Integrated Approach for Building Extraction from InSAR and Optical Image using Object Oriented Analysis Technique

Bikash Kumar Karna[1] & Ashutosh Bhardawaj[2]

[1]Survey Officer, Survey Department, MinBhawan, Kathmandu, Nepal.
[2]Scientist/Engineer, Indian Institute of Remote Sensing, Dehradun, India.

Abstract

Building extraction in built-up area is of great interest for visualization, simulation and monitoring urban landscape which is used for town/city planning as well as regional planning. Building extraction in urban areas based on merely a single high resolution optical data is often hard to conduct and to improve quality of building detection with consistency, completeness and correctness. Optical images are one of the major sources of individual building extraction from orthoimage but most of these do not produce anticipated result especially to building's shape and outlines in dense urban environment. Extraction of objects from InSAR images is a complicated phenomenon for interpretability due to side looking geometry and effects of layover, foreshortening, shadowing and multi bounce scattering. In this study, buildings and building blocks are extracted from fusion of optical and InSAR data using object oriented analysis (OOA) technique. The improvement of building footprint has done with rectangular fit for building hypothesis and building height from normalized digital surface model (nDSM) based on fuzzy membership function. The results of building extraction has found reasonably good and accurate in planned urban layouts. The quality of building extraction has highly dependent on settlement density, contrast and other image characteristics.

Keywords

SAR interferometry, orthoimage, object oriented analysis technique, building extraction.

1. Introduction

Building extraction in built-up area is of great interest for visualization, simulation and monitoring urban landscape which is used for town/city planning as well as regional planning. Building extraction in urban areas based on merely a single high resolution optical data is often hard to conduct and to improve quality of building detection with consistency, completeness and correctness. With the advancement in technology and increasing general awareness among people, the need for using high resolution and Very High Resolution Satellite (VHRS) optical images for various applications is realized and hence the ever escalating demand for high resolution and VHRS optical images. The current and future high resolution earth observations satellites having a spatial resolution of 1m and higher resolution with stereo capabilities can go a long way in the field of mapping in terms of cost, time and accuracy. Nowadays, there is increasing interest of high resolution and VHRS optical images like IKONOS, Geoeye-1&2, Worldview-1/2, Quick bird, Cartosat-1/2, which have opened up the opportunities for exploiting man-made structures in urban area like building, roads are visible in more detail independently; are used in research activities and implementation of development works. Synthetic Aperture Radar (SAR) has become a key remote sensing technique, at present, due to its all weather capability, independence of daylight and its ability to penetrate into the objects. The space borne satellites like TerraSAR-X, TanDEM-X, SAR-Lupe, RISAT-2 and Cosmo-SkyMed-1/2/3/4 are providing VHRS images in fine grid of geometric resolution of one meter. Additionally, the appearance of a particular

buildings in satellite images depends on look angle of the sensor so that building; which has not oriented in azimuth direction with respect to the sensor are often hard to detect, for this, Interferometric SAR (InSAR) image acquisition from two orthogonal flight line direction has been introduced (Thiele et al., 2007). InSAR has the advantages of providing height information of buildings which facilitates for detection for 3D building model. Multi-aspect InSAR data from four aspect is used for model based building detection and reconstruction (Bolter and Leberl, 2000). Building extraction in urban areas based on merely a single high resolution optical data is often hard to conduct. Features of additional sources of data may be needed to improve quality of detection, completeness and correctness. Several works have already dealt with the integration of features derived from high resolution optical data and SAR (InSAR) data with the goal of building detection. Building blocks have been detected and reconstructed with combining high resolution optical and InSAR data. The classification of both data sets separately within a multi-layer neural network for building hypothesis with height of building using rectangles fit. Building detection has reconstructed semi automatically with the combine features of high resolution optical satellite imagery with high resolution SAR data. Building hypothesis of optical data are validated or rejected based on a classification of the SAR image making use of roof textures, bright lines and shadows. Building and its height are derived simultaneously exploiting the different optical and SAR sensor geometries (Khim, N., 2011).

2. Study Area and Data Description

Study Area:

The study area is Dehradun; which is located in the Garhwal region with geographical extent 30°17′49.42″N to 30°18′22.65″N latitude and 78°03′45.04″E to 78°04′21.95″E longitude. It is capital city of Uttarakhand with 236 km north from India's capital city New Delhi and is one of the alternative developed urban centres for population growth. It is situated in the Doon Valley on the foothills of the lesser Himalayas and Churiya Range (Shivalik Range) hill. It is nestled between two of India's mightiest rivers; the Ganges on the east and the Yamuna on the west. The city is famous for its picturesque landscape and slightly milder climate and provides a gateway to the surrounding region. The city is located at altitude of 600 m above MSL. It is a crowded city, more or less having typical characteristics like planned and unplanned built up areas with many buildings, roads, parking areas and some areas with vegetation, such as trees, shrub and tropical forest. It is well connected and in proximity to popular Himalayan tourist destinations. The location diagram of study area is shown in figure 1.

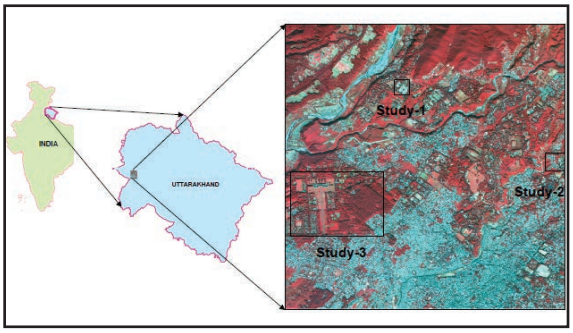

Figure1: Study Area

Data Description:

The following datasets were used in the study; these are listed below in Table 1.

Data Type	Year	Resolution	Source
Remotely Sensed Data:			
Cartost-1 Stereo Image	2006	2.5m	DOS
ALOS PALSAR Image	2010		IIRS
Quick Bird Image	2012	2.4m	IIRS
SRTM DEM	2004	90m	
Base Map:			
Topo map	1996	1:50000	SOI
Field Data:			
Ground Control	2012		Field Works
Ground Truth	2012		Field Works

3. Methodology

For the extraction of building and building blocks, it requires building footprint which has obtained from orthoimage of the very high resolution satellite image then its improvement and refinement has done with building height information and additional different indices such as NDVI, NDWI, SSI etc. For that, firstly orthorectification of the optical images is generated then building height information is estimated from DSM and DEM which has generated from SAR interferometric process. Lastly, building footprint is extracted by the OOA method. By combining the extracted building footprint and the estimated building height information, 2.5D building is modeled and the height of 2.5D building model is assessed with reference building height, generated by using the Cartosat-1 and Aster DSM along with the field visit with a clinometer.

3.1 Orthoimage Generation

Orientation of images is the reconstruction of image object position and its relationship i.e. the model of the earth surface at the time of image acquisition. Interior orientation defines the internal geometry of a camera or sensor as it existed at the time of image capture. In this process, the variables associated with image space are defined and primarily used to transform the image pixel coordinate system to the image space co-ordinate system. Exterior orientation defines the position and angular orientation of the camera that captured an image. The variables defining the position and orientation of an image are referred to as the elements of exterior orientation. The elements of exterior orientation define the characteristics associated with an image at the time of exposure or capture. The angular or rotational elements of exterior orientation describe the relationship between the ground space coordinate system (X, Y, and Z) and the image space coordinate system (x, y, and z). These are defined with roll, pitch and yaw (Grodecki and Gene, 2003). Generally in the satellite image processing, RPC file with ground control point (GCP) collected from differential global position system (DGPS) for refining orientation parameter. Stereo model is displayed and 3D co-ordinate measurement is done with entering the GCPs in the block of stereo model in proper location in all images and corresponding co-ordinate in the ground. The image matching technique with cross correction coefficient and different statistical measures has applied to determine the matching point followed by generating epipolar images which includes: block configuration, tie point extraction, point transfer, mass point, gross error detection, and tie point selection. The ground coordinate of these points is calculated with bundle block triangulation (satellite triangulation) and error is adjusted using by least square adjustment.

Accuracy assessment estimate the probability of error happens in the work for the quality control of work is achieved in whole process. For the accuracy assessment GCPs are collected in sufficiently well distributed for training set of sample to estimate the error occurs in work. Accuracy is usually evaluated in terms of Root Mean Squared Error (RMSE) of residuals between imagery derived coordinates with respect to GCPs coordinates. If the positional accuracy is within the acceptable limit of 1 pixel of RMSE, then DSM has been generated. With GCPs and tie points in the model, image points coordinate into 3D co-ordinates of the ground features is computed as DTM mass points using space forward intersection technique in Leica Photogrammetry Suite (LPS) software and automatic DSM production has generated. DEM has generated from DSM using opening filter. nDSM is generated from DEM subtracted from DSM. Orthoimage production process consists of two steps: orthorectification and resampling. Orthorectification process has done using oriented image with GCPs and DEM of the image area. Resampling process computes the new geometric and radiometric properties of the image after orthorctification (Schenk.T., 1999). Orthoimage is free from relief and tilt displacement, scale distortion and as a like the plannimetric map (Toutin, 2004).

3.2 DSM generation by SAR

Interferometry InSAR processing has done with SARscape in ENVI 4.3 software as the following approach.

Baseline Estimation:

Baseline is an important parameter of the process of InSAR, which directly influences the accuracy of DEM generated and based on data selection in terms of spatial as well as temporal. DSM generation from InSAR images is dependent on the perpendicular baseline in such a way that 2π ambiguity height (m) determines the sensitivity of the interferogram.

Height ambiguity (ha) based on the perpendicular baseline (B) is calculated as:

$$h_a = \frac{\rho\lambda \sin\theta}{2B\cos(\theta-\alpha)} \quad (1)$$

Where λ is the radar wavelength, θ is the look angle from nadir and α is the baseline orientation angle (Xu et al., 2001).

Co-registration:

SAR interferometry requires pixel-to-pixel match between common features in SAR image pairs. Thus co-registration is an essential step for the accurate determination of phase difference and for noise reduction and the alignments of SAR images from two antennas. The entire purpose of the co-registration is to align the samples for phase differencing. The correlation window is used to search for offsets between master and slave images at the sub pixel level co-registration (Sahraoui et al., 2006).

Multi-looking:

Multi-looking processing has been used to reduce phase noise in SAR interferometry processing. Phase noise gives a difficulty to create DEM from an interferogram. If the phase noise, which is mostly caused by radar thermal noise, speckle due to coherent SAR processing, decorrelation, and registration noise, is too strong, some fringes of an interferogram is totally lost which have cause errors in InSAR DEM.

Interferogram Calculation:

After the two single-complex images are co-registered, the interferogram is computed according to:

$$I = S_1 . S_2^* = |S_1| . |S_2^*| . e^{j\varphi} \quad (2)$$

Where S1 and S2* are the corresponding complex values of the co-registered images. Interferometric phase can be clculated as (Yu and Ge, 2010):

$$\varphi = \varphi_f + \varphi_t + \varphi_o + \varphi_d + \varphi_a + \varphi_n \quad (3)$$

where f is the flat earth phase, t is the topographic phase, o is the phase from the objects in the earth surface, d is the deformation phase, a is the atmospheric delay phase and n is the noise from measurement and can be removed by using an Interferogram filtering method and a multi-looking (resampling) process. Since the temporal baseline has 46 days interval in the study is too large, so there is also deformation phase. If the atmospheric delay phase can be ignored, then Equation 3 reduces to:

$$\varphi = \varphi_{flat} + \varphi_{topo} + \varphi_{obj} \quad (4)$$

Interferogram Flattening:

The constant phase (due to the acquisition geometry) and the phase expected for a flat Earth or for a known topography are separated from the interferometric phase. Cartosat DSM has used for interferogram flattening which split the initial interferogram into two components: synthetic phase and flattened interferogram.

Coherence Generation and Adaptive Filtering:

Given two co-registered complex SAR images (S1 and S2), the interferometric coherence (γ) is defined as the absolute value of the normalized complex cross correlation between the two signals (Cloude and Papathanassiou, 1998)

$$\gamma = \frac{\left|\sum_{n=1}^{N} S_1^{(n)} . S_2^{*(n)}\right|}{\sqrt{\left(\sum_{n=1}^{N} |S_1^{(n)}|^2 . \sum_{n=1}^{N} |S_2^{(n)}|^2\right)}} , 0 \le \gamma \le 1 \quad (5)$$

where N is the number of pixels in the moving window for coherence estimation, and which is complex numbers and the complex conjugate.

Phase Unwrapping with Region Growing Method:

The unwrapping phase consists the redistributing with each pixel its absolute phase. Two algorithms (region growing, minimum cost flow) are used for phase unwrapping process; in essence, none of these is perfect and different or combined approaches should be applied on a case by case basis to get optimal results. In two unwrapping algorithms, the accuracy of the result depends on the path chosen to perform the unwrapping. Phase Unwrapping is the process that resolves this 2π ambiguity.

Orbital Refinement:

Orbital refinement is used for registration process of the interferometric geometry which is the most important steps in DSM generation process. The accuracy of the registration is depended on the precision baseline as well as precision of error in the horizontal and vertical obtained in an InSAR DSM. It is thus important to precisely reconstruct the orbital trajectories of the satellites in order to know their position exactly. Also, the orbital parameters of the satellites are not accurate enough to allow a rigorous transformation from the phase to heights, the orbital refinement is needed (Singh, 2003). For optimizing, 14 GCPs from DGPS are used to calculate the correction parameters, was selected on the flattened interferogram. The GCPs must be well distributed throughout the entire scene.

Phase to Height Conversion and Geo-coding:

The absolute calibrated and unwrapped phase combined with the synthetic phase is converted to height with geo-coded into a map projection. Geo-coding was done by considering the WGS-84 datum, UTM 43N map projection.

3.3 Building Foot Print using OOA technique: Visual Image Interpretation:

The visual image interpretation was carried out using the visual image interpretation element such as shape, size, association, shadows, texture and tone. There is difficulty for the extraction of object correctly for the spectral complexity of urban areas where different neighboring objects, such as building, parking areas or road, are constructed with same materials or an object are made from different materials, e.g. a building is built with concrete, metal, bricks, tiles or synthetic materials and mixed to each other.

Multiresolution Segmentation:

Multi-resolution segmentation algorithm is based on a bottom-up region-merging process starting with one-pixel to form image objects. Smaller image objects are subsequently merged into larger ones, forming segmentations with objects in different scales. The merging decision is based on a homogeneity criterion i.e. having similarity among adjacent segments. This homogeneity criterion is defined as a combination of spectral and shape factor. In the homogeneity criterion, three parameters are computed: scale, color and shape. The scale parameter is applied to determine the maximum acceptable heterogeneity for the resulting image objects. Scale parameter are directly proportional to the size of image objects i.e. scale parameter is large size then bigger image objects in size but less in numbers and vice versa. Further, homogeneity criterion depends on a combination of color and shape parameters. The value of the shape factor modifies the relationship between shape and color criteria are modifying by the shape and color criteria. Shape factor is the combination of the compactness and smoothness. The value of the compactness factor modifies the relationship between compactness and smoothness criteria are modifying by the compactness and smoothness (Definens, 2007).

Rule Based Classification:

Different characteristics of these segments like spectral characteristics, shape, orientation, texture, proximity or adjacency to other objects can be calculated to develop rule sets which are used to classify the segments. For classification of distinct objects exactly, some customized rule set were created based on feature properties analysis using spatial, textual, contextual information, nDSM and spectral indices such as NDVI and SSI. Using a threshold value of NDVI, the vegetation feature can be segregated to other feature. Similarly, water bodies can be extracted segregated by using NDWI threshold value. The shadow can be classified by using SSI, which is quite effective to distinguish shadows in high resolution images. After removing the vegetation, water bodies and shadow; the taller objects/ features compare to other features are considered as building. So, the height information from the nDSM can be applied for extraction of the taller objects in the image applying threshold of nDSM. But, the buildings extracted from this way can be roughly distinguished because the nDSM is not accurate enough and the uncertainty of the DSM/ nDSM. But, some parking area, bare ground or road areas may be also classified as building features, whereas some existing building areas may not be classified in the building class as well. For this instance, area (spatial property), brightness (spectral property) and relative border to buildings (contextual property) are used to differentiate buildings from remaining other objects. On the other hand, the actual buildings which were unclassified can also be

reclassified by utilizing the similar spatial, spectral and contextual information or a combination of these properties (Definens, 2007).

4. Results

Orientation (interior and exterior) of Cartosat-1 was done with the RPCs and 14 GCPS (3 GCPs for full control, 5 GCPs for Horizontal control, 2 GCP for vertical control and 4 GCPs for check point). With image matching technique 51 automatic tie points were generated and used for controlling the entire scene for the satellite triangulation process within the bundle block adjustment and adjustment was carried out applying the least square adjustment technique. The satellite triangulation error has found 0.70 pixels which has acceptable limits. DEM generated for single from Cartosat-1. Orthoimage of the Quick bird image were produced by using DEM.

The SAR interferogram was generated by cross-multiplying, sub-pixel by sub-pixel, the first SAR image with the complex conjugate of the second. Its interferometric phase is the phase difference between first and second images. The synthetic phase was subtracted from the interferometric phase, the flattened interferogram (φ remaining) was calculated using Cartosat-1 DEM and coherence image was generated. The flattened interferogram is only modulo 2π; phase unwrapping process was done to this 2π ambiguity of the interferogram and unwrapped phase image was generated.

The orbital refinement was done using 10GCPs in well distributed location for refining orbital inaccuracies and calculating the absolute phase values. RMSEs were computed for the orbital refinement from interferogram output with the ellipsoidal height and different levels of the external Cartosat-1 DEM and found 10.532m. The phase to height conversion was done for reconstruction of InSAR DSM in 10m grid size on the ground range geometry with UTM projection in 43 Zone and vertical height value in meters with reference to WGS-84datumis shown in Figure 2.

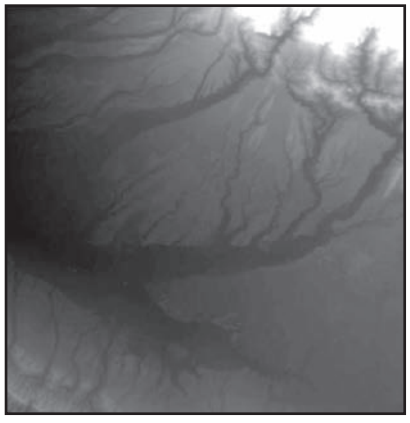

Figure2: DEM from ALOS PALSAR

The buildings in the image have different spectral properties and not pre- defined parameters of scale, shape, color which were equally applicable to get best fitting segments for all building footprints. In this study, a set of parameters was chosen by trial and error method, for the extraction of building footprint which fits practically in the image. Some considerations have taken in mind that the image has not be over or under segmented and over segmentation. The multi-resolution segmentation, scale 20, shapes 0.3 and compactness 0.6 was used as a parameter set (in Figure 3).

Figure 3: Multi-resolution Segmentation

In rule based classification, nDSM was used for separation of taller objects in the image which may be vegetation or building. Vegetation was separated by using NDVI and shadow was extracted using SSI and NDWI. The refinement was done with the membership function of the fuzzy logic with some threshold and optimization of the rule set in feature space of spectral (brightness, standard deviations,

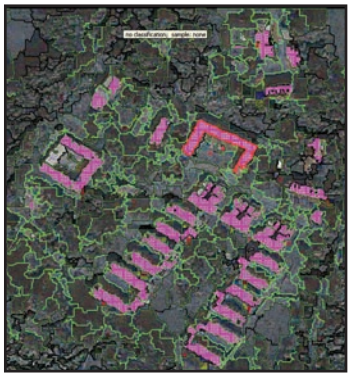

Figure 4: Refinement of Building Footprint

max. diff.), spatial (asymmetry, area, compactness, rectangular fit) and contextual (relative boarder to building, relative boarder to vegetation, relative boarder to open space) properties. The refining building classified result is shown in Figure 4.

The refined building features were export in vector format in polygon feature and finally get the actual extracted building (in Figure 5).

Figure 5: Extracted Building Footprint

The overall accuracy and kappa statistic compared to sample used for classification were directly computed and found 92.3% and 0.829 respectively. For the position based and object based accuracy assessment, reference building in this area were digitized on the ortho-rectified Quick bird image. The position of the extracted buildings was calculated on the basis of building centroid which was identified in both, reference and extracted building. The displacement between reference and extracted building was measured and found that the minimum displacement was 0.671m and maximum up to 2.259m with the average of 1.796m. For object based accuracy assessment, the number of buildings in the reference image and extracted by OOA technique were compared. The number of correctly extracted, missed and wrongly extracted buildings were analyzed and identified in spatial analysis tool within ArcGIS 9.3. The different quality measure such as correctness, completeness and extracted percentage were calculated. The result is shown in Table 2.

Accuracy Measure	No. of Building
Building in Reference Image	48
Extracted Building	42
True Positive(TP) Building	41
False Positive(FP) Building	-
False Negative(FN) Building	1
Split Factor	-
Missing Factor	0.0244
Completeness (%)	97.62
Correctness (%)	100
Error of Omission (%)	2.38
Error of Commission (%)	0
Overall Accuracy (%)	98.81

5. Conclusion

Building extraction in urban area using OOA technique depends upon the spectral, spatial and contextual information. Proper segmentation of image is always helpful for accurate building extraction and quality of extraction is not completely depends upon the initial segmentation. The spectral variability within an object and/or a class and spectral similarity among different class, the extraction of building is not get straightforward in dense urban area in VHRS image. The shape, size and contextual information are useful for distinguish different feature characteristics of image objects and nDSM plays important role in building extraction with comparing height of the image objects. The result of extracted building are fairly good quality having regular structure in dense, complexity of urban structure with less irregular pattern and shadow effect.

References

Bolter, R. and Leberl, F., 2000, "Detection and reconstruction of human scale features from high resolution interferometric SAR Data", Proceedings of the International Conference on Pattern Recognition 2000 (ICPR2000), pp. 291-294.

Cloude S. R. and Papathanassiou K. P., 1998; "Polarimetric SAR interferometry," IEEE Transactions on geoscience and remote sensing, vol. 36, pp. 1551-1565.

Definiens, Developer 7.0 Reference book. Definiens AG, Trappentreustr. 1, D-80339 Munchen, Germany,2007.

Grodecki, j. and Gene, D., 2003; “Block Adjustment of high resolution satellite images described by rational polynomials”, Photogrammetric Engineering and Remote Sensing, 69(1), 59-68.

Khim, N., 2011, “2.5D Reconstruction of building from very high resolution SAR and optical data by using object-oriented image analysis”, M.Sc. Thesis, Geoinformatics, ITC, University of Twente, Enschede.

Sahraoui O., Hassaine B., and Serief C., 2006; "Radar Interferometry with Sarscape Software," Photogrammetry and Remote Sensing.

Schenk,T., 1999; “Digital Photogrammetry”, I. Terra Science, Laureville pp 422.

Singh L. P., 2003; "Application of SAR interferometry in landslide studies with special reference to generation of input data for statistical susceptibility assessment", MSc thesis, ITC, University of Twente, Enschede.

Thiele, A., Wegner, J. and Soergel, U., 2010; "Building reconstruction from multi-aspect InSAR data", Radar Remote Sensing of Urban Areas, vol. 15, pp. 187 - 214.

Toutin, T., 2004; “Geometric processing of remote sensing images: Models, algorithms and methods” International Journal of Remote Sensing, Vol. 25: 1893-1924.

Xu , H., Zhou Y., and. Li, C., 2001; "Analysis and simulation of spaceborne SAR interferometric baseline", in Radar, CIE International Conference on, Proceedings, pp. 639-643.

Yu J.and Ge L., 2010; "Digital elevation model generation using ascending and descending multi-baseline ALOS/PALSAR radar images", presented at the FIG Congress 2010, Sydney, Australia,

Principal Author's Information

Name	:	**Mr. Bikash Kumar Karna**
Academic Qualification	:	P.G. Diploma in Remote Sensing & GIS
Organization	:	Survey Department
Current Designation	:	Survey Officer
Work Experience	:	14 Year
Published Papers/Articles	:	One
Email	:	bikashkumarkarna@gmail.com