ACCURACY ESTIMATION AND FUSION OF DIGITAL ELEVATION MODELS

Haris Papasaika[a], Effrosyni Kokiopoulou[b], Emmanuel Baltsavias[a] and Konrad Schindler[a]
[a]Institute of Geodesy and Photogrammetry
ETH Zurich, Wolfgang-Paulistr. 15, 8093 Zurich, Switzerland
{haris.papasaika|manos.baltsavias|schindler}@geod.baug.ethz.ch

[b]Seminar for Applied Mathematics
ETH Zurich, Rämistrasse 101, 8092 Zurich, Switzerland
effrosyni.kokiopoulou@sam.math.ethz.ch

KEY WORDS: DEMs, data fusion, quality evaluation, learning

ABSTRACT: Nowadays, different sensors and processing techniques provide Digital Elevation Models (DEMs) for the same site, which differ significantly with regard to their geometric characteristics and accuracy. Each DEM contains intrinsic errors due to the primary data acquisition technology, the processing chain, and the characteristics of the terrain. DEM fusion aims at overcoming the limitations of different DEMs by merging them in an intelligent way. In this paper we present a generic algorithmic approach for fusing two arbitrary DEMs, using the framework of sparse representations. We conduct extensive experiments with real DEMs from different earth observation satellites to validate the proposed approach. Our evaluation shows that, together with adequately chosen fusion weights, the fusion yields consistently better DEMs (up to 30%), while blunders are reduced.

1 INTRODUCTION

Digital Elevations Models (DEMs) are one of the most important types of geodata. They are needed in a large number of applications, ranging from virtual globes and visualization to engineering and environmental planning. DEMs of larger areas are usually generated either by photogrammetric processing of aerial and satellite images, SAR (Synthetic Aperture Radar) interferometry, or laser scanning (mainly from airborne platforms). Each sensing technology has its own strengths and weaknesses, and even within one technology the variations in DEM quality are large. DEMs are available at different scales from tailor-made local models to national and even global coverage. We are primarily interested in large-scale national and global products, whose resolution, accuracy, error characteristics, and homogeneity vary a lot. In most cases, the DEM producers provide users with information only on production technology, date of acquisition, and resolution, but only with coarse accuracy measures that fail to capture the local variations in data quality – sometimes only a single global number. DEM fusion – and its necessary prerequisite, fine-grained quality characterization of the inputs – has several benefits: improved accuracy, homogeneity and completeness, as well as fine-grained quality information for the final product. We deal only with $2\frac{1}{2}$-D surfaces in regular grid format, which constitute the vast majority of large-scale DEMs.
In this work we make two contributions:

- we develop a computationally efficient and flexible mathematical method for robust fusion of $2\frac{1}{2}$-D surface models. The formulation is *generic* and can be applied with any two input DEMs, independent of the sensor technology and processing with which they were created, making it useful for practical applications; it takes into account both prior information about plausible terrain shapes (in the form of a dictionary), and the local accuracy of the inputs, controled by interpretable weights; and it poses the complete fusion as a clean, convex mathematical optimisation problem that can be solved to global optimality, and in which the influence of the input DEMs is controlled by an interpretable set of local fusion weights.
- we propose a data-driven method, which allows one to derive local measures of DEM quality (and thus also fusion weights) for each point or segment of a DEM, if no such information is available. To this end we use as input geomorphological characteristics of the terrain (slope, roughness), which are derived directly from the DEMs, as well as optionally semantic information such as land-cover maps. Using existing high-quality ground-truth DEMs as reference, we learn regression functions relating the available geomorphological characteristics to the DEM quality, which then allow one to estimate the local quality of a new DEM.

Surprisingly, there is a relatively small body of work about data fusion for combining DEMs. Schultz et al. (1999) develop a methodology to fuse two stereo-optical DEMs. Honikel (1999) applies two techniques which take advantage of the complementary properties of InSAR and stereo-optical DEMs. In (Damon, 2002) and (Gamba et al., 2003) the specific combination of InSAR and LIDAR data is considered. Roth et al. (2002) describe a technique to combine multi-source DEMs, based on the concept of height error maps. Slatton et al. (2002) combine space-borne InSAR data from the ERS-1/2 platforms with multiple sets of airborne C-band InSAR data using a multi-scale Kalman smoothing approach. Rao et al. (2003) fill holes in an InSAR DEM with height data derived with stereo optical matching. Kääb (2005) combined SRTM and ASTER DEMs to fill the gaps of SRTM. Podobnikar (2005) introduces a fusion technique based on the weighted sum of data sources with geomorphological enhancement.

2 MATHEMATICAL FORMULATION OF FUSION

In the following, we describe the problem statement. Consider two noisy measurements y_l and y_h of a height field x, possibly at different resolutions, e.g., low and high respectively. We assume that the measurements have been produced from the original DEM x by the following model:

$$y_h = x + \epsilon_h \quad \text{and} \quad y_l = Lx + \epsilon_l, \tag{1}$$

where $\epsilon_h,\ \epsilon_l$ are noise vectors and L is an unknown downsampling operator. In the case when the two measurements y_l and y_h are at the same resolution, the operator L equals to 1. The problem addressed in this paper is to fuse the noisy measurements y_l and y_h in order to recover the original DEM x. The hope is that the redundancy of the measurements will offer robustness against noise and result in an accurate estimation of x.

Problem formulation. In order to achieve robustness without oversmoothing, we pose the fusion problem in the framework of sparse representations. Sparse representations have been shown to result in state-of-the-art performance in image denoising (Elad & Aharon, 2006) and super-resolution problems (Yang et al., 2008). To our knowledge, their potential has not been fully exploited in remote sensing problems. We work with local DEM patches to achieve computational efficiency and to ensure a moderately sized dictionary able to capture the variations in terrain shape. In what follows, y_h, y_l, and x thus denote local patches in the corresponding terrain models in (1).

We assume that x can be represented as a sparse linear combination of elements from a dictionary D (i.e., local terrain shapes). The dictionary is a basis set spanning the signal space and is typically overcomplete, that is, it contains more elements than the dimension of the signal space. The elements of the dictionary are called atoms. We say that x is sparsely represented over D if $x = D\alpha_0$, where $\alpha_0 \in \mathbb{R}^N$ is a sparse coefficient vector with most entries zero and very few non-zero entries (Figure 1). N denotes the size of the dictionary whose atoms are organized as columns of D. The sparsity of α_0 implies that already a few atoms are sufficient for obtaining a good approximation of x owing to the overcompleteness of the dictionary. Under this representation, the generative model (1) can be re-written as

$$y_h = \underbrace{D}_{:=D_h}\alpha_0 + \epsilon_h \quad \text{and} \quad y_l = \underbrace{LD}_{:=D_l}\alpha_0 + \epsilon_l, \tag{2}$$

where we have defined a high-resolution dictionary D_h and a low-resolution dictionary D_l. Both dictionaries are coupled via the relation $D_l := LD_h$. The key observation in (2) is that the *same* sparse coefficient vector α_0 is involved in both measured DEMs y_h and y_l. This leads to the following optimization problem in order to recover x from the measured y_l, y_h.

Optimization problem. Assume for a moment that D_h and D_l are available. We postpone the discussion of how to determine these two dictionaries until the end of this section. Given the two dictionaries D_h and D_l and the measurements y_l and y_h, we would like to recover the sparse coefficient vector α_0. Once α_0 has been computed, one can simply recover x by computing $D_h\alpha_0$.

$$\min_{\alpha\in\mathbb{R}^N} \underbrace{\|D_l\alpha - y_l\|_2^2}_{\text{low resolution}} + \underbrace{\|D_h\alpha - y_h\|_2^2}_{\text{high resolution}} + \underbrace{\tau\|\alpha\|_1}_{\text{sparsity term}} \tag{3}$$

The first two (data) terms correspond to the reconstruction error with respect to the observed DEMs y_l and y_h. The third (regularisation) term is associated with the ℓ^1 norm of the candidate solution vector α. It is well known that the minimization of the ℓ^1 norm encourages a sparse solution. Since the true coefficient vector α_0 that we seek to recover is sparse, we would like our estimated solution α to be sparse as well. The parameter $\tau > 0$ controls the trade-off between data fitting and sparsity; its choice is discussed in Sec. 3.2.

The current formulation (3) implicitly assumes that both data terms have the same importance. However, this is not typically the case with DEMs, since the two inputs have, at each point, different accuracy, depending on the sensing technology and processing. It is therefore beneficial to include weights in the problem formulation that will reflect such prior knowledge. We therefore modify the optimization to include positive weight vectors w_l, w_h:

$$\min_{\alpha\in\mathbb{R}^N} \|\sqrt{w}_l \odot (D_l\alpha - y_l)\|_2^2 + \|\sqrt{w}_h \odot (D_h\alpha - y_h)\|_2^2 + \tau\|\alpha\|_1. \tag{4}$$

Here, $\odot$ denotes component-wise multiplication, which offers flexibility in allowing for individual weights at each location. Section 3.1 discusses the choice of these weights, which are crucial for a good fusion.

Consistency among neighbouring patches. Solving (4) for each patch independently would result in blocking artifacts along the patch borders. To remedy this problem, we introduce overlap between patches and impose consistency between neighbouring patches. More specifically, let P denote an operator that extracts the overlap region between the current working patch and the patches that have been computed before. Furthermore let y_p denote a vector that collects the values of the estimated DEM in the overlap region. Minimizing the discrepancy $\|PD_h\alpha - y_p\|_2^2$ between overlapping patches will impose consistency and ensure smooth transitions. Introducing this term into (4), we reach the final formulation of our optimization problem:

$$\begin{aligned}\min_{\alpha\in\mathbb{R}^N} \quad & \|\sqrt{w}_l \odot (D_l\alpha - y_l)\|_2^2 + \|\sqrt{w}_h \odot (D_h\alpha - y_h)\|_2^2 \\ & +\beta\|PD_h\alpha - y_p\|_2^2 + \tau\|\alpha\|_1,\end{aligned} \tag{5}$$

where we have introduced the parameter $\beta > 0$ to control the influence of the patch overlap factor. The choice of this parameter is discussed in Sec. 3.2.

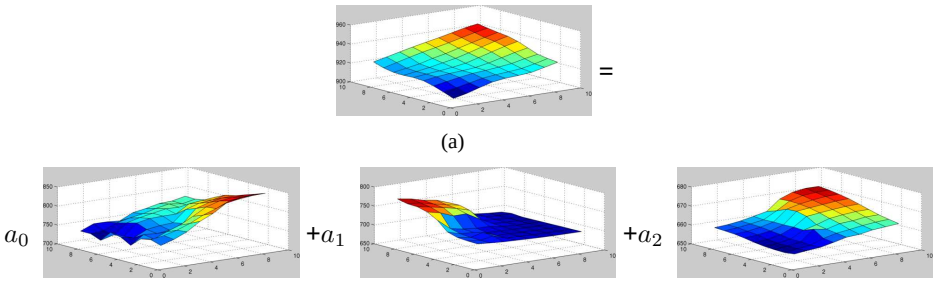


Figure 1: Reconstruction of a 9×9 patch (a) from three non-zero atoms (b)-(d), where $[a_0, a_1, a_2]$ is the sparse non-zero coefficient vector.

Problem (5) is a convex ℓ^1-regularized least-squares problem that can be solved to global optimality. Since optimization problems of this form constitute the main computational kernel of compressed sensing applications, there exists a wide selection of algorithms for their solution. Here, we use *Orthogonal Matching Pursuit* (OMP) (Mallat, 1998), because of its simplicity and computational efficiency. Problem (5) is solved for each patch with OMP. The OMP code reproducing the results in this paper is available for download at (Elad, 2011). Details concerning the processing time are discussed in Sec. 4.

Dictionary construction. The proposed framework requires dictionaries D_h, D_l, which must be acquired from training data. Different learning techniques could be used to obtain a set of atoms from available high-quality DEMs. We have experimented with different methods and found that the best results are obtained by simple random sampling of patches from high resolution DEMs, followed by clustering to remove very similar samples. Hence, for the construction of D_h, we use a training set of high resolution DEMs of high quality (which are of course different from the test DEMs used in the evaluation). If D_l is of lower resolution, its atoms are obtained by downsampling the corresponding atoms in D_h with bicubic interpolation. The preparation of the dictionaries is off-line and needs to be performed only once. Our empirical results in Sec. 4 demonstrate that the dictionaries constructed with the above procedure are well suited for successfully representing real terrain patches using very few (less than ten) atoms.

3 DEM QUALITY AND WEIGHTS

The DEM fusion method used in this research consists primarily of two steps: quality evaluation and fusion. It is assumed that the input DEMs are co-registered into the same coordinate system. The co-registration operates by minimizing 3D separations between a template (master) DEM and a second search (slave) DEM, see Gruen & Akca (2005). After co-registration, the low resolution DEM is resampled at the nodes of the high resolution DEM by bicubic interpolation.

3.1 Fusion Weights

The fusion is accomplished with the support of weight maps that reflect the estimated relative accuracy of the two DEMs at every single grid point. In some cases, DEM providers deliver such error maps, which can then be directly used for fusion. In most, cases, however, these maps are either not available or not reliable, and the weights need to be estimated from the data. We have explored a data-driven strategy to find the weights based on geomorphological characteristics. Geometric properties of the terrain can be derived directly from a given DEM with local characteristics such as slope, aspect, roughness, curvature, etc. We calculate two such parameters, *slope* and *roughness*, and analyze their relation to the co-registration residuals.
Roughness refers to the height variation within a local neighborhood and can be measured in different ways using, e.g., the standard deviation or fractal dimension. We have experimented with several methods and found the entropy to perform best for our purposes. The entropy $E(x, y)$ is defined as $E(x, y) = \sum(p \times \log p)$, where p are the probability densities of the heights, approximated by histogram counts. Each output grid cell contains the entropy value of the surrounding $n \times n$ neighborhood. We point out that the residuals vary (increase or decrease) in a non-random pattern with the two extracted geomorphological parameters. We learn the mapping from a parameter to the expected residual (accuracy) with Gaussian process regression. An example is shown in Figure 2(b). After the mapping, we adjust the expected residuals by linearly scaling the values between 0 and 100, eventually yielding an accuracy map. In the last step, we normalize the resulting accuracy maps of both input DEMs at each overlapping point. The reciprocal values are the weights used for the fusion.

3.2 Fusion of DEMs

The fused DEM covers the common area available in all given DEMs. After merging, a single DEM exists with the same grid spacing as the DEM with the smallest grid spacing. According to the mathematical formulation of the fusion algorithm described in Sec. 2, we have to set the overlapping parameter β, the number of the non-zero atoms used in OMP, the patch size and the number of patches in the dictionary. In order to fine tune these parameters we performed numerous tests using artificial and real world datasets and we compared for each test the produced results with available high quality reference data. Best results have been achieved with the following set of parameters. The overlap parameter β is set within the interval $[0.5, 1.5]$. The number of non-zero atoms used in OMP is set between 7 and 15. Below 7 the results are not reliable and above 15 the processing time increases while the results do not improve any further. The minimum patch size should not be smaller than 3×3 and it should not be bigger than 9×9 because then the processing window becomes too complicated and it is more difficult to find a suitable combination of non-zero sparse atoms to reconstruct it. A “good”

Table 1: (a) Co-registration results. σ_0 is the σ a posteriori, and T_x, T_y, and T_z are the three translations. (b-c) Statistical results of the ALOS-SPOT, ERS-SPOT and ALOS-ERS fusion for the complete area. All units are in meters.

(a)

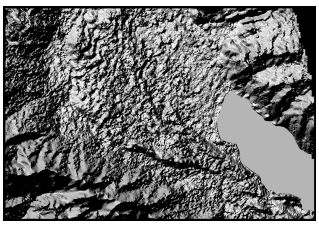
Master	Slave	σ_0	T_x	T_y	T_z
L	A	13.4	18.6	6.4	1.0
L	E	8.9	-16.9	13.8	1.2
L	S	8.9	16.6	3.2	2.9

(b)

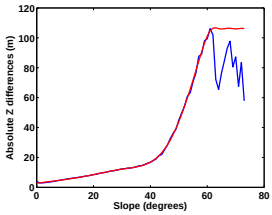
	Input DEMs		
	MEAN	RMSE	MAD
L-A	-1.0	19.4	6.6
L-S	-0.8	15.5	4.4
L-E	-1.2	10.7	3.1

(c)

	Fusion DEMs		
	MEAN	RMSE	MAD
L-F1	-0.7	11.0	4.2
L-F2	-1.2	9.7	2.7
L-F3	-0.7	10.9	3.1



(a) ALOS DEM



(b) SPOT Weights

Figure 2: (a) Overlap area. (b) Absolute Z differences between the Lidar and the SPOT DEM versus slope (mean every degree). The blue line are the original values. The red line is the Gaussians sum that we fit to the original values.

dictionary should contain atoms that describe every possible geomorphological structure, e.g., urban, forest, flat, mountainous areas. This ensures that OMP can find a well fitting combination of atoms for every patch.

4 RESULTS AND DISCUSSION

In this section, we present the results of an experimental validation of the fusion methodology in three realistic DEM fusion examples. The test site is located at Thun, Switzerland, and characterized by areas with different morphology and land cover. We used three DEMs produced with image matching and SAR interferometry. Figure 2(a) shows the area of overlap. The validation was performed by comparing the input and the obtained DEMs after the fusion with a high quality reference lidar DEM provided by Swisstopo.

SPOT Reference 3D DEM (S): 30m grid spacing. Image acquisition date: 30.09.2002. It is produced using image matching by SpotImage. The given absolute elevation accuracy for flat or rolling terrain (slope$\leq$20%) is 10m, for hilly terrain (20 % $\leq$ slope $\leq$ 40%) is 18 m and for the mountainous terrain (slope$>$ 40%) is 30m. The Reference 3D DEM is delivered with a High Resolution orthoimage (SPOT 5 sensor) with 5m ground pixel size.

ALOS/PALSAR DEM (A): 15m grid spacing. Image acquisition date: master 19.06.2006 and slave 04.08.2006. L-Band (ca. 23 cm wavelength). It is produced by Sarmap SA. The overall accuracy is 20m and it has been estimated using the Lidar DEM.

ERS DEM (E): 25m grid spacing. Image acquisition date: late autumn to early spring time, obtained from 1995 to 1998. C-Band (ca. 6cm wavelength). It is produced by Sarmap SA. The overall accuracy is 10.7m and it has been estimated using the Lidar DEM.

Lidar DEM (L): 2m grid spacing. The airborne lidar data were acquired for the Swisstopo in 2000 with a mean density of 1-2 points per m^2, depending on the terrain, and with first and last pulse recorded. The accuracy (1 σ) of the derived DEMs is 0.5m and 1.5m for vegetated areas.

The ALOS and the ERS DEMs are delivered with an accuracy map. The values on the accuracy map are derived according to the formula: $\sigma = \mathrm{AF} \cdot (1 - \mathrm{coherence}^2)/(2 - \mathrm{coherence}^2)$, where $\mathrm{AF} = \mathrm{R} \cdot (\sin(\theta)/(B_n \cdot 4\pi/\lambda))$, R is the range, θ is the local incidence angle, B_n is the baseline normal component and λ is the wavelength. According to our experience these accuracy maps do not always depict the real quality of the DEMs.

We tested the fusion method by fusing (a) the ALOS with the SPOT DEM which results to a final F1 DEM, and (b) the ERS with the SPOT DEM which results to a final F2 DEM and (c) the ALOS with the ERS DEM which results to a final F3 DEM. The three DEMs (S, A, and E) were co-registered to the reference DEM (L). Table 1(a) shows the results of the co-registration. A dictionary of 800 patches of size 9×9 was generated with elements drawn randomly from the lidar DEM. The dictionary was filtered using a K-means clustering algorithm with an eulidean distance measure of 10 m. The clustering reduced the dictionary to 720 patches. In all the fusion examples that we describe below we set the overlap parameter β to 1, the number of the non-zeros atoms used in OMP was set to 10 and we used a 9×9 patch size processing window.

After the fusion, error maps were computed by subtracting the individual DEMs (input and output) from the reference DEM (L) and several statistics measures were computed. For this reason, a grid was generated for all the DEMs at 2 m intervals according to the spatial resolution of the reference DEM. The statistics included in all the tables that follow are (a) the mean value (MEAN), (b) the root mean square error (RMSE), and (c) the median absolute deviation of the median value (MAD). All units are in meters. At the end, we performed a more detailed analysis of the results in relation to the slope, and roughness. The slope and the roughness classes were obtained by processing the lidar DEM. For the calculations we used a 5×5 pixel window. The three slope classes are: Slope$\leq$15°, 15°$<$Slope$\leq$45°, Slope$>$45°. The three roughness classes are: Roughness$\leq$10, 10$<$Roughness$\leq$30, Roughness$>$30. The roughness is scaled in the interval $[0, 100]$.

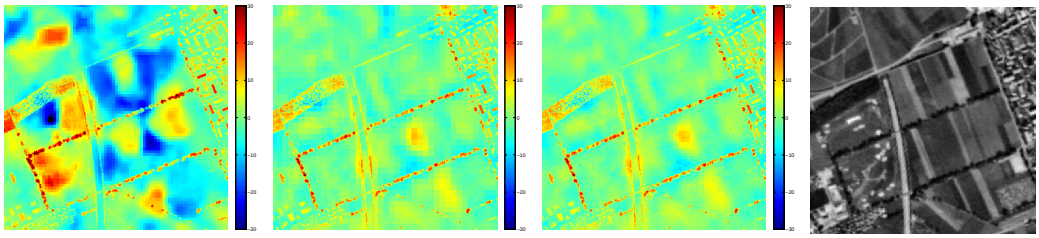
We have run the same tests, instead of using the sparse representation mathematical framework of fusion, with a weighted mean average. The results were very similar, further supporting that the major issue in fusion is the derivation of the weights for each individual input DEM point.

All the tests are done using a computer with Intel Core i7, Q720, 1.6 GHz CPU and 8 GB RAM using only one core, and unoptimized Matlab code.

Table 2: Slope and roughness classes analysis.

Slope	S≤15°, 56.2%			15°<S≤45°, 37.8%			S>45°, 6.0%		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-A	-0.6	11.1	4.1	-0.8	23.6	11	-4.8	39.6	19.4
L-E	-0.9	5.6	1.6	1.1	12.7	6.7	3	24.7	12.1
L-S	-0.2	6.3	3.1	-1.8	17.9	6.6	-13.3	39.9	14.1
L-F1	-0.1	5.4	3	-1.3	12.8	6.1	-7.4	26.2	12.9
L-F2	-0.6	5.2	1.5	-1.4	12.4	6.2	-7.2	24.8	17.2
L-F3	-0.9	5.5	1.6	1	12.7	6.6	2.9	24.7	12.1

Rough.	R≤10, 17.2%			10<R≤30, 25.9%			R>30, 56.9%		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-A	-1.6	9	0.2	0.3	8.5	4.2	-1.3	24.5	10.7
L-E	-1.4	2.4	0.3	-0.8	4.5	1.9	0.9	13.8	6.4
L-S	1.5	3.6	0.9	-0.8	4.4	2.2	-2.9	20.2	6.1
L-F1	1.4	3.4	0.9	-0.5	4.2	2.1	-1.9	14.1	5.7
L-F2	-0.7	1.8	0	-0.5	3.8	1.5	-1.5	11.8	4.6
L-F3	-1.4	2.4	0.3	-0.8	4.5	1.9	0.9	13.8	6.4



(a) (b) (c) (d)

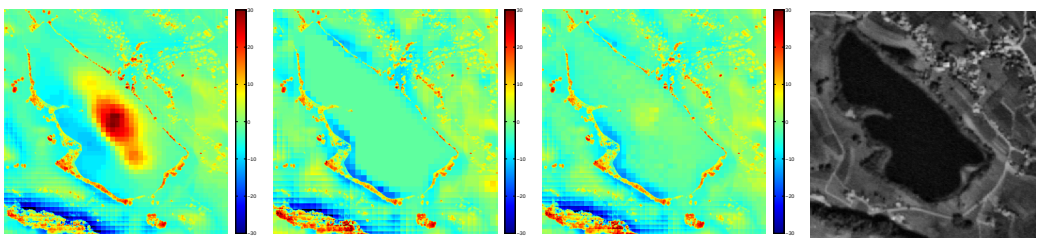
Figure 3: ALOS-SPOT fusion example. (a) Residuals between the L and A DEM, (b) Residuals between the L and S DEM, (c) Residuals between the L and F1 DEM. (d) SPOT orthoimage. The colored Z residuals are mapped in the interval [-30,30]. The bar unit is meters.

4.1 Fusion ALOS-SPOT

A grid was generated for the input DEMs at 15m intervals according to the spatial resolution of the ALOS DEM. After resampling, the input DEMs have a grid size of 800×1167. The processing time required for the fusion was 3.8 minutes. The weights for the ALOS DEM are calculated using the given accuracy map. For the calculation of the weights of the SPOT DEM a fairly good relationship was found between the height differences map of the two input DEMs and the slope. In Figure 2(b) the used weighting function is shown. It is a sum of 6^{th} degree Gaussian functions that we fit to the original values using peak fitting. In Tables 1(b) and 1(c), we can see that compared to the ALOS DEM, the fusion achieved up to 43% improvement in RMSE while maintaining the resolution of 15m. Similarly, as compared to the SPOT DEM, the fusion improved the resolution to 15m from 30m while improving the RMSE by 29%. Figure 3 shows a detail of the Z difference images of the three DEMs. The error of the ALOS DEM is not introduced into the final DEM F1 which supports the choice of the weights. The results of the slope and roughness assessment are presented in the Table 2. We notice that the fusion leads to an improvement especially for medium and high slopes and less for low slopes. The same behavior applies to the three roughness classes but to a lesser degree than for the slope.

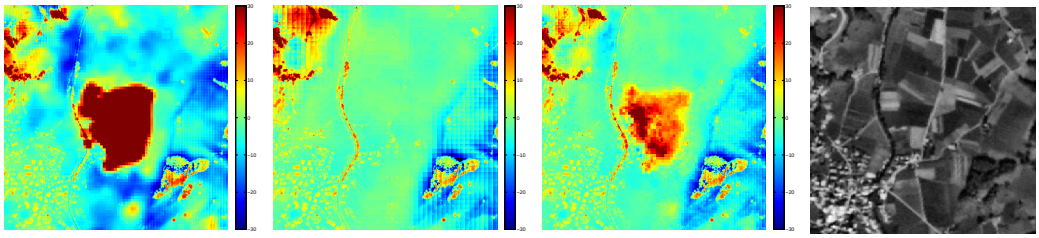
4.2 Fusion ERS-SPOT

A grid was generated for the input DEMs at 25m intervals according to the spatial resolution of ERS. After resampling, the input DEMs have a grid size of 480×700. The processing time required for the fusion was 1.3 minutes. The weights for the ERS DEM are calculated using the given accuracy map. The weights for the SPOT DEM are calculated as described above. In Tables 1(b) and 1(c), we can see that as compared to the ERS DEM, the technique achieved up to 10% improvement in RMSE while maintaining the resolution of 25m. Similarly, as compared to the SPOT DEM, the technique improved the resolution to 25m from 30m while improving the RMSE by 37%. Figure 4 shows a detail of the Z difference images of the three DEMs. The errors of the ERS DEM are not introduced into the final DEM F2. The results of the slope and roughness assessment are presented in the Table 2. The analysis of these results conforms to the analysis for ALOS-SPOT fusion. Here the improvement by fusion is larger for medium and high slopes and roughness than for low ones.



(a) (b) (c) (d)

Figure 4: ERS-SPOT fusion example. (a) Residuals between the L and E DEM, (b) Residuals between the L and S DEM, (c) Residuals between the L and F2 DEM. (d) SPOT Orthoimage. The colored Z residuals are mapped in the interval [-30,30]. The bar unit is meters.



(a) (b) (c) (d)

Figure 5: ALOS-ERS fusion example. (a) Residuals between the L and A DEM, (b) Residuals between the L and E DEM, (c) Residuals between the L and F3 DEM. (d) SPOT orthoimage. The colored Z residuals are mapped in the interval [-30,30]. The bar unit is meters.

4.3 Fusion ALOS-ERS

A grid was generated for the input DEMs at 15m intervals according to the spatial resolution of the ALOS DEM. After resampling, the input DEMs have a grid size of 800×1167. The processing time required for the fusion was 3.8 minutes. The weights are calculated using the given accuracy maps, given the fact that the weights are inversely proportional to the standard deviation values. In Tables 1(b) and 1(c), we can see that as compared to the ALOS DEM, the technique achieved up to 45% improvement in RMSE while maintaining the resolution of 15m. Similarly, as compared to the ERS DEM, the technique improved the resolution to 15m from 25m while not improving at all the RMSE. Figure 5 shows a detail of the Z difference images of the three DEMs. In the ALOS DEM a large blunder exists which does not appear in the ALOS accuracy map, so this blunder is introduced into the final result F3. The results of the slope and roughness assessment are presented in the Table 2. Here for medium and large slopes / roughness there is no impovement by fusion while for low values the improvement is significant.

5 CONCLUSIONS AND FUTURE WORK

We have proposed a methodology for DEM fusion based on sparse representations. First, we have introduced a mathematical framework for the fusion. Next, we have proposed a way to calculate the weight maps for the input DEMs when no prior information is available. We provide experimental evidence using real DEMs that indicates the advantages of the proposed approach after the examination of the post-fusion DEMs. Strategies that take advantage of some complementary factors like the edginess, the land cover or the special attributes of the DEMs production technology are likely to be realized in the near future for the calculation of the weight maps.

ACKNOWLEDGEMENTS

We acknowledge support of this research by the Commission for Technology and Innvovation (CTI), Switzerland within the framework of the project Synergos, in cooperation with the company Sarmap SA, Switzerland, by the Swiss Federal Office of Topography, Bern who provided the lidar DEM and by Sarmap SA, Switzerland who provided the ALOS/PALSAR and the ERS DEMs.

References

Damon, J. (2002). Fusing Lidar and IfSAR DEMs: a seven-step methodology. In *22nd ESRI International Conference*.

Elad, M. (2011). Michael Elad personal webpage, software.

Elad, M. & Aharon, M. (2006). Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. on Image Processing*, 15(12), 3736–3745.

Gamba, P., Dell'acqua, F., & Houshmand, B. (2003). Comparison and fusion of Lidar and InSAR digital elevation models over urban areas. *International Journal of Remote Sensing*, 24(22), 4289–4300.

Gruen, A. & Akca, D. (2005). Least squares 3D surface and curve matching. *ISPRS Journal of Photogrammetry & Remote Sensing*, 59(3), 151–174.

Honikel, M. (1999). Strategies and methods for the fusion of digital elevation models from optical and SAR data. *International Archives of Photogrammetry and Remote Sensing*, 32(7-4-3W6), 83–89.

Kääb, A. (2005). Combination of SRTM3 and repeat ASTER data for deriving alpine glacier flow velocities in the Bhutan Himalaya. *Remote Sensing of Environment*, 94(4), 463–474.

Mallat, S. (1998). *A Wavelet Tour of Signal Processing, 2nd edn.* Academic Press.

Podobnikar, T. (2005). Production of integrated digital terrain model from multiple datasets of different quality. *International Journal of Geographical Information Science*, 19(1), 69–89.

Rao, Y. S., Rao, K. S., Venkataraman, G., Khare, M., & Reddy, C. (2003). Comparison and fusion of DEMs derived from InSAR and optical stereo techniques. In *Third ESA InternationalWorkshop on ERS SAR Interferometry*.

Roth, A., Knöpfle, W., Strunz, G., Lehner, M., & Reinartz, P. (2002). Towards a global elevation product: combination of multi-source digital elevation models. In *Joint International Symposium on Geospatial Theory, Processing and Applications*.

Schultz, H., Riseman, E., Stolle, F., & Woo, D. (1999). Error detection and DEM fusion using self-consistency. In *International Conference on Computer Vision*.

Slatton, K., Teng, S., & Crawford, M. (2002). Multiscale fusion of InSAR data for hydrological applications. In *Symposium on Terrain Analysis for Water Resources Applications*.

Yang, J., Wright, J., Huang, T. S., & Ma, Y. (2008). Image super-resolution as sparse representation of raw image patches. In *IEEE Conf. on Computer Vision and Pattern Recognition*.