Shadow Segmentation and Classification in a Constrained Environment

Caixia Jiang

Department of Electrical Engineering

Worcester Polytechnic Institute, Worcester, MA 01609

and *Matthew O. Ward*

Department of Computer Science

Worcester Polytechnic Institute, Worcester, MA 01609

Running Head: Shadow Segmentation and Classification

Correspondence directed to:

Dr. Matthew O. Ward

Computer Science Department

Worcester Polytechnic Institute

Worcester, MA 01609

phone: 508-831-5671

fax: 508-831-5776

e-mail: matt@cs.wpi.edu

Abstract: A shadow identification and classification method for real images is developed in this paper. The method is based on the extensive analysis of shadow intensity and shadow geometry in an environment with simple objects and a single area light source. The procedure for identifying shadows is divided into three processes: low level, middle level, and high level. The low level process extracts dark regions from images. Dark regions contain both shadows and surfaces with low reflectance. The middle level process performs feature analysis on dark regions, including detecting vertices on the outlines of dark regions, identifying penumbrae in dark regions, classifying the subregions in dark regions as self-shadows or cast shadows, and finding object regions adjacent to dark regions. The high level process integrates the information derived from the previous processes and confirms shadows among the dark regions.

List of Symbols

Δ: Threshold Line Offset

δ: Intensity Threshold

δ_v: Angle Threshold

θ: Angle Between Vectors

$\vec{V_l}$: Left Vector

∞: Infinity

$\int_{y_0}^{y_2}$: Integral from y0 to y2

I_L: Many Subscripted Variables

1 Introduction

Shadows occur in a wide variety of scenes. If shadows are identified in images, they can provide a great deal of information about the shape, relative position, and surface characteristics of objects in a scene [2, 3, 4, 7, 10, 12, 13, 16]. It is not difficult for humans to distinguish shadows from objects. However, identifying shadows by computer involves solving many difficult problems, such as region extraction and knowledge representation/integration. Several methods for identifying shadows have been developed in recent years [17, 14, 7, 4, 3, 1, 15, 8, 11].

Thirion [17] used ray casting to determine shadows in images assuming that 3D models of objects in the image and the position of the camera are known. A ray is cast from each point corresponding to a pixel of the real image toward the light source. If the ray is intercepted, the point is in the shadow. The position of the light source is either known or can be calculated from the corners of buildings and their shadows. Wang, Huang, and Rosenfeld [14] developed a method to detect clouds and cloud shadows on aerial photographs assuming that clouds are at approximately the same height and the direction of the camera is perpendicular to the ground. They first found the sufficiently bright and dark regions as the candidates of clouds and their shadows, and then looked for the pair of parallel lines connecting the centroids of each pair of bright regions and dark regions. Since the clouds are closer to the camera than their shadows, the area of shadows should be smaller than the clouds and the line should be shorter than the line joining the shadows. Liow and Pavlidis's work [7] on shadows involves looking for shadowed lines (which are called shadow making lines in our work). Assuming that the boundaries of buildings can be approximated by straight lines, a line with average gradient greater than a threshold whose dark side is further away from the sun is assumed to be a shadowed line.

Irvin and McKeown [4] assumed that shadows are among the darkest regions in images. They first found the possible object regions using hypotheses regarding shapes (e.g. rectangular components). Then they found the dark regions adjacent to the object regions. The boundaries of the dark regions close to the sun are approximated by straight lines. If the lines form a near 90 degree corner and the corner is concave toward the sun, the dark regions are shadows. Huertas and Nevatia [3] used shadows to detect objects on aerial photographs, assuming also that objects consist of rectangular components. The corners near $90°$ in images were identified and labeled as object corners and shadow corners according to the brightness of the corners. If an object corner matches a shadow corner in the direction of sun light, the lines forming the object corner belong to the object and the lines forming a shadow corner belong to a shadow.

Adjouadi [1] presented a four-step analysis approach to identify shadows based on the fact that a cast shadow on a surface only reduces the surface intensities without changing surface characteristics such as texture. Thus, the method may work well when the surface on which shadows are cast has texture. Witkin [15] classified edges as shadow edges and object edges. If the intensities across an edge are highly correlated with abrupt intensity shifts, the edge is a shadow edge. Lowe and Binford [8] developed a reasoning system which uses some of the constraints on shadow boundaries, object boundaries, and surface marking boundaries to derive a geometric interpretation of an aerial image, which includes the interpretation of shadows. The interpretation of shadows was based on the match of the vertices on shadow-making lines and the vertices on shadow lines for a given light direction (sun light). Nagao, Matsuyama, and Ikeda [11] used the histogram of an image to identify shadows in aerial images. Assuming that the histogram is bimodal, the valley between the two peaks in the histogram was chosen as a threshold to determine shadows.

Most of the above methods assume that the light rays are parallel and the direction of the light is known. Methods employing object models assume rectilinear structure.

In this paper we present details of a new method for identifying and classifying shadows in real images based on the constraints that shadows possess in both intensity and geometry (an abbreviated presentation may be found in [6]. The method is developed under the following assumptions: (1) shadows are cast on a flat or nearly flat, simple, non-textured surface; (2) most image border pixels are on the background surface; (3) occlusion between shadows and objects is minimal; (4) only one area light source illuminates the scene. The procedure to identify shadows has three processes; low level, middle level, and high level.

The low level process extracts regions which are darker than their surroundings. Shadows are assumed to be among the dark regions. The middle level process detects features in dark regions such as the vertices on the outline of dark regions, the penumbrae, self-shadows and cast shadows in dark regions, and the object regions adjacent to dark regions. A *Penumbra Hypothesis*, *Self-shadow/cast shadow Hypothesis*, and *Object Adjacency Hypothesis* can be derived from these features and the high level process integrates these hypotheses and confirms the consistency among the light directions estimated from them. Extracting shadows is not separable from finding the correspondence between shadows and their objects. The correspondence information is then used to verify shadows. Self-shadows and cast shadows are distinguished whenever it is possible.

We show results of processing several images consisting of a small number of objects on a simple background. Our long term goal is to extend the system by making each phase more robust as we gradually relax constraints on the environment.

2 Shadow Geometry and Luminance

Shadows occur when objects totally or partially occlude direct light from a light source. Shadows are visible when a viewer is not coincident with a light source. In this paper, an object which casts a shadow is referred to as the shadow's object and the shadow is referred to as the object's shadow. A shadow has two parts: the *self-shadow* and the *cast shadow*. The *self-shadow* is the part of the object which is not illuminated by direct light. The *cast shadow* is the area projected by the object in the direction of direct light. The part of a cast shadow where direct light is completely blocked by its object is called an *umbra*. The part of a cast shadow where direct light is only partially blocked is called a *penumbra*. A point light source only generates umbrae in shadows. An area light source generates both penumbrae and umbrae in shadows, though a very narrow penumbra may not appear in an image due to digitizing effects.

Shadow boundaries can be classified into four types of segments: *shadow-making lines*, *shadow lines*, *occluding lines*, and *hidden shadow lines* [2]. The lines that separate the illuminated surfaces and nonilluminated surfaces of an object are called *shadow-making lines*. Shadow-making lines are the outlines of an object as seen from a light source. *Shadow lines* are the projection of shadow-making lines in the direction of light rays. *Occluding lines* are the outlines of an object as seen by a viewer. If a shadow-making line is not visible, its corresponding shadow line is called a *hidden shadow line*. Shadow-making lines, shadow lines, occluding lines, and hidden shadow lines are illustrated in Figure 1.

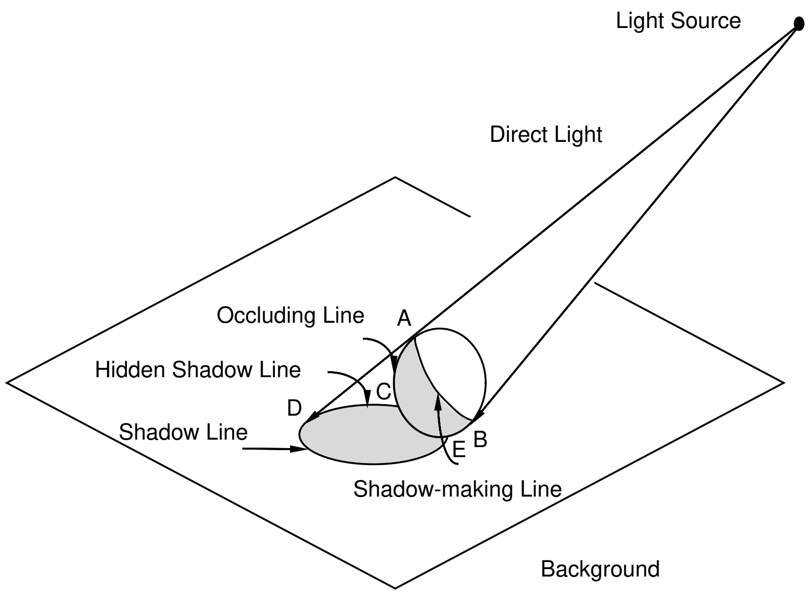

Figure 1: The segmentation of shadow boundaries. Segment AB is the shadow-making line. Segment DE is the shadow line. Segment AC is the occluding line. Segment CD is the hidden shadow line.

The intensities of cast shadows and self-shadows are not absolutely zero because of reflected light and ambient light. When direct light illuminates a background, the background reflects the light to a self-shadow, then the self-shadow reflects the light to its cast shadow. For a single object in a scene, the reflected light illuminating the cast shadow is only from its self-shadow. Thus, the intensities of the cast shadow on the background are usually lower than the intensities of the self-shadow, though this may not be the case if the objects are significantly darker than the background. For a scene with multiple objects, self-shadows and cast shadows may receive reflected light from other sources. If there is another object near a shadow and the illuminated surfaces of that object can reflect light to the shadow, the intensity of the cast shadow may be higher than that of the self-shadow. A self-shadow and its cast shadow are adjacent to each other when the object is placed on a background. The boundary separating the self-shadow and the cast shadow is the boundary where the object touches the background. In most cases, this boundary can be detected because the intensities across it are discontinuous due to the different surface normals and reflective properties of the object and the background. If the boundary cannot be detected, the self-shadow and the cast shadow cannot be distinguished unless the object shape is known.

The intensities in an umbra are lower than the intensities in its penumbra because the umbra does not receive any light directly from a light source. The intensities gradually increase from an umbra to its penumbra. The calculation of the illuminance in a penumbra is similar to that on an object surface, except that only a partial light source needs to be considered.

Figure 2 shows the illuminance calculation in a penumbra for a straight line light source with a constant luminance L_0. The object is a straight line with height h_0 casting its shadow on a flat surface.

For any point z_a in the penumbra, the part of the light source which can illuminate the point z_a is from y_0 to y_2, where $y_0 = \frac{h_0}{z_a - z_0} z_a$. The illuminance generated by an arbitrary small segment dy between y_0 and y_2 at the point z_a is

$$dI(z_a) = \frac{L_o \cos\theta_1 \cos\theta_2}{R^2} dy \tag{1}$$

where $z_1 \leq z_a \leq z_2$ and

$$\begin{aligned} \cos\theta_1 &= \frac{z_a}{R} \\ \cos\theta_2 &= \frac{y}{R} \\ R^2 &= y^2 + z_a^2 \end{aligned}$$

The illuminance generated from y_0 to y_2 at z_a is

$$I(z_a) = \int_{y_0}^{y_2} \frac{L_0 \cos\theta_1 \cos\theta_2}{R^2} dy$$

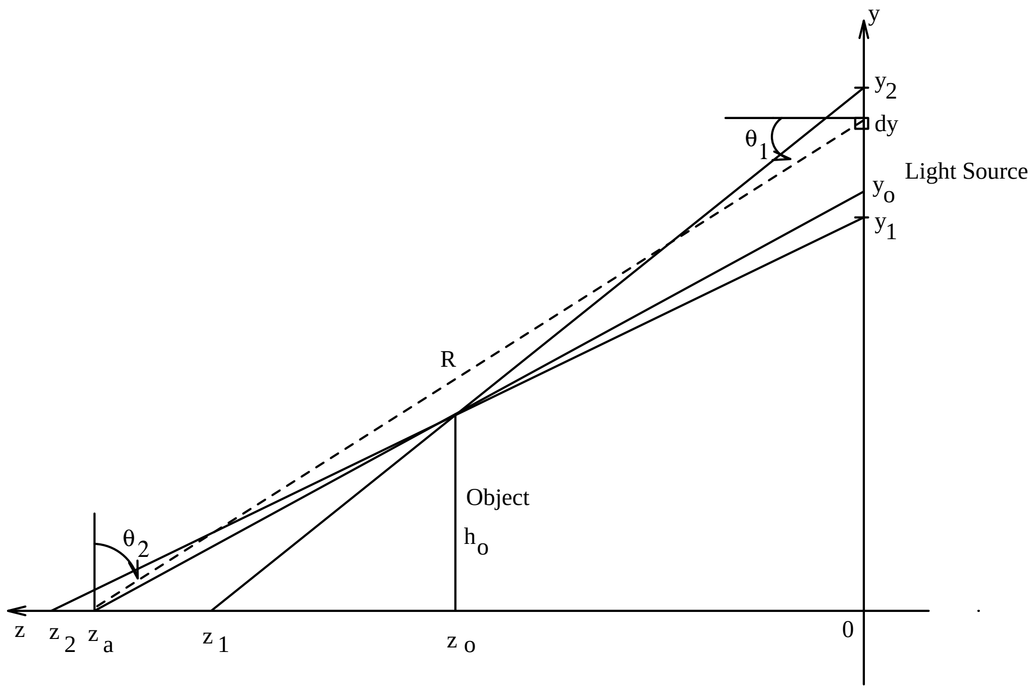

Figure 2: Illuminance calculation in a penumbra for a line light source. h_0 is the height of the object. The line light source is from y_1 to y_2. The umbra is from z_0 to z_1 and the penumbra is from z_1 to z_2. The part of the light source which can illuminate the point z_a is from y_0 to y_2. dy is an arbitrary small segment between y_0 and y_2. θ_1 is the angle between the surface normal of dy and the light ray R from dy to z_a. θ_2 is the angle between the surface normal of the background and the light ray R.

$$= \int_{y_0}^{y_2} \frac{L_0 z_a}{(y^2 + {z_a}^2)^2} y dy$$
$$= \frac{L_0 z_a}{2} \left(\frac{1}{z_a^2 + y_0^2} - \frac{1}{z_a^2 + y_2^2} \right) \quad (2)$$

When z_a increases, the illuminance at z_a increases, that is the intensities increase from the point z_1 (the end of the umbra) to the point of z_2.

From the above example, we see that the variations of intensities in a penumbra are not a simple function of a light source and an object. It is extremely difficult to achieve theoretical formulae of the intensities in a penumbra for an arbitrary object and an arbitrary light.

In general, the calculation of the width of a penumbra is complicated. Figure 3 shows a simple ideal example for calculating the width of a penumbra. The object is a vertical line and the light source is a line light source parallel to the background.

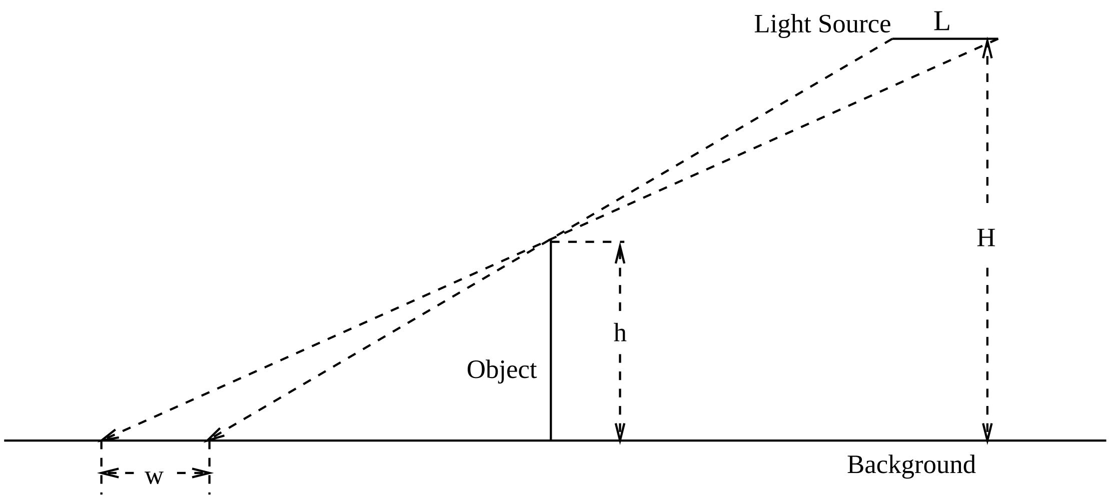

Figure 3: The calculation of the width of a penumbra. The light source is a line light source with length L parallel to the background. The object is a vertical line with length h. The width of the penumbra is w.

The width w of the penumbra in Figure 3 can be calculated by the triangular relation among the light source and the object ([5]). The w is given by

$$w = \frac{Lh}{H - h} \quad (3)$$

where H and L are the height and the length of the light source, respectively. h is the height of the object. For a given h, w depends on the size L and height H of the light source. As

L increases, w increases. Similarly, as H increases, w decreases. For a given H and L, the larger the h, the larger the w. When $h \to H$, $w \to \infty$. If $h > H$, $w < 0$, which means that the umbra becomes infinitely large. Therefore the maximum width of a penumbra generally appears at the shadow boundary cast by the highest point on an object (maximum h) while the minimum width of a penumbra appears at the shadow boundary cast by the minimum height point (minimum h) on an object.

For multiple objects and a single light source, if one object casts its shadow along the direction of light, all visible objects from the light source should also cast their shadows along the direction of light. The effect of shadows cast on a surface is to reduce the intensities on the surface.

3 The Extraction of Dark Regions

The first step toward identifying shadows is to extract dark regions as potential shadows. The constraints on shadows can be used to check if a dark region is a true shadow. The constraints on shadows involve shadow boundaries, junctions on shadow boundaries, self-shadows, and cast shadows. How to use these constraints are of great importance for the success of shadow identification in real images. To use these constraints, the boundaries, junctions, and regions involved in these constraints must first be found, which requires the isolation of dark regions containing both true shadows and areas with low reflectance. This is referred to as the low level process in this paper, and is crucial in determining the effectiveness of the middle level and high level processes. The middle level process depends on reliable isolation of dark regions for feature extraction, which in turn affects the quality and consistency of the the hypotheses derived from these features in the high level process.

Assume that the background of an image is a flat or nearly flat non-textured surface and that the objects and their shadows are within the image. This implies that all image border pixels are on the background. There is only a single light source to illuminate the viewed scene. If a light source emits parallel light rays, the intensities on the background of an image can be approximated by a constant. To find regions darker than the background a single threshold can be set at a value below the intensities of the background. This doesn't guarantee that all shadows will be detected, as it is possible that self-shadows on the objects with high reflectance are brighter than the background, For a light source which does not emit parallel light rays, such as a light source in close proximity, the intensities on the background change smoothly, thus a single threshold to determine dark regions in images does not, in general, give satisfactory results. For a given threshold, dark regions may be too fragmented in one area of an image, while being overmerged with non-dark regions in another area of the image.

A Cartesian coordinate system for an image is chosen with the horizontal direction being the x axis and the vertical direction being the y axis. The intensity on any point (x, y) is denoted by $I(x, y)$ with x in the range from 0 to $n_x - 1$ and y in the range from 0 to $n_y - 1$, where n_x and n_y are the size of the image in the x and y directions, respectively.

If an image is scanned in the x direction at $y = y_0$, the line at $y = y_0$ with the intensities $I(x, y_0)$ for $0 \leq x < n_x$ is called a scanned line. The two points at $x = 0$ and $x = n_x - 1$ of $I(x, y_0)$ can be used to construct a straight line L. The intensities along a line L can be expressed by

$$I_L(x, y_0) = \frac{(I_b - I_a)}{n_x - 1} x + I_a \tag{4}$$

where $I_a = I(0, y_0)$ and $I_b = I(n_x - 1, y_0)$, respectively. Let the intensities on the line L be reduced by an offset value Δ. The new line can be expressed by

$$I_t(x) = I_L(x, y_0) - \Delta \tag{5}$$

where Δ is a positive constant. We call the new line a *threshold line.*

Δ represents the intensity difference between shadows and non-shadows. It varies with different images, being dependent on the variability of background intensities, and is manually determined in this paper. The automatic determination of Δ for a given image is a non-trivial task and is a current topic of our research.

Figure 4 shows a real image with 256 by 256 pixels. Its intensities on the two scanned lines in the x direction at $y = 50$ and 150 are shown in Figure 5 and 6, respectively. The threshold line is constructed for each scanned line at two different offset values, $\Delta = 0$ and $\Delta = 20$.

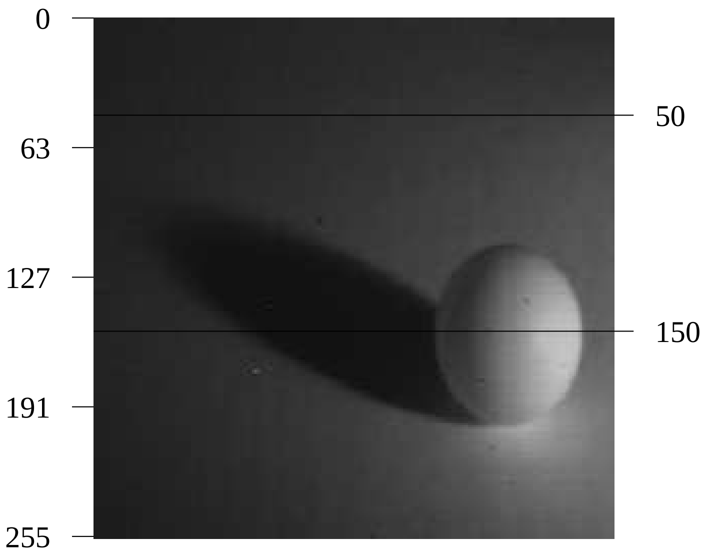

Figure 4: A real image with 256 by 256 pixels.

Any pixel on a scanned line with intensity below the threshold line is labeled as a dark pixel. If we scan an image in the x direction from $y = 0$ to $y = n_y - 1$ and construct a threshold line for each scanned line at a constant offset value, the pixels below the threshold lines will form dark regions. If $\Delta = 0$, some pixels may be labeled as dark pixels just due to the fluctuation of noise on the intensities, as seen in the threshold line 1 shown in Figure 5 and Figure 6, respectively. By choosing an appropriate Δ, the dark pixels resulting from noise may be eliminated. Thus, one of the purposes for using an offset in threshold lines is to reduce the noise effect on determining dark pixels. Different images may require different Δ to compensate for the noise. Since threshold lines are constructed based on the ending points of the scanned lines, they are affected by the size of an image window, even when the light source and the environment do not change.

An image could be scanned in any direction and the ending points of the scanned lines can be used to construct the threshold lines. The reason one might want to scan an image in many directions is to increase the amount of information used in labeling pixels as well

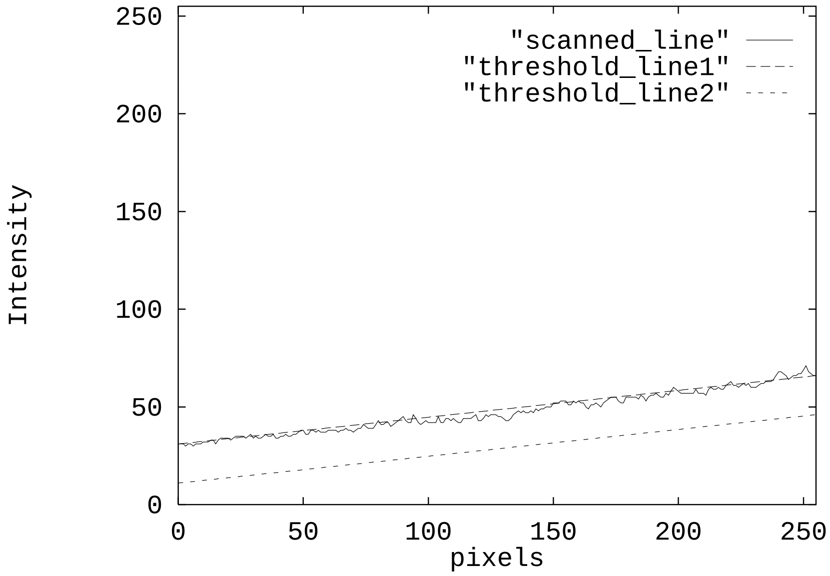

Figure 5: The scanned line and its threshold lines at $y = 50$ of the image shown in Figure 3. The threshold line 1 has zero offset and the threshold line 2 has the offset $\Delta = 20$.

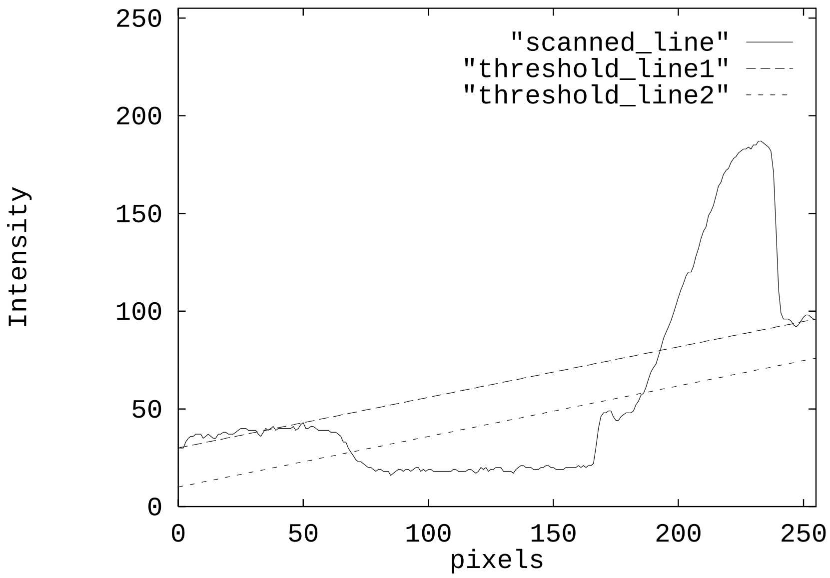

Figure 6: The scanned line and its threshold lines at $y = 150$ of the image shown in Figure 3. The threshold line 1 has zero offset and the threshold line 2 has the offset $\Delta = 20$.

as to attempt to locate dark region boundaries by scanning in directions normal to the boundaries. The intensities I_a and I_b on each scanned line can be the average intensities in the neighborhood of the ending points. The method proposed above to extract dark regions will be demonstrated on several real images in section 6.

The basic threshold line method assumes that image border pixels are on the background, as it uses the image border pixels to construct threshold lines. Obviously, if objects or shadows are on image borders, this assumption is violated. In this case, at the image borders the intensities on the objects or shadows are replaced by a second order polynomial interpolation of adjacent background intensities.

The basic threshold line method fails when the intensities on image borders are lower than those in shadows, which may be caused by a focussed light source. To make the threshold line method work in this case, image border pixels are moved toward the center of the image, but still on the background. The new border pixels should have higher intensities than those in shadows in order for the intensities in shadows to be below the threshold lines constructed by the new border pixels.

A dark region found by the threshold line method may contain several subregions. A dark region can contain cast shadows, self shadows, and dark marks. The number of subregions and the relationship among the subregions in a dark region can provide very important information for detecting shadows. This will be exploited in the later sections.

4 Techniques for Feature Extraction

4.1 Vertices

Vertices on shadow boundaries play an important role in shadow identification. Shadow-making lines and shadow lines are generally terminated at vertices. Vertices are defined as points on a boundary with discontinuous first derivative. We arbitrarily select a starting point on the boundary of a region and walk along the boundary in the clockwise direction. Let a boundary with the number of pixels N_p be represented by parametric functions $\{X(i)$,

$Y(i)$, $0 \leq i < N_p\}$. For any point P_k on the boundary, $X(i)$ and $Y(i)$ between $k-m \leq i \leq k$ are fitted by second order polynomial functions $X_l(i)$ and $Y_l(i)$, respectively, where m is a positive constant. Similarly, $X(i)$ and $Y(i)$ between $k \leq i \leq k+m$ are fitted by second order polynomial functions $X_r(i)$ and $Y_r(i)$, respectively.

Let us denote $\vec{V}_l(k)$ as the tangent vector at the point P_k calculated from $\{X_l(i),\ Y_l(i)\}$ and $\vec{V}_r(k)$ as the tangent vector at the point P_k calculated from $\{X_r(i),\ Y_r(i)\}$, respectively.

$$\vec{V}_l(k) = \left(\frac{dX_l(i)}{di}, \frac{dY_l(i)}{di}\right)_{|i=k} \tag{6}$$

$$\vec{V}_r(k) = \left(\frac{dX_r(i)}{di}, \frac{dY_r(i)}{di}\right)_{|i=k} \tag{7}$$

The angle $\theta(k)$ between the left and right tangent vectors at the point P_k is:

$$\theta(k) = \cos^{-1}\left(\frac{\vec{V}_l(k) \cdot \vec{V}_r(k)}{|\vec{V}_l(k)||\vec{V}_r(k)|}\right) \tag{8}$$

For the point P_k we calculate N pairs of the fitting functions $\{X_l(i),\ Y_l(i)\}$ and $\{X_r(i),\ Y_r(i)\}$, respectively, by varying the value of m. If the smallest $\theta(k)$ among those derived from the N pairs of the tangent vectors $(\vec{V}_l, \vec{V}_r)$ is greater than a threshold δ_v, the point P_k is considered to be a vertex (see [5]).

Figure 7 shows the functions $\theta(i)$, $0 \leq i < 427$ for the outline of the dark region in Figure 4. The solid line represents the function $\theta(i)$ at a fixed $m = 19$. The dashed line represents the minimum of function $\theta(i)$ for $5 \leq m < 20$. The widths of the peaks on the dashed line are narrower than those on the solid line. But the peak values on the dashed line are smaller than those on the solid line. There is a trade-off between the heights and the widths of peaks.

The vertices detected in Figure 7 are shown in Figure 8 with the threshold $\delta_v = 0.5$ radian. A false vertex near point 250 is found which is not shown in Figure 8. This false vertex resulted from the use of an unsmoothed outline and was filtered out due to its proximity to a much stronger vertex.

Since $\theta(i)$ needs to be greater than δ_v at vertices, any vertex with $\theta(i) \leq \delta_v$ is not

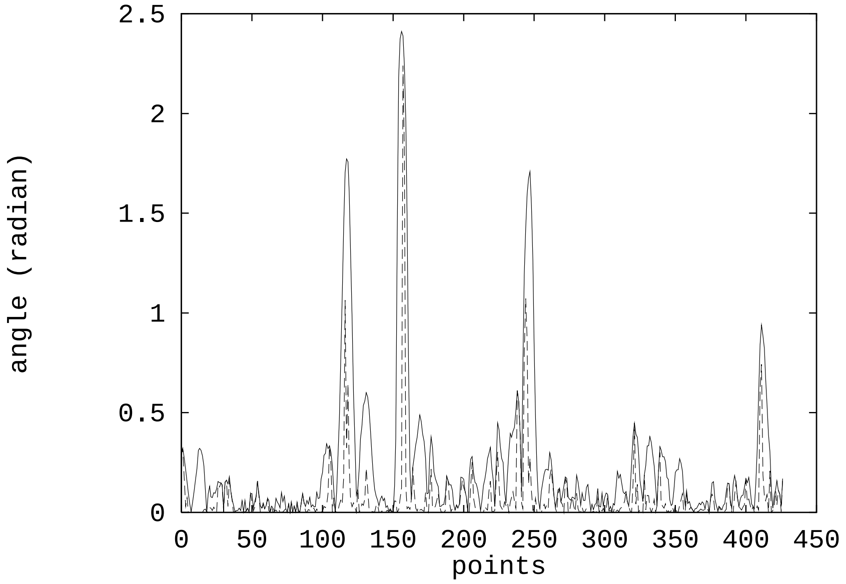

Figure 7: The angles between the left and right tangent vectors of each point on the outline of the dark region in Figure 4. The solid line represents the function $\theta(i)$ at a fixed $m = 19$. The dashed line represents the function of the minimum $\theta(i)$ between $5 \leq m \leq 19$.

detectable. In Figure 8,the diamond symbols on the boundary are the detected vertices. All vertices on the boundary are correctly detected.

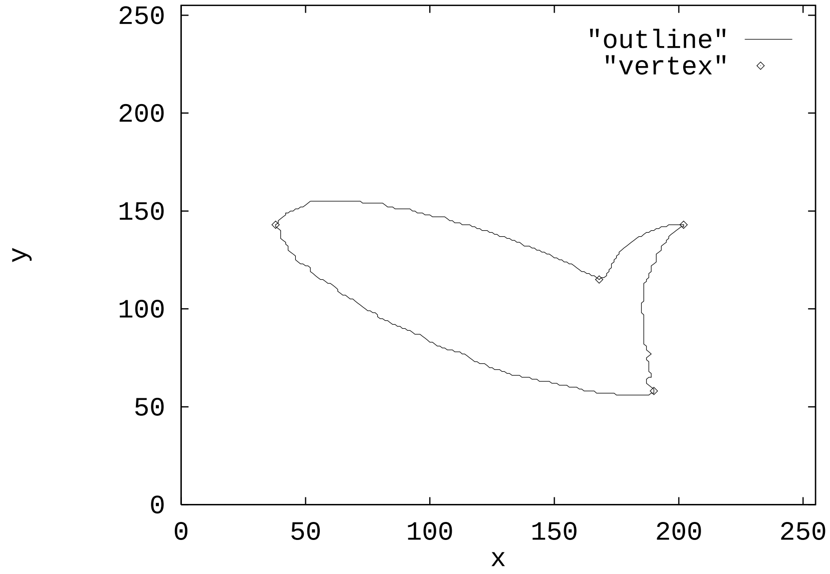

Figure 8: The diamond symbols on the boundary are the detected vertices.

4.2 Penumbrae

Penumbrae in shadows provide one of the most important constraints on shadows. It seems that penumbrae have been generally ignored in identifying shadows, although they have been mentioned as important features in [8, 9]. In some cases, penumbrae are one of only a few available features for to assist in identifying shadows, e.g., an object suspended above a surface and casting its shadow on the surface, where the object is not in the image. In such a scenario the penumbra provides a reliable clue to distinguish the shadow from a dark mark. Even if an object casts a shadow on a surface on which it lies, the presence of a penumbra in the shadow strongly supports the shadow hypothesis. Although penumbrae are strong constraints that shadows possess, penumbrae have not been widely used in identifying shadows. One of the possible reasons is that penumbrae are very difficult to find. The current available edge detection techniques are not able to detect either the boundaries in penumbrae or the width of the boundaries. Another possible reason is that most approaches for identifying shadows use the assumption that a light source emits from a single point. Point light sources do not generate penumbrae. In the same manner as shadows are problems in image analysis if they are not identified, the presence of penumbrae in shadows may make shadow identification very difficult because penumbrae have very smooth gradients.

The intensities from an umbra to a penumbra increase gradually. An umbra should be enclosed by its penumbra. To identify penumbrae from dark regions, we calculate the width of dark region boundaries. When the threshold Δ varies in a certain range (Δ is the offset in the threshold line method described in section 3), the size of the identified dark regions varies accordingly. If we assume that no splitting and merging of dark regions occur when Δ varies, the total number of dark regions remains unchanged. The difference between the biggest size and the smallest size of a dark region is the width of the dark region boundary (see [5]). If a dark region is a shadow, the width of the dark region boundary is the penumbra. Figure 9 shows the outlines of the dark region in Figure 4 found by the threshold line method at $\Delta = 18$ and $\Delta = 23$, respectively. The solid outline is the outline of the dark region at $\Delta = 18$. The dashed outline is the outline of the dark region at $\Delta = 23$. The penumbra is the area between the two lines.

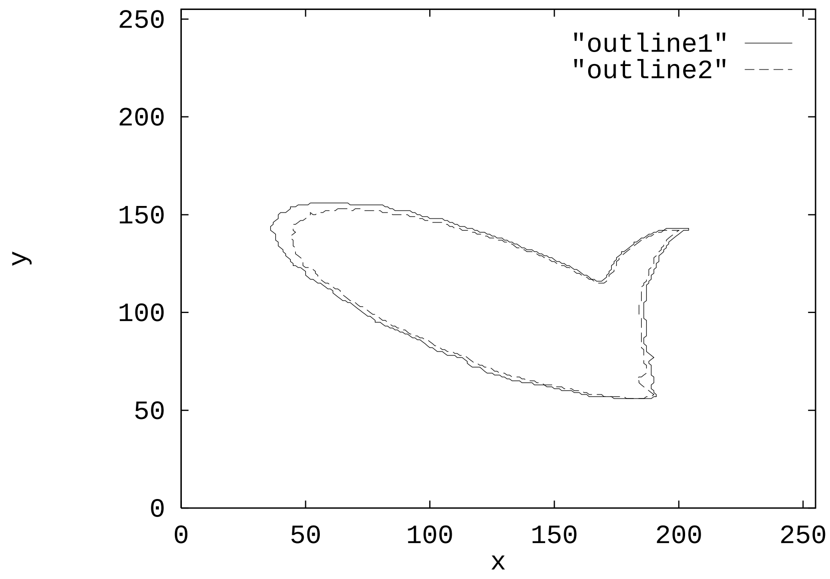

Figure 9: The outlines of the dark region found by the threshold line method at $\Delta = 18$ and $\Delta = 23$ for the image shown in Figure 3. The solid outline corresponds to $\Delta = 18$ and the dashed outline corresponds to $\Delta = 23$.

4.3 Self-shadows vs. Cast Shadows

Each dark region needs to be checked for the presence of any subregions. The reason for finding subregions in each dark region is to separate cast shadows, self-shadows, and dark marks. If a dark region does not contain any subregions, it is called a *simple dark region.* Similarly, if a subregion does not share the boundary with any other subregion or the dark region (the subregion is totally enclosed in the dark region), the subregion is called a *simple subregion.* Simple subregions are considered as neither self-shadows nor cast shadows. Thus, they are assumed to be dark marks. Simple dark regions and simple subregions are sometimes referred to as *simple regions.*

Generally, the intensities in a self-shadow are higher than those in the corresponding cast shadow, since the self-shadow receives reflected light from the background. We assume this relationship when assigning the subregions in a dark region as self-shadows and cast shadows. The errors which occur when cast shadows are brighter than self-shadows (when objects are significantly darker than the background) are generally discovered during the verification process.

A dark region can at most contain one self-shadow and one cast shadow since we have made the assumption that shadows and objects do not occlude each other and themselves. If there are more than 2 subregions in a dark region, an intensity threshold δ is used to distinguish the self-shadow regions and cast shadow regions. δ is chosen as the average of the mean intensities of the darkest subregion and the largest subregion other than the darkest. If the mean intensity of a subregion is greater than δ, it is assigned as a self-shadow subregion, otherwise, it is assigned as a cast shadow subregion. Self-shadow subregions must be adjacent to each other so that they can be merged to be one self-shadow. Similarly, cast shadow subregions must be adjacent to each other to form one cast shadow. The largest self-shadow subregion and the largest cast shadow subregion are retained as a self-shadow subregion and a cast shadow subregion respectively. The other small self-shadow subregions and small cast shadow subregions, if any, are merged with their adjacent subregions. Thus, only one self-shadow region and one cast shadow region are in a dark region. Since the self-shadow and the cast shadow are on two different surfaces, the surface normals across the boundary separating the self-shadow and the cast shadow are usually discontinuous. Thus, the two ending points of the self-shadow and cast shadow boundary should be vertices on the outline of the dark region.

4.4 Object Adjacency

Some of the constraints on shadows rely on the types of regions that surround dark regions. A dark region may be surrounded by background and/or object regions. Note that an object may cast its shadow completely on another object surface. In this case the other object surface may be treated as a "background". In order to find the segments of the outline of a dark region, we need to identify object regions. We already know how difficult it is to detect boundaries and regions. Fortunately, it is not necessary to find the all of the regions for an object. The exact shape of objects is not important because different objects may cast exactly the same shadows on a surface as long as the occluding lines of objects viewed from a light source are the same. We only need to know whether the immediately adjacent area of the outline of a dark region belongs to an object or a background.

To determine the relevant object regions, we detect the edges in non-dark regions by an edge operator. Only the edges that are terminated at dark region boundaries are selected. These object edges are tested to see if they form a closed region with a segment of a dark region boundary. The ending points of the object edges should be the vertices on the outline of the dark region, regardless of the object surface normals. This results from the fact the the object boundary is the outline of the object observered by a viewer and the self-shadow boundary is the outline of the object viewed from a light source. The intersecting points between the two outlines are on the outline of the dark region and have a discontinuous derivative. If the dark region is a shadow, the segment of the dark region boundary is the shadow-making line. A detailed description of this algorithm can be found in [5].

5 Verification of Hypotheses

Shadow hypotheses derived from shadow constraints for each dark region may conflict among each other due to the limitation of the low level process and the complexity of lighting and object shape. The shadow hypotheses derived in this paper consist of the *Penumbra Hypothesis*, *Self-shadow/cast shadow Hypothesis*, and *Object Adjacency Hypothesis*. The occurrence of a dark region in an object implies that the dark region could be a dark mark. However, if this dark region has a penumbra, a self-shadow and cast shadow, or an object adjacent to it, the dark region is likely to be a shadow. Thus, the fact that a dark region is in an object region only implies that it has higher probability of being a dark mark than a dark region not in an object region.

Three shadow hypotheses for each dark region generate eight combinations listed in Table 1.

T and F in Table 1 represent the true and false hypotheses, respectively. A true *Penumbra Hypothesis* for a dark region indicates that a potential penumbra was detected and a

Type of combination	Penumbra	Self-shadow & cast shadow	Object
0	F	F	F
1	T	F	F
2	F	T	F
3	F	F	T
4	T	T	F
5	F	T	T
6	T	F	T
7	T	T	T

Table 1: Eight combinations of three shadow hypotheses for each dark region.

false *Penumbra Hypothesis* for a dark region indicates that the dark region does not have a detectable penumbra. The definitions of true and false for the *Self-shadow/cast shadow* and *Object Adjacency Hypothesis* are similar to that of a penumbra. The hypotheses in each combination may not be correct or they may conflict with each other. Hypotheses are said to conflict if there exists an inconsistency in edge labels or region adjacencies produced by the individual hypotheses. Only the true hypotheses are checked to see if they are correct or conflict with each other. If a true hypothesis is shown to be incorrect, the hypothesis is changed to be false. A false hypothesis is always false.

Let N_t be the total number of correct true hypotheses in each combination after resolving the conflicting hypotheses and identifying incorrect hypotheses, and let C_d be the type of a dark region. There are three types of dark regions: dark marks ($C_d = 0$), cast shadows ($C_d = 1$), and both self-shadows and cast shadows ($C_d = 2$). If N_t is greater than 0 for a dark region, the dark region is identified as a shadow. Whether the dark region is a cast shadow or a self-shadow/cast shadow depends on the value of C_d. Each dark region will be classified as one of the eight combinations. The N_t and C_d in each combination are determined by how consistent the *Penumbra Hypothesis*, *Self-shadow/cast shadow Hypothesis*, and *Object Adjacency Hypothesis* are. The analysis of each type of the combinations for a dark region and strategies for confirmation and inconsistency resolution is summarized in Appendix A.

If more than one dark region in an image is identified to be either a cast shadow or a cast shadow and a self-shadow, the directions of the light rays causing the shadows should be checked to see if they are emitted from the same light source, since one light source is assumed to illuminate the scene in the image.

Methods have been developed to approximate the light source direction for each type of true hypothesis. The light direction derived for a shadow in an image must enter the shadow at a point and exit the shadow at another point (see [2]). If a dark region has a penumbra, the direction of the maximum width from an umbra to its penumbra may be used to approximate the direction of the light rays for the dark region. If a dark region has a self-shadow region and a cast shadow region, we may estimate the light direction as the direction from the middle point of the self-shadow boundary to the middle point of the cast shadow boundary on the outline. If a dark region has an adjacent object region, we may estimate the light direction as the direction from the middle point of the object boundary to the middle point of the cast shadow boundary on the outline. If different light direction estimations are available, the average direction from different estimations are taken as the estimated light direction of the dark region. Generally, the accuracy of the estimated light direction is poor due to the perspective distortion and errors in the low level process. Also for an area light source the light direction may not have a unique definition.

Giving the estimated light direction for each shadow, we need to verify their consistency. All the light rays emitted from a light source converge at the location of the light source. If a light source emits parallel light rays, the light directions for all shadows should be approximately parallel. The light source location estimated from the convergence of the estimated light directions of dark regions should not be in any dark region or object region (see [5]). If a dark region violates the consistency of light directions, it is regarded as a dark mark.

6 Demonstration

We present the results of our method for four real images. All the images are 256 by 256 with a resolution of 1 byte per pixel. For clarity, the identified object regions, self-shadows, cast shadows, and the dark marks in the images are shown with the intensities 230, 140, 50, and 0, respectively, where 255 represents white and 0 represents black as shown in Figure 10.

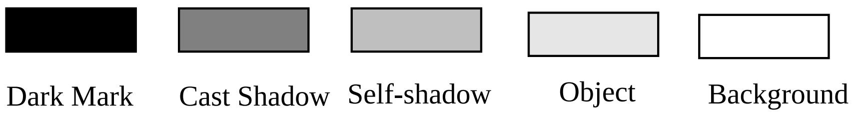

Figure 10: Different brightness to represent the identified dark marks, cast shadows, self-shadows, and objects shown in Figures 11-14.

The demonstration for each real image consists of the original image, the dark regions, and the result of the shadow identification as shown in Figures 11-14, respectively, The subregions in dark regions are marked. Classification errors are circled and explained below. More examples of our method on real images are given in [5].

There are 7 dark regions found in Figure 11. The two dark regions labeled 1 are found on an object. They are identified as dark marks and later associated with object regions, since there is no true hypothesis to prove they are shadows. All the other dark regions have either true Object Adjacency Hypothesis or true Self-Shadow/Cast Shadow Hypothesis to support classification as shadows. The dark region labeled 2 is a self-shadow but misidentified as a part of a cast shadow because the boundary separating the self-shadow and cast shadow is not found. The low intensity contrast between self-shadows and cast shadows in this figure also kept some other self-shadows from being detected. All the dark regions which are recognized as shadows in this figure have consistent light directions (the difference between any two directions is less than 90 degrees). They are -1.0, 25.1, -36.7, -21.8, and 15 degrees, respectively.

Part of the self-shadow in the dark region labeled 1 in Figure 12 is misidentified as a cast shadow because low intensity contrast between the self-shadow and the cast shadow. The dark regions identified as shadows have either true Self-Shadow/Cast Shadow Hypothesis or true Object Adjacency Hypothesis or both. The light directions derived from them are consistent with each other.

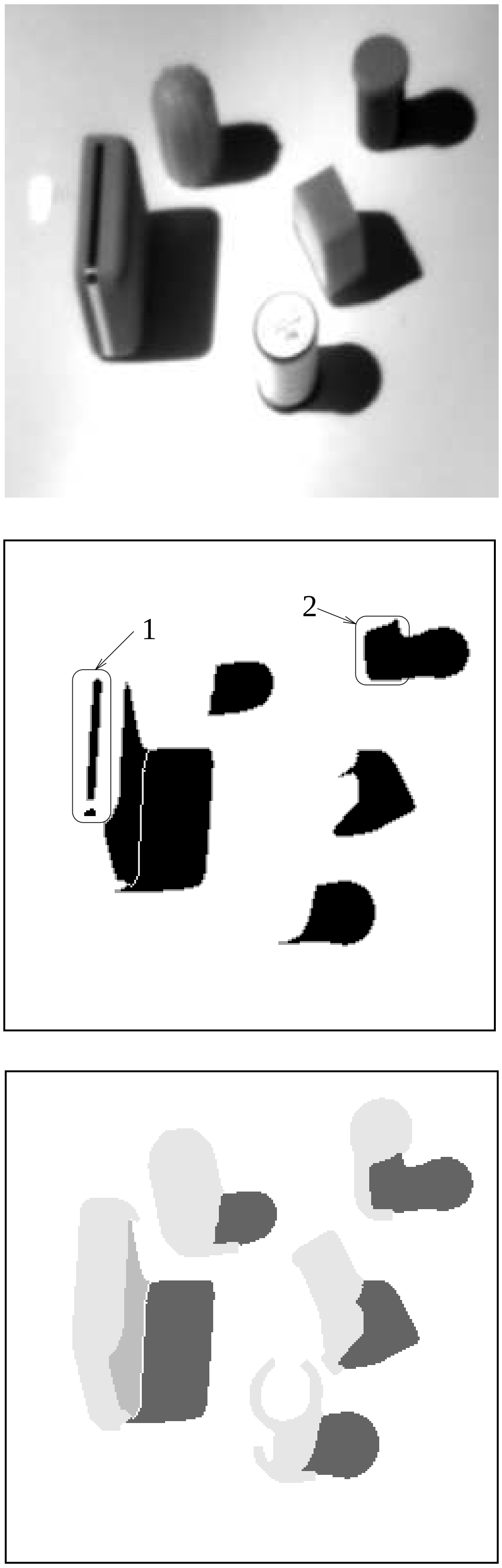

Figure 11: Top: a real image. Middle: the dark regions. Bottom: the objects, self-shadows, and cast shadows.

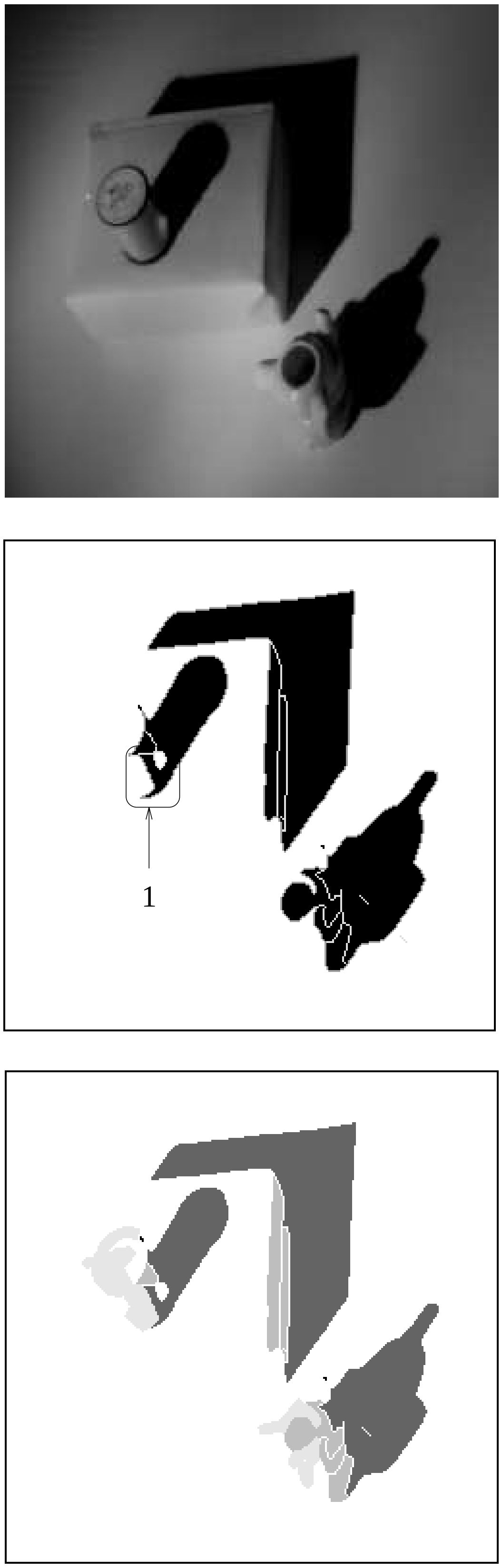

Figure 12: Top: a real image. Middle: the dark regions. Bottom: the objects, self-shadows, cast shadows, and dark marks.

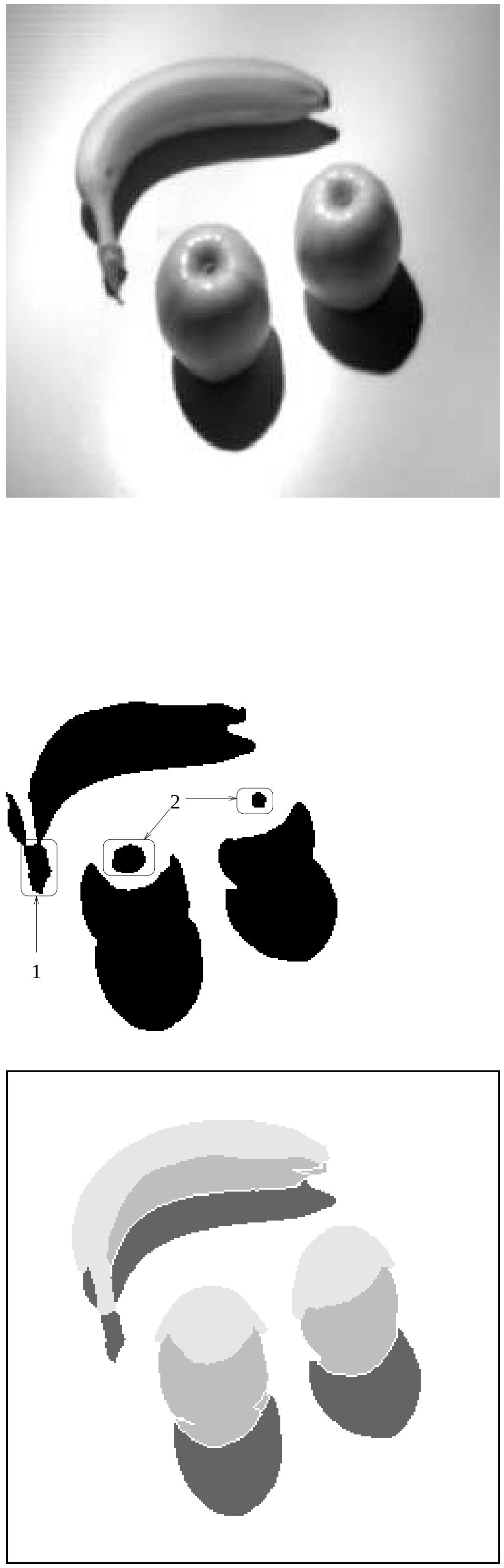

Figure 13: Top: a real image. Middle: the dark regions. Bottom: the objects, self-shadows, and cast shadows.

The dark region labeled 1 is misidentified as a cast shadow because it is segmented into the same dark region as the cast shadow. The left dark region labeled 2 in Figure 13 is identified as a cast shadow because a small object region is found adjacent to it and the light direction derived from it is consistent with the light directions derived from other identified shadows. The right one is identified as a dark mark. Since these two dark regions are on objects, they are not shown in the bottom figure. The dark regions which are identified as shadows have both true Self-Shadow/Cast Shadow Hypothesis and true Object Adjacency Hypothesis.

The dark regions labeled 1 in Figure 14 are identified as dark marks. The objects which correspond to them are too shallow for their cast shadows to be identified. The dark region labeled 2 is assigned as a dark mark since it is a simple subregion. The two dark regions which are identified as shadows have true Self-Shadow/Cast Shadow Hypothesis and true Object Adjacency Hypothesis and the light directions derived from them are consistent with each other(3.4 and -17.1 degrees respectively).

7 Conclusion

A method to identify and classify shadows in real images from a constrained environment is developed. The results shown in the previous section demonstrate the success of the method.

The threshold line method has been developed to extract dark regions in images. The dark regions may include self-shadows, cast shadows, and dark marks. The method performs fairly well on various images. The good performance of the method facilitates the middle level and high level processes in identifying shadows. One parameter in the method, Δ, needs to be predetermined. For different images, Δ varies.

The middle level process extracts features related to dark regions such as vertices, penumbrae, self-shadows and cast shadows, and object regions. The vertices are important in the geometric analysis of shadows. They provide one of the geometric constraints among a self-shadow, its cast shadow, and the shadow's object. One of the unique features in the method

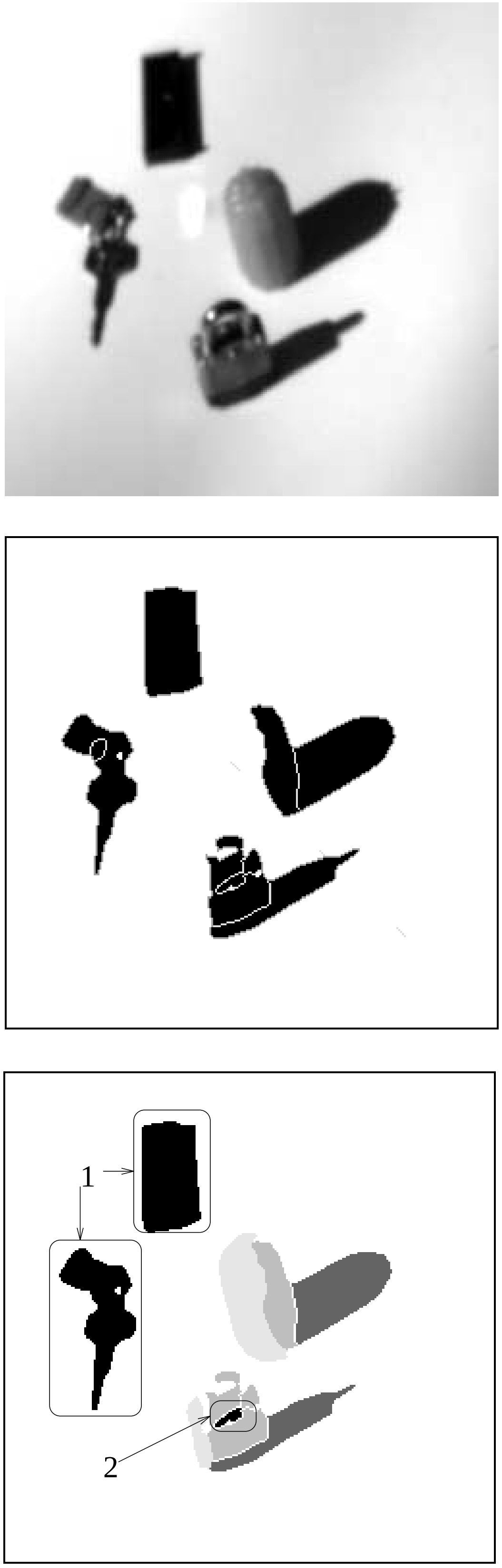

Figure 14: Top: a real image. Middle: the dark regions. Bottom: the objects, self-shadows, cast shadows, and dark marks.

developed to identify shadows in this paper is the use of penumbrae in cast shadows.

The geometrical arrangement of a self-shadow, its cast shadow, its penumbra, and the shadow's object is not arbitrary. The high level process checks the consistency of the features derived from the middle level process and resolves conflicts among them. Shadows in an image should have consistent light directions. Many researchers assume that a light source emits parallel light rays and the direction of the light rays is known in their methods of identifying shadows. In this paper, a light source is an arbitrary area light source with unknown direction. The consistency among the estimated light directions obtained from different regions is also checked. This step further eliminates false shadows.

Our method utilizes both intensity and geometry information in shadow identification. The method has proved to be reasonably effective within the environmental constraints stated in Section 1. Future work will focus on both improving performance under the current constraints as well as gradually relaxing some of these constraints.

Acknowledgements

The authors would like to thank Drs. David Cyganski and Norman Wittels for their assistance and comments on this research.

Appendix A: The Possible Assingments for Each Combination

The detailed analysis of each of the hypothesis combinations for a dark region is presented below.

Type 0: The dark region is assumed to be a dark mark since none of the hypotheses are true. Thus, $C_d = 0$ and $N_t = 0$.

Type 1: The dark region is assumed to be a cast shadow since a penumbra is detected in it. Thus, $C_d = 1$ and $N_t = 1$.

Type 2: A self-shadow and a cast shadow are detected in the dark region. Since the self-shadow and the cast shadow are on two different surfaces, with different reflectivities and normals, there is an intensity discontinuity between the self-shadow and the cast shadow. Thus, the two ending points of the self-shadow and cast shadow boundary should be vertices on the outline of the dark region. If at least one ending point of the self-shadow boundary on the outline is a vertex, we assume that the *Self-shadow/cast shadow Hypothesis* is correct. Thus, $N_t = 1$ and $C_d = 2$. Otherwise, the hypothesis is incorrect and $N_t = 0$ and $C_d = 0$.

Type 3: Each dark region should have at most one object region adjacent to it because we have assumed that objects and shadows do not overlap with each other or with themselves. The ending points of the object boundary should be the vertices on the outline of the dark region, regardless of the object surface normals. The intersecting points between the two outlines are on the outline of the dark region and have discontinuous derivative. Due to the limitation of the low level process of performing edge detection and linking, there may be more than one object region adjacent to the dark region. If at least one of the ending points of the object boundaries is a vertex on the outline of the dark region, the *Object Adjacency Hypothesis* is assumed to be true. Thus, $N_t = 1$ and $C_d = 1$. The longest object boundary among the ones with at least one of their ending points being a vertex is assumed to be the correct object boundary and others are ignored. If no object boundary has a vertex at its ending points, the *Object Adjacency Hypothesis* is assumed to be false. Thus, $N_t = 0$ and $C_d = 0$.

Type 4: The penumbra boundary should be on the cast shadow boundary. If the penumbra boundary is on the self-shadow boundary, the *Penumbra Hypothesis* and the *Self-shadow/cast shadow Hypothesis* conflict. If the length of overlap between the penumbra boundary and the cast shadow boundary exceeds $c_1 L_1$, where c_1 is a positive constant less than 1 and L_1 is the smaller length of the penumbra boundary and the cast shadow boundary on the outline, we label the hypotheses consistent. Thus, $N_t = 2$ and $C_d = 2$. If the overlap is less than or equal to $c_1 L_1$, only one of the hypotheses is assumed to be true. The *Self-shadow/cast shadow Hypothesis* is assumed to be true and is checked as in Type 2. If at least one of the ending points of the self-shadow boundary on the outline is a vertex, $N_t = 1$ and $C_d = 2$. Otherwise, the *Penumbra Hypothesis* is assumed to be true and $N_t = 1$ and $C_d = 1$.

Type 5: The penumbra boundary should be on the nonobject boundary on the outline of the dark region if both the *Penumbra Hypothesis* and *Object Adjacency Hypothesis* are true. Otherwise, the two hypotheses conflict with each other. It is possible that there are multiple object boundaries on the outline for the reasons stated in Type 3. If the overlap between the penumbra boundary and the total nonobject boundaries on the outline is more than $c_1 L_1$, where L_1 is the smaller length of the penumbra boundary and the total nonobject boundaries, the hypotheses on both penumbrae and objects are consistent. Thus, $N_t = 2$ and $C_d = 1$. If the length of overlap between the penumbra boundary and the total nonobject boundaries is less than or equal to $c_1 L_1$, the longest object boundary that overlaps with the penumbra boundary less than or equal to $c_1 L_1$ is selected and is assumed to be a correct object boundary. Other object boundaries are ignored. Thus, $N_t = 2$ and $C_d = 1$. If the overlap between any object boundary and the penumbra boundary is greater than $c_1 L_1$, only one of the two hypotheses is correct. First, let us assume that the *Penumbra Hypothesis* is false. The Type 5 combination becomes the Type 3 combination. If none of the ending points of the object boundaries is a vertex on the outline of the dark region, we then assume that the *Penumbra Hypothesis* is true and the *Object Adjacency Hypothesis* is false. Then the Type 5 combination become the Type 1 combination.

Type 6: The object boundary should overlap with the self-shadow boundary. Otherwise, the *Object Adjacency Hypothesis* and the *Self-shadow/cast shadow Hypothesis* conflict with each other. If the overlap between the object boundaries and the self-shadow boundary is larger than $c_1 L_1$, the *Object Adjacency Hypothesis* and the *Self-shadow/cast shadow Hypothesis* are consistent. This results in $N_t = 2$ and $C_d = 2$. If the overlap is smaller than $c_1 L_1$, the longest object boundary that overlaps with the self-shadow boundary greater than $c_1 L_1$ is assumed to be the correct object boundary and others are ignored. If none of the object boundaries overlaps with the self-shadow boundary greater than $c_1 L_1$, the *Object Adjacency Hypothesis* on the objects is assumed to be false. Thus, the Type 6 combination becomes the Type 2 combination. If at least one of the ending points of the self-shadow boundary is a vertex on the outline of the dark region, $N_t = 1$ and $C_d = 2$. Otherwise, the *Self-shadow/cast shadow Hypothesis* is assumed to be false. The Type 6 combination becomes the Type 3 combination. If at least one of the ending points of the object boundaries is a vertex on the outline of the dark region, $N_t = 1$ and $C_d = 1$. Otherwise, both hypotheses are false and $N_t = 0$ and $C_d = 0$.

Type 7: When the *Penumbra Hypothesis*, *Object Adjacency Hypothesis*, and *Self-shadow/cast shadow Hypothesis* are all true, the consistency among them needs to be verified. If all three hypotheses are consistent, every two hypotheses should be consistent. If the overlap between the penumbra boundary and the cast shadow boundary is larger than c_1L_1 as in Type 4, let $C_1 = 1$. If the overlap between the penumbra boundary and the object boundaries is smaller than or equal to c_1L_1 as in Type 5, let $C_2 = 1$. If the overlap between the self-shadow boundary and the object boundary is larger than c_1L_1 as in Type 6, let $C_3 = 1$. It is impossible that two of C_1, C_2 and C_3 be 1 and one of them be 0. If C_1, C_2 and C_3 are all equal to 1, all three hypotheses are consistent. Thus, $N_t = 3$ and $C_d = 2$. If only C_1 is equal to 1, the Type 7 combination becomes Type 4. If only C_2 is equal to 1, the Type 7 combination becomes the Type 5. If only $C_3 = 1$, the Type 7 combination becomes Type 6. If all C_1, C_2, and C_3 are not equal to 1, we assume that no hypothesis is true. Thus, $N_t = 0$ and $C_d = 0$. The reason to assume that no hypothesis is true is that each hypothesis conflicts with the other two hypotheses.

References

[1] M. Adjouadi, Image Analysis of Shadows, Depressions, and Upright Objects in the Interpretation of Real-World Scenes, *Proc. Int. Conf. on Pattern Recognition*, 1986, pp. 834-838.

[2] L. N. Hambrick, Entry-Exit Method of Shadow Boundary Segmentation, D.Sc. Dissertation, Dep. Elec. Eng. Comput. Sci., George Washington Univ., Washington, DC, 1985.

[3] A. Huertas and R. Nevatia, Detecting Buildings in Aerial Images. *Comput. Vision, Graphics and Image Processing*, **41**, 1988, pp. 131-152.

[4] R. B. Irvin and D. M. Mckeown, Methods for Exploiting the Relationship between Building and their Shadows in Aerial Imagery, *IEEE Tran. on System, Man and Cybernetics*, **19**, 1989, pp. 1564-1575.

[5] C. Jiang, Shadow Identification, D.Sc. Dissertation, Dept. Elec. Eng., Worcester Polytechnic Institute, Worcester, MA, 1991.

[6] C. Jiang and M. O. Ward, Shadow Identification, *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1992.

[7] Y. Liow and T. Pavlidis, Use of Shadows for Extracting Buildings in Aerial Images, *Computer Vision, Graphics and Image Processing* **49**, 1990, pp. 242-277.

[8] D. G. Lowe and T. O. Binford, The Interpretation of Geometric Structure from Image Boundary, *Proc. DARPA Image Understanding Workshop*, 1981, pp. 39-41.

[9] D. Marr, Vision, *W. H. Freeman and Company, New York*, 1982, pp . 88-89.

[10] G. Medioni, Obtaining 3-D from Shadows in Aerial Images. *Proc. Conference on Computer Vision and Pattern Recog.*, 1983, pp. 73-76.

[11] M. Nagao, T. Matsuyama, and Y. Ikeda, Region Extraction and Shape Analysis in Aerial Photographs. *Computer Graphics and Image Processing*, **10**, 1979, pp. 195-223.

[12] S. A. Shafer and T. Kanade, Using Shadows in Finding Surface Orientations, *Comput. Vision, Graphics, Image Processing*, **22**, 1983, pp. 145-176.

[13] W. B. Thompson, M. T. Checky, and W. F. Kaemmerer, Shadow Stereo–Locating Object Boundaries Using Shadows, *National Conf. on Art. Intel.*, 1987, pp. 761-766.

[14] C. Wang, L. Huang, and A. Rosenfeld, Detecting Clouds and Cloud Shadows on Aerial Photographs, *Patt. Recog. Letters*, 1991, pp. 55-64.

[15] A. P. Witkin, Intensity-Based Edge Classification. *Proc. Nat. Conf. Art. Intell.*, 1982, pp. 36-41.

[16] J. R. Kender and E. M. Smith, Shape from Darkness: Deriving Surface Information from Dynamic Shadows. *Proc. Nat. Conf. Art. Intell.*, 1986, pp. 664-669.

[17] J. B. Thirion, Realistic 3D Simulation of Shapes and Shadows for Image Processing. *CVGIP: Graphical Models and Image Processing*, **54**, 1992, pp. 82-90.