster Assumption even if the assumption is not true for the whole scene. In which case, color could then be used to find any shadow. Alternative strategies for initiating shadow recognition from geometric cues are surely possible. In an ideal system, shadow recognition would be initiated by both color and geometric cues and those shadow candidates for which there is a consensus would be accepted as shadows.

Acknowledgments

We strongly thank S. W. Lee for his insight on color. We thank C. Thorpe and T. Jochem for making images from the Carnegie Mellon University Navlab available. This research was supported in part by: AFOSR Grants 88-0244, 88-0296; Army/DAAL 03-89-C-0031PRI; NSF Grants CISE/CDA 88-22719, IRI 89-06770; ARPA Grant N0014-88-K-0630; and Dupont Corp.

References

[Funka-Lea and Bajcsy, 1993] Gareth Funka-Lea and Ruzena Bajcsy. Active color image analysis for recognizing shadows. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, pages 1573–1578, Chambéry, France, August 1993.

[Funka-Lea, 1994] Gareth D. Funka-Lea. *The Visual Recognition of Shadows by an Active Observer.* PhD thesis, GRASP Laboratory, Dept. of Computer and Information Science, University of Pennsylvania, December 1994.

[Gershon *et al.*, 1986] Ron Gershon, Allan D. Jepson, and John K. Tsotsos. Ambient illumination and the determination of material changes. *Journal of the Optical Society of America A*, 3(10):1700–1707, October 1986.

[Horn, 1986] B. K. P. Horn. *Robot Vision.* MIT Press, Cambridge, Mass, 1986.

[Huertas and Nevatia, 1988] A. Huertas and R. Nevatia. Detecting buildings in aerial images. *Computer Vision, Graphics, and Image Processing*, 41(2):131–152, February 1988.

[Irvin and McKeown, 1988] R. B. Irvin and David M. McKeown. Methods for exploiting the relationship between buildings and their shadows in aerial images. *IEEE Transactions on Systems, Man and Cybernetics*, 19(6):1564–1575, November 1988.

[Jiang and Ward, 1992] Caixia Jiang and Matthew O. Ward. Shadow identification. In *Proc. Computer Vision and Pattern Recognitin*, pages 606–612, Champaign, Illinois, June 1992.

[Klinker *et al.*, 1990] Gudrun J. Klinker, Steven A. Shafer, and Takeo Kanade. A physical approach to color image understanding. *International Journal of Computer Vision*, 4(1):7–38, January 1990.

[Lee, 1991] Sang Wook Lee. *Understanding of Surface Reflections in Computer Vision by Color and Multiple Views.* PhD thesis, GRASP Laboratory, Dept. of Electrical Engineering, University of Pennsylvania, August 1991.

[Leonardis *et al.*, 1990] Aleš Leonardis, Alok Gupta, and Ruzena Bajcsy. Segmentation as the search for the best description of the image in terms of primitives. In *IEEE International Conference on Computer Vision*, pages 121–125, Osaka, Japan, December 1990.

[Liow and Pavlidis, 1990] Yuh-Tay Liow and Theo Pavlidis. Use of shadows for extracting buildings in aerial images. *Computer Vision, Graphics, and Image Processing*, 49(2):242–277, February 1990.

[Nagao *et al.*, 1979] Makota Nagao, Takashi Matsuyama, and Yoshio Ikeda. Region extraction and shape analysis in aerial images. *Computer Graphics and Image Processing*, 10(3):195–223, July 1979.

[Novak and Shafer, 1992] Carol L. Novak and Steven A. Shafer. Anatomy of a color histogram. In *Proc. Computer Vision and Pattern Recognitin*, pages 599–605, Champaign, Illinois, June 1992.

[Rubin and Richards, 1988] John M. Rubin and W. A. Richards. Color vision: Representing material categories. In Whitman Richards, editor, *Natural Computation.* MIT Press, 1988.

[Scanlan *et al.*, 1990] Joseph M. Scanlan, Douglas M. Chabries, and Richard W. Christiansen. A shadow detection and removal algorithm for 2-D images. In *IEEE Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, volume 4, pages 2057–2060, Albuquerque, NM, April 1990.

[Shafer, 1985a] Steven A. Shafer. *Shadows and Silhouettes in Computer Vision.* Kluwer Academic Publishers, Boston, Mass., 1985.

[Shafer, 1985b] Steven A. Shafer. Using color to separate reflection components. *COLOR Research and Applications*, 10(4):210–218, Winter 1985.

[Waltz, 1975] David Waltz. Understanding line drawings of scenes with shadows. In Patrick Henry Winston, editor, *The Psychology of Computer Vision.* McGraw-Hill Book Co., 1975.

[Wang *et al.*, 1991] Chengye Wang, Liuqing Huang, and Azriel Rosenfeld. Detecting clouds and cloud shadows on aerial photographs. *Pattern Recognition Letters*, 12:55–64, January 1991.

[Witkin, 1982] Andrew P. Witkin. Intensity-based edge classification. In *Proceedings of the National Conference on AI (AAAI-82)*, pages 36–41, Pittsburgh, PA, August 1982.