

True Orthophoto Generation of Built-Up Areas Using Multi-View Images

Jiann-Yeou Rau, Nai-Yu Chen, and Liang-Chien Chen

Abstract

Hidden areas and shadow effects are major causes of information loss in large-scale aerial photos of built-up areas. Both types of loss severely degrade the interpretability of orthophotos. An abrupt change of surface height is the primary cause of these defects. Thus, surface discontinuities, the orientation parameters of the sensors, and solar orientation are all key factors in determining the extent of defects. We thus propose an ortho-rectification scheme, which will compensate for hidden areas and shadow effects in built-up areas, by using multi-view images. The proposed scheme utilizes projection geometry to detect hidden and shadowed areas. For hidden areas, lost information is recovered from corresponding images. A seamless mosaic technique, utilizing gray-value balance, is suggested to reduce gray value discontinuity. For shadowed areas, dimmed features are enhanced using the local histogram matching method to improve image interpretability. Experimental results indicate that the proposed scheme can significantly reduce hidden and shadow defects. Both radiometric and geometric aspects of the proposed product and process are investigated.

Introduction

The generation of orthophotos from remotely sensed images is important and has applications in many fields, such as cartography, environmental monitoring, city planning, etc. In addition, with the maturity of geographic information systems (GIS) technology, the need for timely spatial information is becoming greater (Joshua, 2001). Thus, ortho-rectified images have become important due to their low cost and short production time.

A number of investigations on the generation of orthophotos from digitized aerial photos (Wiesel, 1985; Hohle, 1996), airborne multispectral scanner images (Chen *et al.*, 1997), and satellite images (O'Neill and Dowman, 1988; Chen and Lee, 1993) have been reported. The image rectification procedure includes (1) orientation modeling, (2) image-object reprojection, and (3) image resampling. Ground control points (GCPs) are needed during the modeling process for determining orientation parameters, while digital terrain models (DTM) are required in the image-object reprojection to compensate for relief displacements.

Orientation Modeling

Remote sensing imagers may be grouped into three categories in terms of imaging geometry. The first group is the 2D array imager, of which aerial photos are the major example. Each scene is associated with a set of six orientation parameters,

which may be modeled by photo triangulation or space resection. The second group utilizes the pushbroom imager. Many medium- to high-resolution satellites, including SPOT, IKONOS, and EROS, are of this kind. Each CCD line is associated with a set of six parameters, which may be modeled as functions of time. Chen and Chang (1998) reported that SPOT images have a good relative accuracy; thus, only a small number of GCPs might be needed to refine the ephemeris data. A similar approach could be applied to high-resolution satellite images, for instance IKONOS. The third group utilizes imagers with mechanical scanning devices. Each image pixel is associated with a set of orientation parameters. The Landsat series satellites are of this kind. Airborne multispectral scanners (AMSS) have a similar imaging geometry, but with a much larger field of view. The steadiness of the platform is also much lower. Thus, a large number of GCPs are needed (Chen *et al.*, 1997).

There is a growing trend to integrate GPS and inertial navigation system (INS) data with remotely sensed images to directly model orientation parameters. Accordingly, the number of GCPs required may be reduced.

DTM vs. DSM

The major geometric computation in image ortho-rectification is the correspondence between an image point and its counterpart on the ground. This computation might be done through a bottom-up or a top-down approach. For aerial photos and pushbroom images, the bottom-up approach is favored (Chen and Rau, 1993; Hohle, 1996). On the other hand, for mechanical scanning images, e.g., AMSS top-down reprojection is more appropriate (Chen *et al.*, 1997). No matter which approach is taken, a DTM is required to correct for relief displacements. For those images with medium or low resolutions, for example 10 m or coarser, the value for each image pixel represents an averaged reflectance over a relatively large area on the ground. Thus, the features of buildings are smoothed on an image plane. During image ortho-rectification, it might be assumed that the surface on the ground is continuous. Accordingly, a DTM that does not delineate buildings could be used without the introduction of significant model errors.

On the other hand, aerial photos and high-resolution satellite images can provide a great deal of ground information, including man-made structures. The surface discontinuity caused by buildings should thus be taken into account during the ortho-rectification procedure. Under such circumstances, a digital surface model (DSM), delineating the canopy of the object surface, is preferable. Such a model not only represents the terrain relief but also the things on the ground. Among these

J.-Y. Rau and L.-C. Chen are with the Center for Space and Remote Sensing Research, National Central University, Chung-Li, Taiwan, R.O.C. (jyrau; lcchen@csrr.ncu.edu.tw).

N.-Y. Chen is with the National Space Program Office, Taiwan, R.O.C. (nychen@nspo.gov.tw).

Photogrammetric Engineering & Remote Sensing
Vol. 68, No. 6, June 2002, pp. 581-588.

0099-1112/02/6806-581\$3.00/0

© 2002 American Society for Photogrammetry
and Remote Sensing

things, buildings are the most important, so they should be considered first.

Related Works

The generation of true orthophotos is more complicated than that of traditional ones. In order to restore as much information as possible, both geometric and radiometric aspects should be considered. Amhar *et al.*, (1998) and Schickler and Thorpe (1998) considered the hidden effect introduced by abrupt changes of surface height (e.g., buildings and bridges). Schickler and Thorpe (1998) further consider seamless mosaicking around the fill-in areas in order to reduce gray-value discontinuity. In the radiometric aspect, shadow enhancement to restore information within a building's shadow areas has not been reported on.

Image Occlusion

Image occlusion due to buildings not only influences image quality but also the accuracy of orthophotos (Skarkatos, 1999; Joshua, 2001). To consider occlusions caused by buildings in aerial photos, Amhar *et al.* (1998) proposed using the Z-buffer technique to detect the hidden areas and an index map to differentiate roofs and terrain. Both the Z-buffer and the index map are defined on the image plane. In the ortho-rectification process, terrain relief and buildings were treated separately. The two results were then merged using a simple logical operation to obtain an intermediate true orthophoto. For each raw aerial image, one intermediate true orthophoto is generated. In Schickler and Thorpe (1998), they also created an intermediate digital orthophoto image (DOI) for each of the raw aerial images. Those hidden areas in each of the intermediate DOIs are marked. A combining procedure is then performed for merging those DOIs. In order to alleviate the computation of generating each intermediate DOI and perform data merging at the same time, we will propose an index map on the ground plane instead of the image plane.

Data Merging

In order to obtain completeness, Amhar *et al.* (1998) also use a logical operation to merge all intermediate true orthophotos to get the final product. Some problems still remained, such as artifacts along the fill-in boundaries. Schickler and Thorpe (1998) proposed an innovative method to "pick the best" pixel from all intermediate true orthophotos. Three purposes were considered in the approach: (1) to minimize the perspective distortion, (2) to avoid the smearing problems coming from severe over-sampling, and (3) to alleviate the "disharmonious" effect near the blind spot areas. The seam line is optimized by establish a path with the highest correlation, and a feathering technique is applied to reduce the gray-value discontinuity along the seam line. It is observed that the method could introduce many small patches together with many seam lines.

Shadow Enhancement

A minimization of the shadow effect is considered in our approach. This treatment enhances the image radiometry using published methods. Due to the multiple scattering effects of illumination during image acquisition, shadow areas are illuminated weakly. Thus, a suitable enhancement technique could restore the information within building shadow areas.

The Proposed Scheme

Figure 1 shows a flowchart of the proposed scheme, which includes (1) image-object reprojection, (2) hidden areas detection, (3) seamless mosaicking, and (4) shadow treatment, provided that the orientation parameters and the DSM are available. Selecting one of the multi-view images as the master

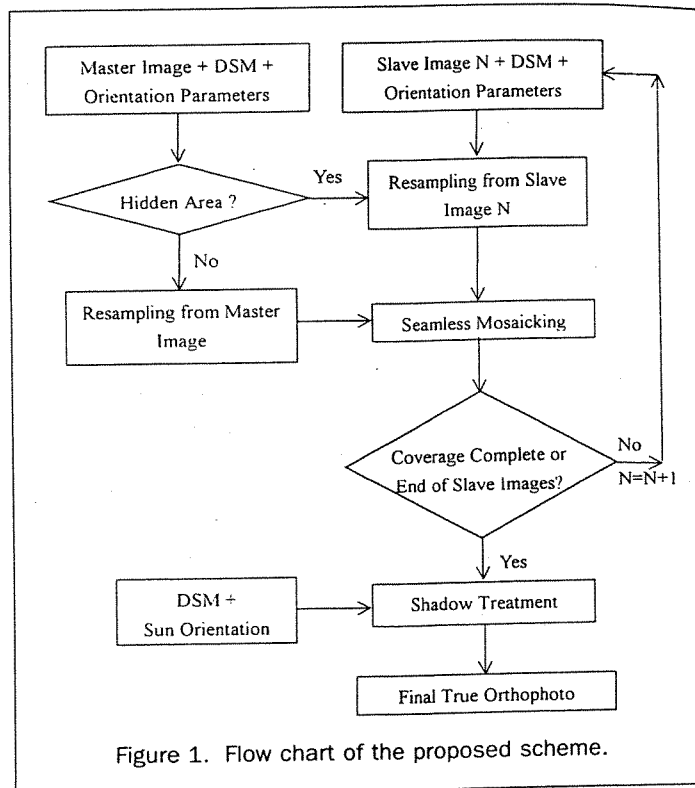


Figure 1. Flow chart of the proposed scheme.

image, the remaining images are treated as slave images. Starting from the master image, the DSM and the orientation parameters are used to calculate the corresponding image location of a ground element (groundel), by using the collinearity condition equations. Hidden areas will be detected and then compensated for by their counterparts in the slave images. The visibility of the groundel in the slave images will be checked, and then a suitable combination among the slave images will be selected. The image pieces will then be stitched together through a seamless mosaicking procedure. Finally, a shadow treatment is included to finish the process. Referring to Amhar *et al.* (1998) and Schickler and Thorpe (1998), the term "true" orthophotos will also be used in this investigation when occlusions caused by buildings are considered.

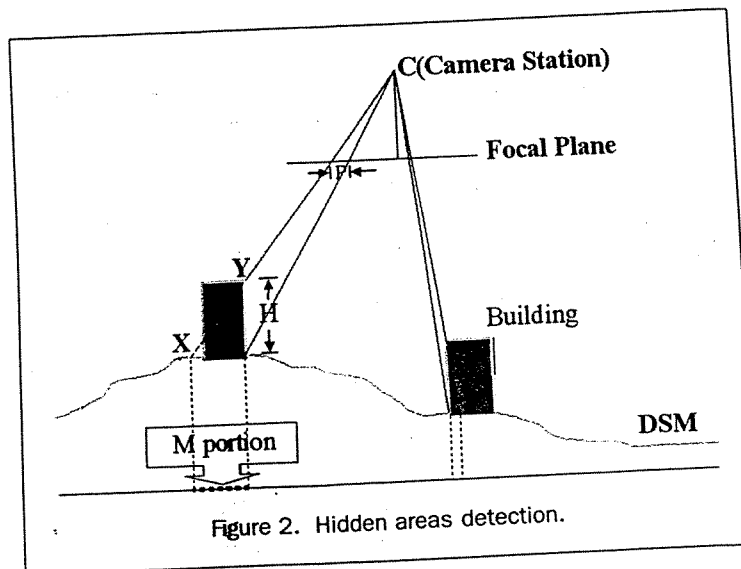
Image-Object Reprojection

The bottom-up approach for reprojection is selected in this investigation. This calculation has been widely used in image ortho-rectification. This calculation has been widely used in image ortho-rectification. Once a DTM and orientation parameters for an image are available, the collinearity condition equations may be used to calculate the image pixel corresponding to a groundel.

Hidden Area Detection

The Z-buffer technique is used here to detect hidden areas. The Z-buffer is a set of matrices. One matrix records the location of an object point. The other one stores the distance between the camera station and the object point. Both of them are defined on the image plane. Accordingly, we propose using an index map on the ground plane with the same size as the true orthophoto. The map provides necessary information for the successive processing, i.e., seamless mosaicking. The index map indicates four regions: (1) *master-visible region*, (2) *slave-visible-only region*, (3) *feathering region*, and (4) *fully invisible region*. Thus, in the proposed scheme we do not have to create in advance an intermediate DOI for each of the aerial images.

By applying the collinearity condition equations, the projected image coordinates for each groundel can be calculated



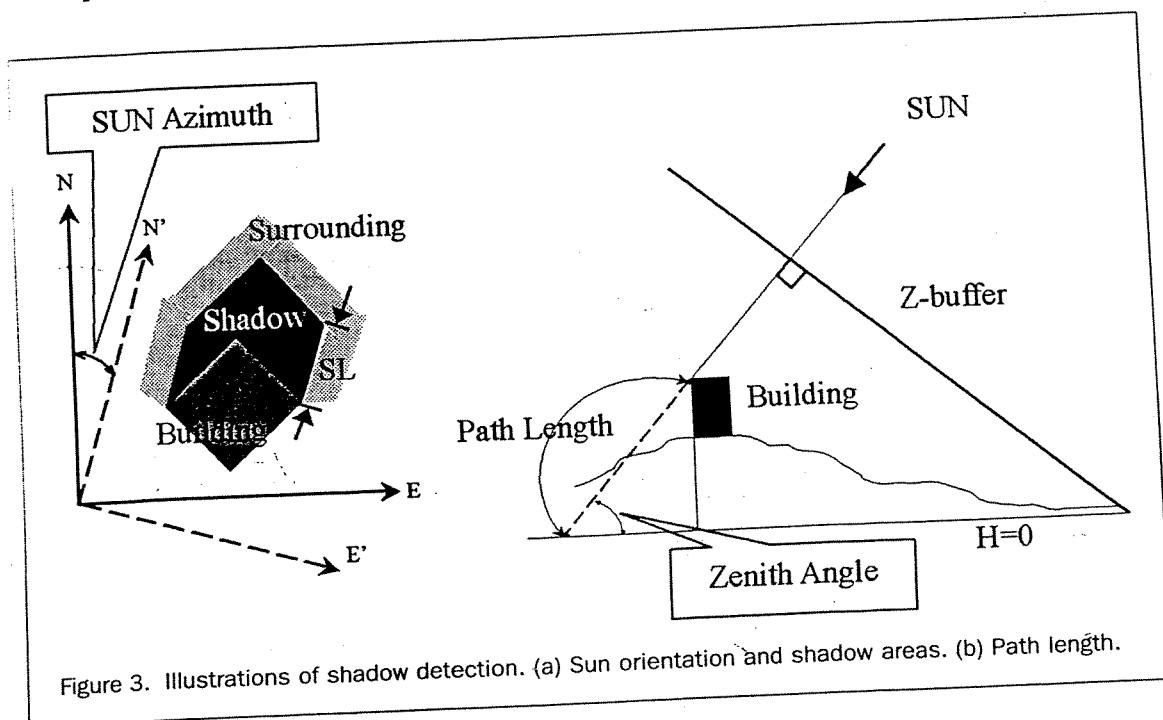
and our projection calculated. If a calculation, the index map with the areas. As Camera S, respective index map, the information regarding Y will be stored in the Z-buffer. Because CY is smaller than CX, the cell on the corresponding Y will be flagged as hidden, and the foreground detection method searches for hidden cell-by-pixel, from the raster DSM, certain hidden very steep surface elements (e.g., building walls) will cause a double mapping problem. Such a problematic area is the M portion in Figure 2. To solve this problem, pseudo groundels are introduced to determine the number of pseudo pixels. The height difference and

the projection length are estimated locally. For instance, the left-hand building in Figure 2 has a height of H and its projection length on the image plane is P. Thus, on the vertical walls of the building, pseudo pixels are placed at a spacing of no greater than H/P . We then apply the detection rules again, as described in the previous section, for these pseudo pixels.

Due to numerical rounding errors, isolated points or lines may occur in the index map, which could cause noisy results. Thus, a post-process for noise removal is implemented by using a majority rule.

Seamless Mosaicking

The calculation of the camera location is a least-squares triangulation procedure or something similar. Random errors may be introduced due to measurement errors of ground control points and image points, as for the building model. In addition, the viewing conditions for multi-view images are different. This is especially true for those photos on adjacent strips and for color photos. The radiometric responses for a ground point



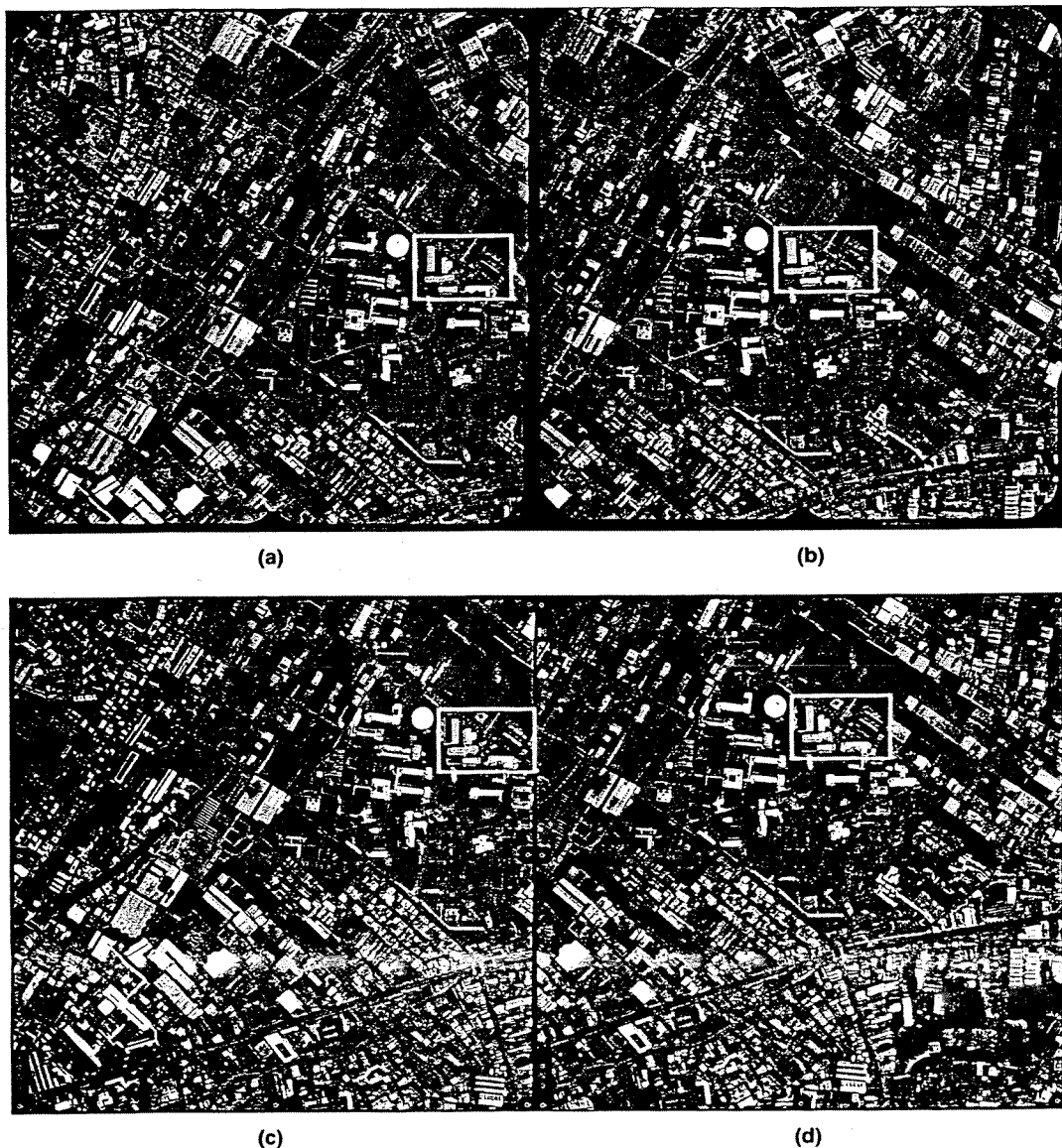


Figure 4. Test images. (a) Image 1.(b) Image 2.(c) Image 3. (d) Image 4.

are different in different images. So, to directly fill in the gray values, for hidden areas in the master image from its slave images, may cause radiometric differences.

The index map created in the previous section provides a basis for this mosaicking process. The gray value of the true orthophoto, corresponding to the *master-visible region*, is resampled from the master image. The gray value of the true orthophoto, corresponding to the *slave-visible-only region*, is resampled from the slave image. The *feathering region* stores the weighting factor for the seamless mosaicking of the master and the slave image. Finally, the *fully invisible region* is the area invisible in both the master and the slave image. Once a new slave image is combined, the index map is updated. This process is done successively for the remaining slave images.

The purpose of seamless mosaicking is to reduce discontinuities for gray values around the fill-in areas. For the boundary of each hidden area, a *feathering region* is built. This feathering region is built using a "dilation" morphology operator. Within the feathering region, the gray value is calculated from

the master and the slave image by means of a weighted average calculation. The weighting factors are determined according to the distance between the pixel and the boundary of the hidden area.

Shadow Treatment

This treatment is designed to enhance the interpretability of shadowed areas caused by buildings. It is understood that differences in the gray values, between a shadow area and its surroundings, are large. Hence, we need to locate the shadowed area, and then provide an image enhancement of these shadow areas. However, in order to minimize the gray-value differences between a shadow area and its surroundings, a histogram matching technique is applied for image enhancement (Richards, 1986).

The shadow detection technique illustrated in Figure 3 is similar to that for hidden area detection. An index map on the ground plane is constructed after shadow area detection. The index map used here is to denote the shadowed areas and the

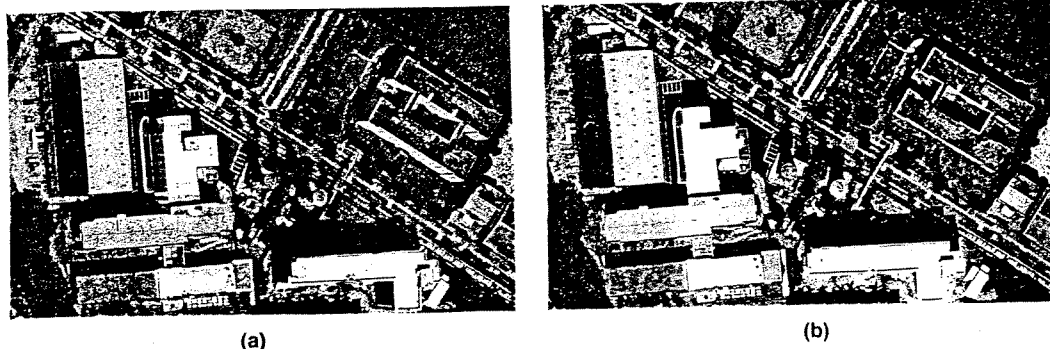


Figure 5. Results of ortho-rectification using DTM. (a) Image 1. (b) Image 2.



Figure 6. Enlargement of true orthophotos generated from a stereo-pair. (a) Without shadow enhancement. (b) With automatic shadow enhancement.

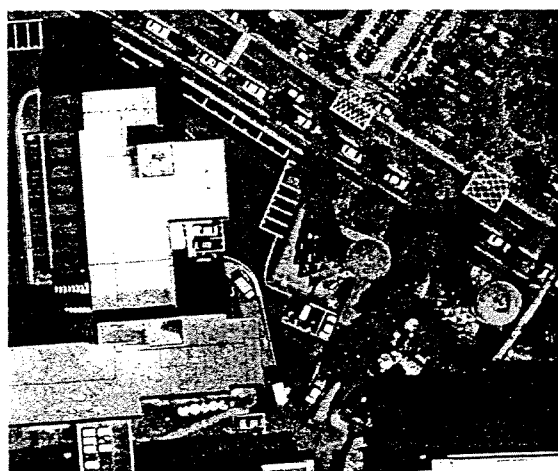
buffer zone around the shadowed areas. The buffer zone is also a feature for histogram matching. The zone is also a feature to avoid gray-value discontinuity between the areas and the original image. The generation of a buffer zone may be done by the use of a dilation operator. The shadow detection technique assumes parallel projection of the sun's rays. The Z-buffer plane is built perpendicular to the direction of the sun. In cases where such information is not available, the direction can be estimated from the shadow of tall buildings. As shown in Figure 3a, the azimuth angle of the sun is estimated by the direction of the shadow. The zenith angle is calculated from the ratio of the building height to the shadow length (SL). As shown in Figure 3b, the distance value of the buffer is the path length between the surface and the shadow. It is obvious that the searching algorithm used for shadow area detection, described in the previous section, is not the one we need here. Thus, we may repeat the process by simply changing the value of the Z-buffer from the shadow distance to the path length. To enhance illumination differences around shadowed areas, image enhancement technique should be applied. Around buildings, the characteristics of ground features are often dominated by human activities, and are usually quite different from place to place. Hence, image enhancement in a local-

ized way is better than a global one. In creating a buffer zone, if a shadow appears, rooftops are further excluded from both the shadowed area and the buffer zone, in order to maintain the similarity of image contents in the shadowed area and that in the buffer zone on the ground, rather than on the roof. The histograms for a particular buffer zone and shadow area pair are calculated accordingly. Considering the former as the reference histogram, a gray-value transformation table for the shadowed area is then created, by using a histogram-matching method. Thereafter, shadow enhancement may be accomplished locally.

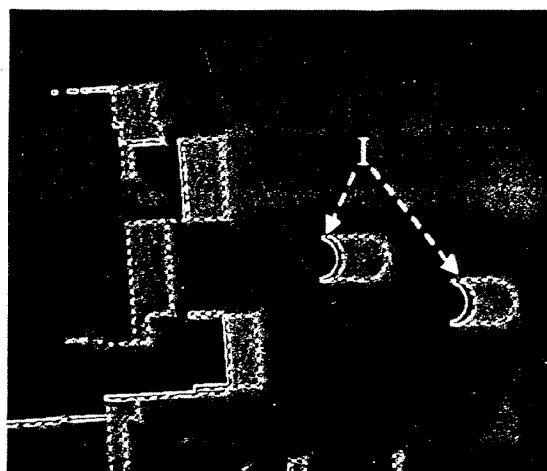
Experimental Results

To evaluate the performance of the proposed scheme, a fully automatic orthographic rectification procedure was tested. Along with the source images, the input data included a DSM, the pre-calculated orientation parameters, and the orientation of the sun. A set of four digitized aerial photos forming a quadruplet was analyzed. Test images overlapped 60 percent at two perpendicular directions, as shown in Figure 4. Experiments were only performed on those areas in the white boxes. We compared the processed results between a two-image pair and a four-image quadruplet.

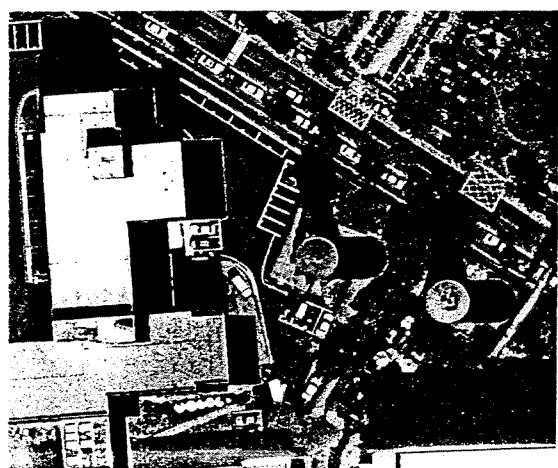
The scale of the aerial photos was 1:5,000. By scanning at a resolution of 25 μm , the test images had a nominal resolution



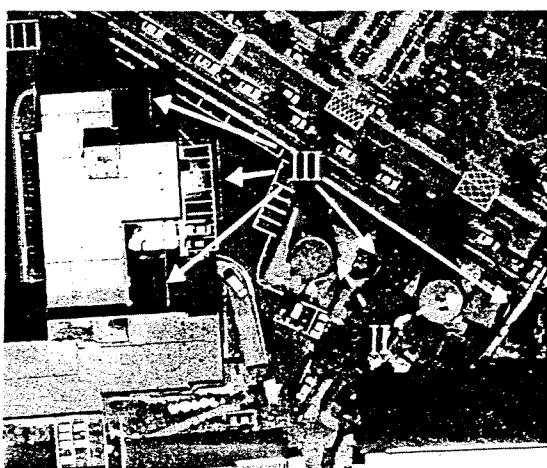
(a)



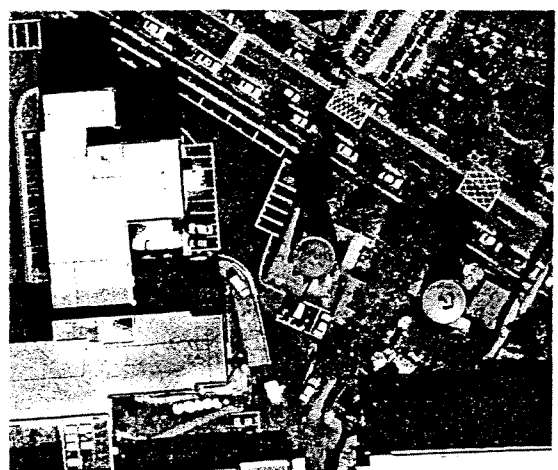
(b)



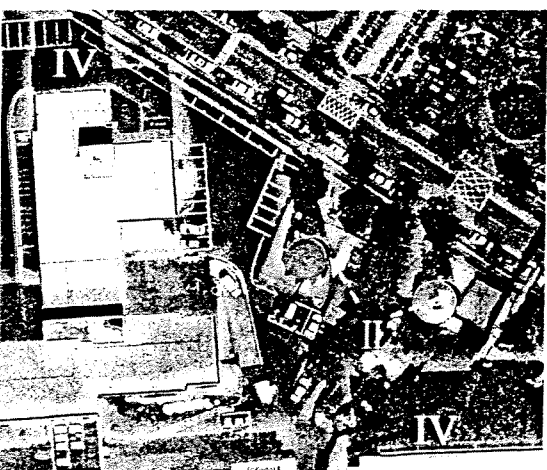
(c)



(d)



(e)



(f)

Figure 7. Intermediate results and final product of true orthophoto generation. (a) Ortho-rectification using DTM. (b) Index map for seamless mosaic. (c) With relief displacement correction only. (d) With hidden compensation applied. (e) With hidden compensation using a seamless mosaic technique. (f) With shadow enhancement.

of 12.5 cm on the ground, and the ground sample distance of the generated true orthophotos was chosen to be 12.5 cm. The DSM describing the building was constructed geometrically of three-dimensional feature points, which were manually measured on a Digital Photogrammetric Workstation (Socet Set of LH-Systems).

Although the ultimate goal of this investigation was to produce true orthophotos from multi-view aerial images, we included two supplements. The first gave the results of orthorectification using a DTM rather than a DSM. The second supplement gave the results for a two-view stereopair using a DSM in order to compare the improvements in information content when a quadruplet is applied.

Ortho-Rectification Using a DTM

To demonstrate the process, we selected a portion of the test area for observation. Figures 5a and 5b show the results for Image 1 and Image 2. It is obvious that the relief displacements due to buildings were not corrected. The walls of the buildings may be observed in the figures. Without knowledge of hidden areas, there is no reason for data merging.

Orthoimage from a Stereopair

Figure 6 is an enlargement of the areas indicated in Figure 4, but with corrections for the building's relief displacement and compensation for hidden areas. After merging images 1 and 2, the *fully invisible regions* are marked as black (Part A for instance). Notice that the information loss for occluded areas in both images may be reduced when images viewed from other directions are available. Figures 6a and 6b give a comparison of automatic shadow enhancement. It may be observed that features within the shadowed areas become distinguishable after shadow enhancement.

Orthoimage from a Quadruplet

In cases where a quadruplet is used for true orthophoto generation, the hidden areas will be minimized after data merging. With Image 1 as the master image, a true orthophoto is generated using the developed procedure. The intermediate results and the final product, for the highlighted area, are shown in Figure 7. Figure 7a, enlarged from Figure 5a, represents the master image with terrain relief correction only. Figure 7b illustrates the index map for seamless mosaicking. The black background indicates the *master-visible region*. The dash-line regions indicate the *slave-visible-only regions* that the fill-in is possible from slave Image 1. The feathering effect indicates the weighting factor of the *feathering region* for seamless mosaicking data that come from the master and the slave image. Finally, the thick-line areas are the *fully invisible regions*, as depicted as Part I in Figure 7b. Once a new slave image is combined, the index map is updated. Accordingly, the *fully invisible regions* are reduced. Figure 7c shows the result for relief displacement correction. The hidden areas are marked as black. As shown in Figure 7d, the defect due to hidden areas is significantly reduced after data merging using three slave images, but a few *fully invisible regions* still exist (Part II for instance). In comparison to the two-view situation indicated in Figure 6c, the hidden areas are minimized. Taking a closer look at Figure 7d, an abnormal gray-value discontinuity can be observed along the fill-in boundary (e.g., Part III), and it may be misinterpreted as a ground feature. This has been improved in Figure 7e by the seamless mosaicking technique. By further applying local enhancement for shadow defects, the final true orthophoto product, shown in Figure 7f, is improved. This is especially true in terms of interpretability in shadowed areas (e.g., Part IV). Compared to Figure 7a, the interpretability of Figure 7f is significantly improved.

TABLE 1. THE CHECKPOINT ERROR STATISTICS (UNITS: CM)

Category	A	B	C	Total
Number of Checkpoints	69	16	76	161
Delta X { Mean	4.3	3.5	-0.1	4.9
Standard Deviation	9.3	3.9	6.2	7.6
Delta Y { Mean	-8.3	-4.5	-2.4	-5.1
Standard Deviation	9.8	7.9	3.5	7.7

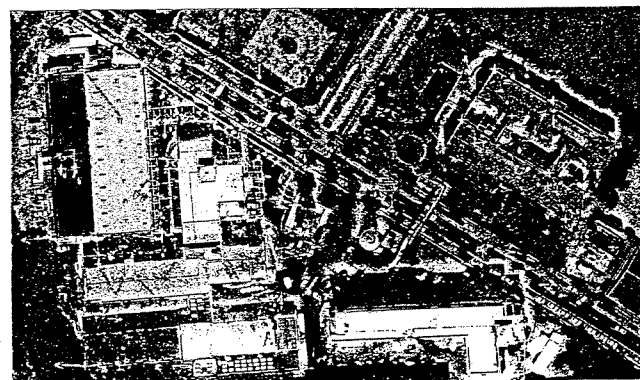


Figure 8. Error vector plots (+:Category A; □:Category B; ×:Category C).

Assessment of Geometric Quality

This is an assessment of the geometric accuracy of the final true orthophoto after data merging and shadow enhancement. Considering the error sources of (1) orientation parameters in four images and (2) the DSM, we will demonstrate the accuracy performance of the proposed scheme.

The reference data set is acquired manually using a Digital Photogrammetric Workstation. We analyzed three categories of planimetric checkpoints. Category A, on the building roof, was to check the relief displacement error. Category B, within the hidden areas, was to evaluate the geographic error within the hidden areas after data merging. The final one, category C, consisted of randomly distributed feature points on the ground, for checking the terrain relief errors.

Table 1 shows the checkpoint error statistics for the three categories. It reveals that every category has a standard deviation of less than 10 cm, less than the ground sample distance of 12.5 cm. The mean errors in the X and Y directions are also small. This indicates that the bias might be ignored. Figure 8 shows the error vector plots for each category, denoted by different symbols. It illustrates that there is no significant systematic error for the whole product.

Assessment of Radiometric Quality

Here we are concerned with the radiometric quality inside the shadowed area. We investigated changes in the dynamic range, before and after histogram matching. Figures 9a and 9b show the histogram distributions, before and after applying histogram matching, as observed from Figures 7e and 7f. The dynamic ranges before and after image enhancement are between 49 and 130 and 49 and 175. After automatic image enhancement, it is easier to interpret ground features.

Conclusion

In this study, an orthographic rectification scheme is proposed for the generation of large-scale true orthophotos of built-up

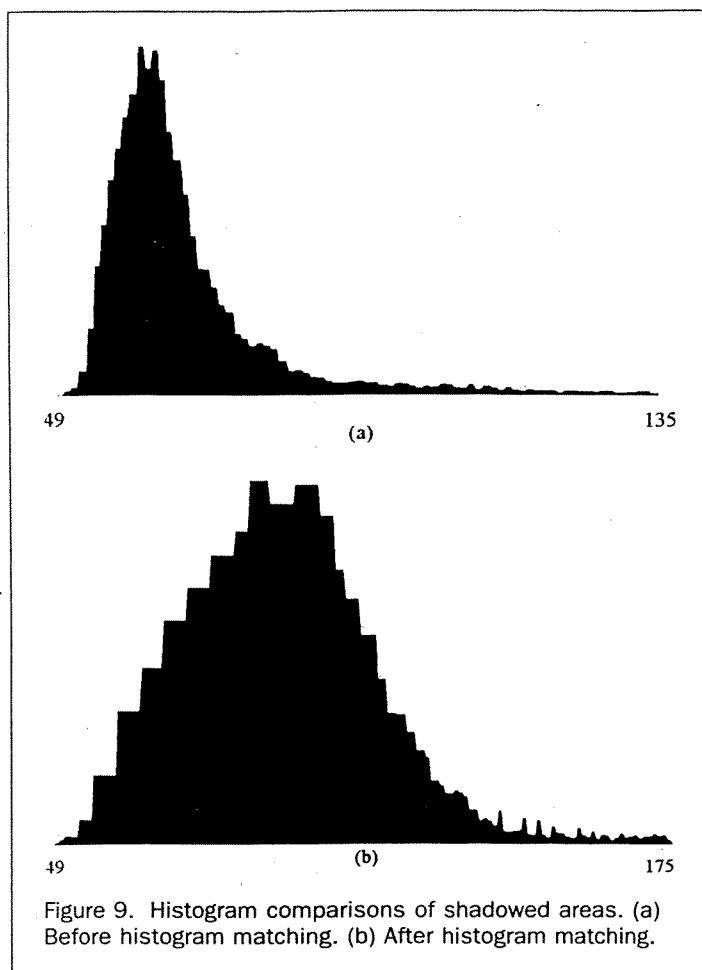


Figure 9. Histogram comparisons of shadowed areas. (a) Before histogram matching. (b) After histogram matching.

areas. Provided that a DSM is available, the scheme minimizes both hidden areas and shadow defects, and improves the interpretability of true orthophotos. The proposed scheme achieves

- the correction of relief displacement for buildings and terrain, in order to increase the accuracy of orthophoto products;
- data merging, in order to preserve information within the hidden areas from corresponding photos;
- the seamless stitching of multi-view images in order to reduce artifacts along fill-in boundaries; and
- the reduction of illumination deterioration, leading to an increase in the interpretability of shadowed areas.

A fully automatic procedure has been developed. Experimental results indicate that the information loss and degradation can be significantly reduced, thus the products retain good radiometric and geometric quality.

Acknowledgment

The authors would like to thank Prof. S.C. Wang of the Department of Surveying Engineering, National Cheng Kung University, Taiwan, for providing the test data set. Special thanks go to the anonymous reviewers for their valuable comments and suggestions.

References

- Amhar, F., J. Josef, and C. Ries, 1998. The generation of true orthophotos using a 3D building model in conjunction with a conventional DTM, *International Archives of Photogrammetry and Remote Sensing*, 32(Part 4):16–22.
- Chen, L.C., and L.H. Lee, 1993. Rigorous generation of digital orthophotos from SPOT images, *Photogrammetric Engineering & Remote Sensing*, 59(3):655–661.
- Chen, L.C., and J.Y. Rau, 1993. A unified solution for digital terrain model and orthoimage generation from SPOT stereopairs, *IEEE Trans. on Geoscience and Remote Sensing*, 31(6):1243–1252.
- Chen, L.C., and L.Y. Chang, 1998. Three dimensional positioning using SPOT stereo strips with sparse control, *ASCE Journal of Surveying Engineering*, 124(2):63–72.
- Chen, N.Y., H.T. Wang, and M.M. Lin, 1997. Orthographic correction of airborne scanner imagery for mountainous areas, *Proceeding of the 3rd International Airborne Remote Sensing Conference and Exhibition*, 07–10 July, Copenhagen, Denmark, 2:293–299.
- Hohle, J., 1996. Experiences with the production of digital orthophotos, *Photogrammetric Engineering & Remote Sensing*, 62(10):1189–1194.
- Joshua, G., 2001. Evaluating the accuracy of digital orthophotos quadrangles (DOQ) in the context of parcel-based GIS, *Photogrammetric Engineering & Remote Sensing*, 67(2):199–205.
- O'Neill, M.A. and I.J. Dowman, 1988. The generation of epipolar synthetic stereo mates for SPOT images using a DEM, *Proceeding of XVI ISPRS Congress 1988, Commission III*, 01–10 July, Kyoto, Japan, B3:587–598.
- Richards, J.A., 1986. *Remote Sensing Digital Image Analysis: An Introduction*, Springer-Verlag, Berlin, Heidelberg, New York, 281 p.
- Schickler, W., and A. Thorpe, 1998. Operational procedure for automatic true orthophoto generation, *International Archives of Photogrammetry and Remote Sensing*, 32(Part 4):527–532.
- Skarlatos, D., 1999. Orthophotograph production in urban areas, *Photogrammetric Record*, 16(94):643–650.
- Tudor, G.S., and L.J. Sugarbaker, 1993. GIS orthographic digitizing of aerial photography by terrain modeling, *Photogrammetric Engineering & Remote Sensing*, 59(4):499–503.
- Wiesel, J.W., 1985. Digital image processing for orthophotos generation, *Photogrammetria*, 40(2):69–76.

(Received 21 March 2001; revised and accepted 16 November 2001)