

Active Color Image Analysis for Recognizing 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.upenn.edu

Abstract

Many existing computer vision modules assume that shadows in an image have been accounted for prior to their application. In spite of this, relatively little work has been done on recognizing shadows or on recognizing a single surface material when directly lit and in shadow. This is in part because shadows cannot be infallible recognized until a scene's lighting and geometry are known. However, color is a strong cue to the presence of shadows. We present a general color image segmentation algorithm whose output is amenable to the recovery of shadows as determined by an analysis of the physics of shadow radiance. Then, we show how an observer that can cast its own shadows can infer enough information about a scene's illumination to refine the segmentation results to determine where the shadows in the scene are with reasonable confidence. Having an observer that can actively cast shadows frees us from restrictive assumptions about the scene illumination or the reliance on high level scene knowledge. We present results of our method on images of complex indoor and outdoor scenes.

1 Shadows and Image Understanding

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 identifying shadows for image understanding, attempts to date have been overly simplistic. In aerial image interpretation, for instance, shadows are considered to be those parts of an image whose values are below a threshold and are adjacent to a rectangular shape (a building) on the side away from the direction of illumination [Nagao *et al.*, 1979;

Huertas and Nevatia, 1988; Irvin and McKeown, 1988; Liow and Pavlidis, 1990].

In this work, we present a color image segmentation algorithm that recovers image regions amenable to shadow recognition. In addition, we sketch how this technique fits in with a larger system to enable an active observer to recognize shadows in images of a scene.

We begin with a physical model for color image formation for a single surface material directly lit and in shadow. We discuss the limitations of previous color image work for shadows and present our assumptions and techniques for addressing the problem. In particular, we use an observer that can actively cast a shadow into the scene in order to examine the illumination conditions. Finally, we present results and a discussion of our work.

2 Spectral Model of Shadows

2.1 Shadows Without Other Reflectance Effects

Let $D(\lambda, p)$ be the amount of energy emitted at each wavelength λ by a direct source of illumination D as measured at the back-projection of pixel p onto a given surface. In addition to the light $D(\lambda, p)$, the light $A'(\lambda, p)$ that has been reflected or scattered in the environment and any other direct sources of illumination $L_1(\lambda, p), \dots, L_n(\lambda, p)$ also strike the surface. Assume now that an object is brought between the light source D and the surface. The reflected light in the scene changes due to reflections off the obstructing object, call it now $A(\lambda, p)$. So, the illumination striking the surface is

$$\alpha D(\lambda, p) + A(\lambda, p) + \sum_{i=1}^n L_i(\lambda, p)$$

where $\alpha \in [0 \dots 1]$ indicates that the light source D will be only partially obstructed at some locations on the surface if D is not a point light source. The partial obstruction of an extended light source results in the penumbra of a shadow.

Let $S(\lambda, p)$ be the surface reflectance (albedo). For the moment, assume that $S(\lambda, p)$ is independent of the viewing direction and direction of illumination; and hence, that there are no specularities (high-lights) nor shading across the surface. Let $Q_j(\lambda)$ be the weighting function of the observer's camera system for the j th filter

($j \in [1, \dots, m]$). Then, the camera measurement from the surface directly lit and in shadow for one filter is

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

where $E(\lambda, p) = A(\lambda, p) + \sum_{i=1}^n L_i(\lambda, p)$, and Λ is the range in which $Q_j(\lambda)$ is non-zero.

We will use the following notation:

$$\mathbf{D}(p) = \begin{bmatrix} \int_{\Lambda} D(\lambda, p) S(\lambda, p) Q_1(\lambda) d\lambda \\ \vdots \\ \int_{\Lambda} D(\lambda, p) S(\lambda, p) Q_m(\lambda) d\lambda \end{bmatrix}$$

$$\mathbf{E}(p) = \begin{bmatrix} \int_{\Lambda} E(\lambda, p) S(\lambda, p) Q_1(\lambda) d\lambda \\ \vdots \\ \int_{\Lambda} E(\lambda, p) S(\lambda, p) Q_m(\lambda) d\lambda \end{bmatrix}$$

where $\mathbf{D}(p)$ and $\mathbf{E}(p)$ are m element vectors. From the above equations it follows that the image of the surface directly lit and in shadow is

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

Equation 2 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: red, green, and blue. Some images were converted to the form $(S_0, S_1, S_2) = (\text{brightness}, \sin \lambda, \cos \lambda)$, where λ ranges over the visible wavelengths [Lee, 1991]. This space is convenient for describing hue and color saturation. Hue is defined as $\tan^{-1}(\frac{S_2}{S_1})$. Color saturation is defined as $\frac{\sqrt{S_1^2 + S_2^2}}{S_0}$. Equation 2 holds in both color spaces.

2.2 Shading, Inter-Reflections, and Specularities

Variations in the amount of light striking a surface due to a change in geometry (*shading*) can complicate the model we have described above. For a scene with a single light source, no inter-reflections, and a uniformly colored Lambertian surface which receives varying amounts of illumination due to surface curvature or varying distance from the light source, the reflection from the surface will describe a linear cluster in color space [Shafer, 1985b; Lee, 1991]. This linear cluster will lie along the linear cluster resulting from a shadow penumbra under the same illumination conditions.

In the case where a uniformly colored Lambertian surface is illuminated by light in varying amounts from two distinctly colored lights, the reflection from the surface will describe a planar cluster in color space due to changes in the term $\mathbf{E}(p)$ in Equation 2. If the surface is illuminated by light in varying amounts from multiple, distinctly colored lights, the reflection from the surface will describe a volume in color space [Lee, 1991]. For multiple, differently colored light sources, the reflection

distortion in color space due to a shadow will be superimposed on the volumetric cluster due to shading.

The light reflected from one surface onto a second surface serves as a source of irradiance for the second surface. As such, inter-reflections complicate our shadow model in the same way that multiple light source do.

If a surface is not perfectly Lambertian, then the image of that surface may include specularities. Specularities from the obstructed light source cannot occur in the umbra of a shadow. Specularities from the obstructed light source that fall within the penumbra are visible but their shape will be truncated at the boundary of the umbra. Diffuse specularities due to a rough, microfaceted surface [Torrance and Sparrow, 1967] can, however, become dimmer in a penumbra because the specular irradiance at a point on the image can be due to microfacets in the umbra as well as in the penumbra of a shadow.

3 Assumptions for Shadow Recognition

Many researchers assume that the light illuminating a shadow and the obstructed light are spectrally proportional. Hence, that $\mathbf{E}(p) = \beta \mathbf{D}(p)$ [Rubin and Richards, 1988]. And, the color cluster for a single material directly lit and in shadow is a linear cluster that if extended would pass through the origin. Gershon et al. [1986] refer to such shadows as *ideal shadows*. They also considered the somewhat more general case of $\mathbf{E}(p) = \beta \mathbf{D}(p) + \mathbf{K}(p)$ where $\mathbf{K}(p) = [k_1, \dots, k_m]^T$ is constant. This case is referred to as that of *non-ideal shadows*.

Rather than require a strong, global constraint on the scene illumination as in [Rubin and Richards, 1988; Gershon et al., 1986], we prefer to use a weaker, local constraint during our initial color image segmentation. We want to insure that we can segment an image into regions such that if shadows are present, a uniformly colored surface directly lit and in shadow will be represented by a single region when the scene is lit by a compact light source or a cross section of the penumbra of a shadow will be represented as a single region when multiple compact light sources are present. Only those segmented regions with a linear model will be shadow candidate regions. The shadow candidate regions will be further analyzed for evidence to support or refute the hypothesis that a shadow is present. In general this will require bringing both geometric and spectral information to bear. Hence we concentrate on spectral information.

In order to meet our goal for image segmentation, we will make the following assumption:

Linear Color Cluster Assumption for Penumbrae: We assume that the light irradiating a penumbra, with the exception of the the partially obstructed light, does not vary or varies insignificantly.

Consequently, the variation in reflection in a penumbra on a uniformly colored surface is due entirely to the obstruction of a direct light source. And, in order to find regions of an image that could represent the same surface lit and in shadow across a penumbra we present a scheme in which an image is segmented into line-like or uniform color clusters.

However under our assumption, 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 [Klinker *et al.*, 1990; Lee, 1991]. We will examine later how to distinguish the case of shadows.

The analysis discussed so far was done strictly in color space and ignores image or scene locality. Because all the pixels in an image of a complicated scene taken together may result in many line-like color clusters, we introduce local image continuity as a constraint in our color image segmentation. So, we will only be looking for contiguous sets of pixels in an image that form line-like color clusters. This restricts our image interpretation to those shadows for which the same surface can be seen directly lit and adjacently in shadow. Note also that the boundary of the shadow must include a penumbra for our system to work. This means that we cannot handle scenes with point light sources. Fortunately, true point light sources are rare in non-laboratory environments. The importance of penumbra in detecting shadows has also been noted by [Jiang and Ward, 1992].

4 Color Image Segmentation

Our color image segmentation is founded on three ideas. First, to find line-like models in color space which are our shadow candidate regions. Second, to dove-tail the processing between color-space and image-space in order to take into account aspects of each. And finally, the realization that segmentation should be the search for the best description of an image in terms of primitives [Leonardis *et al.*, 1990]. Figure 1 provides an outline of the control structure of the algorithm.

The bulk of the work is accomplished through a segmentation by region growing and pruning algorithm [Leonardis *et al.*, 1990]. Which in outline is:

WHILE change DO LOOP

- Grow each region based on the region model.
- Update each region to fit the new and old data.
- Consider a higher order model.
- Prune away regions based on overlap, size, goodness of fit, and model order.

END LOOP

Our region models are uniform or linear functions in color space, while [Leonardis *et al.*, 1990] uses bivariate polynomials functions of pixel location.

Region growing is initiated by finding seed regions. Seed regions are found based either on strong color samples found in a color histogram of the image, or based on uniformly colored image areas within a grid of 7x7 pixel squares which is overlaid on the image. 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 (see [Funka-Lea, 1992] for the algorithm).

We have used a two dimensional color space for the histogramming. The 2D color spaces used were

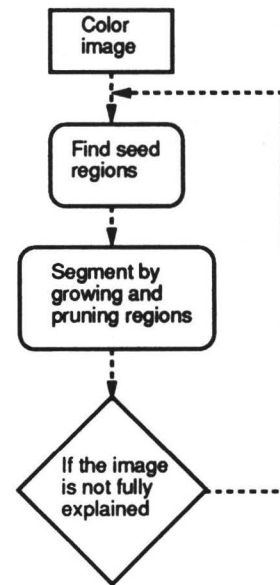


Figure 1: Color Image Segmentation Outline

$(\frac{Green}{Red}, \frac{Blue}{Red})$ and $(\frac{S_1}{S_0}, \frac{S_2}{S_0})$. We believe that any one or two dimensional color space that de-emphasizes color brightness is a suitable choice for the initial histogramming. (See Figure 2 for a sample 2D color histogram.)

The initial 3D color models for seed regions are either uniformly colored or correspond to lines in 3D color space that pass through the origin. If a uniform region does not grow during an iteration then a linear model is tried. The linear model is accepted if the error is relatively small and the region can be grown by a considerable amount. Through the process of choosing a new model during updating, linear models need no longer pass through the origin. Consequently, we can recover shadows that are not *ideal shadows*. See [Leonardis *et al.*, 1990] for the details of model updating and pruning.

The tolerance criteria for region growing is determined based on the variance of the peak in the histogram or on the fit error of a model from the grid method. The tolerance is varied uniformly with pixel intensity because greater color variation is possible for brighter pixels.

Our image segmentation decouples the region models from pixel coordinates. This insures that a single material lit by a single light source on which numerous shadows are cast is still recovered as one image region. Our post-segmentation processing will concentrate on analyzing individual regions to see if they contain shadows and will not need to compare regions that are not adjacent.

5 Employing an Active Observer

Because line-like color clusters can result from physical events besides shadow, our observer must examine the shadow candidate regions found by our color segmentation procedure to determine if a shadow is present.

The simplest shadow criteria for the linear color clusters follows from the fact that shadows are darker than the same surface directly lit. Consequently, for a color space whose bases are band-limited functions (such as

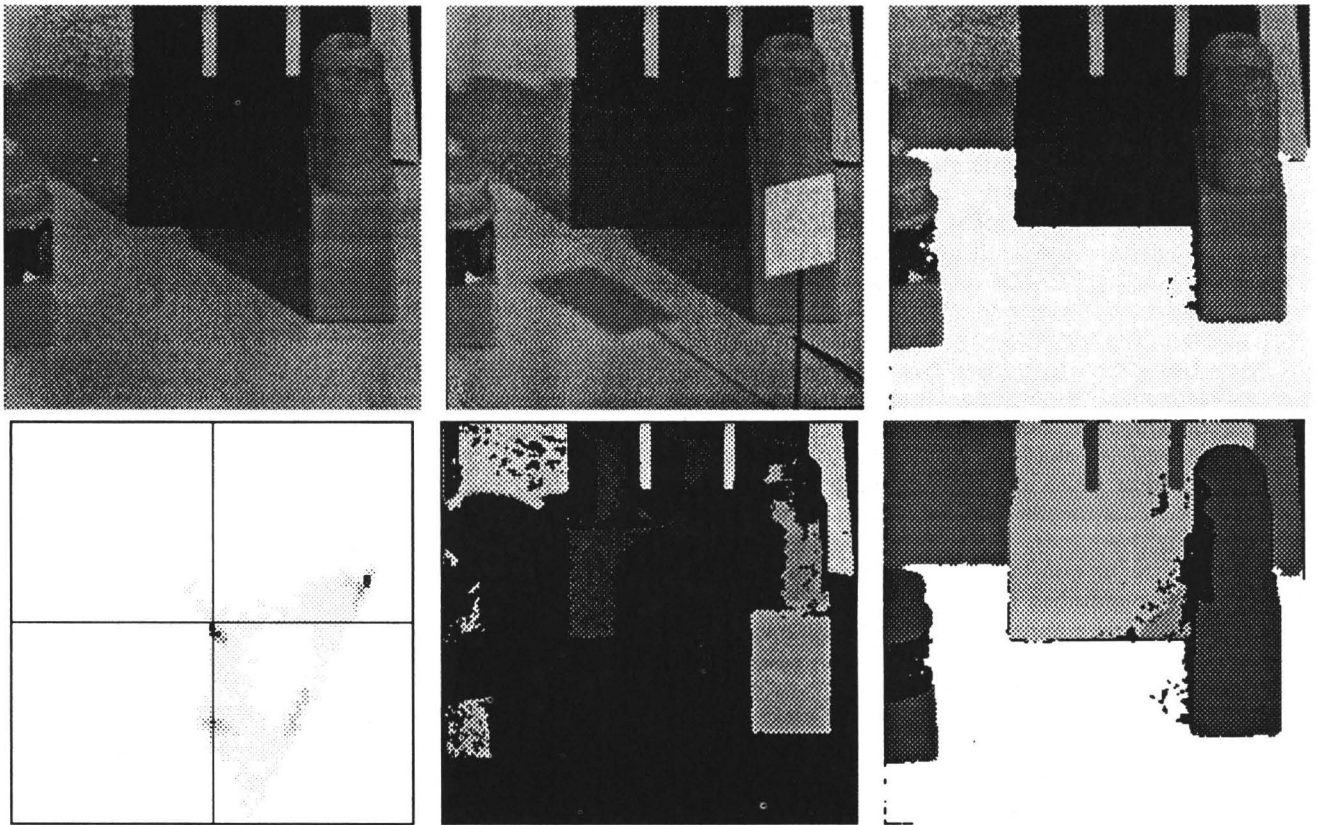


Figure 2: Color Image Segmentation: Scene 1. Top Left: The original image. Top Center: The same scene after introducing the shadow probe. Top Right: The partial segmentation based on the probe. Bottom Left: The 2D color histogram of the remaining parts of the image. Bottom Center: The seed regions from the color histogram. Bottom Right: The results of segmentation. (See the text for a complete explanation.)

red, green, blue), the linear color cluster for a shadow must not get brighter along any of the bases in one direction along the line.

If all the shadows in a scene are illuminated by similarly colored light, we expect the shadows to show a similar relative change in brightness, hue and color saturation. We judge any trends in the color of shadows based on experiments the observer does by actively casting its own shadows by extending a probe into the environment. If we assume the environment is unchanging during the extension of the shadow probe then any shadows that are cast and are visible can be easily detected by taking the difference of images. We assume the observer is mobile and that in time the observer will image a shadow that it has cast, if shadows are possible.

The first use of the probe shadow is to provide a partial segmentation of the image. We segment the image area of the probe shadow using our color segmentation algorithm. Then, we allow the results to grow into the surrounding parts of the image. Thus, we assume that the shadow we cast on a material will look like other shadows cast on the same material.

Next, we examine the situational information gained from the probe shadow. If no shadow is cast, we infer that there is no compact light source. If the shadow has no detectable penumbra, we infer that there is a point

light source and that our shadow recognition system is not applicable. If we find disjoint shadows cast by the probe, then there must be multiple compact light sources present. If only one compact light source is indicated, then shadows cast on a single material will lie within image regions. However, if there are multiple compact light sources, then a shadow cast on a single material may be found across several contiguous regions. In future work we will address the problem of recognizing that a shadow that is split over several regions is a single shadow.

Note that the observer can test the validity of the Linear Color Cluster Assumption for Penumbrae for those shadows that it casts. The penumbra of these shadows can be found independently of our segmentation algorithm (see [Funka-Lea, 1992]) and the penumbra color values can be tested against a linear color model.

We now examine the spectral characteristics of the scene illumination inferred from the probe shadow. First, the observer tests if the ratio of the brightness of the probe umbra to the same surface directly lit is constant for a variety of surfaces (i.e. *ideal shadows*). If it is, then the observer can use this ratio as a criteria for detecting shadows. If the ratio is not tightly bounded, we assume that the color trends we see in the probe shadow will hold for all the shadows in the scene. For instance, if all the probe shadows show a bias towards blue then

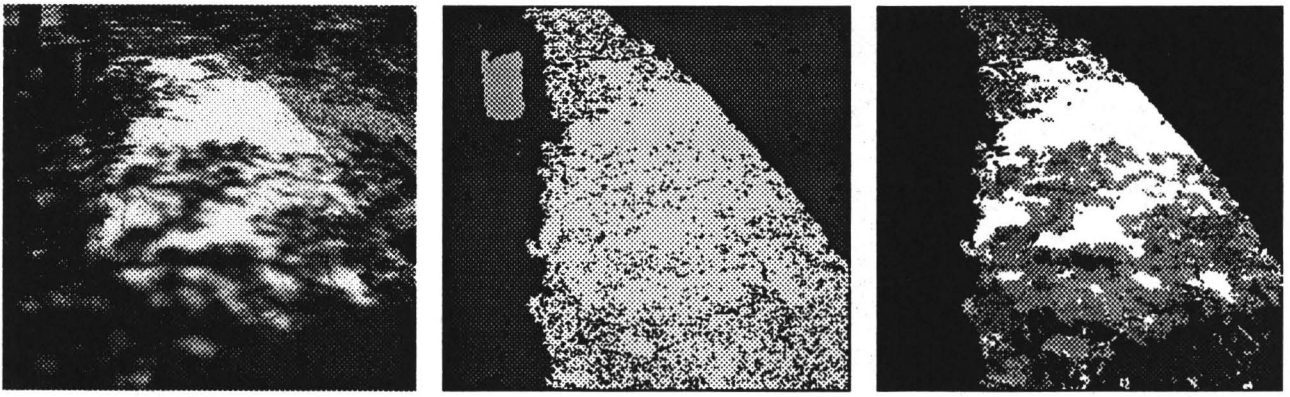


Figure 3: Color Image Segmentation: Scene 2. Left: An image of a road courtesy of the Carnegie Mellon Navlab project. Center: The full segmentation of the original image. The different regions are indicated by changes in the gray-scale. Black indicates no region was found. Right: The recovered umbra and penumbra of the shadow on the road. The umbra is dark gray, the penumbra light gray, and the road directly lit is white.

we will expect all shadows to follow this rule under the present lighting conditions. We look for trends along the criteria of hue and saturation among the probe shadows. (This is similar in spirit to the *pull factor* for *non-ideal shadows* of Gershon et al. [1986].)

6 Results

Figure 2 shows the complete processing of one image. In the image, a red block is at the right and a green block is in the center. The coordinates of the pictured 2D color histogram are $(\frac{S_1}{S_0}, \frac{S_2}{S_0})$. Unsaturated colors are near the origin and saturated colors are at the periphery. Red is to the right, green at the bottom, and blue is at the upper left corner. Strong responses can be seen for the white background (near the center) and for the red block (near the right). There is a weaker response for the green block not in shadow (at the bottom center). The green block in shadow is a line from green to red.

The top right picture in the figure shows in white the segmented region found from the probe shadow. At the bottom of the figure are shown the gray-scale coded regions from segmentation. Some image locations remain unexplained after segmentation. In areas where no seed region is fit and no region is grown, the image is better explained by texture, edges, or higher order color models. In the area where the two blocks meet, there are strong inter-reflections. This results locally in a non-line-like color cluster which our system was not designed to recover. (See [Klinker et al., 1990; Lee, 1991] for recovering inter-reflections.) The shadows in the scene are recognized as compatible with the probe results of a nearly *ideal shadow* with a slight bias to red.

In the next two figures no probe was used and only the success of the segmentation algorithm is tested.

Figure 3 demonstrates that our method can handle complex, natural shadows. This example also demonstrates the recovery of the umbra, penumbra, and directly lit parts of a surface (the road). Some small parts of this image remain unexplained due to camera noise, cracks on the road, and variations in the ground cover. Note that the partially shadowed exterior of the green

garbage can is successfully recovered against the grass. However, the shadowed tree trunks and the interior of the garbage can are not distinguished from the grass in shadow due to the lack of hue data with low brightness.

In Figure 4 the scene is illuminated by two compact light sources. This is evident in the two specularities on the plastic ball in the center of the image. On the right side of the image is an arch of colored wooden blocks which casts a shadow across the ball and the ground. Because of the illumination situation, there is shading within the shadow cast on the ball. Consequently, we cannot expect to recover the ball with one line-like color cluster model. However, if our assumptions hold, we should be able to recover at least the penumbra of the shadow as one region. In fact, the penumbra and most of the ball is recovered as a single region. This demonstrates how weak a constraint the Linear Color Cluster Assumption is on the illumination. Again, small parts of the image remain unexplained; this is mainly on account of specularities and color variations on the wood blocks.

7 Conclusion

We have presented a general algorithm for segmenting color images into regions that form uniform or linear clusters in color space. Under our Linear Color Cluster Assumption for Penumbrae, shadows are represented as piece-wise linear color clusters and the width of a penumbra will always be a single segmented region. Consequently, all linear color clusters become shadow candidate regions. Many of these regions can be discounted as shadows because they do not show a darkening simultaneously along red, green, and blue. Other regions can be discounted because they show color trends not compatible with the results from the probe shadow. For a complete shadow recognition system, geometric constraints would also need to be brought to bear [Funk-Lea, 1992].

Rather than rely on restrictive or untestable assumptions, or high level knowledge about the scene illumination [Gershon et al., 1986], we have made use of an observer than can actively cast shadows. For relatively uncluttered scenes where shadows are present, the observer

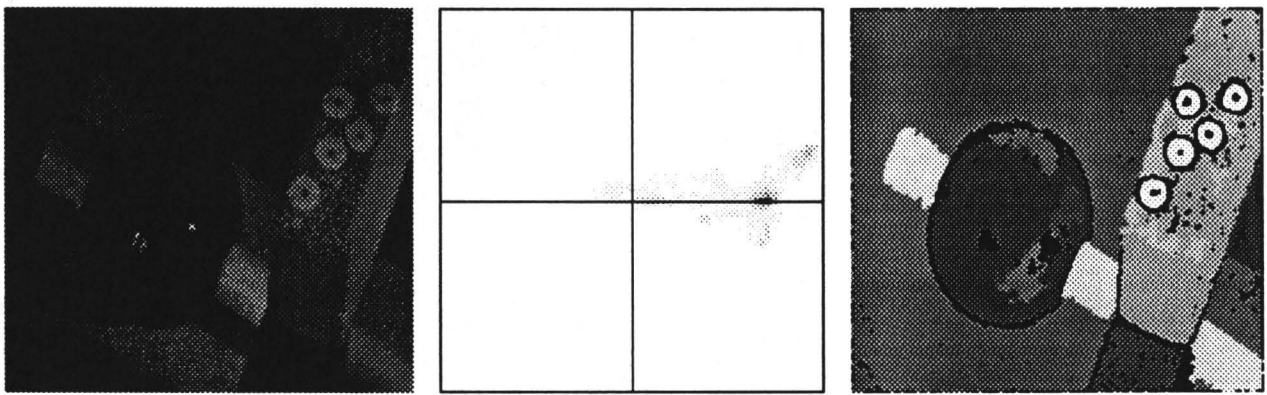


Figure 4: Color Image Segmentation: Scene 3. Left: A scene illuminated by two strong source of illumination. Center: 2D color histogram of the entire image. The dark spot at the right corresponds to the color of the ground. Right: Gray-scale coded regions from the color segmentation algorithm. Black indicates no region was found.

will be likely to cast a visible shadow. By addressing the problem of shadow recognition prior to recovering high level scene knowledge, we have a system that meets the needs of a large number of visual modules for geometric recovery that assume that shadows have been accounted for. In addition, because of our need for only relatively weak a priori assumptions, our system can handle scenes with complex geometry and illumination.

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, 1992] Gareth D. Funka-Lea. A proposal concerning the analysis of shadows in images by an active observer. Tech. report MS-CIS-92-78, GRASP Lab 335, University of Pennsylvania, Philadelphia, Oct. 1992.
- [Gershon *et al.*, 1986] Ron Gershon, Allan D. Jepson, and John K. Tsotsos. Ambient illumination and the determination of material changes. *J. of the Optical Society of America A*, 3(10):1700-1707, Oct. 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, Feb. 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, Nov. 1988.
- [Jiang and Ward, 1992] Caixia Jiang and Matthew O. Ward. Shadow Identification. In *Proc. Computer Vision and Pattern Recognition*, Illinois, Dec. 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, Jan. 1990.
- [Lee, 1991] Sang Wook Lee. *Understanding of Surface Reflections in Computer Vision by Color and Multiple Views*. PhD thesis, GRASP Lab., Dept. of Electrical Engineering, University of Pennsylvania, Aug. 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 Conf. on Computer Vision*, Osaka, Dec. 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, Feb. 1990.
- [Nagao *et al.*, 1979] Makoto Nagao, Takashi Matsuyama, and Yoshio Ikeda. Region extraction and shape analysis in aerial images. *Computer Graphics and Image Processing*, 10(3):195-223, Jul. 1979.
- [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.
- [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*, 10:210-218, 1985.
- [Torrance and Sparrow, 1967] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflections from roughened surfaces. *J. of the Optical Society of America*, 57:1105-1114, 1967.
- [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.