

Height Determination of Extended Objects Using Shadows in SPOT Images

V.K. Shettigara and G.M. Sumerling

Abstract

Building heights were estimated with relatively high accuracy using shadows in a set of single-look SPOT panchromatic and multispectral images taken from the same satellite simultaneously. Shadows cast by rows of trees in SPOT images were first used to estimate mean heights of trees. Calibration lines were then constructed to relate the actual mean heights of rows of trees to the estimated heights. Using these calibration lines, heights of some industrial buildings in the image were estimated using their shadows with sub-pixel accuracy. The accuracy achieved is better than three metres, or one-third the pixel size of the SPOT panchromatic image. One of the important challenges involved in the process was to determine an appropriate threshold for delineating shadow zones in the images. A technique is provided for this problem.

The technique is useful for estimating heights of extended objects situated in flat terrains. The type of resampling used for overlaying a multispectral image over a panchromatic image changes the accuracy of height estimation. However, the change is tolerable if the heights to be estimated are within the ground-truth data range used for deriving calibration lines.

Introduction

In the last two decades enormous gains have been made in processing civilian satellite images. This paper presents a quantitative application of remote sensing in building height estimation in the sub-pixel range using SPOT images. The procedure has significant applications in defense for wide-area surveillance, urban monitoring, and environmental studies.

Photogrammetrists have been able to extract heights of objects from aerial photographs using parallax in stereo-pair photographs. The lengths of the shadows cast by the objects are also used to determine heights (Huertas and Nevatia, 1988; Liow and Pavlidis, 1990). If the sun and sensor geometry are known, it is fairly simple to establish a relationship between shadow lengths and the heights of objects. The above usages, however, are confined to high resolution photographs, with pixel resolutions much better than the object heights being measured.

In a related area, coarse resolution Synthetic Aperture Radar (SAR) images have been used for mapping terrain elevations using shape from shading techniques (Willey, 1986; Frankot and Chellappa, 1987; Guindon, 1989). These are extensions of shape from shade studies in optical images. In these techniques, the relationship between local shading variations and local slope variations is used to determine terrain shape. The technique assumes that the scattering coefficient is a constant for the area and also that the scattering cross section does not vary with the variation of local incidence angle.

Although shadows have been used successfully for estimating heights of objects in high resolution aerial photographs, the use of civilian satellite images, such as SPOT, for height estimation of buildings reported in the literature is limited. There have been some reports in this direction lately (Cheng and Thiel, 1995; Hartl and Cheng, 1995; Meng and Davenport, 1996). The main reason for the scarcity in applications is that the resolution of the civilian satellite images is much coarser than that of the aerial photographs and shadows are not well defined for short and commonly occurring objects. This causes problems in determining shadow widths that are needed for estimating heights. The solution to this problem is a major issue considered in this paper.

A Summary of Related Reports

The papers of Cheng and Thiel (1995), Hartl and Cheng (1995), and Meng and Davenport (1996) are closely related to the contents of this paper. Cheng and Thiel (1995) have used a SPOT panchromatic image to estimate building heights using shadows. They have achieved a root-mean-square error of 3.69 m in height estimation of 42 buildings. It should be noted that the mean of their building shadow widths is on the order of 7.7 pixels and that building heights varied from 8 to 66 m. In the present study, the average shadow width is less than 2 pixels. Hence, the challenges involved in achieving sub-pixel accuracy are greater in our study. However, we have achieved an accuracy better than that achieved by Cheng and Thiel (1995). Our estimation of height is indirect, using calibration curves, while Cheng and Thiel (1995) have computed heights directly. The number of buildings in our study is much smaller than theirs.

In this paper we have discussed the role of thresholding and have developed a procedure for selecting an optimal threshold for measuring shadow width with sub-pixel accuracy. After the selection of the threshold, the shadows are delineated by segmenting the whole image. Cheng and Thiel (1995) have used an adaptive procedure for thresholding each shadow.

Hartl and Cheng (1995) is the continuation of Cheng and Thiel (1995). Here they have reported a root-mean-square error of 6.13 m in height estimation. The study involved height measurement of 77 buildings.

Meng and Davenport (1996) have developed a procedure to fix shadow boundaries, which is crucial for height estimation, with sub-pixel accuracy. They claim that the edges can be fixed to 1/100th of a pixel. Unfortunately, they have not provided any ground-truth data for accuracy assessment. For accurately locating the shadow boundaries, they create an edge-image template using the point-spread function of the

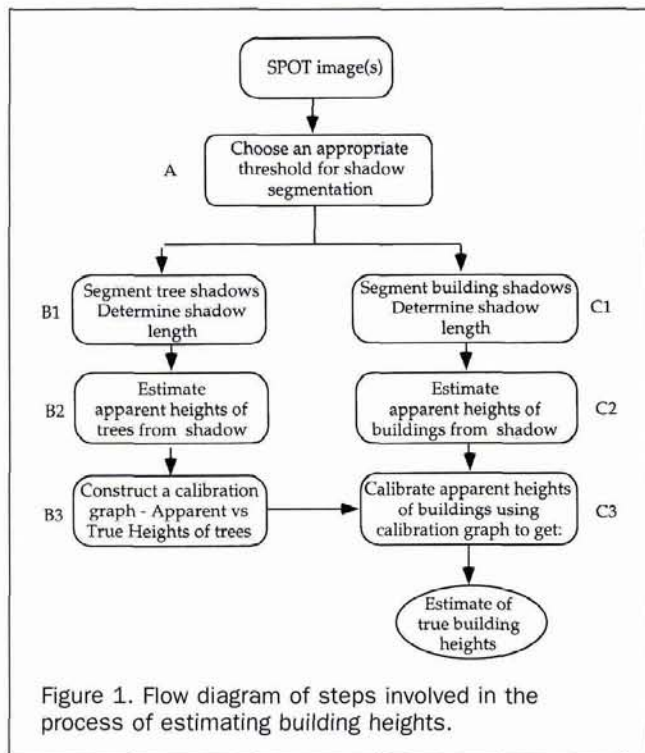
Photogrammetric Engineering & Remote Sensing,
Vol. 64, No. 1, January 1998, pp. 35-44.

0099-1112/98/6401-0035\$3.00/0

© 1998 American Society for Photogrammetry
and Remote Sensing

Defence Science and Technology Organisation, Salisbury, SA 5082, Australia.

G.M. Sumerling is presently with the Intergraph Corporation, Kew, Victoria 3101, Australia.



An indirect approach is used to estimate the heights of buildings in this study. The method has three major steps (Figure 1):

- (A) Selection of an optimum threshold for shadow segmentation.
- (B) Segmentation of shadows and estimation of apparent heights of rows of trees using sun-sensor-shadow geometry (B1, B2); construction of a calibration graph to relate apparent heights and true heights of rows of trees (B3).
- (C) Segmentation of shadows and estimation of apparent heights of buildings using sun-sensor-shadow geometry (C1, C2); estimation of true heights of buildings using the calibration graph derived in B (C3).

Shadows cast by a number of rows of trees in the area were initially used for computing a calibration line. The reason for using rows of trees, instead of buildings, is that the buildings were either not tall enough or long enough to cast consistent shadows. It will be shown later that the estimated heights are a function of the threshold used for delineating shadows, the pixel size, the azimuth of the shadow, and the actual heights of objects casting the shadow.

In principle, one should be able to estimate the heights of objects directly using sun-shadow geometry. This is not necessarily the case when dealing with images that have pixel sizes comparable to object heights. A particular issue that plays a major role in such cases is the selection of an appropriate intensity threshold for shadow delineation. This issue will be addressed in a separate section below.

Details of the Study Area and the Data

The study area (Figure 2) is a flat terrain in Australia, about 25 kilometres north of the city of Adelaide. The main area has a number of rows of trees, probably planted at different times, generally in the same direction. Trees in each row are of fairly uniform heights. The rows of trees surround open fields and factory-like buildings and the trees cast clear shadows onto the fields. On the southeastern side of the area (Figure 2) a major car manufacturing plant is located. This area has tall buildings with heights varying from 10 to 15 metres; the buildings are surrounded by asphalted parking areas or dark greasy drive ways. Only three buildings cast clear shadows that could be separated from their backgrounds.

The image set used for the present investigation consisted of a SPOT panchromatic band image with a nominal pixel resolution of 10 m and a multispectral image with a nominal resolution of 20 m, simultaneously taken from the same satellite. Both the images are standard level 1B products of SPOT, which include radiometric corrections and geometric corrections for systematic distortions caused by the Earth's rotation and panoramic effects. The images were collected on 21 April 1987. In principle, the process described in this study requires only one image that has shadows, and the objects casting shadows must be well defined. The reasons for using two images are as follows:

From the point of view of convenience, it would be better if the shadow zones could be determined from just a SPOT panchromatic band image. However, we found that, in the panchromatic band, both trees and shadows appear equally dark, and a clear separation of trees and shadows was not possible. On the other hand, in the infrared band of a multispectral image (Figure 2), trees show a significantly higher reflectance than the surroundings and the shadow zones. The infrared band could be used to determine the shadow zones, but with lesser precision due to its coarser resolution.

In order to improve the accuracy of the estimate of shadow widths, both the panchromatic and infrared bands were used together. Another reason for using the combina-

sensor. The actual edge is located at the best matching location of the template. The process is semi-automatic because the initial rough locations of the edges have to be provided manually.

Sun Shadow — An Important Part of Images

Sun shadows form a unique part of any image. They are easily detectable as they show a low intensity in all the multi-spectral bands. Shadows play a dominant role in image statistics which affect image enhancement and pattern recognition. Shadows contain three-dimensional (3D) information of objects and they have been used in object identification, terrain classification, and geological mapping (Curran, 1985).

Shadows have been used by many workers in the photogrammetric and computer vision communities for detecting buildings in high resolution aerial photographs and estimating their heights (Venkateswar and Chellappa, 1990; Irvin and McKeown, 1989; Huertas and Nevatia, 1988). All have utilized the nature of shadows for modeling buildings.

Organization of the Paper

In the next section, the objectives of the study and an outline of the methodology are provided. In the section to follow the details of the study area and the data are introduced. The basic equations for sun-shadow geometry are presented in the subsequent section. Following this, a relationship is established between estimated height, actual height, pixel size, and intensity threshold used for delineating shadows. A procedure for selecting an optimum threshold for shadow segmentation is presented in the following section. This will be followed by sections on Results and Discussion, and Conclusions. In this section, the effect of the type of resampling used for overlaying the infrared band onto the panchromatic band on accuracy is included.

Objectives and Outline of the Methodology

The objective of the study is to investigate the capabilities of single look (i.e., non-stereo) SPOT images in determining heights of buildings using shadows.

number will increase in real situations where the manual selection of ground control points in the reference image can be totally outside the four correct pixels. The number will also increase if the rotation of the input image is involved or if geometric distortions exist in the input image.

Although an RMS residual of 0.81 pixel appears large, its effect is expected to be uniform and systematic across the image because an affine transformation is used. As will be explained in detail later, least-squares regression lines are used as calibration lines to estimate the heights of objects from shadows. The systematic effects of mis-registration will be automatically taken into account in the calibration procedure. The high accuracy results obtained support this argument.

In this study, nearest-neighbor resampling (NNR) was used for the following reasons: Trinder and Smith (1979) have compared the performance of NNR with the cubic-convolution resampling (CCR) technique. They have found that, for the measurement of cartographic details, the two resampling techniques performed equally well. For visual interpretation purposes, CCR provided better images. Derenyi and Saleh (1989) have studied the effect of resampling techniques on geometric and radiometric fidelity. They found that CCR had influenced radiometric values and image statistics quite significantly and in turn affected quantitative analysis of the images. In contrast, NNR created no new radiometric values and the changes in statistics were not significant. It is also known that CCR causes edge overshoots and edge shifts in some cases (Atkinson, 1988). Because the preservation of edge characteristics is crucial in this study, it was decided to use NNR as the resampling technique.

Although there are good reasons to choose NNR, the influence of CCR on the final outcome needs to be analyzed. The details of the analysis are provided in the discussion.

Sun-Satellite Geometry and the Equations

Figure 3 shows the sun-satellite geometry as an end view (2D display of 3D geometry for simplicity) for the set of SPOT images used in this study. A row of trees, running normal to the page, is casting the sun shadow. As illustrated in the figure, the shadow part seen by the sensor is dependent upon its location with respect to the sun and the trees. The width of the sun shadow, measured along a normal to the rows of trees, depends on the azimuths of the sun, the image scan line,



Figure 2. Infrared SPOT image of the study area. The rows of trees used in the study are marked and the buildings whose heights are estimated using the tree shadows are shown at the lower right-hand corner.

tion is for shadow detection accuracy. A low value in the infrared band does not necessarily mean that it is a shadow zone. For example, a bitumen surface gives a low reading in infrared, but a relatively higher value in a panchromatic image. A combination of panchromatic and infrared bands would reduce false detection of shadow areas.

From a practical point of view, the use of a combination of bands described above is expected to be easier in the future because the SPOT-4 sensor suite, to be launched in 1998, is expected to have onboard registered 10-m and 20-m resolution bands.

In this study, the infrared band (20-m pixels) was registered to the panchromatic band (10-m pixels) using an affine transformation and nearest-neighbor resampling. The root-mean-square (RMS) residual of registration was 0.81 pixels. Because the sensor viewing geometries of the two images were similar, it was not difficult to co-register the two images with moderate accuracy. The RMS residual is reasonable in view of the theoretical limits explained below.

Consider a simple case of registering two geometrically undistorted images with pixel resolutions of 10 m and 20 m. For every pixel in the coarse resolution input image, there are exactly four pixels in the fine resolution reference image. Assume also that, for a coarse pixel, any one of the four corresponding pixels will be chosen with equal probability. If the coarse resolution image is registered to the fine resolution image using only translation (no rotation), it may be easily shown that the RMS residual will be 0.5 of a pixel. This

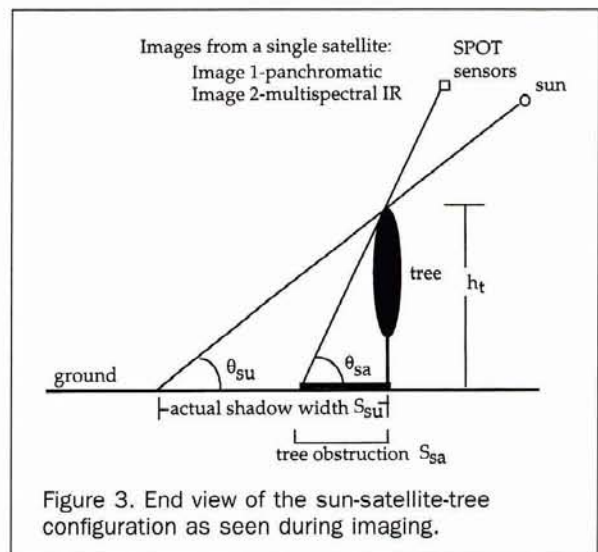
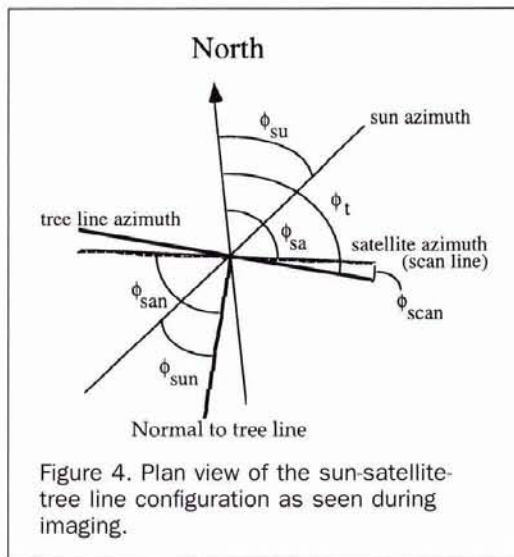


Figure 3. End view of the sun-satellite-tree configuration as seen during imaging.



and the object, as will be described below using Figure 4. From here on we will reserve the term "length of shadow" to indicate the length of shadow along the row of trees.

For relating shadow width to object (tree) height, a few assumptions were made (Figure 3). First, the object is assumed to be vertical, that is, the object is perpendicular to the Earth's surface which is flat. Second, the shadows are cast directly onto the ground. Third, it is also assumed that the shadow starts from the tree-trunk line on the ground. Finally, it is assumed that, if the sensor and the sun are on opposite sides of the object, the sensor is able to see the entire shadow.

Using Figure 3, we can write the shadow width along the sun azimuth as

$$s_{su} = h_t / \tan(\theta_{su}). \quad (1)$$

The shadow width obstructed along the azimuth of the sensor (Figure 3) by the object (tree) in the sensor's field of view is

$$s_{sa} = h_t / \tan(\theta_{sa}). \quad (2)$$

The component of the shadow of the object along the normal to the tree line (Figure 4) is

$$s_{sun} = s_{su} \cdot \cos(\phi_{sun}). \quad (3)$$

The component of the shadow obstructed by the object along the normal to the tree line (Figure 4) is

$$s_{san} = s_{sa} \cdot \cos(\phi_{san}) \quad (4)$$

where

$$\phi_{sun} = \phi_{su} + 90 - \phi_t \quad \text{and} \quad (5)$$

$$\phi_{san} = \phi_{sa} + 90 - \phi_t \quad (6)$$

and in which ϕ_{su} , ϕ_{sa} , and ϕ_t are azimuths of the sun, the image scan lines, and the tree lines, respectively, as shown in Figure 4.

The shadow width that is seen by the satellite along the normal to the tree line is given by

$$s = s_{sun} - s_{san} = h_t \{ [\cos(\theta_{sun}) / \tan(\theta_{su})] - [\cos(\phi_{san}) / \tan(\theta_{sa})] \} \quad (7)$$

Therefore,

$$h_t = s / \{ [\cos(\phi_{sun}) / \tan(\theta_{su})] - [\cos(\phi_{san}) / \tan(\theta_{sa})] \} \quad (8)$$

If the tree (shadow) line is at an angle to the image scan line, an additional correction needs to be applied to Equation 8, as explained in the Appendix. The shadow width s in Equation 8 needs to be multiplied by $\sec(\phi_{scan})$, where ϕ_{scan} is the angle between the scan lines and the tree line (Figure 4).

The terms other than s in Equation 8 may be combined into a function. The only variable in this function, for a given image, is the tree line azimuth ϕ_t . Thus, Equation 8 can be simplified to

$$h_t = s \cdot K_{\phi_t} \quad (9)$$

where

$$K_{\phi_t} = \sec(\phi_{scan}) / \{ [\cos(\phi_{sun}) / \tan(\theta_{su})] - [\cos(\phi_{san}) / \tan(\theta_{sa})] \}. \quad (10)$$

K_{ϕ_t} is a function of ϕ_t . Note that $s_{san} = 0$ if the satellite is on the opposite side of the object from the sun.

Relationship between Estimated Height, Actual Height, Pixel Size, and Intensity Threshold

The above relationships assume that the shadow zones are sharply defined and that the pixel size is infinitesimally small. However, in practice, neither the shadow zones are well defined nor are the pixel sizes negligibly small in relation to shadow widths. The following discussion presents a simple relationship between pixel intensities and shadow boundaries.

If the brightness of the shadow is I_0 and that of the background is I_b , then the brightness of a pixel, I_p , may be approximated by

$$I_p = f I_0 + (1 - f) I_b \quad (11)$$

where f is the area fraction of shadow in the pixel.

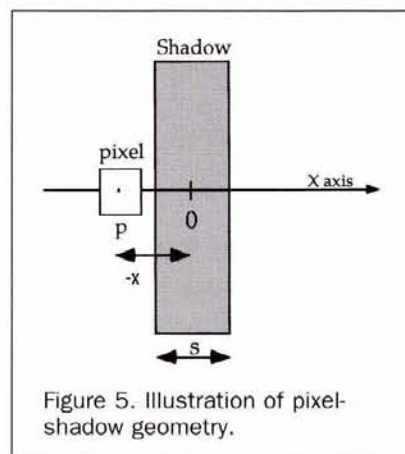
If I_0 and I_b are known from pure shadow and background pixels, then

$$f = (I_p - I_b) / (I_0 - I_b). \quad (12)$$

There will normally be a threshold f_{min} such that if $f < f_{min}$, then the pixel is classed as shadow. If not, it is classed as background. Clearly, the representation of shadow in the image depends on the threshold chosen.

The Role of Pixel Dimension

As we are considering the shadows cast by linear objects of length much greater than that of the pixel dimension, we will not lose any generality if we consider the relationships only along a profile, say the X axis, which is orthogonal to the shadow. If the shadow of width "s" (Figure 5) of uniform intensity is centered at the origin of the profile and the pixel



size "p" is smaller than the shadow width, then the intensity of pixels along the profile for various "x" is given by

$$f = \begin{cases} 0 & \text{for } x \leq \frac{-(s+p)}{2} & \text{(pixel outside object)} \\ \frac{(x+(s+p)/2)}{p} & \text{for } \frac{-(s+p)}{2} < x \leq \frac{-(s-p)}{2} & \text{(pixel partially on the object)} \\ 1 & \text{for } \frac{-(s-p)}{2} < x \leq \frac{(s-p)}{2} & \text{(pixel wholly within the object)} \\ \frac{(-x+(s+p)/2)}{p} & \text{for } \frac{(s-p)}{2} < x \leq \frac{(s+p)}{2} & \text{(pixel partially on the object)} \\ 0 & \text{for } x \geq \frac{(s+p)}{2} & \text{(pixel outside object)} \end{cases} \quad (13)$$

If the threshold is f_{\min} , then the estimated width s' of the shadow of width s will be

$$s'(p, f_{\min}) = s + p(1 - 2f_{\min}) \quad (14)$$

or, in terms of estimated tree heights h_{est} and actual tree heights h_t , by multiplying Equation 14 by K_{dt} and comparing to Equation 9,

$$h_{\text{est}} = h_t + K_{\text{dt}} p(1 - 2f_{\min}). \quad (15)$$

which is a linear relationship between h_{est} and h_t with an intercept $K_{\text{dt}} p(1 - 2f_{\min})$ and a slope of one. Note that "p" is a constant for a given image and f_{\min} is a constant for a given threshold. If f_{\min} is very small, the intercept is roughly $K_{\text{dt}} p$ and if f_{\min} is very high (near one), the intercept is roughly $-K_{\text{dt}} p$. The intercept is zero for $f_{\min} = 1/2$, which means that the boundary between the shadow and the background is halfway between their respective intensities. In that case, the estimated height is equal to the actual height.

A choice of proper threshold would ensure that the term $(1 - 2f_{\min})$ is near zero in Equation 15. This would make the effect of K_{dt} , which is a function of only the shadow azimuth for a given image, less significant. However, it is important for the investigator to assess the impact of K_{dt} on the objects being investigated. We have investigated and discussed the impact of K_{dt} for the study area in the Result and Discussion section. We have found that, for a range of shadow azimuths, K_{dt} remains fairly constant and, for that range of azimuths, the linear relationship between h_{est} and h_t in Equation 15 is valid.

Shadow Segmentation for Shadow-Width Measurement

Choosing the Threshold for Delineating Shadow Zones

In high resolution images, one might find a separate peak in the histogram corresponding to shadow areas, as described by Otsu (1979). In such cases, it is easy to separate the shadow zones. In lower resolution images, it is not common to find a separate peak for the shadow areas. The current study required a general procedure for delineating shadows falling on various types of backgrounds. The backgrounds included lush grass, barren soil, white concrete, and black asphalt. Experiments were conducted by choosing various thresholds by varying "n" in the equation

$$f_{\min} = \mu - n\sigma \quad (16)$$

where μ and σ are the mean and standard deviation of the image. Note that, as n increases, the threshold decreases. The weight n was varied from 1.6 to 0.8 in steps of 0.1. These two limits were found adequate because, at a threshold of $(\mu - 1.6\sigma)$, most of the shadow areas were mis-classified as non-shadow areas and, when a threshold of $(\mu - 0.8\sigma)$ was used, most of the non-shadow areas were classified as shadow areas.

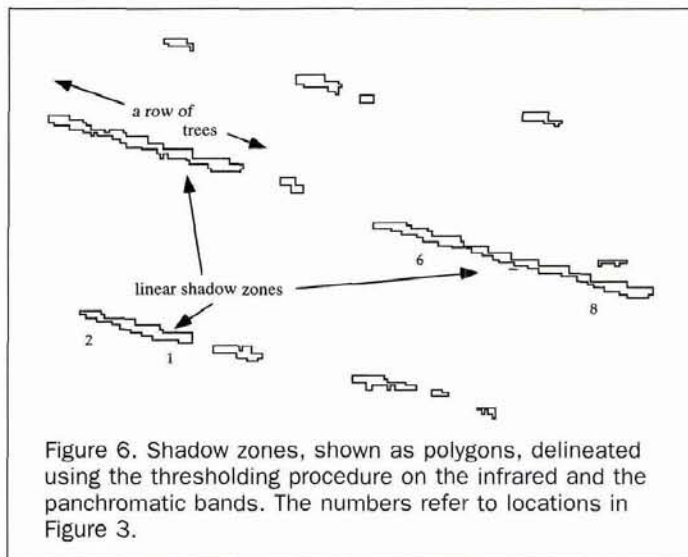


Figure 6. Shadow zones, shown as polygons, delineated using the thresholding procedure on the infrared and the panchromatic bands. The numbers refer to locations in Figure 3.

Measurement of Shadow Width

Using a single f_{\min} value at a time, shadow areas common to both the infrared band and the panchromatic band were determined. Figure 6 shows an example of the shadow segments delineated from the images. Using the segmented shadow areas, an average shadow width is measured for each row of trees. As can be seen from the figure, the edges of the shadow areas are not straight. This is because of the alignment of the rows of trees at an angle to the scan lines and is also due to the coarse pixel resolution. The jagged nature of the shadow boundaries makes the estimation of the width complicated. The average width was estimated by measuring the area covered by each shadow zone and then dividing it by the length of the zone. Shadows of length less than 50 m were ignored as they are considered not long enough. The justification for using this cut off will be given in the Discussion section. Using the shadow-width estimates, the tree heights were computed using Equation 9.

Most of the shadow zones shown in Figure 6 are correctly identifiable as being cast by the rows of trees. However, false identification of shadow areas has also occurred. The dark bitumen in the west of the image and some wet market garden areas in the southwestern part of the image (Figure 1) were mis-classified as shadows. The false shadow areas were eliminated from further computation. In this study, no automatic procedures have been developed to isolate "false shadow" zones.

Results and Discussion

Data for the Study Area

The relevant data required for the use of Equation 9 for the study area were collected from an astronomical almanac and satellite ephemeris supplied with the images. From the almanac, the sun angle (Figure 3) during imaging was $\theta_{\text{su}} = 55.5^\circ$ and the sun azimuth (Figure 4) was $\phi_{\text{su}} = 39.2^\circ$. From the satellite ephemeris, the incidence angle (Figure 3) was $\theta_{\text{sa}} = 23.4^\circ$. The scan-line azimuth was computed from the image orientation supplied with the image header. For the SPOT images used, the orientation was 12.5° . The scan-line azimuth is determined by adding 90° to the image orientation. That is, $\phi_{\text{sa}} = 102.5^\circ$. The tree-line azimuths ϕ_t are around 116° for the area. Using Equations 5 and 6 and the tree-line azimuths, the values of ϕ_{sun} and ϕ_{san} were calculated for each row of trees. The tree-line azimuths were obtained in the field by direct measurement. However, they can also be determined

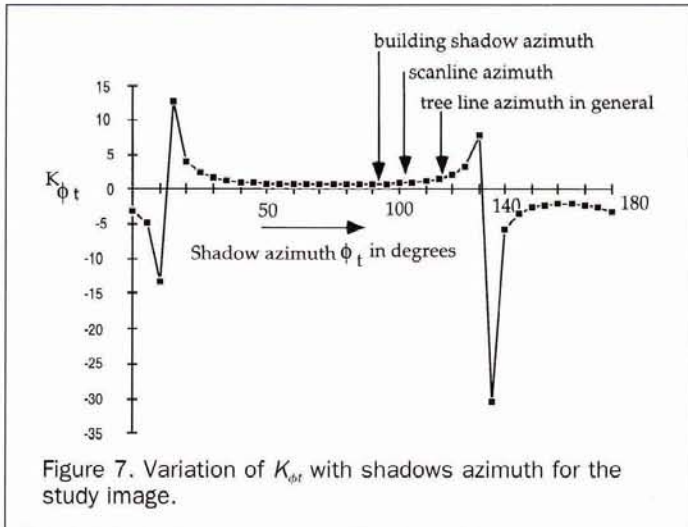


Figure 7. Variation of K_{ϕ_t} with shadows azimuth for the study image.

by registering the images to an area map and then measuring the direction. For estimating the building heights, the shadows cast on the southern side of the buildings were used (Figure 2). The azimuth of shadows of the buildings marked in Figure 2 is 93° .

Analysis of the Impact of Shadow or Tree Azimuth ϕ_t in the Area

Using the above parameters and Equation 10, K_{ϕ_t} values were computed for ϕ_t ranging from 0 to 180° in steps of 5° . Figure 7 shows the variation of K_{ϕ_t} values for the study image. The graph will be the same for ϕ_t values ranging from 180° to 360° . On the figure, the azimuths of shadows of buildings and tree lines are shown. As can be seen, the values of K_{ϕ_t} are low and fairly constant for the range of ϕ_t values encountered in the area. This assures that the linear relationship assumed between the estimated heights and the true heights in Equation 15 is valid. The graph also reveals that shadows in the range of azimuths 0° to 25° and 120° to 145° should be avoided for height estimation if the impact of K_{ϕ_t} is to be minimized.

Estimated tree heights were computed using the above parameters and Equation 9 for shadow widths determined for each row of trees using a threshold level. Figure 8 shows the regression lines of estimated heights, computed from shadow widths determined by using different thresholds, and actual tree heights. For the sake of clarity, regression lines and data points for only two thresholds are plotted.

Notice that the correlation between the observed heights and the estimated heights deteriorate drastically when the threshold is increased from $(\mu - 1.2\sigma)$ to $(\mu - 1.1\sigma)$. The variation of height estimates is shown for all the sites with these two thresholds. The height estimates increase systematically, as expected, for all the sites except Site 5. The increase in height estimate for Site 5 is disproportionately high. This is due to the sudden inclusion of the background areas, caused by the low contrast between the shadow and the background, in the shadow estimates below a certain threshold. This site dramatically demonstrates the need to have an optimum threshold. Notice that the regression lines for any of the cases involving thresholds less than $(\mu - 1.1\sigma)$ do not pass through the origin as, ideally, we would have liked them to. To achieve an intercept of zero, the threshold has to be varied between $(\mu - 0.9\sigma)$ and $(\mu - 0.8\sigma)$. However, as we increase the threshold above $(\mu - 1.2\sigma)$, background areas are increasingly classed as shadow areas. Thus, the increase of threshold beyond $(\mu - 1.2\sigma)$ is not justifiable.

The coefficient of determination (r^2) for thresholds

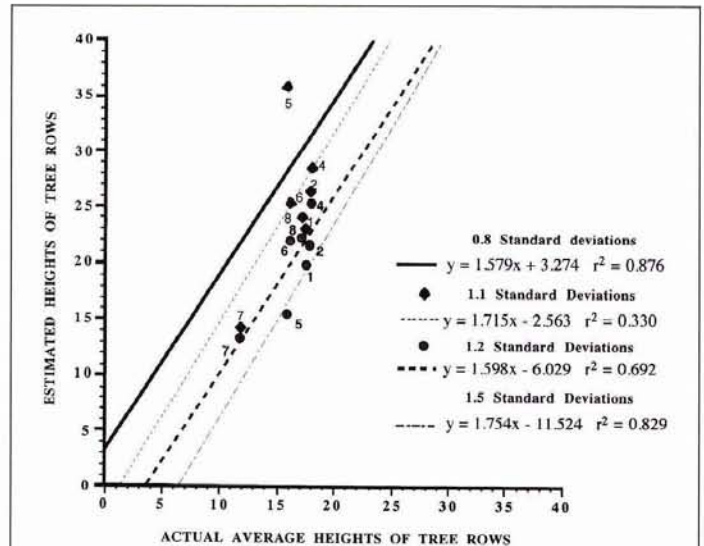


Figure 8. Regression lines for different thresholds. Notice that the estimated heights vary systematically for all the sites except for Site 5. The drastic variation for Site 5 is attributed to the inclusion of background areas when the threshold is increased from $(\mu - 1.2\sigma)$ to $(\mu - 1.1\sigma)$.

smaller than $(\mu - 1.2\sigma)$ is high (Figure 8). This indicates that the shadow estimates are very well correlated (correlation coefficient 0.832 for threshold $\mu - 1.2\sigma$) to the measured tree heights. One point worth considering is whether the number of points used for the regression is large enough. The probability of getting a correlation of above 0.8 when the two variables are uncorrelated, and the number of measurements involved is more than six, is less than 6 percent (Taylor, 1982). This suggests that the results of regression are reliable.

Figure 9 shows how estimated heights vary with various thresholds. Each line in the plot corresponds to a row of trees and a location. The plot reveals two interesting characteristics: (a) As the threshold increases, the height estimate increases. The rate of height variation is uneven in the range of thresholds used. Nearly all the estimates vary sharply between thresholds $(\mu - 1.1\sigma)$ and $(\mu - 1.2\sigma)$ compared to other threshold intervals. This point is vividly illustrated by

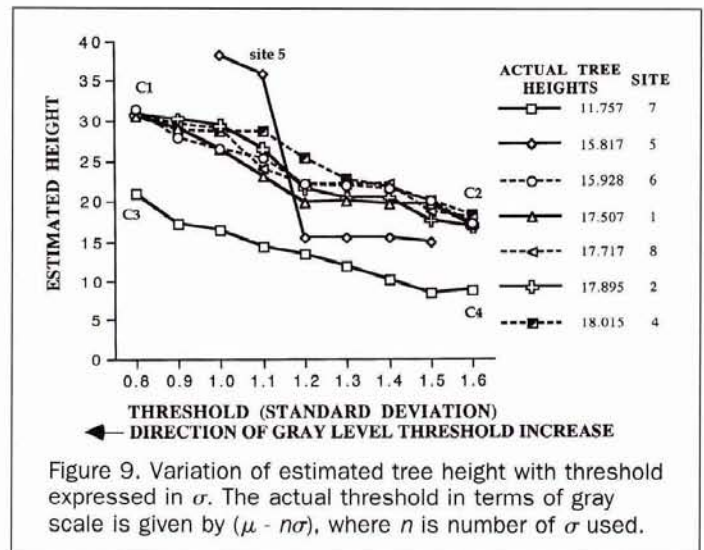


Figure 9. Variation of estimated tree height with threshold expressed in σ . The actual threshold in terms of gray scale is given by $(\mu - n\sigma)$, where n is number of σ used.

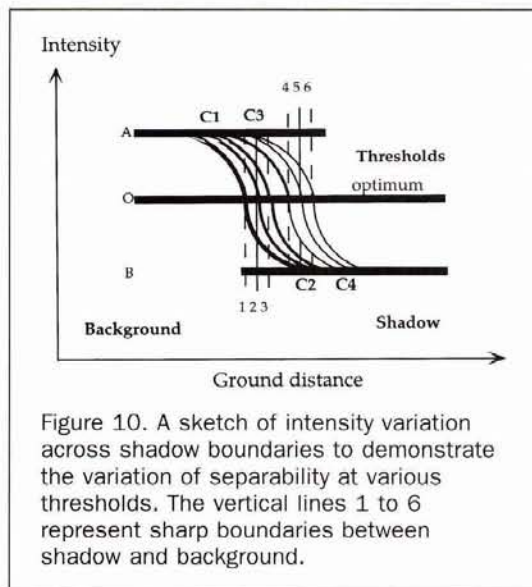


Figure 10. A sketch of intensity variation across shadow boundaries to demonstrate the variation of separability at various thresholds. The vertical lines 1 to 6 represent sharp boundaries between shadow and background.

the estimate for Site 5. (b) It can be seen in the figure that the height estimates converge to common values at the two extreme thresholds applied. Sites 7 and 5 are two exceptions. The case of Site 5 is already explained above. In the case of Site 7, the whole line is shifted in parallel, along the height axis, by roughly the pixel dimension of 10 m. It is important to note that the tree height for Site 7 is significantly different from the other sites. This site will be further discussed below.

The convergence characteristics mentioned above can be explained with the help of Figures 10 and 11. In Figure 10, intensity curves are sketched for shadow-background boundaries at different locations along the ground. The six curves represent observed intensity variation across six boundary locations marked by the dashed vertical lines. The curves mimic the smoothing of sharp boundaries due to the point-spread functions of the sensors. As can be seen at A and B, on the intensity axis, the intensity curves tend to overlap and at O they show maximum separation. Considering the inherent presence of both random and systematic noise in the images, it may be expected that the closely spaced intensity curves, marked by C1-C2 in Figure 10, can be inseparable at thresholds A and B. The effect of noise in the data in resolving the individual curves at A and B is represented by the thick horizontal lines. This explains the asymptotic convergence of estimated tree heights at two extreme thresholds in Figure 9.

The set of curves 1 to 3, shown in bold in Figure 10, form a set with limited shadow width (height) variations and they may be compared to the set of curves C1-C2 in Figure 9 representing limited variation in shadow width. Similarly, the curves 4 to 6 in Figure 10, which may be compared to the curve C3-C4 in Figure 9, representing a significantly different shadow width compared to the set C1-C2.

In figure 11, the lines E1-E2 and E3-E4 represent the two edges of a shadow zone. Within the shadow zone some pixels are fully covered by the shadow, which is indicated by dark gray color. Along the boundary the pixels are partially covered by the shadow and those pixels are indicated by the dotted pattern. These pixels account for the variation with threshold of shadow widths and the estimated heights in Figure 9. As the threshold decreases, fewer pixels will be included in the shadow segment. Ultimately, the segment will converge to the dark gray areas, causing the convergence of height estimates at C2 in Figure 9. The convergence at C2 in

Figure 9 can be explained through the family of curves at C2 in Figure 10. Conversely, as the threshold increases, more partially covered areas will be included in the shadow segment. If the threshold is increased beyond a certain level, background will be included in the shadow segment, causing a sharp increase in the shadow-width estimate or height estimate. After this stage, the values converge to a point like C1 in Figure 9, for various heights within a range. As before, the convergence of curves in C1 can be explained through the family of intensity curves at C1 in Figure 10.

The convergence of values occurs only for shadow widths within a range. For example, if the lower shadow boundary is E5-E6 instead of E3-E4, then the set of dark gray pixels, which are the pixels fully covered by the shadow, will be significantly smaller. Probably, the Site 7 curve in Figure 9 corresponds to another family of curves which converge at points C3-C4 in Figures 9 and 10. We propose that the shift from C1 to C3 and C2 to C4 is on the order of one pixel. We cannot prove this with examples because we do not have a family of rows of trees with heights in the range of 11 m. However, we can argue the case thus: Assume that the lower boundary of the shadow zone has moved roughly one pixel width to E5-E6. This line corresponds to the mean position of the boundary for a family of intensity curves shown by C3-C4 in Figure 10. This would cause the reduction of the fully covered pixels (dark gray pixels) by blocks of ΔX and ΔY throughout the length of the shadow as shown in Figure 11. Consequently, the reduction in width is

$$\Delta W = \Delta X \cdot \Delta Y / \Delta R = \tan^{-1}(\Phi_{scan}) \Delta Y^2 / \Delta R \quad (17)$$

$$= \tan^{-1}(\Phi_{scan}) / \Delta R \quad \text{for } \Delta Y = 1 \text{ pixel}$$

where Φ_{scan} is the angle between the tree line or shadow line and the scan line. In the study area for most of the tree lines, Φ_{scan} was 15°. This provides for the shift in convergence points on the order of 0.98 pixels, which is 9.8 metres. This, in turn, will reduce the height estimation by 7.7 metres as per Equation 9. The shifts in convergence points C1 to C3 and C2 to C4 in Figure 9 are 9.5 and 8 metres, respectively. This closely agrees with the theoretical estimate of 7.7 metres.

The above argument can be extended to model the difference in height estimates between C1 and C2 or C3 and C4 in Figure 9. Here we have to consider the effects of thresholding on both sides of the shadow zone. Thus, the differ-

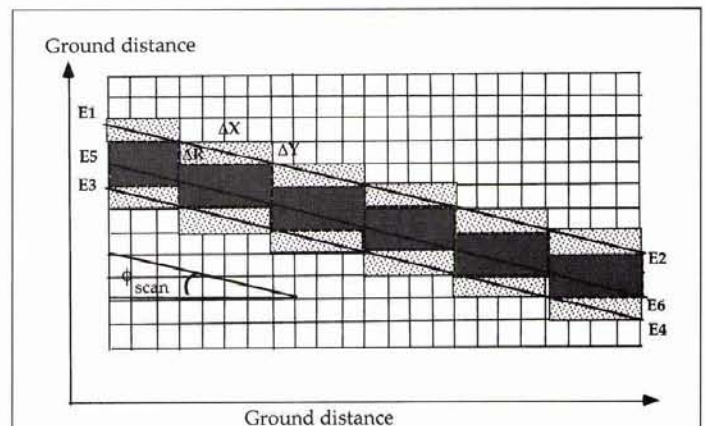
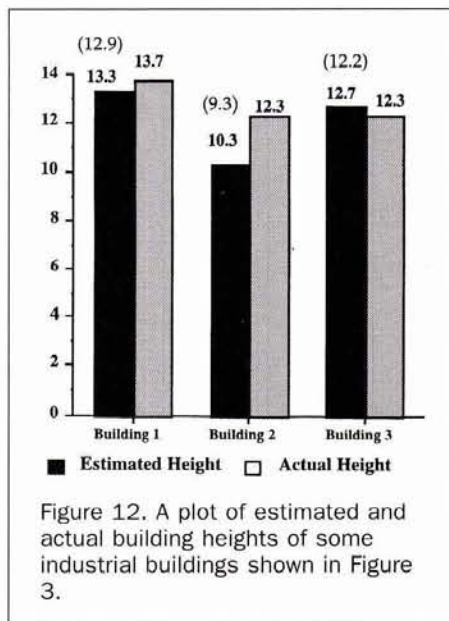


Figure 11. Effect of thresholding on shadow width estimation. E1-E2 and E3-E4 are the actual boundaries of the shadow. The pixels in dark gray are fully covered by the shadow. The dotted pixels are partially covered by shadow.



ence should be roughly twice 7.7 metres, or 15.4 metres. This estimate again closely agrees with the observed difference between C1 and C2, and C3 and C4, on the order of 13 metres in Figure 9.

Minimum Shadow-Length Requirement

Figure 11 also assists us in deciding the minimum length of the objects casting shadows needed to derive a reliable calibration line. The pattern of pixels fully and partially covered by shadows is cyclical in nature. For one cycle, the length of the object is ΔR for $\Delta Y = 1$ pixel. We have discussed above that the pixels that are partially covered with shadows are responsible for the variation in height estimates while using different thresholds. In order to even out this effect across all tree-line shadows involved, it was thought the tree lines should be at least one cycle in length. Assuming that $\Phi_{\text{scan}} = 15^\circ$, in order to achieve a $\Delta Y = 1$ in Figure 11, the length ΔR should be approximately four pixels. In order to accommodate all tree lines, it was decided that the minimum shadow length should be 5 pixels or 50 m.

Building Height Estimation

From Figure 9, it appears that the height estimates are best resolved in the range of $(\mu - 1.1\sigma)$ and $(\mu - 1.2\sigma)$. This suggests that the optimum threshold for estimating shadow widths exists between $(\mu - 1.1\sigma)$ and $(\mu - 1.2\sigma)$. The curve

TABLE 1. EFFECT OF RESAMPLING TECHNIQUE

Shadow Number (Figure 2)	Number of Pixels in Segmented Shadows		Percent Difference in Width and Height Estimate between NNR and CCR
	Nearest-Neighbour Resampling (NNR)	Cubic-Convolution Resampling (CCR)	
1	73	70	-4.1
2	23	21	-8.7
4	54	55	1.9
5	85	88	3.5
6	116	113	-2.6
7	30	32	6.7
8	105	99	-5.7

corresponding to Site 5 clearly demonstrates that the threshold of $(\mu - 1.1\sigma)$ tends to incorporate background as shadow, causing a sudden increase in the estimated height. For this reason, a threshold of $(\mu - 1.2\sigma)$ was assumed to be the appropriate threshold. The regression line corresponding to $(\mu - 1.2\sigma)$ was taken as the optimum calibration line (Figure 8) to relate the estimated height from Equation 9 to the actual height. Using the widths of shadows cast by the industrial buildings in the southeastern part of the image (Figure 2), the building heights were computed using Equation 9. The true heights were then estimated using the calibration line of Figure 8. The results are shown in Figure 12. As can be seen, the estimated heights closely follow the measured heights. A high sub-pixel accuracy, better than one-fifth of the pixel size, was achieved in this experiment.

Effect of Resampling Technique

In the section on the "Details of the Study Area and the Data," it was explained why the nearest-neighbor resampling technique was used for registering the infrared band of the multispectral image to the panchromatic band image. In order to ascertain the effect of the other popular resampling technique, cubic convolution, an experiment was conducted. The infrared band was registered to the panchromatic band using the same set of ground control points that were used for the nearest-neighbor resampling. Assuming that the optimum threshold is $(\mu - 1.2\sigma)$, for reasons given earlier, the shadow widths are again determined. Table 1 summarizes the percentage differences in shadow width estimation using the two resampling procedures. As can be seen, in most cases the difference is not significant.

Although the errors in width estimates of individual shadows is not significant, the overall effect on calibration lines can be considerable. For this reason, the calibration lines relating estimated heights to actual heights, for the two resampling techniques, are shown in Figure 13. The calibration lines are similar to the lines shown in Figure 8, but for a threshold of only $(\mu - 1.2\sigma)$. As can be seen, the difference

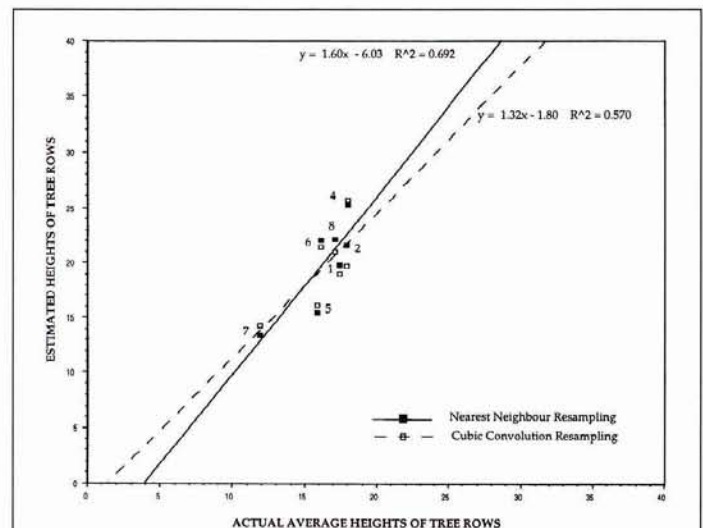


Figure 13. Regression lines relating estimated tree heights to actual tree heights for a threshold of $(\mu - 1.2\sigma)$. The solid line is for nearest-neighbor resampling used for overlaying the infrared band onto the panchromatic band (also shown in Figure 8) and the dashed line is for cubic-convolution resampling. The figure shows the effect of the two resampling procedures on the accuracy of height estimation.

in estimates of actual building heights (X values) from the two lines for an apparent height estimated from shadow widths (Y values) is not significant within the calibration (ground truth) data ranges used. However, the errors can be significant, as can be expected, outside the ground truth data ranges.

This point is substantiated in Figure 12. The building height estimates derived using the cubic-convolution resampled image are provided in brackets. These figures are comparable to the original figures obtained from nearest-neighbor resampled images.

Some Limitations of the Technique

The technique is useful for estimating heights of only extended objects. The objects should be at least five pixels long. The technique cannot be used if the object does not cast a shadow, for example, when the extended side of the object is parallel to the principal plane of the sun.

The shadows should not be interrupted by other objects. In this study, Site 3 (Figure 2) was not used, because it was found early in the investigation that the shadow width delineated by the process was exceptionally large compared to the height of the trees. At that site, two rows of trees were found. The probable reason for such an anomalous shadow width is that one row of trees is casting shadow on the other row of trees and making the second row of trees and its shadow appear as a large shadow. Due to the coarse resolution of the SPOT image, it is often difficult to ascertain the above situation. However, it is important to eliminate such sites from the regression analysis.

The mathematical relationships presented in this study assume that the terrain is flat and that objects are standing normal to the surface. The procedure may not be useful for undulating terrain.

The type of resampling technique used for overlaying the infrared band over the panchromatic band has some effect on the accuracy of height estimation. The effect appears to be particularly significant for ranges where the calibration lines are extrapolated.

It must also be realized that the trees used in this project were of a limited height range. A larger range of heights would have been more suitable for such a study. At this stage, it is not clear what would be the effect of extrapolating the calibration lines (Figure 8) beyond the ground-truthed heights.

The procedure described assumes that in Equation 15, which relates estimated heights to true heights, the effect of azimuths of shadows is constant. As described with the help of Figure 7, this is true only for certain ranges of azimuths. For each image, such ranges have to be determined before using the technique.

As illustrated in the Appendix, the shadows will be discontinuous if they are less than a pixel wide. As the procedure requires long shadows, of at least five pixels in length, it appears reasonable to assume that the objects should cast at least a one-pixel-wide shadow in order to measure their heights. Equations 11 to 15 also make the assumption that the shadow width is larger than the pixel width.

In principle, a single panchromatic image will suffice for estimating heights of objects, provided that the object casting the shadow and the shadow are clearly separable. In the present study, that was not the case and, for that reason, an infrared image of coarser resolution was overlaid on the panchromatic image.

Conclusion

Building heights can be estimated with sub-pixel accuracy using shadows in a set of SPOT panchromatic and infrared images taken from a single satellite. In the current study,

shadows cast by rows of trees were used to estimate building heights. The accuracy achieved is better than one-third of the panchromatic image pixel size. The key to the success of the technique was in determining an optimal threshold for delineating shadow zones. A procedure has been provided in this study for determining an optimal threshold.

The procedure described has the following limitations. The technique is useful for estimating heights of only extended objects situated in flat terrain. The type of resampling used for overlaying multispectral imagery over panchromatic images changes the accuracy of height estimation. However, the difference in accuracy, when cubic-convolution resampling is used instead of nearest-neighbor resampling, is not significant within the ground-truth data range used for deriving the calibration lines.

Acknowledgments

Dr. David Jupp of CSIRO and Mr. Geore Poropat of Electro Optics Systems, Australia, have provided useful insight into this investigation. Their assistance is gratefully acknowledged.

Appendix

Pixelization Effect on Shadow-Width Measurement

In analytical terms, the determination of width of a linear object can be achieved in many ways. One simple approach is to divide the area covered by the object by its length. This approach cannot be directly implemented in image processing for computing shadow widths.

In Figure A1, shadows are drawn as the area between pairs of lines. Widths are indicated opposite the pairs of lines. The surface is divided into 10- by 10-m grids to simulate SPOT image pixels. The black pixels are those in which 50 percent or more of the pixel area is covered by shadow. It

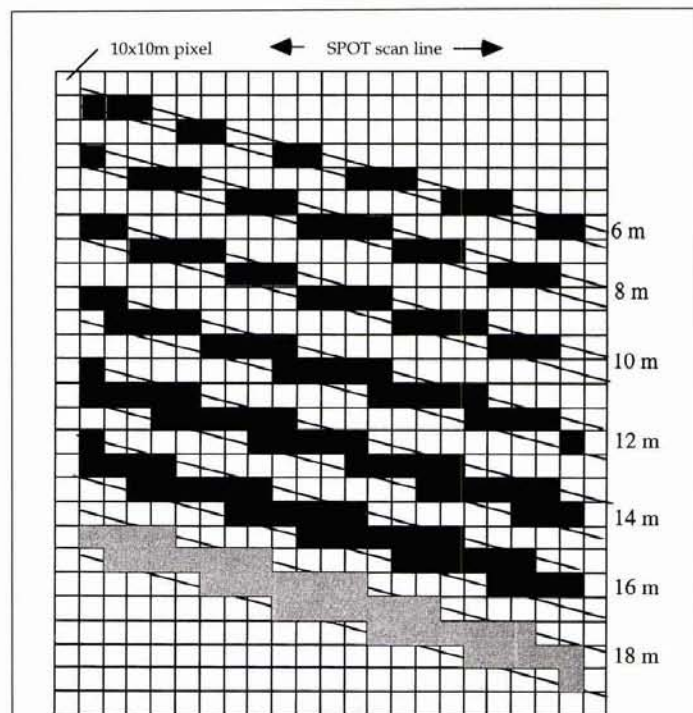


Figure A1. A sketch to show the pixelization effect on shadow detection and shadow-width measurement. If the shadow width is smaller than the pixel dimension, fragmentation of the shadow zone may be expected.

is assumed that a perfect technique is available to threshold such pixels and label them as shadows.

Due to the discrete pixel sizes involved in images, the shadow zones may appear as continuous areas or as fragments. Under perfect thresholding conditions, as described above, one can expect continuous shadow zones only if the shadows have widths equal to or greater than the pixel dimensions.

From the figure, it is also clear that the shadow width computed by dividing the area of shadow pixels by the length needs to be corrected. For example, the area covered by the 10-m-wide shadow is 20 pixels or 2000 m². The length of the shadow is 208.8 m. The computed width of the shadow is 9.58 m, which is less than 10 m. In order to get the correct width, the estimate must be multiplied by the secant of the angle between the shadow zone and the scan line.

References

- Atkinson, P., 1988. The Interrelationship between Resampling Method and Information Extraction Technique, *Proc. IGARSS'88*, Edinburgh, pp. 521-527.
- Cheng, F. and K.H. Thiel, 1995. Delimiting the Building Heights in a City from the Shadow in a Panchromatic SPOT-Image - Part 1 - Test of Forty-Two Buildings, *Int. J. Remote Sensing*, 16:409-415.
- Curran, P.J., 1985. *Principles of Remote Sensing*, John Wiley & Sons Inc., New York, 282 p.
- Derenyi, E.E., and R.K. Saleh, 1989. Effect of Resampling on the Geometric and Radiometric Fidelity of Digital Images, *Proc. IGARSS'89*, Vancouver, pp. 620-623.
- Hartl, Ph., and F. Cheng, 1995. Delimiting the Building Heights in a City from the Shadow in a Panchromatic SPOT-Image - Part 1 - Test of a Complete City, *Int. J. Remote Sensing*, 16:2829-2842.
- Huertas, A., and R. Nevatia, 1988. Detecting Buildings in Aerial Images, *Computer Vision, Graphics and Image Processing*, 41:131-152.
- Irvin, R.B., and D.M. Mckeown, 1989. Methods for Exploiting the Relationship between Buildings and their Shadows in Aerial Imagery, *IEEE Trans. on Systems, Man and Cybernetics*, 19:1564-1575.
- Liow, Y.T., and T. Pavlidis, 1990. Use of Shadows for Extracting Buildings in Aerial Images, *Comput. Vision Graphics Image Process*, 49:242-277.
- Meng, J., and M. Davenport, 1996. Building Assessment with Sub-Pixel Accuracy Using Satellite Imagery, *Proceedings SPIE*, 2656: 65-76.
- Otsu, N., 1979. A Threshold Selection Technique from Grey-Level Histograms, *IEEE Trans. on Systems, Man and Cybernetics*, 23: 258-272.
- Taylor, J.R., 1982. *An Introduction to Error Analysis*, University Science Books, Oxford University Press, Mill Valley, 270 p.
- Trinder, J.C., and C.J.H. Smith, 1979. Rectification of Landsat Digital Data, *Aust. J. Geod. Photo. Surv.*, 30:15-30.
- Venkateswar, V., and R. Challappa, 1990. A Framework for Interpretation of Aerial Images, *Proceedings of the 10th International Conference on Pattern Recognition*, 1:204-206.

(Received 24 June 1996; accepted 10 March 1997)



**Fifth International Conference
Remote Sensing for Marine and Coastal Environments**

5-7 October 1998



Call for Papers

You are invited to attend the Fifth International Conference on Remote Sensing for Marine and Coastal Environments in San Diego, California, *America's Finest City*. This unique conference focuses on applying remote sensing technologies to solve real-world problems in marine and coastal environments.

Paper Submission

Interested contributors should submit a 250 word summary by **27 February 1998**. Include conference topic addressed. State whether general or interactive session presentation is preferred. Accepted summaries received electronically via the ERIM website or can be accessed on the World Wide Web. Do not submit to more than one address.

Electronic submissions:

E-mail: marine@erim-int.com

web: <http://www.erim-int.com/CONF/marine/MARINE.html>

Written and faxed summaries:

ERIM/Marine Conference

P.O. Box 134008

Ann Arbor, MI 48113-4008, USA

Telephone: 1-313-994-1200, ext. 3234

Fax: 1-313-994-5123

Please provide your complete mail/delivery address and facsimile number on all correspondence

Conference Topics

*Natural Resource
Management*

Coastal Hazards

*Oceanographic
Applications/Processes*

*Economic Benefits of
Remote Sensing: Case
Studies*

*New Data Sources,
Sensors, and
Measurement
Techniques*

*Data Access,
Integration, and
Dissemination*

97-20745 R2