BUILDING IMAGES ACQUISITION AND EXTRACTION FROM AERIAL IMAGERY USING HYBRID APPROACH

Prof. V.K. Shandilya*

Miss. Snehali M. Khakse**

Miss. Bhavna B.Khajone**

ABTRACT-

Automated update of building information in maps from high-resolution aerial imagery is one of the most important and challenging researches in the field of photogrammetric and remote sensing. Up-to-date building information is necessary for many practical applications such as fixed assets inventory, city planning, GIS application analysis, etc. Urban land used to study and focused on building extraction and height estimation from space borne optical imagery. The advantage of such methods is a 3D visualization of urban areas, digital urban mapping, and GIS databases for decision makers. In particular, for efficient building extraction from optical multi-angular imagery, first need to remove the presence of shadows in very high resolution (VHR) images, because shadow in VHR can represent a serious obstacle for their full exploitation. This paper proposes to face this problem as a whole through the proposal of a complete processing chain for Shadow Detection and Reconstruction in VHR images. After, a template matching algorithm is formulated for automatic estimation of relative building height, and the relative height estimates are utilized in conjunction with a support vector machine (SVM)-based classifier for extraction of buildings from non-buildings. The final results are presented as a building map

* Computer Science and Engineering, Sipna College of Engineering, Amravati, Maharashtra, India.

** Computer Engineering, Sipna College of Engineering, Amravati, Maharashtra, India.

A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories
Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A.

International Journal of Management, IT and Engineering
http://www.ijmra.us

371

and an approximate 3D model of buildings building extraction. The building detection accuracy of the proposed method is improved to 88%, compared to 83% without using multi-angular information.

Index Terms—Building extraction, height estimation, shadow detection, shadow reconstruction, support vector machines (SVMs), very high resolution (VHR) images.

INTRODUCTION

Recently, very high resolution (VHR) satellite images   opened a new era in the remote sensing field. Because of the increase of spatial resolution, new analysis, classification, and change detection techniques are required. Unfortunately, VHR images high spatial resolution entails also some drawbacks like the unsought presence of shadows, vegetation particularly in urban areas where there are larger changes in surface elevation due to the presence of buildings, bridges, towers, etc. To attenuate these drawbacks and, thus, to increase image exploitability, two steps are necessary: 1) shadow detection and 2) shadow reconstruction. Regarding height estimation from spaceborne optical images, various approaches, such as template matching, dynamic programming (DP), and a variant version of DP, were analyzed [1]. In the template matching method, an image is searched for the best correlated area for a given template from another image.

Early research on building extraction was focused on aerial imagery due to its high spatial resolution Region-based ACM Driven geometric snakes (GS) are more suitable for aerial imagery since they rely on intensity, texture, and statistical features of the pixels. Another advantage of GS is their ability To detect topological changes without noticeable edge data [2].The existing GS approach for building detection was Improved bytraining the model with gray level values from target regions.

The traditional snake-based building detection was improved through the incorporation of digital surface model (DSM) and height data [3]. The DSM was generated from existing LIDAR data and a DEM.Later, approximate contours and initial points of interest were identified and the

snakes were optimized to fit the building contours.The standard approach for DEM-based building extraction depends on functional minimization with Regularization [4], this method was computationally improved by modification of the energy-based objective function in the original formulation. In another study [5], void and elevated Areas were extracted from LIDAR data. Using normalized difference vegetation index (NDVI) data and edge detection of void areas, building shape extraction was improved.The limitation of ACM-based methods described in [3] is that they perform well on isolated buildings, but further research is required to analyse their performance on closely spaced and connected buildings. Additionally, the images have nearly uniform background [4], [6], thus there are no non-building objects with edges or elevated areas. In contrast, the images of commercial urban scenes have complex backgrounds requiring a more sophisticated approach. For instance, methods in [7] and [8] attempt to detect shadows using a space color transformation and an automatic threshold estimator. Other algorithms rely on the idea of adding features capable to better discriminate shadow areas (e.g., normalized difference vegetation index [9], normalized saturation-value difference index (NSVDI) [10]). In this paper, we study a hybrid building extraction and relative building height estimation methodology for an urban scene using the multi-angular characteristics of optical imagery. Our major contributions are:

1) Detection and reconstruction of shadow areas, to recognize building position and to estimate their height and other useful parameters.2) A relative height estimation methodology based on a template matching is developed and 3) A hybrid approach is proposed to significantly improve building extraction

II. METHODOLOGY

1. Shadow Removal for height estimation

A flowchart with the principal steps of the proposed methodology In brief shown below, let us consider a VHR image I of dimensions $m \times n$, composed of N bands and characterized by the presence of shadow areas,\

1. Ground truth information is collected by selecting different regions of interest (ROIs) in order to discriminate the ("clean," nonshadow) classes present in the image as well as their shadow counterpart.

2. The resulting ground truth will allow performing first a binary classification in order to distinguish between shadow and nonshadow regions.

3. To deal with noise, which may result in the obtained binary mask M, two mathematical morphological operators are applied, namely, opening and closing by reconstruction. Because the binary mask does not handle the border B between them is created which are exploited in the last step of the processing chain for interpolation purposes.
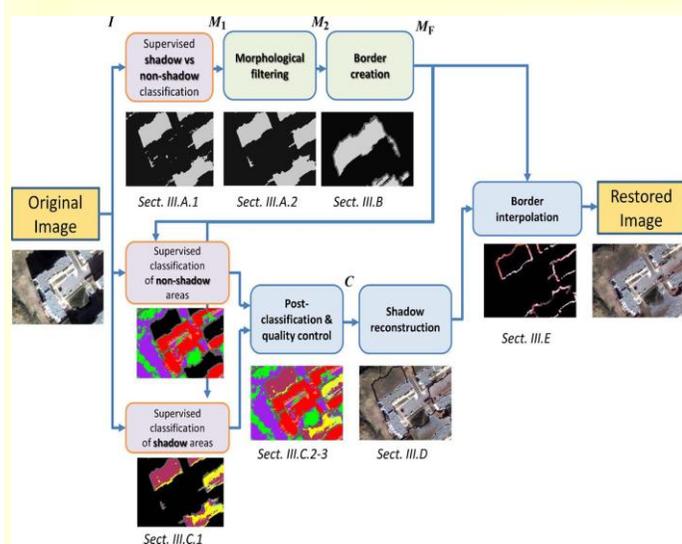

Fig1. Flowchart of the proposed method.

4. In a successive step, shadow and nonshadow classesare classified separately with the same initial ROI. Such classification allows the localization of the available couples of shadow and nonshadow related to the same object and, thus, to define the spectral relationship between them as a means to perform the reconstruction of the shadow areas.

5. Finally, the border between the reconstructed shadow and the nonshadow areas undergoes a linear interpolation operation to yield a smooth transition between them.

Denoting the shadow class as $X \sim N(\mu_s, \sigma^2_S)$ and the corresponding nonshadow class as $Y \sim N(\mu s^-, \sigma_s^{-2})$ the reconstruction of the shadow class will be reduced to a simple random variable transformation

$$X \sim N(\mu_s, \sigma^2_S) \rightarrow X' \sim N(\mu s^-, \sigma_s^{-2}) \qquad \text{.... (1)}$$

1.1. Mask Construction

The shadow versus nonshadow mask is created in two steps, namely, binary classification followed by a postprocessing.

1.1.1) Binary Classification-: The binary classification procedure [see $M1$ in fig.1] is implemented in a supervised way by means of a support vector machine (SVM). Classes SVMs have often been found to provide better classification results.

1.1.2) Postprocessing-: The binary image $M1$ may be characterized by a "salt and pepper" effect due to the presence of noise in the image. The choice of morphological filters to deal with this problem is motivated by their effectiveness, better shape preservation capability and the possibility to adapt them according to the image filtering requirements as is the case in the border creation (described in the next section). Both morphological operators are needed in order to remove isolated shadow pixels in a nonshadow area and also isolated nonshadow pixels in a shadow area.

1.2. Border Creation

`The transition in between shadow and nonshadow areas can raise problems such as boundary ambiguity, color inconstancy, and illumination variation. A border between the shadow and nonshadow classes is defined in order to appropriately handle the border pixels. These lasts are not processed within the shadow reconstruction procedure as is, but separately. The border region is constructed by means of morphological operators. The mask c_imgB2 is dilated (δ) and eroded (ε).Then, the difference between these two images is computed to form the border image.

$$B[x, y] = \delta(c\_imgB2[x, y]) - \varepsilon(c\_imgB2[x, y]) \qquad \text{.......(2)}$$

A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories
Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A.

International Journal of Management, IT and Engineering
http://www.ijmra.us

375

In order to better track the actual shadow direction the final mask image becomes [see example in fig. 1(c)]

$$c\_imgBNEW[x, y] = \begin{cases} B[x, y], & \text{if } B[x, y] = 1 \\ c\_imgB2[x, y] & \text{if } B[x, y] = 0 \end{cases} \quad\quad …… (3)$$

1.3 Classification Maps

1.3.1) Multiclass Classification: The previously obtained mask is exploited to guide a further level of classification applied separately to the shadow and nonshadow areas. The aim is to distinguish between the different predefined nonshadow classes on the one side and the corresponding shadow classes on the other side. The result is a final classification map *C*, which is important to define the spectral relationship between the shadow and nonshadow versions of the same object (class) and thus, to perform customized reconstruction of shadow areas. For such purpose two multiclass SVMs are trained in the feature space described earlier for the shadow and nonshadow classifications, respectively. After the training phase, to generate *C*, both are applied to predict the label of each pixel of the corresponding areas, shadow and nonshadows respectively, defined in *MF*.

1.3.2) Post classification: In order to improve the classification map *C* before exploiting it for the reconstruction of shadow areas, post classification is applied by adopting a simple 3×3 majority filter for removing isolated labels and thus, smoothing the map.

*1.3.3) Quality Control: T*he reconstruction process is directly based on the classification map; it becomes necessary to control its quality in order to decide if compensation is a particular shadow class. For such purpose, a confusion matrix

For the shadow classes is computed on the basis of the available ground truth. If one of these accuracies is lower than a predefined value, shadow reconstruction is carried out only wherever a sufficient guaranty of correct shadow recognition is available in order to avoid error propagation in the processing chain.

A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories
Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A.
International Journal of Management, IT and Engineering
http://www.ijmra.us

376

1.4. Shadow Reconstruction

Image reconstruction is one of the most important steps in our methodology. As done in the literature [8], [9], [10], [11], for the sake of getting a simple but satisfactory reconstruction model, we assume that the underlying relationship between the

Nonshadow class (*Y)* and the corresponding shadow classes

(*X*) Is of the linear type. Regarding the statistical model of the classes, three estimation ways may be envisioned: 1) histogram estimation by box counting; 2) kernel density estimation; or 3) parametric estimation. In our case, we will adopt the last method by assuming that the classes follow a Gaussian distribution. This is motivated by the need to derive an analytically tractable and easy-to-implement reconstruction method. The two distributions are assumed linearly correlated, *x* and *y* may be linked by

$$y = Kx + c \ldots\ldots\ldots\ldots (4)$$

$$\begin{cases} \mu_s = K\mu_s + c \\ \Sigma_s = K\Sigma_s K^T \ldots\ldots\ldots\ldots (5) \end{cases}$$

Where *K* is a transformation matrix, K^T is its transpose, and *c* a bias vector. To estimate *K* and *c*, the Cholesky factorization is applied

$$\begin{cases} \\ c = \mu_s - K\mu_s \end{cases}$$

$$K = U_s V_s^{-1} \ldots\ldots\ldots\ldots (6)$$

Where U_s and V_s are the lower and upper triangular Cholesky matrices related to the nonshadow and shadow classes, respectively. Once *K* and *c* are estimated, (4) is applied to compensate the pixels of the shadow class. When applying the shadow compensation, the restored area may

appear noisy. This is due to the fact that the initial shadow distribution is much more concentrated than the nonshadow one. To mitigate this effect, the coefficient of variation (CV) defined as

$CV[i] = \sigma_i/\mu_i$ (σ_i and μ_i are the standard deviation and the mean along the ith image band, respectively) is used to weigh the target variance of the reconstructed area [31]. In particular, we compute the CV ratio between the nonshadow and shadow classes, i.e., $NCV[i] = CV^-S[i]/CVS[i]$, for the ith original image band ($i = 1, N$). Then, if $NCV[i] \geq 1$, the covariance matrix is corrected to reduce the variability of the nonshadow class.

$$CovNEW[i, k] = \frac{covNEW[i, k]}{NCV[i] \cdot NCV[k]}*$$

with $(i, k) \in [1, N]^{2}$ .... (7)

1.5. Border Reconstruction

After the reconstruction of the shadow areas, the processing is not completely finished since thin borders between nonshadow and reconstructed shadow areas still remain with their original aspect, which may be in contrast with the two adjacent areas. In order to smooth such contrast, pixels of the borders undergo an easy-to-implement and fast contextual linear interpolation. In greater detail, a sliding window of predefined $S \times S$ size is adopted and within which four directional linear interpolations are considered. Among the directions, namely, North–South, West–East, NE–SW, and NW–SE, just those crossing the given reconstructed shadow and nonshadow areas are retained. For instance, in Fig. 2, only three directions are used to perform the linear interpolations (i.e., N–S, W–E, and NW–SE). The NE–SW direction is removed since it does not cross the reconstructed shadow area within the window. A given direction, the linear interpolator is a simple affine transformation defined as follows:

$$z = m \cdot i + q .... (8)$$

Where i stands for the coordinate along that direction while m

A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories
Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A.
International Journal of Management, IT and Engineering
http://www.ijmra.us

378

And *q* is interpolation parameters. In order to estimate these lasts, we will make use of the least square estimator, also known under the name of pseudo inverse method,
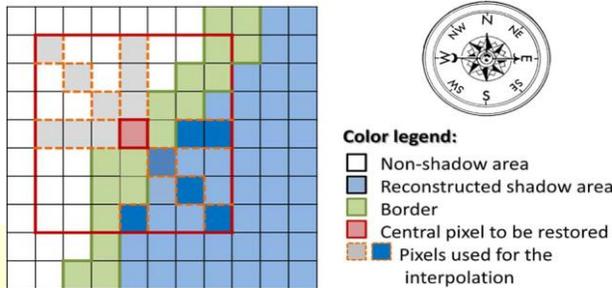

Fig. 2 Illustration of the reconstruction of a border pixel with a 7×7 size window.

Which consists in our case in collecting the set of points which are as follows: 1) within the window; 2) outside the border; and 3) along the considered direction crossing the central pixel of the window:

$$\left| \begin{matrix} Z1 \\ ... \\ ZN \end{matrix} \right| = \left| \begin{matrix} i1 & 1 \\ .... & 1 \\ iN & 1 \end{matrix} \right| \left| \begin{matrix} m \\ \\ q \end{matrix} \right| \quad \leftrightarrow \quad Z = \Gamma \cdot \beta$$

…….. (9)

An interpolator is derived for each available direction. Different strategies could be adopted for merging the estimates provided by the obtained pool of interpolators. We found by visual inspection that the best fusion rule among four common rules, which are the MIN, MAX, average, and median rules, is achieved by retaining as estimate for the central pixel the largest value yielded by the pool (i.e., the MAX rule). This can be explained by the fact that the MAX rule provides values which are brighter and, thus, less contaminated by shadow.

2. Vegetation detection

Vegetation detection helps avoid the challenges that can be encountered in a template matching wit objects such as tall trees. In addition, this detection enables reducing significant

percentage of non target information load for further processing. Finally, after the masking of shadows and vegetation, the resulting image consists of only buildings, and other man-made structures such as roads, parking lots, and pavements, just to mention a few. The NDVI (normalized difference vegetation index) [10] is calculated by using the near Infrared (NIR) and Red bands of the pan-sharpened image as

$$NDVI = \frac{PS(NIR3)-PS(red)}{PS(NIR3)+PS(red)} \qquad \ldots\ldots\ldots (11)$$

The vegetation mask is obtained by using a threshold from the NDVI image. This threshold can be determined with the maximum between-class variance (MBCV) criterion. In the MBCV method, it is assumed that the histogram of an image would have two major modes, and the threshold is selected such that the inter-class variance between these two modes is maximized so as to maximize the separability of these two modes. Previous studies show that the normalized shadow index (NSI) and the normalized spectral shadow (NSS) produce very low index values for shadows but often mix up with the surrounding water bodies and vegetation. Similarly, shadow regions are processed by transforming the histogram for grey scale compensation to discriminate from non-shadow regions.

3. Template Matching for Potential Building Extraction:

The template matching is commonly used in several image processing applications, such as object tracking for forward looking infrared imagery.In this study the method to use this technique for relative height estimation of a given image scene based on the template displacements. Consider two k-th band images PS (θ) (k) and PS (Φ) (k) at different view angles θ and Φ. In the PS (θ) (k) image, each template tmp(c) is tracked in the PS (Φ) (k) image to estimate the relative height, where is the center pixel. We consider the template is a spatial window of size L×L pixels in the PS (θ) (k) image. The displacement D_t (p_1) corresponding to a template tmp $(p1)$ in the θ image is assumed to be

$$D_t(p1) = \left| p_1p_2 \right| \qquad \ldots\ldots\ldots\ldots (12)$$

Where tmp (p_2) is the matching area in the second PS (Φ) (k) image, centered at pixel P2. As illustrated in Fig3.

The height of a building <u>hB</u> can be related to this displacement with the cosine rule:

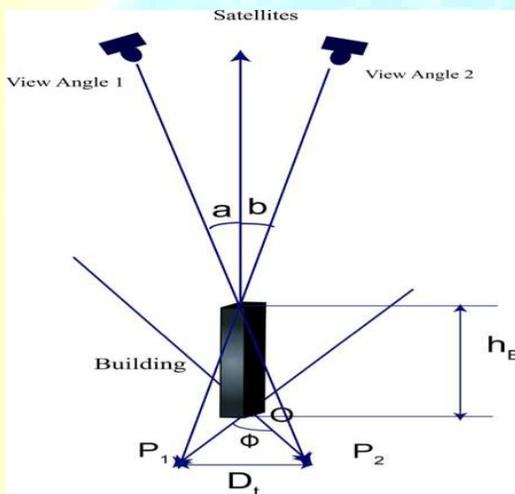$$hB = \frac{Dt}{\sqrt{\tan^2 a + \tan^2 b - 2\tan a . \tan b . \cos\Phi}} \qquad \ldots\ldots\ldots\ldots (13)$$

fig3: An illustration of the relation between template displacement and object height.

Where and are two different incidence angles, and is the

Angle between ground projections of sensor lines of sight. Here, the angle Φ depends only on the view angles of the satellite. The angles and depend on the building height. However, compared to the altitude of the satellite, the influence of the building height variation is negligible. Hence, from (4), the building height is linearly related to the displacement D_t, and D_t can be used as a relative height. Initially, for a given template, a corresponding search area, called frame, is

selected from the PS (Φ) (k) image. This frame has to be large enough to contain even the farthest displacement from, which depends on the height of the building. The frame size depends on the difference between view angles of the two images. The template size and frame size are not usually equal. When computing the FFT-based correlation, the frame and template are padded with zeros to the nearest larger power of size 2.

4. Building Extraction:

At this stage, first need to extract those portions which were covered by shadow, for that implementation object-based classification. The spatial information extracted from objects can help to decrease the number of Misclassifications between the spectrally similar *road/building, water/shadow* and *Tree/grass* classes. Feature extraction is crucial for object-based image classification. The PS(θ)(k) image is divided into a spatial block whose the statistical features, such as mean , variance ,skewness , kurtosis ,energy and are extracted.2D-wavelet-based features are also computed as follows. Initially, a 2D wavelet decomposition is applied to each spatial block using the following recursive filtering equations at the

$(j + 1)$ desired level of decomposition with initial cAj (u, v) =A:

$$cA_{j+1}(w)=\sum_u g(u) \sum_v g(v)cAj (u - 2w, v - 2w)$$
$$cH_{j+1}(w)=\sum_u h(u) \sum_v h(v)cAj (u - 2w, v - 2w)$$
$$cV_{j+1}(w)=\sum_u g(u) \sum_v h(v)cAj (u - 2w, v - 2w)$$
$$cD_{j+1}(w)=\sum_u h(u) \sum_v h(v)cAj (u - 2w, v - 2w) \qquad \text{.......(14)}$$

where cA, cH, cV, cD represent the approximation, horizontal, vertical, and diagonal coefficients, respectively, and h(**.**) and g(**.**) denote highpass and lowpass filters of the selected wavelet, respectively. The statistical features in the wavelet domain, such as energy "we", entropy "wh", maximum coefficient "wmax ", and mean "wµ ", are computed. A feature vector for the n-th block in the k-th band is constructed as:

$$x_{nk}=[\mu \log(\sigma^2),\gamma_1,k,et,we,wh,w \max,w\mu]^T \qquad \text{..........(15)}$$

After the minimum building height is detected, SVM is adopted to classify buildings from the nonbuilding. SVMs are very attractive for the classification of remotely sensed data it classify

buildings from the nonbuilding classes, where building detection has been formulated into a binary classification problem. It provides high classification accuracy is achieved with small training sets. The SVM classification, a building map is yielded for the image scene. The areas of individual segments of this map are computed and the smaller segments are discarded based on the minimum building area. This threshold can be detected using MBCV for the above segment areas. Finally, the resulting map

Consist of buildings only. The building map is overlaid on the relative height map to obtain the height data for individual buildings. Then, by averaging this height data separately for each building, an approximate 3D model can be generated.

Example illustrating the complete processing chain of building images extraction:

1. Shadow Removal for height estimation result:

It contains large shadows in a suburban area [see Figure.4]

2. Template Matching for Potential Building Extraction result:

Visual inspection of Fig. 5(a) and (b) show that planar objects such as roads and other non-elevated structures are located at the same co-ordinates in both images. For example, there rectangular template shown in Figure 5(a) and (b) is such a planar template. On the other hand, we can clearly see that all structures such as buildings roofs (yellow boxes in Fig. 5) are displaced considerably in the second image. Thus, this is about similarity comparison between two non-planar templates at point and. Moreover, the displacement of the non-planar templates depends on the building height. Since the two images have different view angles, the apparent areas of the

A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories
Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A.
International Journal of Management, IT and Engineering
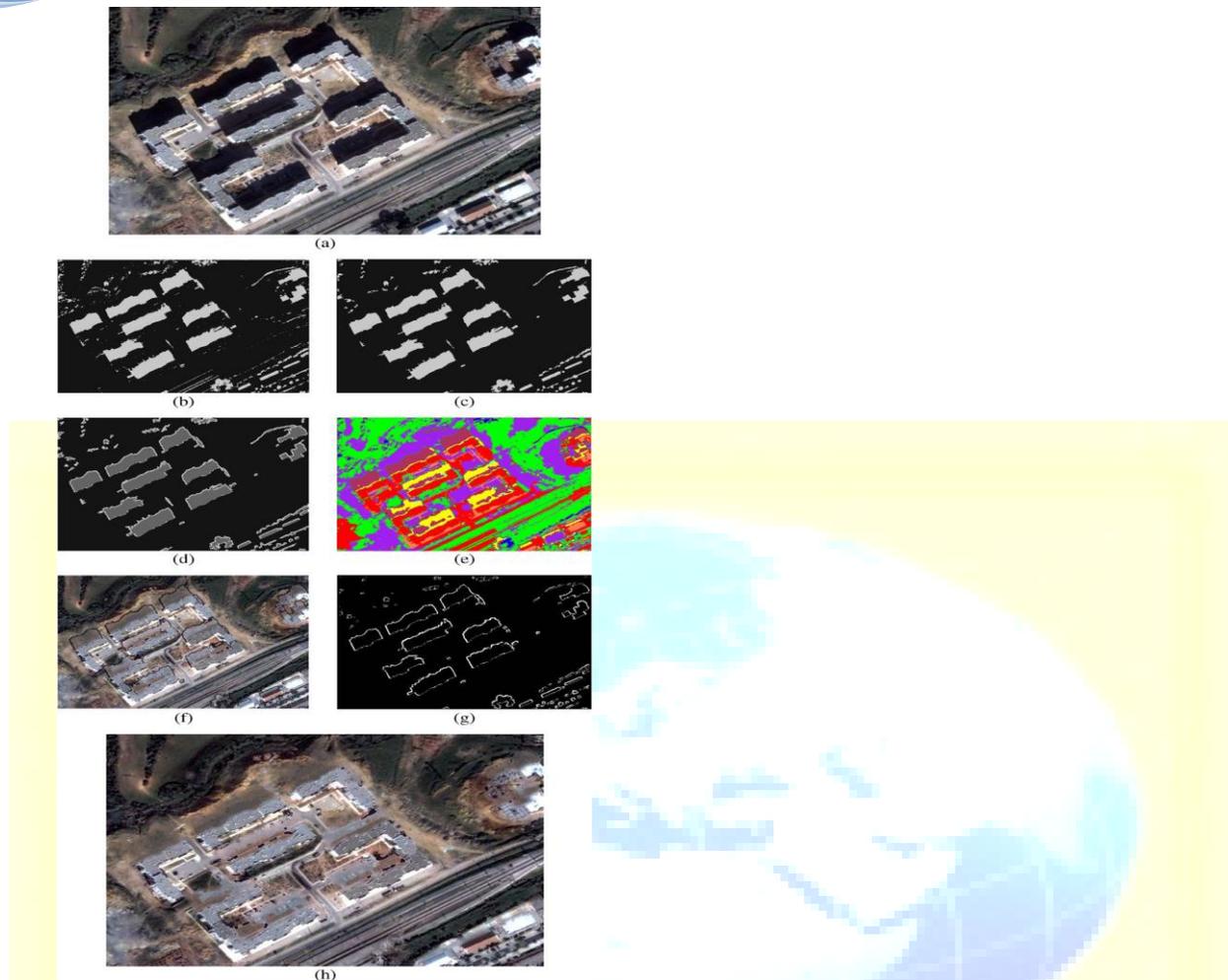**http://www.ijmra.us**

383

figure4: result for reconstruction image. (a) Original image. (b) Binary mask. (c) Post processed mask. (d) Mask with borders. (e) Multiclass classification. (f) Shadow reconstruction. (g) Border interpolation. (h) Final output image

regions representing shadows and vegetation vary between the images.

A relative height map was generated as in Fig. 6(a).In this process, L=21 were used for the template block size. Thus, the total number of templates in the first image was 625. The template size 21×21 was found to be optimal because a smaller template may have correlation with other regions in the search area that are not necessarily the best matches. In addition, if the template size is much larger, then it may have mixed categories of planar and non-planar regions. From Fig. 7(a), we can clearly see that the regions representing buildings showed high relative height values. However, a region corresponding to one building had several different height values even though the actual height was constant for a given building. This discrepancy can be attributed to a

high degree of similarity between different sections of buildings. Most of the templates belonging to non-buildings have a very low height output which is justified since they are planar objects. Based on a trial-and-error approach, the optimal threshold for a minimum correlation value for a match is found to be 0.6. In the histogram of the log-relative- height data in Fig. 7(b), the section shown in red represents the non-building regions in the image and the green section represents the regions that are mostly buildings.
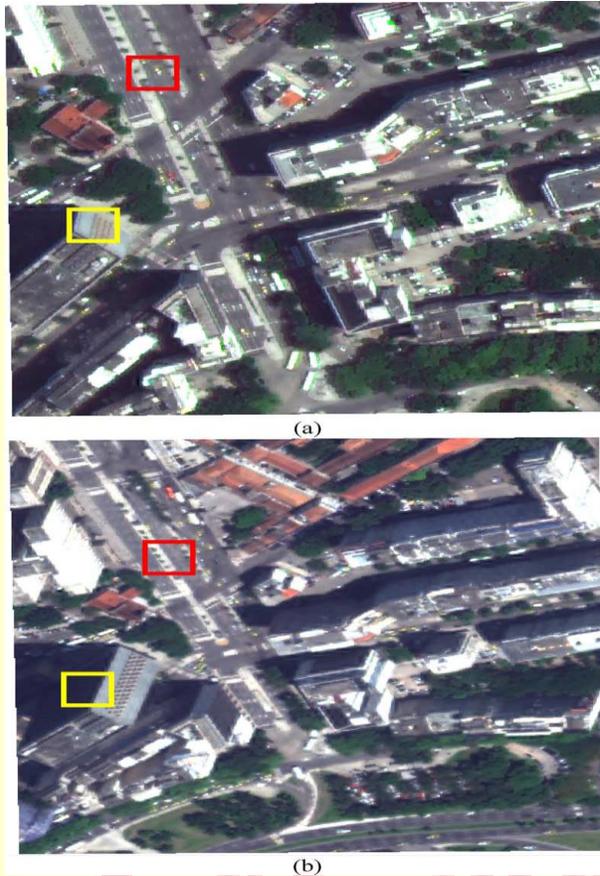
Fig5: Pan-sharpened images using the Gram Schmidt method. (a) Pan-sharpened image from view angle 3 (scene 1). (b) Pan-sharpened image from view angle 2 (scene 1)
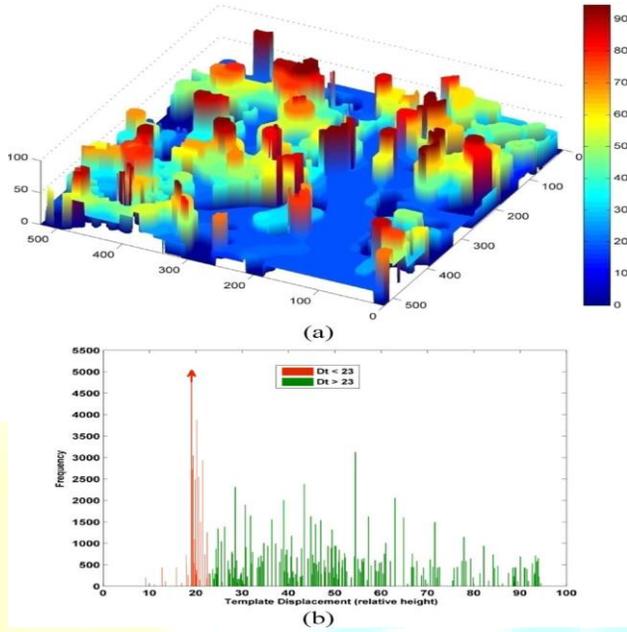
Fig6: Template matching output

The template displacement calculated for each template in view 3 is assigned as the relative height for each 21×21 block.(a) Relative height of scene 1 from view angle 3. The color bar indicates the relative height value (template displacement). (b) Histogram of the relative height output.
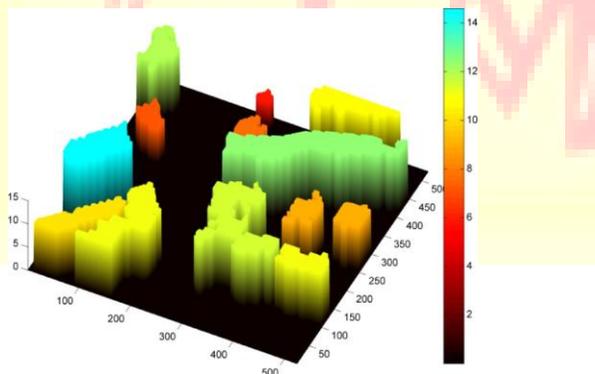
3. Building Extraction result:

Fig7: An approximate 3D model of buildings computed from the relative height data (displacement) for scene 1

A better visualization of the building classification output is also presented using the relative building height derived from the template matching as in Figure5. This approximate 3D building model is created using the shaded surface plots. It provides a more appealing visualization, compared to the previous height output in Fig.6 where the height within a building was inconsistent and the shapes of the buildings were not well preserved. Averaging the building segments from the classification output resulted in smooth heights and shapes of complete buildings. This could also be used to illustrate the relative height variations of tall buildings in the scene. The tallest building in the 3D building model (shown in a sky blue color) corresponds to the actual tallest building (based on visual interpretation) in the scene, followed by the second tallest building (in green), and small buildings shown in other colors (e.g., red, etc.) shown in Fig.7.

CONCLUSION:

In this study, a hybrid building extraction methodology is developed and applied to multi-angular imagery. Initially, the Gram-Schmidt orthogonalization based pan-sharpening method is used to fuse the panchromatic and multispectral imagery. Then, vegetation and shadow areas are not only detected but also classified so as to allow their customized compensation. For a shadow removal the classification tasks are implemented by means of the state-of-the-art SVM approach. Borders are explicitly handled by adaptive morphological filters and linear interpolation for the prevention of possible border artifacts in the reconstructed image.

Later, a 2D Fourier transform-based template matching is used to estimate the relative height model of the image scene. In the final stage, features from the pan-sharpened data along with this height data are exploited for building extraction. A relative3D building model is generated from the relative height data and building classification map. Some of the key parameters for the hybrid approach include:

1) The recognition of the couples of nonshadow and shadow classes can help in a more accurate compensation but involves a classification process which can create false shadow areas and, thus, overload the image with mis-reconstructions.

2) The threshold for vegetation detection from NDVI data, which has to be, decided from the histogram of the NDVI values.

3) The block size L in the template matching, whose selection influences efficient omission of non-building blocks without losing any building blocks.

4) The minimum height threshold, which separates buildings from non-buildings.

REFERENCCES:

[1] A.Alobeid, K. Jacobsen, and C. Heipke, "Building height estimation in urban areas from very high Resolution satellite stereo images," *Proc. ISPRS*, vol. 38, 2009.

[2] U. Soergel, E. Michaelsen, A. Thiele, E. Cadario, and U. Thoennessen, "Stereo analysis of high-resolution SAR images for building height estimation in cases of orthogonal aspect directions," *ISPRS J. Photogramm.*vol. 64, pp. 490–500, Feb. 2009.

[3] S. Ahmadi, M. J. V. Zoej, H. Ebabi, H. A. Moghaddam, and A. Mohammadzadeh, "Automatic urban building boundary extraction from high resolution aerial images using an innovative model of active contours," *Int. J. Appl. Earth Obs.*, vol. 12, pp. 150–157, Feb. 2010.

[4] M. Kabolizade, H. Ebadi, and S. Ahmadi, "An improved snake model for automatic extraction of buildings from urban aerial images and LiDAR data," *Computers, Environment and Urban Systems*, vol. 34, pp. 435–441, Apr. 2010.

[5] M. Ortner, X. Descombes, and J. Zerubia, "Building outline extraction from Digital Elevation Models using marked point processes," *Int. J. Comput. Vision*, vol. 72, no. 2, pp. 107–132, 2007.

[6] M. Awrangjeb, M. Ravabaksh, and C. S. Fraser, "Automatic detection of residential buildings using LIDAR data and multispectral imagery," *ISPRS J. Photogramm.*, vol. 65, pp. 457–467, Jul. 2010.

[7] G. Finlayson and S. Süsstrunk, "Optimization for hue constant RGB sensors," in *Proc. IS&T/SID 10th Color Imag. Conf.*, 2002, vol. 10, pp. 343–348.

[8] E. Salvador, A. Cavallaro, and T. Ebrahimi, "Shadow identification and classification using invariant color models," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2001, vol. 3, pp. 1545–1548.

[9] H. Y. Yu, J. G. Sun, L. N. Liu, Y. H. Wang, and Y. D. Wang, "MSER based shadow detection in high resolution remote sensing image," in *Proc. ICMLC*, 2010, pp. 780–783.

[10] D. Cai, M. Li, Z. Bao, Z. Chen, W. Wei, and H. Zhang, "Study on shadow detection method on high resolution remote sensing image based on HIS space transformation and NDVI index," in *Proc. 18th Int. Conf. Geoinf.*, Jun. 2010, pp. 1–4.

[11] H. Ma, Q. Qin, and X. Shen, "Shadow segmentation and compensation in high resolution satellite images," in *Proc. IEEE IGARSS*, Jul. 2008, vol. 2, pp. 1036–1039.

[12] A. Makarau, R. Richter, R. Müller, and P. Reinartz, "Adaptive shadow detection using a blackbody radiator model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2049–2059, Jun. 2011.

[13] T. Kasetkasem and P. K. Varshney, "An optimum land cover mapping algorithm in the presence of shadow," *IEEE J. Select. Topics Signal Process.* vol. 5, no. 3, pp. 592–605, Jun. 2011.

[14] B. Abdel Latif, R. Lecerf, G. Mercier, and L. Hubert-Moy, "Preprocessing of low-resolution time series contaminated by clouds and shadows," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2083–2096, Jul. 2008.