SATELLITE STEREO BASED DIGITAL SURFACE MODEL GENERATION USING SEMI GLOBAL MATCHING IN OBJECT AND IMAGE SPACE

S. Ghuffar

Geospatial Research and Education Lab (GREL), Department of Space Science,
Institute of Space Technology, Islamabad Highway 1, Islamabad, Pakistan, - sajid.ghuffar@grel.ist.edu.pk

Commission I, WG I/4

KEY WORDS: Stereo, Semi-Global Matching, DSM, High resolution Satellite Images

ABSTRACT:

This paper presents methodology and evaluation of Digital Surface Models (DSM) generated from satellite stereo imagery using Semi Global Matching (SGM) applied in image space and georeferenced voxel space. SGM is a well known algorithm, used widely for DSM generation from airborne and satellite imagery. SGM is typically applied in the image space to compute disparity map corresponding to a stereo image pair. As a different approach, SGM can be applied directly to the georeferenced voxel space similar to the approach of volumetric multi-view reconstruction techniques. The matching in voxel space simplifies the DSM generation pipeline because the stereo rectification and triangulation steps are not required. For a comparison, the complete pipeline for generation of DSM from satellite pushbroom sensors is also presented. The results on the ISPRS satellite stereo benchmark using Worldview stereo imagery of 0.5m resolution shows that the SGM applied in image space produce slightly better results than its object space counterpart. Furthermore, a qualitative analysis of the results on Worldview-3 stereo and Pleiades tri-stereo images are presented.

1. INTRODUCTION

The availability of very high resolution (VHR) satellite stereo imagery continues to increase due to launch of various VHR stereo imaging capable satellites over the past decade. This stereo imagery is now utilized in various applications including 3D city modelling and topographic mapping. In addition to typical two image stereo, satellite vendors deliver tri-stereo imagery as a standard product. These satellites are further capable of acquiring multi-view images from a single pass (d'Angelo and Kuschk, 2012). Stereo imagery with up to 30 cm GSD (Ground Sampling Distance) is now commercially available. To exploit the potential of VHR satellite stereo imagery, robust and accurate dense image matching and 3D reconstruction algorithms are required. This paper, presents and evaluates methods for generating digital surface models from VHR satellite stereo data.

This paper has two main components. Firstly, the methodology of in-house implementation of automatic 3D reconstruction algorithm from satellite pushbroom sensors is presented, which uses semi global matching as the dense matching algorithm. Secondly, a technique to apply SGM directly in the georeferenced voxel space, without the need of stereo rectification and triangulation is presented. The minimum cost at each grid point directly gives the height at each grid element. The evaluation of the proposed methods is performed on the ISPRS stereo matching benchmark, which uses the DSM generated from LiDAR data as ground truth for quantitative evaluation.

2. LITERATURE REVIEW

The three major components of the pipeline for 3D reconstruction from stereo imagery are: stereo rectification, dense matching and triangulation, which generates a 3D point cloud. In satellite stereo imagery applications, this point cloud is often interpolated in to a 2.5D raster or a digital surface model (DSM). In comparison to stereo images from frame cameras, there are two major differences when dealing with satellite stereo imagery. Firstly, the

satellite imaging systems typically acquire images from a pushbroom sensor, which sequentially captures images line by line following the satellite motion. Secondly, the transformation between the image and the object space is given in terms of Rational Polynomial Coefficients (RPC) (Fraser et al., 2006), which should be incorporated in the stereo reconstruction pipeline.

The first step in the 3D reconstruction pipeline is the stereo rectification or epipolar resampling of the satellite stereo images. The process of stereo rectification transform the images in such a way that the corresponding points between the two images lie on the same image row. This simplifies the search of corresponding points and off-the-shelf stereo matching algorithms can be used. The standard procedure of stereo rectification in frame cameras involves estimation of Fundamental matrix using automatic detection of corresponding points like SIFT (Lowe, 2004) and then based on this Fundamental matrix a homomorphy is estimated for each image. In contrast to frame cameras, the epipolar geometry of the satellite pushbroom sensors shows different behaviour. The epipolar curves of the pushbroom cameras are not straight and the conjugate pair does not exist (Morgan et al., 2004; Oh et al., 2010; Habib et al., 2005). Therefore, the commonly used techniques for stereo rectification of frame cameras can't be directly applied for rectification of images from pushbroom sensors. Several techniques have been proposed for pushbroom rectification with the affine camera model approximation being the most popular solution for pushbroom rectification. The affine assumption is suitable due to the following reasons: the height variations in the object space are small compared to the satellite altitude, the angular field of view is small and the satellite is assumed to travel in a straight line with a constant velocity (Morgan et al., 2004; Habib et al., 2005) during the short image acquisition time. An affine transformation is estimated using corresponding points between the two images and the resulting transformation causes the corresponding points to lie along the same image rows. In De Franchis et al. (2014) this affine rectification model is applied on small image tiles of size e.g. 500×500, which results in epipolar error of less than 1/10th of a pixel. Similarly, the results in Wang et al. (2011), show that the vertical parallax from

the affine approximated stereo rectification method is less than 0.4 pixels (RMS value) using data from different satellites. Some authors (Gruen, 1985; Hirschmüller, 2008) use the information from epipolar geometry directly in the dense matching without rectifying the original stereo images.

The second and the most essential component is the dense image matching or the stereo matching. Semi global matching is the most widely used dense image matching algorithm for 3D reconstruction from aerial (Haala and Rothermel, 2012) and satellite stereo (d'Angelo and Reinartz, 2011) imagery and is also used in applications like driver assistance system (Hermann and Klette, 2013). SGM is ranked high on KITTI (Geiger et al., 2012) and Middlebury (Scharstein and Szeliski, 2002) stereo benchmarks. The widespread use of SGM is due to its accuracy, time efficiency and ease of implementation (Drory et al., 2014). In computer vision, there is a large collection of publications on the topic of stereo matching, while most of the work on satellite stereo utilizes SGM with a few exceptions. In Kuschk et al. (2014) a variational stereo matching algorithm is used for DSM generation from satellite stereo data. In Pierrot-Deseilligny and Paparoditis (2006) a multi-resolution maximum flow algorithm is used for dense image matching.

Several authors have presented automatic workflows for generation of DSM from satellite stereo data ((De Franchis et al., 2014; Kuschk et al., 2014; Wohlfeil et al., 2012). The automatic satellite stereo workflows mentioned above use SGM as the dense matching algorithm.

A different class of multi view stereo matching algorithms often referred as volumetric reconstruction algorithms, apply dense matching in object space (Vogiatzis et al., 2007; Roy, 1999). These methods first discretize the object space in to voxels and then project the images on to these voxels and apply some photo-consistency measure. A satellite based multi view reconstruction in object space using SGM has been presented in d'Angelo and Kuschk (2012). Another example of application of SGM in object space is presented in Bethmann and Luhmann (2014). A similar scheme of matching in object space is also implemented in software MicMac developed at French National Survey (IGN) (Deseilligny, 2015), which uses a graph based min cut max flow algorithm (Roy and Cox, 1998) for optimization. The object space variant of SGM presented in this work is similar to the methodology used in d'Angelo and Kuschk (2012) and Bethmann and Luhmann (2014).

3. METHODOLOGY

This section presents all the steps of the 3D reconstruction pipeline from satellite stereo imagery. The whole process is automatic and doesn't require any ground control points (GCP). The rectification and triangulation process relies on the RPC models provided along with the stereo data. The direct RPC model gives the forward projection from image space *(x,y)* to object space *(lat,lon,alt)*, while the inverse model RPC^{-1} gives the backward projection from object space to the image space. Each model consists of 78 coefficients. The Worldview imagery is provided with only RPC^{-1} model coefficients and the direct model coefficients have to be computed.

In this work, a tile based approach for 3D reconstruction is employed as also done in the stereo pipeline for pushbroom images presented in De Franchis et al. (2014). There are a couple of reasons for using this tile based. Firstly, due to memory requirement of SGM algorithm only a limited size of images can be processed at a time. Secondly, the affine approximation fits well to a smaller

image tile as compared to the whole satellite image as it is shown in De Franchis et al. (2014) that the epipolar error increases with the size of the image tile. Therefore, the processing is done on image tiles of size e.g. 2000×2000 pixels. Hence, given a satellite stereo pair, one image tile is defined as the left stereo image (I_L) and the second one, the right stereo image (I_R).

3.1 Relative RPC Error

There is an error associated with the image-ground transformation provided by the RPC models. This error arises due to the errors in the satellite attitude and Ephemeris determination as well as due to the limited accuracy of sensor orientation described by the RPC models as compared to the rigorous sensor model (Fraser and Hanley, 2005; Grodecki and Dial, 2003). In this work, only the relative RPC error (De Franchis et al., 2014) is removed so that the corresponding points between the two images lie on the corresponding epipolar lines. This relative RPC error can be modelled as a translation for a small image tile, while for the whole image an affine transformation may me required. The translation or the shift to reduce relative RPC error is estimated as follows; First the corresponding points between the images are found by matching SURF features (Bay et al., 2006). Then the epipolar lines in (I_R) are computed for each feature point in (I_L). The distance between the corresponding SURF point (x) and its epipolar line ($l=(l_1, l_2, l_3)$) gives the relative RPC error (Luong and Faugeras, 1996).

$$d(x,l) = \frac{x^T l}{\sqrt{l_1{}^2 + l_2{}^2}} \qquad (1)$$

Now the shift to compensate this error is estimated as the median of the translation between the SURF points and its epipolar lines.

3.2 Rectification

This section describes the rectification procedure applied to each corresponding image tile of the stereo images. The well known stereo rectification algorithm given in Hartley and Zisserman (2003) is adopted with minor adjustments. The first step is to compute the affine fundamental matrix and sending the right epipole to infinity along x-axis. The process requires computation of corresponding points between the two stereo images. As the RPC models of the two images are available, virtual corresponding points between the two images can be computed following the strategy of De Franchis et al. (2014). For the corresponding points (x, x'), the affine fundamental matrix can be written in the form (Hartley and Zisserman, 2003):

$$F_A = \begin{bmatrix} 0 & 0 & a \\ 0 & 0 & b \\ c & d & e \end{bmatrix} \qquad (2)$$

$$x^T F_A x' = ax + by + cx' + dy' + e = 0 \qquad (3)$$

As the epipolar lines are parallel in a given image tile, a simpler solution is to compute the epipolar lines and then rotate the image according to the mean slope of these lines. Using the RPC and RPC^{-1} models, epipolar line for any pixel of I_L can be drawn in I_R by choosing different altitude values. The rotation matrix R for rotating image I_R can also be computed from the affine fundamental matrix as:

$$R = \begin{bmatrix} \frac{b}{norm(a,b)} & \frac{a}{norm(a,b)} \\ \frac{-a}{norm(a,b)} & \frac{b}{norm(a,b)} \end{bmatrix} \qquad (4)$$

An affine transformation is then estimated between the transformed right image and the left image using virtual corresponding points. The left image is then transformed using this affine transformation. The resulting images are now rectified and any dense matching algorithm can be applied on this rectified image pair. When multiple images are available, the rectification process is done for each image pair. The rectification algorithm is summarized below.

Algorithm: Satellite Stereo Rectification

Data: I_L, I_R, RPC_L, RPC_R
1. Draw epipolar lines in I_R and calculate rotation matrix **R**
2. Rotate I_R according to **R**
3. Compute affine transformation (**H**) between RI_R and I_L
4. Transform I_L according to **H**

3.3 Dense Matching

The dense stereo matching can be formulated as a minimization of an energy function as given in Eq. 5. The first term is the data term, while the second term enforces 2D smoothness constraint across neighbouring pixels. The global minimum of such an energy function is NP-hard for a general stereo matching problem (Boykov et al., 2001). However, the minimization of the energy along 1D path can be computed efficiently using dynamic programming. It is well known that applying such constraints along 1D paths (e.g. along the rows) paths only, leads to streaking artefacts. To overcome this problem, SGM computes minimum by aggregating costs from several (typically 8 or 16) 1D paths along different directions. SGM is actually a general purpose heuristic to optimize such energy functions.

$$E(D) = E_{data} + E_{sm} = \sum_p C(p) + \sum_{q \in N_p} \lambda C(q, p) \quad (5)$$

The data term constitutes the pixel-wise matching cost between the pixels of the stereo images. Several data terms like mutual information, census and normalized cross correlation have been evaluated in earlier studies (Hirschmüller and Scharstein, 2009; Vogel et al., 2013). The results of these studies have shown that census matching cost is more robust than the other evaluated date terms. Therefore, in this work, census matching cost is used as the data term. The Census transform (Zabih and Woodfill, 1994) encodes intensity pattern around a pixel in terms of a binary string. The Census cost between two pixels of the stereo images is then computed as the hamming distance between the two binary strings.

$$E(D) = \sum_p C(\mathbf{p}, D_p) + \sum_{q \in N_p} (P_1 T[|D_p - D_q| = 1])$$
$$+ \sum_{q \in N_p} (P_2 T[|D_p - D_q| > 1]) \quad (6)$$

The 2D smoothness term enforces similar disparity values for neighbouring pixels (N_p). Here, T is a logical operator, which outputs 1 when the condition is satisfied and 0 otherwise. The small disparity changes are penalized by a constant cost P_1 and the higher disparity changes are penalized by cost P_2. As suggested by Hirschmüller (2005) the value of P_2 is adopted according to the intensity difference between the neighbouring pixels

to allow disparity change across possible object boundaries. The subpixel disparity is estimated by fitting a quadratic curve to the neighbouring cost values around the minimum and then computing the maximum of a quadratic curve as done in Hirschmüller (2005). A left right consistency check is then applied to filter out the occluded and unreliable pixels.

3.4 Triangulation and DSM

The dense matching algorithm gives the corresponding pixels between the two rectified stereo images. The process of triangulation computes the object space coordinates of each pixel. Using the RPC models, the goal is to estimate the height of each pixel in the left image such that the forward projection from left image and then the backward projection to the right image maps to the corresponding point of each pixel. This process is applied iteratively, to compute the required heights for each pixel. The resulting 3D point cloud is then median filtered using the nearest neighbours of each 3D point. The filtered point cloud is then interpolated in to a DSM using the moving planes interpolation in software OPALS (Pfeifer et al., 2014).

3.5 Object Space Semi Global Matching (OSGM)

In the second method the SGM algorithm is applied in the object space instead of the image space. The SGM is applied directly in the object space, therefore, the rectification and the triangulation steps are not required. The transformation between object space and the image space is given by the RPC models as described previously. Firstly, the object space is defined as a 2D grid (*lat,lon*) and an associated height value. Hence, the object space consists of a voxels having a (*lat,lon,alt*) value. Here, the voxel space can also be defined using UTM coordinates instead of WGS84 coordinates. The grid size is chosen to be the mean GSD of the nadir image, while the height of each voxel is chosen using a predefined height value ΔZ. A smaller ΔZ will help to achieve better height resolution, however, it also increases the memory requirement. Similar to the tile based approach described in previous section, the object space is divided in to tiles of size e.g. $1km \times 1km$ and the OSGM is applied to each tile separately due to limited available memory.

First, the relative RPC error is corrected using the method described in the previous section. Now using the RPC, the center of each voxel is back projected to original images to compute the pixel coordinates for each voxel. The original images are then interpolated at these pixel coordinates, creating an image cube for each original image. Hence for each voxel there is a corresponding pixel for each image in the image cube. The census matching cost can now be computed for each pixel by computing the hamming distance of the binary strings encoding intensity pattern around the corresponding points. In contrast to the work of Bethmann and Luhmann (2014), it is observed that Census transform works well for matching in object space as depicted by the results. The voxel with the correct height value, ideally produces the best match and thus the lowest cost between the pixels of the two images.

The SGM energy function given in Eq. 6 is modified to compute height at each grid element (Bethmann and Luhmann, 2015).

$$E(Z) = \sum_g C(g, Z) + \sum_{q \in N_p} (P_1 T[|Z_g - Z_q| = 1])$$
$$+ \sum_{q \in N_p} (P_2 T[|Z_g - Z_q| > 1]) \quad (7)$$

Similar, to the SGM computation performed earlier. Cost from 8 paths is aggregated at each grid point and the minimum of the aggregated cost gives the height value of the grid point. The thresholds P_1 and P_2 penalizes height differences of one voxel and more respectively.
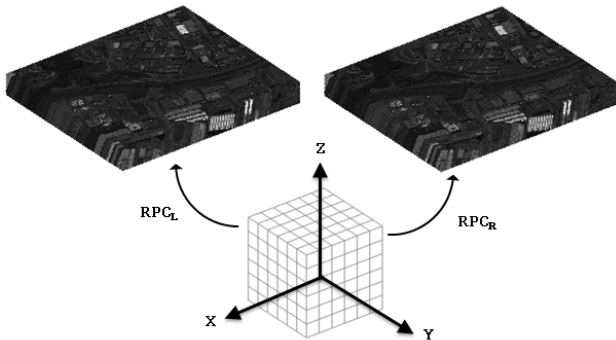

Figure 1: Voxel space to 3D image arrays of stereo pair

Similar, to the previous section, the sub voxel height can be computed by fitting a quadratic curve around the neighbours of the minimum cost voxel at each grid point. Here, the penalty P_2 is kept constant as it the matching is now done in object space. Furthermore, the left right consistency check to find occlusions is not applicable here as there is only one cost array computed in OSGM.

One of the main benefits of the OSGM is that multi-image matching can be performed efficiently, directly in the voxel grid. One can even define the pixel wise cost like Census for multiple images i.e. computing the hamming distance of binary strings from multiple images. Hirschmüller (2005) has also suggested that one can compute pixel wise cost with multiple images. However, in satellite imagery, only three images (tri-stereo) of one scene are typically available at maximum. In the current work, pairwise DSMs from tri-stereo images are computed and then median filtered using the height values at each grid location.

4. DATASET

The first dataset used for evaluation is the Worldview-1 stereo imagery from Terrasa, Spain, which is available through the IS-PRS Commission I stereo matching benchmark (d'Angelo and Reinartz, 2011). The Nadir image has a mean GSD of 0.50 m and it was acquired in August, 2008 (see Figure 2). The Li-DAR data of the same area was acquired in 2007, with a point density of 0.3 $points/m^2$. The RPC provided with the stereo images have already been block adjusted using the GCPs derived from the available orthophotos and LiDAR data. Therefore, the georeferencing information of stereo images is well in accordance with the LiDAR data and no further transformation is required for the registration of DSM from stereo and LiDAR data. More details on the benchmark can be found in d'Angelo and Reinartz (2011) and Kuschk et al. (2014). The second dataset is the Worldview-3 stereo images of Muscat, Oman with a resolution of 0.4m. The third dataset is the Pleiades tri-stereo imagery of Melbourne, Australia, with 0.7m resolution. The relative RPC error in the Worldview-3 and Pleiades data is removed by estimating a 2D translational component.

5. RESULTS AND DISCUSSION

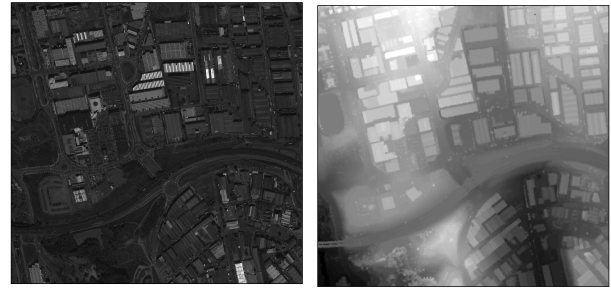The results of the DSM computed form SGM and OSGM are shown in Figures 3 and 4. The DSM from LiDAR is shown for

Figure 2: Left: Nadir Image of the AOI Right: LiDAR DSM

	SGM* (m)	OSGM* (m)	(Kuschk et al., 2014) (m)
Mean Abs. Error	1.08	1.28	1.11
Median Abs. Error	0.49	0.62	

Table 1: Error values from DSM differences. (* The error values have been computed only for AOI)
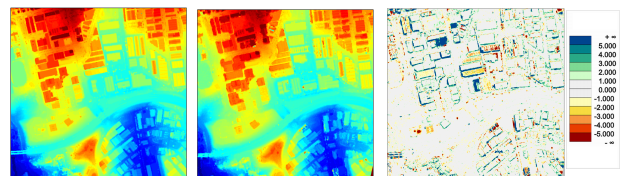

Figure 3: Left: DSM from Laser scanning Middle: DSM computed from SGM Right: Difference between the two DSM
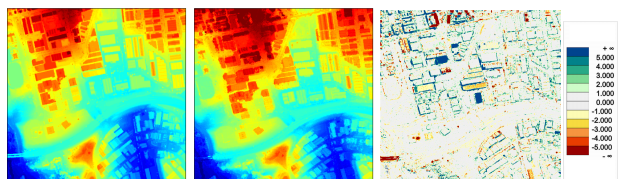

Figure 4: Left: DSM from Laser scanning Middle: DSM computed from object space SGM Right: Difference between the two DSM

comparison. Table 1 shows the error values (mean absolute error) derived from the difference of the LiDAR and stereo DSM. The results show that similar OSGM can achieve similar accuracies as the the SGM. The Census transform with 9×9 window was used for computing the data term. The SGM algorithm is insensitive to changes in the parameters and the results don't alter significantly by changing the penalty thresholds P_1 and P_2.

Here, it is important to mention that height of the trees are derived very well in LiDAR data, while the robust derivation of tree heights from satellite stereo still remains a challenge. A smaller baseline satellite stereo configurations would probably be more suitable for dense matching over forested area. It is suggested that there should be a mask for vegetated areas in the ISPRS stereo benchmark so that these areas can be filtered out when required.

The OSGM algorithm is further tested on stereo data from Worldview-3 and tri-stereo data from Pleiades. The Figure 5 shows the DSM generated from 40 cm Worldview-3 stereo images. As no ground truth data is available for this scene, only qualitative analysis can be performed. It can be seen that good quality DSM can be generated from satellite stereo images using OSGM. The Figure 6, shows the DSM generated from tri-stereo images. From three

Figure 5: Left: Worldview-3 Image (Muscat, Oman) Right: DSM generated from OSGM

images three DSMs were computed using different pair combinations. The fused DSM is generated by taking the median of the three DSMs. Consequently, the reliability of the DSM is increased by fusing these individual DSMs. Furthermore, pixel-wise matching using all three Pleiades images at once was investigated, i.e. computing the hamming distance of three binary strings, however, the current results don't show significant improvement in the resulting DSM.

In OSGM, the accuracy of the matching is strongly dependent on the relative RPC error of the two images in the same way as the accuracy of the dense matching depends on the remaining y-parallax of the rectified images. In (De Franchis et al., 2014) it is shown that the epipolar error of less than 0.1 pixels can be achieved. Therefore, it is essential that the relative RPC error should be removed to achieve good accuracy. The RMS RPC error in the Worldview-3 and Pleiades data after RPC correction was less than 0.5 pixels, which resulted in a good overall quality of the results.

One of the disadvantage of the current implementation of OSGM is that it lacks a robustness measure similar to the left-right consistency check used in dense stereo matching algorithms. Consequently, the DSM generated from OSGM has more gross errors. Another disadvantage of OSGM is that the penalty term P_2 is kept constant, while in SGM the penalty term P_2 is varied with respect to the brightness gradient to allow discontinuity in the disparity across edges in the image. To overcome these limitations, ideas like explicit occlusion detection, edge contour matching and height discontinuity estimation used in volumetric multi view approaches should be incorporated in the OSGM methodology. Therefore, the future work is directed towards incorporating these techniques. In a recent work (Drory et al., 2014), the relation between SGM and belief propagation is derived and it is shown that SGM is closely related to the min-sum belief propagation over a union jack shaped graph. Furthermore, a novel uncertainty measure for SGM have been proposed in (Drory et al., 2014) and it is envisaged that future progress in SGM would come through these insights.

References

Bay, H., Tuytelaars, T. and Van Gool, L., 2006. Surf: Speeded up robust features. In: Computer vision–ECCV 2006, Springer, pp. 404–417.

Bethmann, F. and Luhmann, T., 2014. Object-based multi-image semi-global matching. concept and first results. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 40(5), pp. 93–100.

Bethmann, F. and Luhmann, T., 2015. Semi-global matching in object space. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 1, pp. 23–30.

Boykov, Y., Veksler, O. and Zabih, R., 2001. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(11), pp. 1222–1239.

d'Angelo, P. and Kuschk, G., 2012. Dense multi-view stereo from satellite imagery. In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 6944–6947.

d'Angelo, P. and Reinartz, P., 2011. Semiglobal matching results on the isprs stereo matching benchmark. In: ISPRS Hannover Workshop 2011: High-Resolution Earth Imaging for Geospatial Information.

De Franchis, C., Meinhardt-Llopis, E., Michel, J., Morel, J.-M. and Facciolo, G., 2014. An automatic and modular stereo pipeline for pushbroom images. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

Deseilligny, M. P., 2015. Micmac, apero, pastis and other beverages in a nutshell. http://logiciels.ign.fr/micmac.

Drory, A., Haubold, C., Avidan, S. and Hamprecht, F. A., 2014. Semi-global matching: a principled derivation in terms of message passing. In: Pattern Recognition, Springer, pp. 43–53.

Fraser, C., Dial, G. and Grodecki, J., 2006. Sensor orientation via rpcs. ISPRS Journal of Photogrammetry and Remote Sensing 60(3), pp. 182–194.

Fraser, C. S. and Hanley, H. B., 2005. Bias-compensated rpcs for sensor orientation of high-resolution satellite imagery. Photogrammetric Engineering & Remote Sensing 71(8), pp. 909–915.

Geiger, A., Lenz, P. and Urtasun, R., 2012. Are we ready for autonomous driving? the kITTI vision benchmark suite. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 3354–3361.

Grodecki, J. and Dial, G., 2003. Block adjustment of high-resolution satellite images described by rational polynomials. Photogrammetric Engineering & Remote Sensing 69(1), pp. 59–68.

Gruen, A., 1985. Adaptive least squares correlation: a powerful image matching technique. South African Journal of Photogrammetry, Remote Sensing and Cartography 14(3), pp. 175–187.

Haala, N. and Rothermel, M., 2012. Dense multi-stereo matching for high quality digital elevation models. Photogrammetrie-Fernerkundung-Geoinformation 2012(4), pp. 331–343.
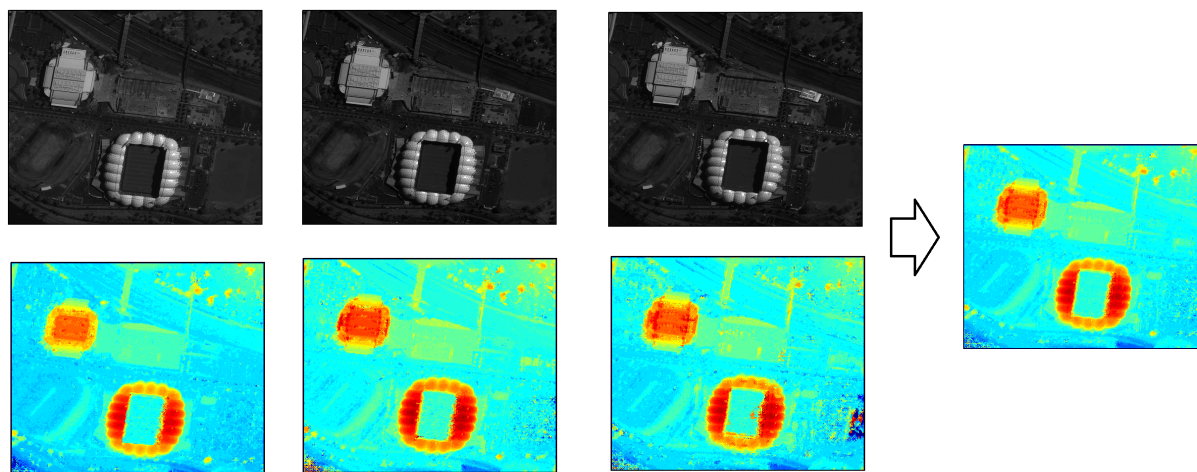
Figure 6: Top Row: Pleiades Tri-stereo images (Melbourne, Australia) Bottom: DSM generated from image pairs: 1-2, 2-3, 1-3, Right: Median DSM

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(109), pp. 27–47.

Hartley, R. and Zisserman, A., 2003. Multiple view geometry in computer vision. Cambridge university press.

Hermann, S. and Klette, R., 2013. Iterative semi-global matching for robust driver assistance systems. In: Computer Vision–ACCV 2012, Springer, pp. 465–478.

Hirschmüller, H., 2005. Accurate and efficient stereo processing by semi-global matching and mutual information. In: Computer Vision and Pattern Recognition, Vol. 2, IEEE, pp. 807–814.

Hirschmüller, H., 2008. Stereo processing by semiglobal matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence 30(2), pp. 328–341.

Hirschmüller, H. and Scharstein, D., 2009. Evaluation of stereo matching costs on images with radiometric differences. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(9), pp. 1582–1599.

Kuschk, G., d'Angelo, P., Qin, R., Poli, D., Reinartz, P. and Cremers, D., 2014. Dsm accuracy evaluation for the isprs commission i image matching benchmark. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 40, pp. 195–200.

Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), pp. 91–110.

Luong, Q.-T. and Faugeras, O. D., 1996. The fundamental matrix: Theory, algorithms, and stability analysis. International Journal of Computer Vision 17(1), pp. 43–75.

Morgan, M., Kim, K., Jeong, S. and Habib, A., 2004. Epipolar geometry of linear array scanners moving with constant velocity and constant attitude. ISPRS, Istanbul, Turkey.

Oh, J., Lee, W. H., Toth, C. K., Grejner-Brzezinska, D. A. and Lee, C., 2010. A piecewise approach to epipolar resampling of pushbroom satellite images based on rpc. Photogrammetric Engineering & Remote Sensing 76(12), pp. 1353–1363.

Pfeifer, N., Mandlburger, G., Otepka, J. and Karel, W., 2014. Opals–a framework for airborne laser scanning data analysis. Computers, Environment and Urban Systems 45, pp. 125–136.

Pierrot-Deseilligny, M. and Paparoditis, N., 2006. A multiresolution and optimization-based image matching approach: An application to surface reconstruction from spot5-hrs stereo imagery. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences.

Roy, S., 1999. Stereo without epipolar lines: A maximum-flow formulation. International Journal of Computer Vision 34(2-3), pp. 147–161.

Roy, S. and Cox, I. J., 1998. A maximum-flow formulation of the n-camera stereo correspondence problem. In: Sixth International Conference on Computer Vision, IEEE, pp. 492–499.

Scharstein, D. and Szeliski, R., 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision 47(1-3), pp. 7–42.

Vogel, C., Roth, S. and Schindler, K., 2013. An evaluation of data costs for optical flow. In: J. Weickert, M. Hein and B. Schiele (eds), Pattern Recognition, Lecture Notes in Computer Science, Vol. 8142, Springer Berlin Heidelberg, pp. 343–353.

Vogiatzis, G., Hernández, C., Torr, P. H. and Cipolla, R., 2007. Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(12), pp. 2241–2246.

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(3), pp. 347–355.

Wohlfeil, J., Hirschmüller, H., Piltz, B., Börner, A. and Suppa, M., 2012. Fully automated generation of accurate digital surface models with sub-meter resolution from satellite imagery. Int. Arch. Photogramm. Rem. Sens. Spatial Inf. Sci pp. 34–B3.

Zabih, R. and Woodfill, J., 1994. Non-parametric local transforms for computing visual correspondence. In: ECCV'94, Springer, pp. 151–158.