# A novel technique for mapping the disparity of off-terrain objects

**A Suliman[1], Y Zhang[1] and R Al-Tahir[1]**

University of New Brunswick, Department of Geodesy and Geomatics Engineering
15 Dineen Drive, Fredericton, New Brunswick, E3B 5A3, Canada

E-mail: A.Suliman@unb.ca

**Abstract.** Third-dimension information is of a great importance for several remote sensing applications, such as building detection. The main data-source for these applications is very high resolution (VHR) satellite images which allow detailed mapping of complex environments. Stereo VHR satellite images allow the extraction of two correlated types of third-dimension information: disparity and elevation information. While the disparity is measured directly, the elevation information is derived computationally. To measure the disparity information, two overlapped images are matched. However, for the backward and forward off-nadir VHR stereo images, building facades occlude areas and hence create many data gaps. When the disparity is required to represent only the off-terrain objects, interpolation and normalization techniques are typically used. However, in dense urban environments, these techniques destroy the quality of the generated data. Therefore, this paper proposes a registration-based technique to measure the disparity of the above-ground objects. The technique includes constructing epipolar images and registering them using common terrain-level features to allow direct disparity mapping for the off-terrain objects. After the implementation, the negative effects of occlusion in the off-nadir VHR stereo images are mitigated through direct disparity mapping of the above-ground objects and bypassing the interpolation and normalization steps.

Key words: Disparity information, Interpolation, Normalization, Off-nadir VHR images, Off-terrain objects, Registration.

## 1. Introduction

Third-dimension (3D) information is of a great importance for several remote sensing applications, such as 3D city modeling, image classification, and building detection. The currently available remote sensing satellite systems provide very-high-resolution (VHR) images with sub-meter ground sampling distance, broad coverage, and rich information content. Additionally, these systems can capture stereo images using along- and across-track off-nadir acquisition capabilities. All of these properties make VHR images an ideal data source for modeling and mapping applications. However, off-nadir VHR images acquired over urban areas suffer from severe building lean and occluded areas.

Extracting the third-dimension information from remote sensing images is an essential goal for the photogrammetric approaches. These approaches apply image measurements to multiple stereo images to generate elevation models that describe the visible surface, which are known as digital surface models (DSMs). Normally, these measurements are made automatically using image correlation techniques that match corresponding pixels in a pair of stereo images. To increase the efficiency of

these techniques, the image stereo-pairs are epipolar rectified to linearize the search areas and measure the disparity for only one direction. Epipolar images have all the corresponding pixels lying on the same rows in both images. Therefore, the x-disparity (also known as the x-parallax) of a pixel is simply the difference between its column locations in the right and the left epipolar images. The generated surface disparity models (SDMs) have similar representation to their corresponding DSMs. However, when off-nadir VHR images of urban areas are used, the measured SDMs suffer from many no-data regions due to building lean and occlusion, which make it impossible to find point matches in these hidden areas.

The third-dimension information that represents only the off-terrain objects could be calculated by normalizing the extracted stereo information (i.e., DSM/SDM) of the visible surface that includes both the terrain and off-terrain objects. The normalization process is classically made by removing the terrain variation from the surface model. Thus, the terrain information models need to be extracted first and then subtracted from the corresponding surface information models. The normalized elevation (or disparity) surface models (nDSM/nSDM) resulting from this normalization process represent the third-dimension information for only the off-terrain objects. The normalized stereo information is critical for urban mapping applications. However, when the generated stereo information has many data gaps, due to the use of off-nadir VHR images of urban areas, the building detection quality will be destroyed. This is due to the misleading information resulting from interpolating the data gaps in complex urban environments.

For the purpose of normalizing the surface disparity models of urban regions, a literature review was conducted on the topic of terrain information extraction in urban areas. Krauß, Arefi and Reinartz [1] compared and evaluated the steep-edge-detection algorithm [2] against the most successful algorithms for terrain model extraction in urban areas. These algorithms are the classical morphological approach [3] and the Geodesic dilation [4]. Although the steep-edge-detection algorithm is the most efficient in urban areas, it assumes that the roads/streets are adjacent to the building steep edges. This assumption, unfortunately, is not guaranteed in off-nadir VHR images over dense urban areas. A more general terrain extraction algorithm was developed by [5]. This algorithm is based on a moving search window of constant size for detecting local-minima points. The algorithm assumes that the local minima represent the bare earth points within that search area, which in the case of large building roofs falsely detects off-terrain points and hence destroys the constructed terrain information. Besides the extra computation, all of the compared algorithms are based on assumptions that limit their effectiveness in extracting terrains from incomplete urban disparity maps when the occlusion in off-nadir images is severe.

In summary, the occlusion resulting from the building lean and facades in off-nadir VHR stereo images acquired over dense urban areas is usually severe. The occluded areas will result in data gaps which typically require executing interpolation to fill them. However, interpolating gaps in dense urban areas will result in misleading terrain disparities, over narrow streets and between buildings, which destroy the quality of the subsequent processes. Therefore, if the original stereo images are epipolar rectified and co-registered for the corresponding ground objects (e.g., roads), the remaining measurable x-disparity should represent only the off-terrain objects (i.e., nSDM). Thus, the aim of this research is to develop a technique to directly measure normalized disparity information.

The remainder of this paper is organized as follows. Section 2 describes the proposed technique. Section 3 provides the experimental results and discussion. Finally, Section 4 gives the conclusions.

## 2. The Proposed Technique

The key concept of the proposed technique is to prepare the original VHR stereo images in a way that allows the matching techniques to measure the disparity of only above-ground objects. This technique begins by projecting both images onto an object-space plane that minimizes the effects of the terrain variation and allows co-registration of terrain-level objects. Then, the projected images are re-oriented to construct an epipolar stereo pair. The epipolar geometry of the linear array sensors is comprehensively discussed in [6] and [7]. The two images of the epipolar pair are co-registered using

terrain features such as roads to eliminate their disparity in the x-direction. Finally, an efficient matching technique must be used to map the disparity of the off-terrain features. As investigated in [8], semi-global matching, as introduced by Hirschmüller [9], is strongly recommended. The whole process of the proposed technique to directly generate normalised disparity maps is flowcharted in four steps in figure 1.
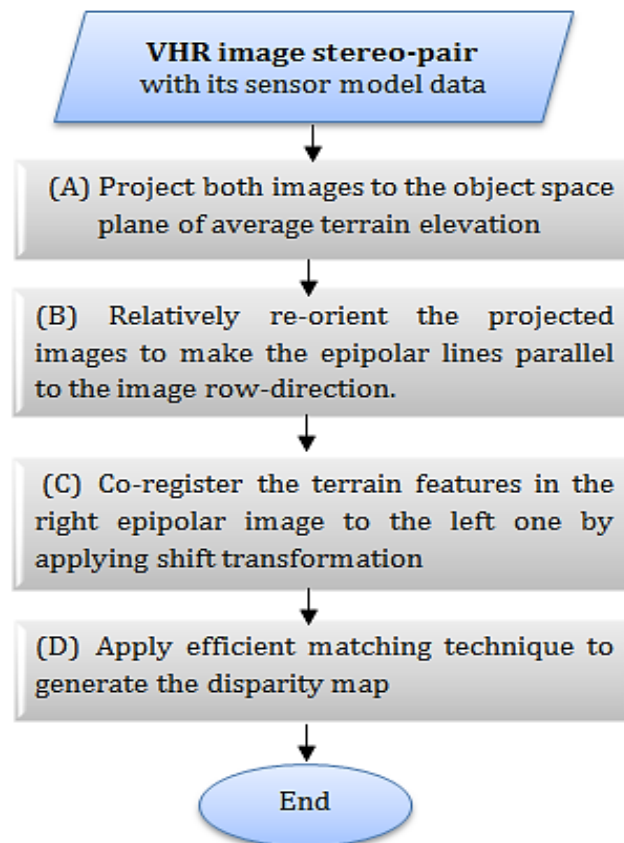


**Figure 1.** The proposed registration-based technique for direct disparity map generation of above-ground objects.

## 3. Experimental Results and Discussion

### 3.1.  Dataset and Study Area
The optical data used in this work are a subset of stereo VHR images acquired by a linear array sensor with push-broom scanning mode. These stereo images were taken by the Pleiades-1A satellite sensor over a dense urban area of Melbourne, Australia. The two opposite acquisition angles of the forward and backward off-nadir images are approximately +15 and -15 degrees respectively. The product ground resolution of the pan-sharpened VHR image bands is 0.5 m/pixel.

### 3.2.  Results of the Proposed Technique
The result of the co-registration of the stereo epipolar pair is illustrated in figure 2. The evaluation of this co-registration was made visually relative to straight terrain-level roads. The leaned buildings in the left epipolar image do not appear in the right one. a1, b1, c1, and  d1 are four road intersections in Ep1 co-registered with their corresponding ones a2, b2, c2, and d2 in Ep2.  For visual assessment of

the co-registration achieved, straight lines (yellow thin lines) representing the road center lines (i.e., terrain-level features) are drawn in Ep1 and transfered to Ep2. The intersections of these lines were used to calculate the average shifts applied to register the two epipolar stereo images. By overlaying the co-registered epipolar images on top of each other, visual inspection at the indicated road intersections reveals a co-registration accuracy of less than ± 4 pixels. This acceptable quality of the terrain-level co-registration is mainly the result of minimizing the terrain-relief displacement.
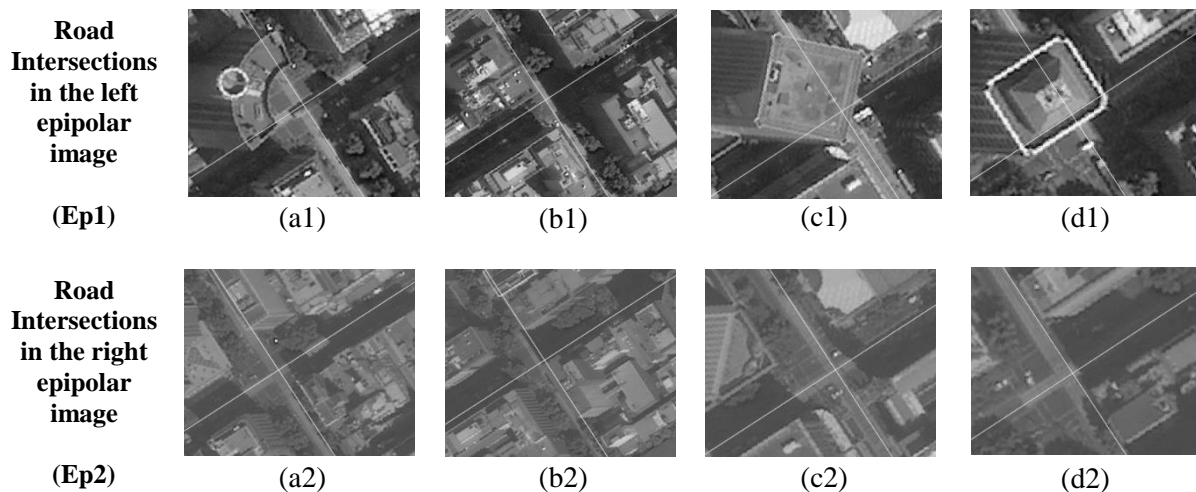
| | | | | |
|---|---|---|---|---|
| **Road Intersections in the left epipolar image**<br><br>**(Ep1)** | (a1) | (b1) | (c1) | (d1) |
| **Road Intersections in the right epipolar image**<br><br>**(Ep2)** | (a2) | (b2) | (c2) | (d2) |

**Figure 2.** Results of co-registering the terrain objects (i.e, roads) of the right epipolar image (Ep1) with the corresponding one in the left epipolar image (Ep2).

Figure 3 represents the generated disparity map using the SGM algorithm. The pixels of zero values (black pixels) between buildings indicate the removed terrain objects. From visual inspection, the terrain-level roads and streets between buildings are all filtered out. The pixels of non-zero values represent the measured disparity of the above-ground objects. This demonstrates for the proposed technique that the terrain-relief distortion was minimized, and therefore almost normalized disparity values can be measured.

It is worth noticing that the measured disparity map represents the building roofs only. This resulted from the image acquisition geometry of the off-nadir stereo VHR images. The off-nadir stereo images were of opposite angles (i.e., +15$^o$ and -15$^o$) acquired using a linear array sensor with push-broom scanning mode. This acquisition geometry does not allow building facades in one image to appear in the other one. This occlusion type is commonly considered to be a problem. However, it became an advantage in our proposed registration-based technique.

## 4. Conclusion

For the backward and forward VHR stereo images, the same building facades cannot be seen in both images. Thus, many data gaps will result due to the matching failure. Hence, results of the mapping applications that require the third-dimension information of the above-ground features will be degraded in quality. Therefore, this paper developed a registration-based technique to directly measure the disparity of the above-ground objects. This technique mitigates the negative effects of building lean, facades, and occlusions in the off-nadir VHR stereo images acquired over dense urban areas.

The developed technique produces a disparity map only for the above-ground features. It is based on co-registered epipolar images using the terrain-level features to be filtered out, allowing direct measurement of the normalized disparity data.

After the implementation, the technique was found to be efficient, as the terrain relief distortions were dramatically reduced after the average level-based rectification. This allowed acceptable co-registration for the terrain-level objects in both epipolar stereo images. The generated disparity map was found to represent mainly the building roofs, which indicates the success of the proposed technique.

The future work of this research is to investigate the validity of the generated map in a disparity-based building detection application. To demonstrate the applicability of the proposed approach, a building detection procedure should be designed, implemented, and validated.
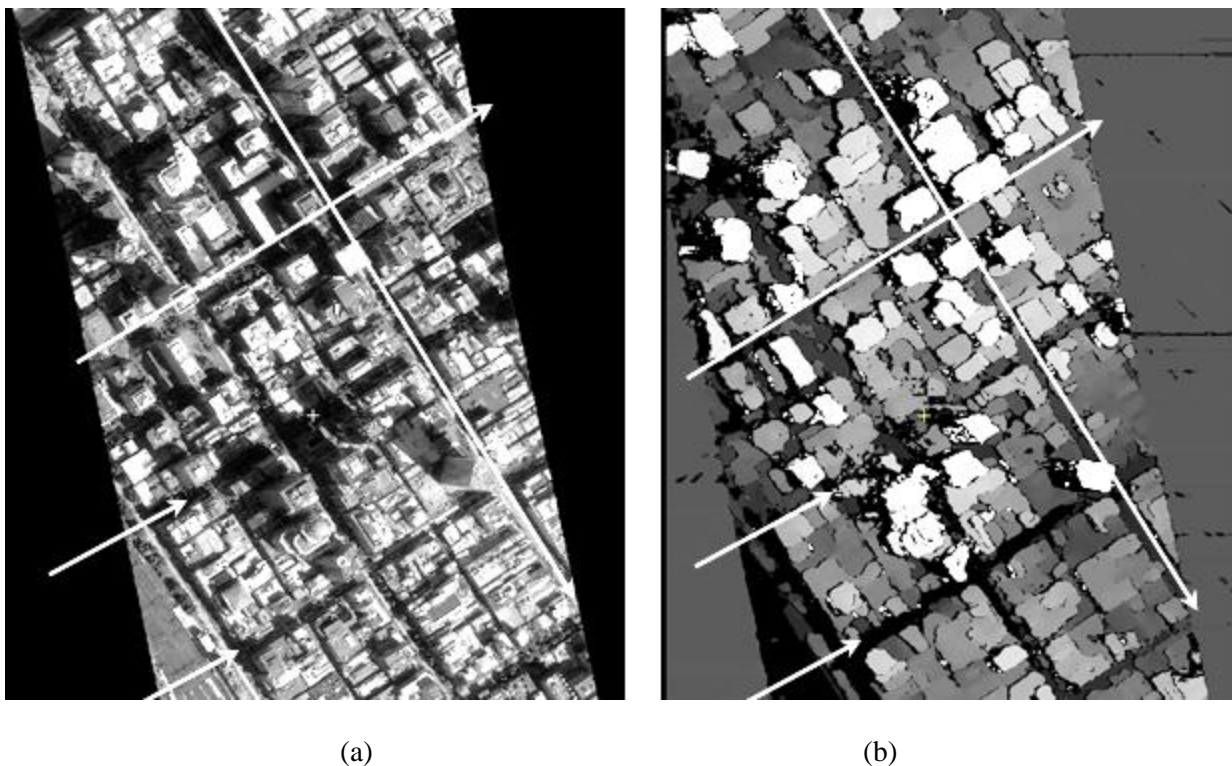


(a)　　　　　　　　　　　　　　　　　　　　(b)

**Figure 3.** (a) The left VHR epipolar imagery. (b) The disparity map generated by executing the SGM matching algorithms on the processed image using the proposed technique. The roads on the ground have minimal disparity values.

**Acknowledgements**

**References**
[1]　Krauß T, Arefi H and Reinartz P 2011 Evaluation of selected methods for extracting digital terrain models from satellite born digital surface models in urban areas. *Int. Conf. on Sensors and Models in Photogrammetry and Remote Sensing (SMPR 2011)*
[2]　Krauß T and Reinartz P 2010 Urban object detection using a fusion approach of dense urban digital surface models and VHR optical satellite stereo data. *ISPRS Istanbul Workshop,* (Istanbul , Turkey: Citeseer)

[3]  Weidner U and Förstner W 1995 Towards automatic building extraction from high-resolution digital elevation models *ISPRS journal of Photogrammetry and Remote Sensing* **50** 38-49

[4]  Arefi H, Engels J, Hahn M and Mayer H 2007 Automatic DTM Generation from Laser-Scanning Data in Residential Hilly Areas. *ISPRS Stuttgart 2007 workshop,* ed J Schiewe and M Hahn (Stuttgart, Germany: ISPRS  International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences)

[5]  Zhang Y, Tao C V and Mercer J B 2004 An initial study on automatic reconstruction of ground DEMs from airborne IfSAR DSMs *Photogrammetric Engineering & Remote Sensing* **70** 427-38

[6]  Habib A F, Morgan M, Jeong S and Kim K-O 2005 Analysis of Epipolar Geometry in Linear Array Scanner Scenes *The Photogrammetric Record* **20** 27-47

[7]  Wang M, Hu F and Li J 2011 Epipolar resampling of linear pushbroom satellite imagery by a new epipolarity model *ISPRS Journal of Photogrammetry and Remote Sensing* **66** 347-55

[8]  Alobeid A, Jacobsen K and Heipke C 2010 Comparison of matching algorithms for DSM generation in urban areas from Ikonos imagery *Photogrammetric Engineering and Remote Sensing* **76** 1041-50

[9]  Hirschmüller H 2008 Stereo Processing by Semiglobal Matching and Mutual Information *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30** 328-41