

FUSION OF TANDEM-X AND CARTOSAT-1 DEMS USING TV-NORM REGULARIZATION AND ANN-PREDICTED WEIGHTS

H. Bagheri, M. Schmitt, XX. Zhu

Signal Processing in Earth Observation, Technical University of Munich, Munich, Germany
Remote Sensing Technology Institute, German Aerospace Center, Oberpfaffenhofen, Wessling, Germany

ABSTRACT

This paper deals with TanDEM-X and Cartosat-1 DEM fusion over urban areas with support of weight maps predicted by an artificial neural network (ANN). Although the TanDEM-X DEM is a global elevation dataset of unprecedented accuracy (following HRTI-3 standard), its quality decreases over urban areas because of artifacts intrinsic to the SAR imaging geometry. DEM fusion techniques can be used to improve the TanDEM-X DEM in problematic areas. In this investigation, Cartosat-1 elevation data were fused with the TanDEM-X DEM by weighted averaging and total variation (TV)-based regularization, resorting to weight maps derived by a specifically trained ANN. The results show that the proposed fusion strategy can significantly improve the final DEM quality.

Index Terms— Data fusion, L_1 norm total variation, weight map, Artificial Neural Network, TanDEM-X DEM, Cartosat-1 DEM

1. INTRODUCTION

In the recent years, the new global Digital Elevation Model (DEM) produced from InSAR data acquired by the German TanDEM-X mission set a new standard regarding global elevation maps by bistatic SAR interferometry, offering a nominal pixel spacing of 12 m and an accuracy better than 2 m for slopes lower than 20% [1]. In spite of its high HRTI-3 quality standard, it is less accurate than optical stereoscopic DEMs in urban areas. The main reason for lower quality in complex terrains like urban areas returns to SAR-inherent imaging effects such as layover and shadowing. One possibility for an improvement of the quality of TanDEM-X elevation data in urban areas is to fuse it with another kind of elevation data which are derived from stereo optical imagery like Cartosat-1. Cartosat-1 is an Indian satellite equipped with a dual-optics 2-line along-track stereoscopic push broom scanner with a stereo angle of 31° and a resolution of 2.5 m. It is specifically designed for high-resolution DEM generation with relatively large areal coverage. An initial evaluation quantified the quality differences between the Cartosat-1 DEM and the TanDEM-X DEM in urban areas [2].

This paper first quickly evaluates the raw TanDEM-X and Cartosat-1 DEMs over urban areas and then will focus on the fusion of both DEMs to achieve a final global DEM with higher accuracy in urban areas. For this purpose, different methods such as weighted averaging and an approach incorporating L_1 norm-Total Variation regularization are employed for DEM fusion. In addition, weight maps generated with the support of a specifically trained artificial neural network will be applied throughout the DEM fusion procedure to assist the fusion algorithm to reach an optimal solution.

2. TANDEM-X AND CARTOSAT-1 DEM PREPARATION

Before the actual fusion process, the elevation data must be properly aligned and homogenized, as each original DEM dataset has its own unique pixel spacing and datum. Thus, all elevation data must be converted and transferred to a specified target datum with equidistant pixel spacing. For the experiments in this paper, we chose to use UTM as joint reference system with heights measured above the WGS84 ellipsoid. All data were resampled to 5 m pixel spacing which is nearly similar to TanDEM-X pixel spacing (6 m in the equator).

After datum homogenization, the DEMs must be precisely aligned to the reference DEM in order to remove any rotational and horizontal translations. In this study, the Iterative Closest Point (ICP) algorithm [3] was used for DEM coregistration. After alignment, the DEMs can be input into the fusion framework to produce the final improved DEM in urban areas.

3. DEM FUSION OVER URBAN AREAS

An uncertainty assessment of the TanDEM-X and Cartosat-1 DEMs [2] illustrates that the Cartosat-1 DEM is more accurate than the TanDEM-X DEM in urban areas while the accuracy of both DEMs is nearly same over non-urban areas. The results of this assessment of the relative accuracies in different land types over urban and non-urban areas are briefly summarized in Table 1 for sake of completeness. The RMSE as a metric for relative accuracy assessment is calcu-

Table 1. TanDEM-X and Cartosat-1 relative accuracy assessment (in meter) over urban and non-urban areas

Areas		Cartosat-1	TanDEM-X
		RMSE	RMSE
Urban	Industrial	1.993	2.830
	Inner city	4.075	5.390
	High building	4.957	8.652
	Residential	2.407	2.992
Non-Urban	Forested	3.474	3.567
	Agricultural 1	1.227	0.953
	Agricultural 2	0.993	1.285
	Lake	1.736	10.486

lated after getting rid of the vertical bias and also discarding the outliers which were detected based on the LE90 margin (1.645*STD). As mentioned in Section 1, a feasible solution for the improvement of TanDEM-X elevation data over urban areas is the fusion with Cartosat-1 DEM data, which were shown to provide much better accuracy in urban areas. Versatile approaches have been formulated for DEM fusion [4][5][6][7] as an important application case of data fusion in the field of remote sensing [8]. The simplest but most popular one is weighted averaging which has been used widely for DEM fusion. In this investigation, it will be proven that using proper weights for DEM fusion can efficiently increase the performance of the fusion results. Two approaches, weighted averaging and TV-regularized fusion, are carried out for the TanDEM-X and Cartosat-1 fusion task. In the weighted averaging method, each DEM with its own weight map will be finally fused according to:

$$DEM_{fused} = \frac{W_{TDX}DEM_{TDX} + W_{CS1}DEM_{CS1}}{W_{TDX} + W_{CS1}}, \quad (1)$$

where W_{TDX} and W_{CS1} are the weights of the TanDEM-X and the Cartosat-1 heights, respectively. DEM_{TDX} and DEM_{CS1} are the height values taken from the TanDEM-X and Cartosat-1 DEMs.

In addition to weighted averaging, an advanced technique based on TV regularization can be formed for DEM fusion. The general idea was first proposed for image restoration [9], and later combined with a robust data term modeled by L_1 -norm (TV- L_1) [10]:

$$\min_{u_f} \{ \alpha \| \nabla u_f \|_1 + \sum_{i=1}^m \| u_f - d_i \|_1 \} \quad (2)$$

In this equation, the regularization term (left term in the formula) guarantees to achieve a smooth solution (whereas using the L_1 -norm instead of the L_2 -norm prevents smoothing of edges and details) and the L_1 data term (right term) forces the solution to be similar to each of the individual input DEMs. α is a trade-off parameter balancing the two terms of the energy functional. In the energy functional, d_i refers to the input DEMs (here: Cartosat-1 and TanDEM-X DEMs) and u_f

is the unknown fused DEM which will be solved during the minimization procedure of the energy functional. A specifically adapted version of TV- L_1 was also proposed for DEM fusion in which the appropriate weight maps w_i relevant to each input DEMs will assist the fusion algorithm [6]:

$$\min_{u_f} \{ \alpha \| \nabla u_f \|_1 + \sum_{i=1}^m w_i \| u_f - d_i \|_1 \} \quad (3)$$

The main advantage of using TV- L_1 for fusion is that the functional energy is convex, thus ensuring an optimal solution. In this paper, the dual-primal algorithm was followed for solving the convex optimization problem.

Regardless of the specific fusion method, proper weights need to be found that correlate to the local expected accuracy of each DEM. In the next section, more details of the framework for weight map generation by a specifically trained ANN are presented.

4. ANN-BASED WEIGHT MAP PREDICTION

The main idea of weight map prediction for DEM fusion was presented in [2], in which it was shown that the height error maps (HEMs) delivered with the DEM products cannot always support the fusion process properly. Thus, a more sophisticated strategy was proposed aiming at a more precise weight map production. The proposed framework detailed in [2] employs several features that can describe the height residual performance and uses them as input to an artificial neural network (ANN) to predict the weights used in the subsequent fusion procedure. Its training stage consists of three main steps: 1) feature extraction and height residual computation, 2) data preprocessing, and finally 3) training of an artificial neural network to generate a suitable weight predictor. The output of the network is a model that can approximately predict the height residuals in the target areas where two DEMs are supposed to be fused.

For the training of the ANN, training data representing different relevant land types are used. From those, features such as slope, aspect, edginess, different types of roughness, the anisotropic coefficient of variation (ACV) and height error map (HEM) are extracted [11]. In addition, height residual maps are calculated from the corresponding DEM patches and LiDAR ground truth data. Before building the ANN structure, pre-processing the height residuals related to feature values is essential to remove outliers and reduce the noise influence. For this purpose, a two-step mean filtering process is carried out. After pre-processing, the data used to train the ANN for weight map prediction. The framework is implemented for both kinds of DEMs to generate an individual weight map for each DEM separately. The predicted residual maps can then be used as weight maps associated with each kind of DEMs in the fusion process. The NN is trained using the filtered

feature vectors as inputs and the modified height residuals as outputs. 70% of the data are allocated to training, 15% for validation, while the rest is reserved as testing data to monitor the performance of the network during the training.

5. EXPERIMENTS AND RESULTS

5.1. Study Area and DEMs

The Cartosat-1 DEM used in this study has been produced by the XDibias image processing system of DLR, using the pipeline described in [12]. It has a nominal pixel spacing of 5 m. The TanDEM-X raw DEM used in this study was produced by DLRs Integrated TanDEM-X Processor (ITP) from data with an incidence angle of about 38° and a height of ambiguity of about 46 m and has a pixel size of 0.2 arc seconds.

For training the NN, training data representing versatile land types, e.g. inner city area (densely packed, relatively high buildings; subset A), residential areas (single family homes and detached buildings; subset C), agricultural areas, and forested areas (subset B) are used in order to ensure generalizability. The performance of the full DEM fusion chain was then evaluated on a completely independent subset as target area (subset D). Figure 1 shows the locations and descriptions of the training subsets in addition to the target area used to evaluate the full DEM fusion chain. All these study areas are located in the area of Munich, Bavaria, so that for each of them a highly-accurate LiDAR point cloud (with a density of 1 point per m^2) provided by the Bavarian surveying administration is available as reference.

5.2. DEM fusion results

The TanDEM-X and Cartosat-1 DEMs of a subset of the target area – with a size of $238 \times 298 = 70924$ pixels – were fused by employing the methods described in Sections 3 and 4. Then the ability of the weighted versions of these methods were examined in the target area. The weight map for each DEM was produced according to framework that was mentioned in Section 4, supported by the ANN-predicted weights. The size of the training data taken from the Cartosat-1 DEM and the TanDEM-X DEM was 91272 and 89574 pixels, respectively, taken from the three subsets A, B, and C as described above. In Table 2, the fusion results adopting the simple versions of the afore-mention fusion approaches (i.e. averaging and $TV-L_1$) are compared to the versions equipped with ANN-predicted weight maps (weighted averaging: WA, weighted $TV-L_1$: WTV- L_1). The quality of the DEM fusion results is evaluated by different metrics like median, RMSE (Root Mean Square Error), MAE (Mean Absolute Error) and NMAD (Normal Median Absolute Deviation). The green shaded values determine the best results while the blue ones show the improvement of values by WA and WTV- L_1 fusion.

Figure 2 depicts the absolute height residual maps of input DEMs and fused ones by using different DEM fusion methods

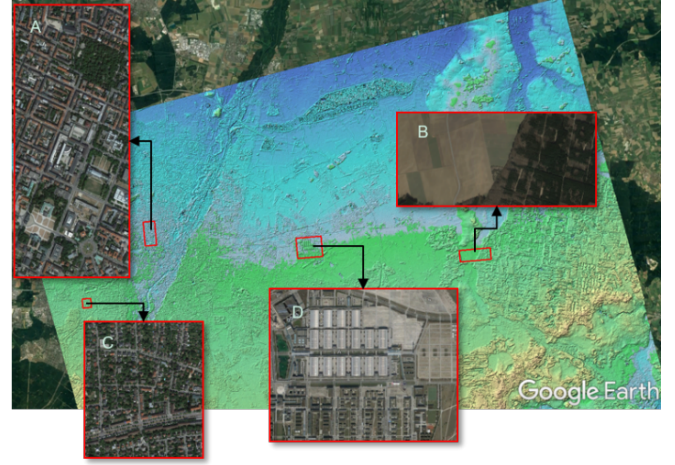


Fig. 1. Locations of training subsets (A, B, C) and target subset (D)

Table 2. Results of fusion of TanDEM-X and Cartosat-1 DEMs (in meter). WA: Weighted Averaging, WTV- L_1 : Weighted $TV-L_1$

DEM	Median	RMSE	MAE	NMAD
Cartosat-1	0.212	4.041	2.407	1.847
TanDEM-X	0.337	5.122	3.412	2.950
Averaging	0.291	3.768	2.539	2.258
ANN-WA	0.236	3.534	2.253	1.815
$TV-L_1$	0.245	3.819	2.487	1.957
ANN-WTV- L_1	0.218	3.695	2.270	1.669

in comparison to LiDAR data.

6. DISCUSSION

The DEM fusion results illustrate that appropriate weights play a vital role, especially when elevation data derived from different systems are to be combined. Both simple averaging and $TV-L_1$ decrease the RMSE value of the final DEM in comparison to the original DEMs.

Using appropriate weights that approximately describe the patterns of errors related to each kind of DEM can improve the values of all metrics. These weights are predicted via a dataset-specific ANN for the target area. The results achieved by using the weighted averaging and weighted $TV-L_1$ fusion method proved the significance of weights for the Cartosat-1 and TanDEM-X DEM fusion task. While simple averaging and $TV-L_1$ -based fusion just improve the RMSE value, their weighted versions lead to a DEM with better quality in comparison to the input DEMs in all metrics. In addition, $TV-L_1$ fusion achieves the greatest improvement for the NMAD metric, which is reasonable since NMAD is a robust metric insensitive to small errors. Generally, the

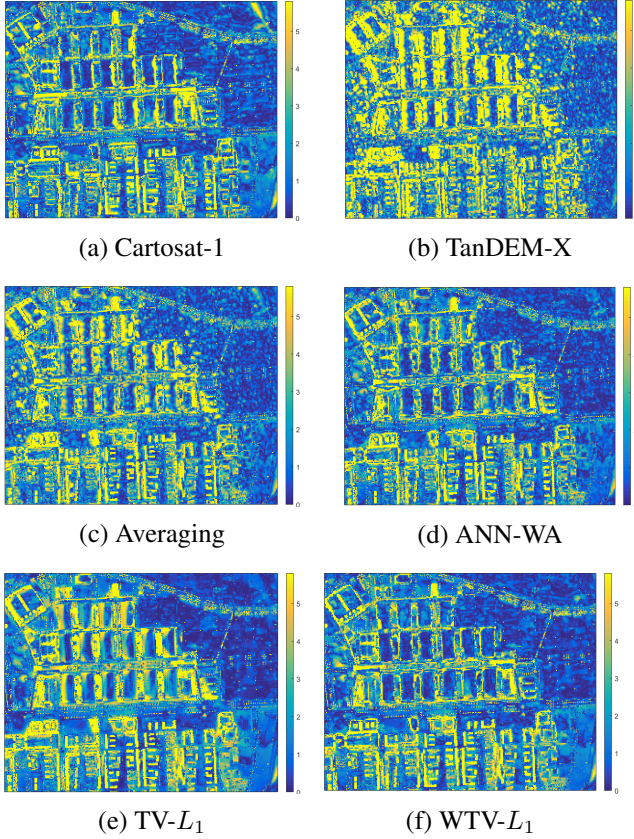


Fig. 2. Absolute height residual maps (meter) before and after DEM fusion.

weighted TV- L_1 fusion approach improves all metrics without creating any systematic bias after fusion process.

7. CONCLUSION AND OUTLOOK

The fusion results shown in this paper demonstrate the importance of using suitable weights for Cartosat-1 and TanDEM-X DEM fusion. Appropriate weights can be predicted for the target area using an ANN pre-trained on a suitable training dataset. The simplest method for DEM fusion is weighted averaging whereas better results can be achieved using a fusion framework based on weighted TV- L_1 regularization.

8. ACKNOWLEDGMENT

The authors would like to thank Prof. Reinartz and Dr. d'Angelo of DLR for providing the Cartosat-1 DEM; Dr. Fritz of DLR for providing the TanDEM-X raw DEM; and the Bavarian Surveying Administration for providing the LiDAR data.

9. REFERENCES

- [1] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, and M. Zink, "TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 11, pp. 3317–3341, 2007.
- [2] H. Bagheri, M. Schmitt, and X. Zhu, "Uncertainty assessment and weight map generation for efficient fusion of TanDEM-X and Cartosat-1 DEMs," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, in press.
- [3] M. Ravanbakhsh and C. S. Fraser, "A comparative study of DEM registration approaches," *Journal of Spatial Science*, vol. 58, no. 1, pp. 79–89, mar 2013.
- [4] P. Reinartz, R. Müller, D. Hoja, M. Lehner, and M. Schroeder, "Comparison and fusion of DEM derived from SPOT-5 HRS and SRTM data and estimation of forest heights," in *Proc. EARSeL Workshop on 3D-Remote Sensing, Porto*, vol. 1, 2005.
- [5] H. Papasaïka, E. Kokiopoulou, E. Baltsavias, K. Schindler, and D. Kressner, "Fusion of Digital Elevation Models Using Sparse Representations," pp. 171–184, 2011.
- [6] G. Kuschik, P. D'Angelo, D. Gaudrie, P. Reinartz, and D. Cremers, "Spatially Regularized Fusion of Multiresolution Digital Surface Models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 3, pp. 1477–1488, 2017.
- [7] C. E. Fuss, A. A. Berg, and J. B. Lindsay, "DEM Fusion using a modified k-means clustering algorithm," *International Journal of Digital Earth*, vol. 9, no. 12, pp. 1242–1255, 2016.
- [8] M. Schmitt and X. X. Zhu, "Data Fusion and Remote Sensing: An ever-growing relationship," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 4, pp. 6–23, dec 2016.
- [9] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, nov 1992.
- [10] T. Pock, L. Zebedin, and H. Bischof, *TGV-Fusion*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 245–258.
- [11] V. Olaya, "Chapter 6 Basic Land-Surface Parameters," in *Geomorphometry Concepts, Software, Applications*, T. H. Science and H. I. R. B. T. D. in Soil, Eds. Elsevier, 2009, vol. Volume 33, pp. 141–169.
- [12] P. D'Angelo, A. Uttenthaler, S. Carl, F. Barner, and P. Reinartz, "Automatic Generation of High Quality Dsm Based on Irs-P5 Cartosat-1 Stereo Data," in *Proc. of ESA Living Planet Symposium*, 2010, pp. 3–7.