

# IMAGE MATCHING AND OUTLIER REMOVAL FOR LARGE SCALE DSM GENERATION

Pablo d'Angelo

German Aerospace Center (DLR), Remote Sensing Technology Institute, D-82234 Wessling, Germany  
email: Pablo.Angelo@dlr.de

Commission I/4

**KEY WORDS:** Spaceborne Scanner Systems, Digital Elevation Models (DEM), Image Matching, CARTOSAT-1

## ABSTRACT:

Digital surface models can be efficiently generated with automatic image matching from optical stereo images. Detailed reconstruction from very high resolution stereo images requires the use of dense stereo matching algorithms, such as Semiglobal Matching. However, for cases where only a single stereo pair is available, the results sometimes still contain some outliers, mostly due to changing reflectance behaviour, clouds, water and moving objects. This is especially critical for large scale and fully automated DSM generation with dedicated stereo satellites, such as Cartosat-1 and ALOS/PRISM. An effective outlier removal algorithm for dense stereo matching algorithms, based on region based consistency checks between to matching results with slightly different parameters is presented in this paper. Ground truth has been manually extracted on 66 test areas, covering different land cover types, seasons and elevation heights. On this diverse and demanding dataset more than 99% of all outliers are removed automatically, with minimal influence on the correctly matched regions.

## 1 INTRODUCTION

Digital surface models can be efficiently generated with automatic image matching from optical stereo images. Detailed reconstruction of small structures from very high resolution stereo images requires the use of dense stereo matching algorithms that are not based on correlation of image windows. Very good results with a reasonable runtime have been achieved with the Semiglobal matching (SGM) algorithm (Hirschmüller, 2008). However, for cases where only a single stereo pair is available, the results sometimes still contain a small number of outliers, mostly in scenes that are dominated by agricultural areas or near cloud shadows and waves. Image matching fails in these areas and a small number of mismatches remain, mainly due to strongly changing reflectance behaviour and movement of object borders such as clouds and cloud shadows.

Errors in the DEM need to be cleaned by manual editing, which is very costly for continent wide DSM generation. Classical DSM filtering methods based on the evaluation of local statistics are not suitable for these kinds of outliers, as they affect many well recovered man made objects such as large buildings and skyscrapers. As SGM and other global, pixel based matching algorithms consider the relationship between neighbouring matching hypothesis, they typically produce outlier regions, and not outlier points. This is already exploited by the original SGM algorithm, where a left right check (consistency of matching the first image with the second image and vice versa) is performed, and the resulting disparity map is segmented into regions with similar disparities. Small, isolated disparity regions are removed (Hirschmüller, 2008). The choice of a generally applicable minimal region size is critical as a larger size will remove many outlier regions in problematic areas, but will also erase buildings and smaller objects.

This paper presents an improved outlier detection and removal method for dense stereo matching algorithms. A second matching with a different cost function and parameters is performed. Areas where the matching becomes unstable result in different

disparities. Still, some mismatches remain, and are removed using a region based consistency check. Only regions that share a significant area with regions in the second matching result are kept.

## 2 LARGE SCALE DSM GENERATION

The outlier removal developed in this paper is part of a larger processor for the generation of a high resolution, European DSM from Cartosat-1 stereo scenes (d'Angelo et al., 2008). Several tens of thousands stereo scenes need to be processed as part of this project. To limit the amount of manual edits, the stereo matching needs to be as robust as possible. The full DSM generation process includes bundle adjustment to improve the RPCs of large Cartosat-1 stereo scene blocks, dense stereo matching, reprojection and merging of the DSMs derived from each stereo pair and closing remaining holes in the DSM with delta surface fill (Grodecki et al., 2004) and interpolation.

### 2.1 Stereo Matching

Stereo image matching is the central technique for the DSM generation from stereo images. To archive a dense reconstruction, for each pixel position in the first image the corresponding position in the second image should be found. Several approaches for dense matching exist in the literature (Otto and Chau, 1989, Scharstein and Szeliski, 2002, Hirschmüller, 2008). In this work, the Semiglobal Matching algorithm (Hirschmüller, 2008) is used, but the proposed outlier removal procedure can be applied to other dense stereo methods, too.

For high quality matching results, a high redundancy in the image matching step is desirable. Multiple viewing directions help to eliminate outliers and minimise areas that cannot be reconstructed due to occlusions. For many satellite based stereo reconstruction tasks, only stereo pairs or triplets are available, often with a fixed geometry. This mostly applies to satellites designed for stereo acquisition, for example, Cartosat-1, ALOS/PRISM and ASTER. Agile single camera satellites, such as GeoEye-1,

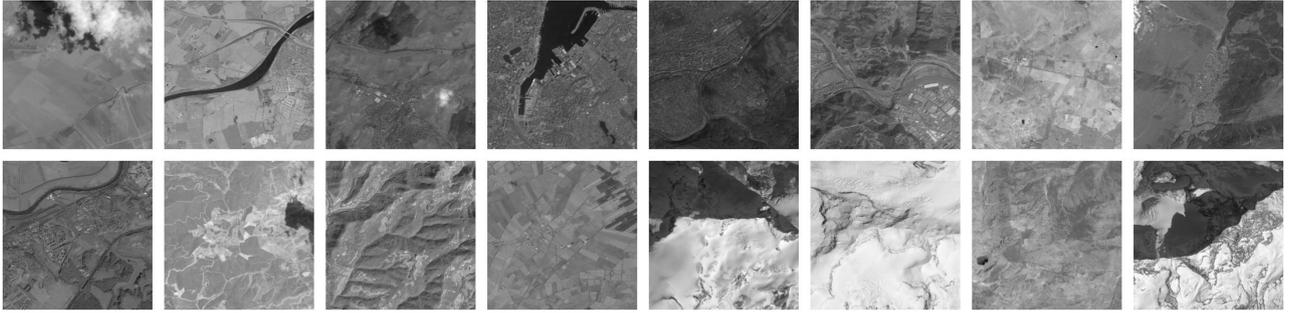


Figure 1: 16 of 74 test sites used in the evaluation.

Worldview-1 & 2, can acquire a large number of views from the same orbit, however this is seldom done due to increased data acquisition cost.

If only stereo pairs are available, the redundancy of the matching step is very low, and dense matching will produce some mismatches or blunders. Manual editing and outlier removal needs to be limited for processing of large areas, and it is therefore essential that the stereo matching does not produce outliers, even if the input data is difficult or impossible to match, for example due to large shadows, clouds and water areas.

SGM avoids using matching windows, and is thus able to reconstruct sharp object boundaries. Instead of strong local assumption on the local surface shape, a global energy function  $E$  is minimized for all disparities (local shifts between the stereo pair)  $D$ . SGM performs a semi global optimization by aggregating costs from 16 directions, and finds a image  $D$  which leads to a low energy  $E$ :

$$E(D) = \sum_p \left( C(p, D_p) + \sum_{q \in N_p} D_c = \begin{cases} p_1, & |D_p - D_q| \leq 1 \\ p_2, & \text{otherwise} \end{cases} \right) \quad (1)$$

The function  $C$  defines the matching cost between the image pixels for each pixel location  $p$  and possible disparity  $D_p$  in the first image. Both the Mutual Information (Hirschmüller, 2008) as well as the Census cost function (Zabih and Woodfill, 1994) are used in this paper. These cost function adapt to brightness changes in the stereo images and allow matching of images with large viewing angle differences. The second and third terms of  $E$  penalize disparity changes in the neighbourhood  $N_p$  at each position  $p$ . The penalty  $p_1$  is added for all disparity changes equal to one pixel. At larger discontinuities (disparity change  $> 1$  pixel), a fixed cost  $p_2$  is added. This cost function favours similar or slightly changing disparities between neighbouring pixels, and thus stabilizes the matching in image areas with weak contrast, but also allows large disparity jumps in areas with high contrast.

A left/right consistency check invalidates inconsistent disparities in occluded and mismatched regions. Especially at hard to match areas, some small, isolated regions with bad disparity values remain. The disparity map is segmented into regions of similar disparity and regions smaller than a threshold are eliminated. The choice of a maximum region size is demanding, as a larger values will remove both outliers and small, well reconstructed structures, such as houses or mountaintops.

## 2.2 Outlier filtering

The DSMs of some Cartosat-1 stereo scenes still contain outliers, especially in agricultural fields, snowy and mountainous terrain with large occlusions, water areas, clouds and cloud shadows.

Manual treatment of these areas is not feasible for continent wide DSM generation. A robust matching algorithm that can handle these situations and still produce a highly detailed results is required.

To archive a robust reconstruction, two matchings with different cost function and matching parameters are performed. The resulting disparity maps  $D_1$  and  $D_2$  are fused using a region based consistency check. The second matching provides important information, as the matching becomes unstable in ambiguous areas, and small changes in matching parameters do not affect the disparity values in well reconstructed areas. A consistent disparity image  $D_c$  is computed, by accepting only disparities where the corresponding in  $D_1$  and  $D_2$  differ by less than  $t_d$ . As SGM and other global, pixel based matching algorithms consider the relationship between neighbouring disparity values, they typically produce outlier regions, and not outlier points. Thus, a region based outlier treatment is required for good performance.

Both disparity images  $D_1$  and  $D_2$  are segmented into regions where the neighbouring disparity changes less than 1, similar to the region based outlier removal of Semiglobal Matching. The regions derived from both disparity images should be in good agreement. In “unstable” matching areas, SGM typically produces smaller regions whose shape and size is sensitive to small changes of the matching parameters. This can be used to distinguish outliers from valid small matching regions, such as trees or houses, whose shape is not significantly influenced by small changes to the parameters. Smaller regions that overlap only partly with the consistent disparity image  $D_c$  can be considered to be part of an unstable matching region. The size  $s_i$ , and the amount of consistent disparities  $q_i = (i \cap D_c)/s_i$  is computed for each region  $i$ . Regions are classified as outliers, if  $s_i \leq t_s$  and  $q_i \leq t_q$ . The choice of thresholds  $t_s$  and  $t_q$  is less critical than the single minimum region size parameter, as it only applies to inconsistent regions.

In some case, outlier regions are consistent between  $D_1$  and  $D_2$ . This especially happens in large cloud or water areas. As the majority of the water or cloud areas are correctly detected as invalid, regions smaller than  $t_s$  can be removed, if they share a border with a void region larger than  $t_v$ . As larger voids can also occur due to occlusions, for example in mountainous terrain, this criteria might also remove a larger amount of good matches, depending on the type of the scene. It should therefore be applied with caution.

## 3 EXPERIMENTAL EVALUATION

The proposed outlier removal process has been applied to several hundred full Cartosat-1 stereo scenes. The results for two 1500 by 1500 pixel cut outs are shown in Fig. 2. The two disparity images shown in Fig. 2c and d contain different void areas, mostly

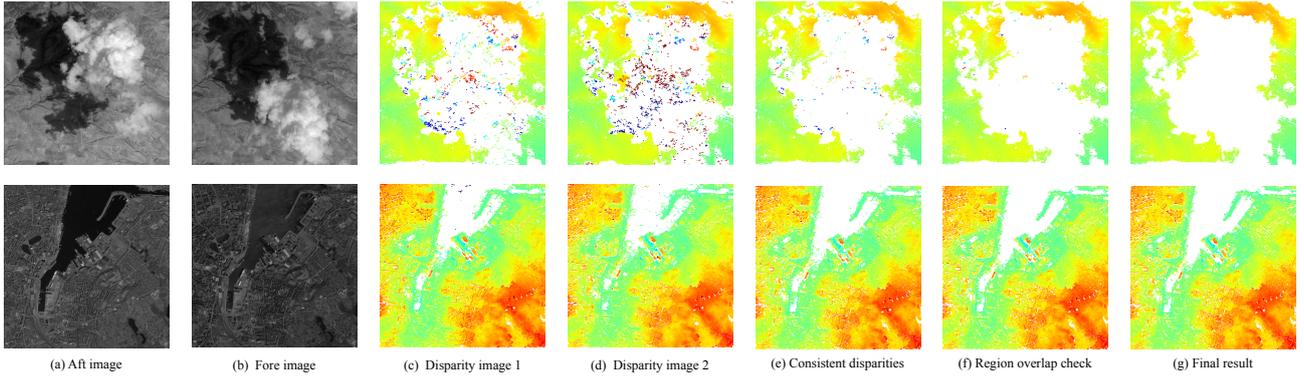


Figure 2: Outlier removal procedure applied to a subset of a Cartosat-1 stereo pair. (a) and (b) show the Cartosat-1 cut out, (c),(d) the disparity images obtained with different matching parameters and (e) the consistent disparities in (c),(d). (f) disparity image after removing inconsistent regions, (g) Final disparity, after removal of small regions in large void areas. White colour indicates invalid disparity.

in areas with weak texture, or differing scene content. However, strong outliers seem to be located in approximately the same areas, but with different shapes. Some outliers remain after the consistency check (Fig. 2e). After rejecting small, inconsistent disparity regions (Fig. 2f), the outliers are mostly removed. The remaining outliers can be effectively treated by removing small regions in large holes (Fig. 2g). This step also removes some good regions in areas that cannot be matched very densely, and experience has shown that it is only required for cloudy scenes.

A quantitative evaluation has been performed on 66 test areas, extracted from various Cartosat-1 stereo image pairs, covering different land cover types, seasons and height ranges. Some example areas are shown in Fig. 1. We tried to extract quantitative results by comparison against various reference datasets (SRTM, DTMs provided by mapping agencies (where available)). Due to various differences between the reference datasets (resolution, large time difference, surface model vs. ground model, different acquisition strategy) and the generated DSMs, a detailed analysis by comparison with reference data did not result in easily interpretable numbers. Thus, bad disparity regions in the first disparity image have been labelled manually as either inlier or outlier and are used as reference for the evaluation of the outlier removal.

Parameter tuning has been performed by evaluating the outlier removal algorithm with a large set of different parameters on all test areas. The resulting working points of the algorithm are shown in a receiver operator characteristics diagram in Fig. 3. Each point represents a specific parameter set.

The ROC diagram shows the performance of each parameter set. For the remainder of this paper, the parameter set shown by the red circle in Fig. 3 is used. The corresponding parameters are  $t_m = 200$ ,  $t_d=2$ ,  $t_s = 2500$ ,  $t_q = 0.2$ ,  $t_v = 30000$ . The ROC plot shows that this combination removes more than 99% of all mismatches and keeps more than 99.9% of all correctly matched pixels.

The performance of the algorithm on each of the 66 test areas is shown in Fig. 4 and Fig. 6. As Fig. 4 shows, only a small number of valid disparities is removed from most test areas. Exceptions are test areas 54, 55 and 62. These test areas contain steep snow and ice covered alpine terrain in Zermatt, with many occlusions due to the steep mountains. Fig.5 shows these scenes. Large occlusions and steep terrain creates a relatively fragmented disparity map with large void regions. This results in removal of many regions. For this kinds of scenes, smaller values for  $t_s$ , and no removal of regions in large areas might result in a better performance.

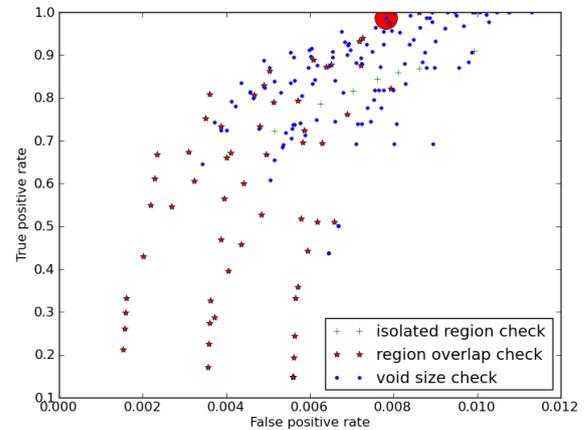


Figure 3: ROC plot for various parameter combinations. Each point represents a parameter set  $t_m, t_d, t_s, t_q, t_v$ . Red crosses indicate the performance of the isolated region removal used by SGM ( $t_m$  varies between 50 and 1000), blue stars the region consistency check ( $t_m, t_s, t_q$ ), and green dots region consistency check and removal of regions in large voids. The parametrisation highlighted by the red circle has been used in this paper.

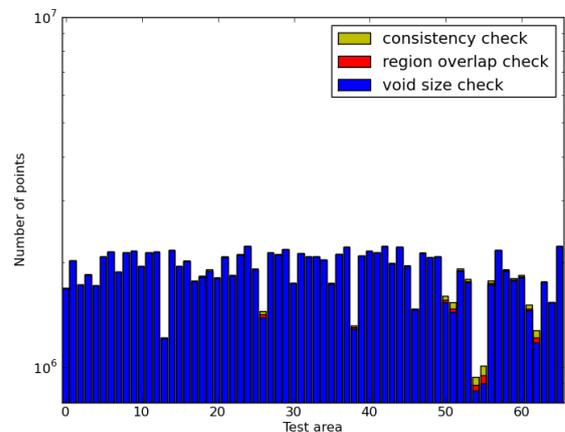


Figure 4: Amount of outlier disparities remaining after the consistency, region overlap and regions in large voids check. For better visibility of a small number of outliers, a logarithmic scale has been chosen.

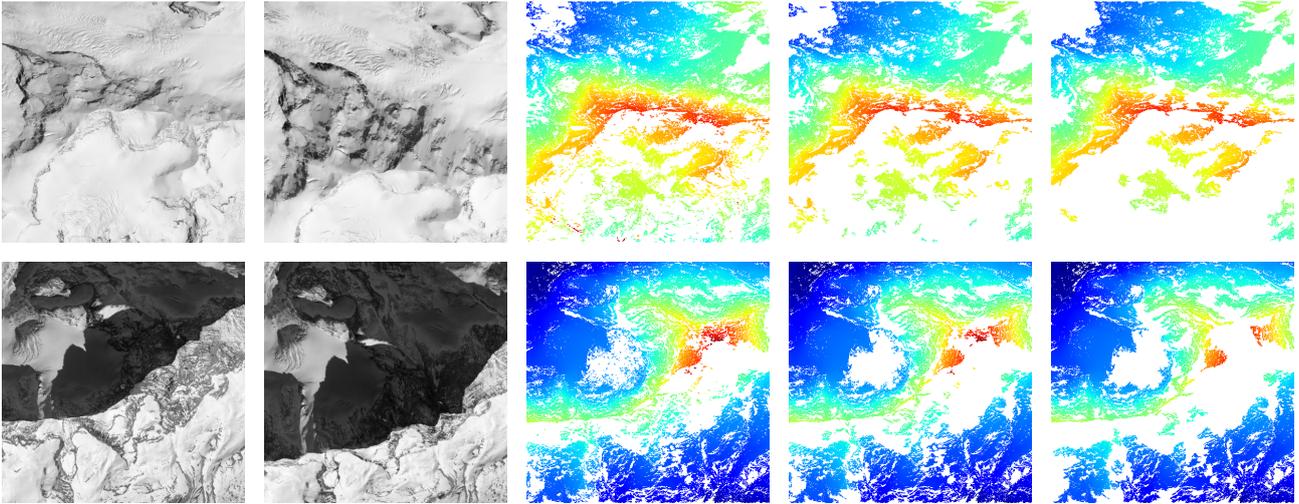


Figure 5: Test areas 54 and 62. Left to right: Aft image, Fore image, consistent disparity, disparity after overlap check, disparity after removal of regions in large voids. These are challenging scenes, where the algorithm removes a larger amount of correct matches.

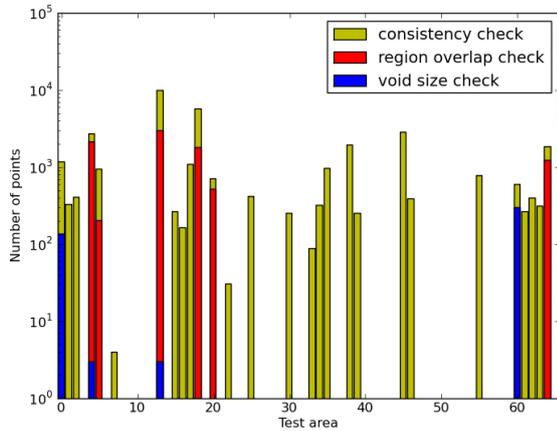


Figure 6: Amount of outlier disparities remaining after the consistency, region overlap and regions in large voids checks.

The corresponding plot for the manually labelled outlier is shown in Fig. 6. No bar is shown for test areas where the consistency check removed all outliers, the yellow bar shows the number of remaining outlier disparities after the consistency check. Test areas where all outliers could be removed with the region overlap check are shown with a yellow bar only. The red bar shows the amount of outliers remaining after the region overlap check. The blue bar shows the amount of remaining outliers after regions in large voids have been removed. After all 3 removal steps, outliers remain only in 4 test areas. Only test area 0 and 60 show a larger amount of outliers that could not be removed with the proposed method. Test areas 4, 5, 13, 18, 20 and 64 needed the the large regions in void outlier check.

The results for test areas 0, 4, 13, 45 and 60 are shown in Fig. 7. Visually, the result for test areas 0 and 13 do not contain any major outliers after the region overlap check. Further investigation of these scenes is needed. Test area 4 contains a big outlier in the cloud shadow area that can only be removed by the combined region check, and 3 undetected outlier pixels. Area 45 is another example, where the region overlap check removed all outliers. Test area 60 is again a complex alpine scene with a large snow and ice area, which SGM handles very well. Due to the large height range, the remaining 200 outlier pixels might not be visi-

ble in the overview plot.

## 4 CONCLUSIONS

This paper presents a region based outlier removal procedure for dense stereo matching, such as Semiglobal matching. It has been designed for applications where only stereo pairs from a fixed viewing geometry are available, for example Cartosat-1 and ALOS PRISM. These systems are designed for large scale 3D reconstruction, an application where a very robust stereo matching algorithm is required to avoid costly manual editing of huge datasets. A region based consistency check between two stereo reconstructions performed with slightly different matching parameters is used to identify and eliminate regions where the dense stereo matching becomes unstable and results in invalid disparity measurements. For cloudy scenes, it is additionally helpful to eliminate small regions inside big void areas.

A comprehensive evaluation is performed on 66 diverse test areas extracted from Cartosat-1 stereo scenes. Ground truth for the evaluation was established by manual labelling of the SGM output into inlier and outlier regions. The procedure removes more than 99% of all mismatches and keeps more than 99.9% of all correct matches, and works especially well for agricultural and water areas. The elimination of small regions inside large voids works very well for cloudy scenes, but removes good matches in alpine terrain.

Future work will include using image features for better handling of cloud areas and machine learning for optimal parametrisation of the outlier removal step. Nevertheless, the current algorithm has been successfully applied to hundreds of Cartosat-1 stereo scenes.

## 5 ACKNOWLEDGEMENTS

We thank Euromap GmbH for providing the Cartosat-1 stereo scenes and Paul Czioska for labelling the test areas.

## REFERENCES

d'Angelo, P., Lehner, M., Krauss, T., Hoja, D. and Reinartz, P., 2008. Towards automated DEM generation from high resolution stereo satellite images. In: ISPRS Conference 2008, Vol. XXXVII, pp. 1137 – 1342.

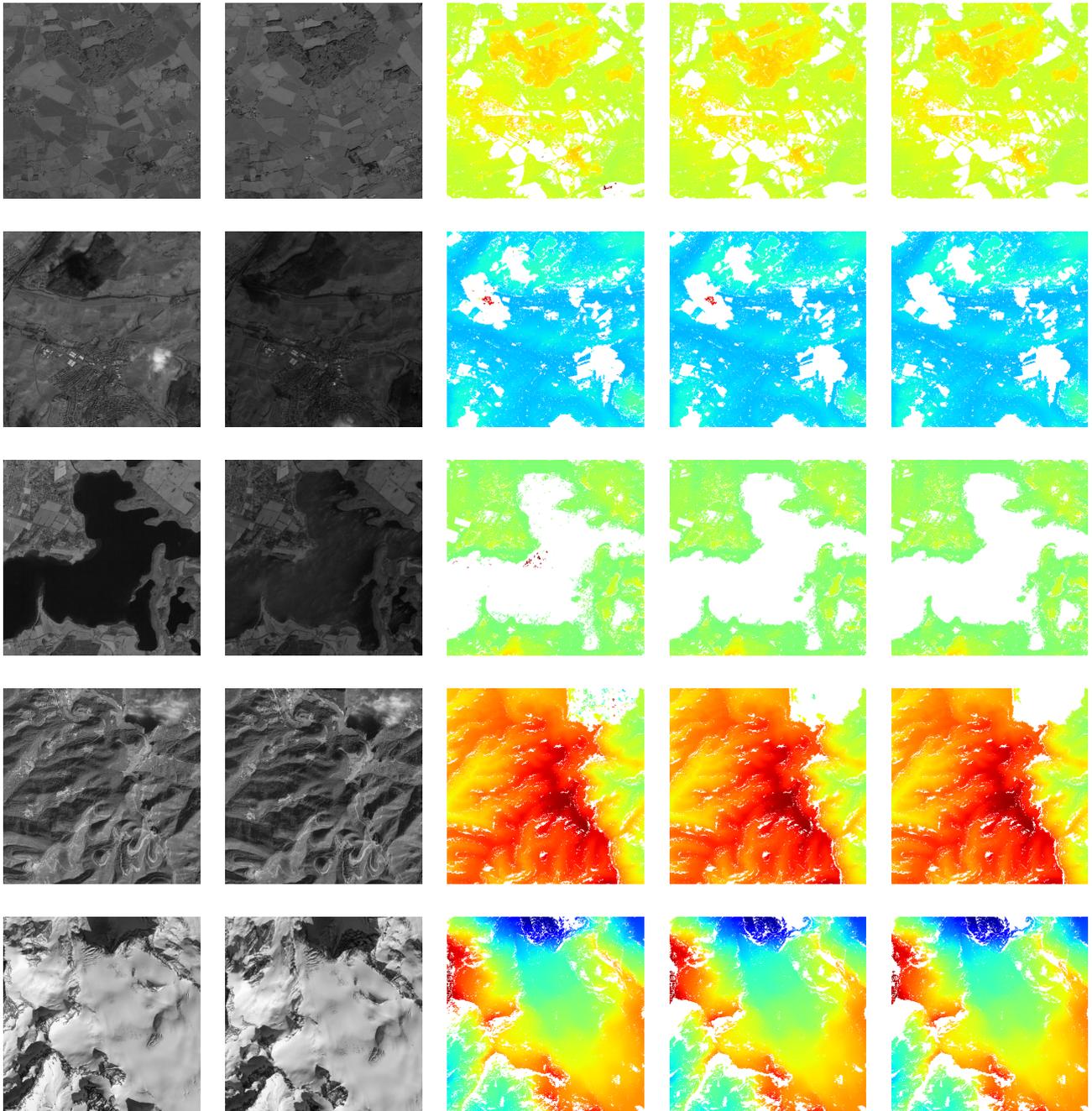


Figure 7: Top to bottom: test areas 0, 13, 45 and 60. Left to right: Aft image, Fore image, consistent disparity, disparity after overlap check, disparity after removal of regions in large voids.

Grodecki, J., Dial, G. and Lutes, J., 2004. Mathematical model for 3D feature extraction from multiple satellite images described by RPCs. In: ASPRS Annual Conf. Proc.

ceedings of the Third European Conference-Volume II on Computer Vision, Springer-Verlag, London, UK, pp. 151–158.

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

Otto, G. P. and Chau, T. K. C., 1989. Region-growing algorithm for matching of terrain images. *Image Vis. Comput.* 7, pp. 83–94.

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.

Zabih, R. and Woodfill, J., 1994. Non-parametric local transforms for computing visual correspondence. In: *ECCV '94: Proceedings of the Third European Conference-Volume II on Computer Vision*, Springer-Verlag, London, UK, pp. 151–158.