Sparsity-seeking fusion of digital elevation models

H. Papasaika*, E. Kokiopoulou, E. Baltsavias*, K. Schindler*
and D. Kressner

*Institute of Geodesy and Photogrammetry, ETH Zurich, Wolfgang-Paulistr. 15, 8093
Zurich, Switzerland

SPARSITY-SEEKING FUSION OF DIGITAL ELEVATION MODELS

Haris Papasaika[a], Effrosyni Kokiopoulou[b], Emmanuel Baltsavias[a], Konrad Schindler[a] and Daniel Kressner[b]

[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|daniel.kressner}@sam.math.ethz.ch

ABSTRACT:

Nowadays, different Digital Elevation Models (DEMs) are typically available for the same geographic area, which are provided by different sensors and/or processing techniques. Each DEM contains measurement errors due to the acquisition technology, the atmospheric conditions, the processing chain and the characteristics of the observed terrain. DEM fusion aims at overcoming the limitations of individual DEMs by merging them in an intelligent way that exploits the available complementary information and discards the redundancies. In this paper we present a generic algorithmic approach for fusing two arbitrary DEMs of the same geographical area, using the framework of sparse representations. A signal is called sparse when most or all of its information can be captured by a linear combination of a few elementary signals, also known as atoms. The atoms are chosen from a dictionary, i.e. an over-complete basis set, from which only a small subset is required to effectively represent each actual signal. We provide experimental results on real elevation datasets from different earth observation satellites to validate the proposed approach. Our evaluation shows that the proposed algorithm, along with carefully chosen fusion weights, yields consistently better DEMs.

1 INTRODUCTION

1.1 Motivation and Aims

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 (as an example, consider the characteristics of image matching algorithms). 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.

In an ideal world, one would of course obtain the raw measurements and sensor models from all sensors, and merge them by fitting a single DEM to the entire set of heterogeneous observations, along the way computing quality measures for every single height value. Unfortunately, this is usually not feasible in practice. Thus, one resorts to the next best solution, namely to fuse DEMs from different providers into a higher-quality product, and estimate its quality in the process from the available redundancy.

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 (although our framwork could in principle be extended to TINs).

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.

The proposed method is evaluated in detail with three different satellite datasets, and shows a significant improvement in DEM quality, consistently over all combinations of inputs.

1.2 Related Work

Surprisingly, there is a relatively small body of work about data fusion for combining DEMs. Schultz et al. (1999) developed a methodology to fuse two stereo-optical DEMs. The techniques discussed in this paper are based on the concept of using self-consistency to identify potentially unreliable points. Honikel (1999)

applies two techniques which take advantage of the complementary properties of InSAR and stereo-optical DEMs. First, the cross-correlation of phases, respectively the gray values, are determined for each DEM point and used as fusion weights. The second technique takes advantage of the fact that errors of the DEMs are of different nature, and attempts to find regions where each DEM is less correct than the other, and to replace the data with the one from the more correct counterpart. 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 which is based on the concept of height error maps. The pre-condition is the availability of a height error description and the fusion is done by a weighted averaging. Slatton et al. (2002) combined 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, and then used the resulting DEM to derive glacier flow in the mountains of Bhutan. Podobnikar (2005) introduces a fusion technique based on the weighted sum of data sources with geomorphological enhancement. A DEM is modelled through the averaging of individual datasets but considering their quality. The main step of geomorphological enhancement is the generation of trend surfaces as low frequency functions.

2 MATHEMATICAL FORMULATION OF FUSION

We describe now 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 x by the following model,

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

where ϵ_h, ϵ_l are noise vectors and L is an unknown downsampling operator. In case 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. The problem, as stated above, can be seen as a denoising problem – in the case of different resolutions, with simultaneous super-resolution for the coarser signal.

Problem formulation. In order to achieve robustness without oversmoothing, we pose the fusion problem in the framework of sparse representations. Sparse representations have resulted in state-of-the-art performance in image denoising Elad and Aharon (2006) and super-resolution problems Yang et al. (2008) and, to our knowledge, their potential has not been fully exploited in remote sensing problems. We work with local DEM patches for 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 i.e. contains more elements than the dimension of the signal space. The elements of the dictionary are called atoms. When x is sparsely represented over D it means that $x = D\alpha_0$, where $\alpha_0 \in \mathbb{R}^N$ is a sparse coefficient vector

whose most of the entries are zero and very few entries are non-zero (Figure 1). N denotes the size of the dictionary whose atoms are organized as columns of D. The sparsity of α_0 implies that only a few atoms are sufficient towards obtaining a good approximation of x thanks to the overcompleteness of the dictionary. Under this representation, the generative model (1) can be re-written as

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

where we have further defined a high-resolution dictionary D_h and a low-resolution dictionary D_l, which is coupled with D_h 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 optimisation problem in order to recover x from the measured y_l, y_r.

Optimization problem. Assume for a moment that D_h and D_l are available. We postpone the discussion of how we 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 $D_h\alpha_0$.

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

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[1] (see, e.g., Tibshirani (1996)). 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 formulation (3) in its current form 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 such weights w_l, w_h:

$$\min_{\alpha \in \mathbb{R}^N} \|\sqrt{w_l} \odot (D_l\alpha - y_l)\|_2^2 + \|\sqrt{w_h} \odot (D_h\alpha - y_h)\|_2^2 + \tau\|\alpha\|_1 \quad (4)$$

In the above, $\odot$ denotes component-wise multiplication, which is more flexible since it allows 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

---

[1] In fact it is the basis of the sparse representation framework that the computationally inconvenient number of non-zero elements (ℓ^0-norm) can be replaced by the ℓ^1 norm.

smooth transitions. Introducing this term into (4), we reach the final formulation of our optimization problem:

$$\min_{\alpha \in \mathbb{R}^N} \quad \|\sqrt{w_l} \odot (D_l\alpha - y_l)\|_2^2 + \|\sqrt{w_h} \odot (D_h\alpha - y_h)\|_2^2$$

$$+\beta\|PD_h\alpha - y_p\|_2^2 + \tau\|\alpha\|_1, \tag{5}$$

where we have introduced $\beta > 0$ to control the influence of the patch overlap factor. We discuss its choice in Sec. 3.2.

Equation (5) can be written in the following form:

$$\min_{\alpha \in \mathbb{R}^N} \|\tilde{D}\alpha - \tilde{y}\|_2^2 + \tau\|\alpha\|_1, \qquad \text{where}$$

$$\tilde{D} = \begin{bmatrix} \sqrt{w_l} \odot D_l \\ \sqrt{w_h} \odot D_h \\ \sqrt{\beta}PD_h \end{bmatrix} \quad \text{and} \quad \tilde{y} = \begin{bmatrix} \sqrt{w_l} \odot y_l \\ \sqrt{w_h} \odot y_h \\ \sqrt{\beta}y_p \end{bmatrix}. \tag{6}$$

Problem (6) is a convex ℓ^1-regularized least-squares problem that can be solved to global optimality. Optimization problems of this form constitute the main computational kernel of compressed sensing applications. Thus, 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 (6) is solved for each patch with OMP. Due to lack of space we omit details on OMP and refer the interested reader to the original publication. The OMP code reproducing the results in this paper is available for download at Elad (2011). Details concerning the processing time are discussed in 4.

Dictionary construction. The proposed framework requires dictionaries D_h, D_l, which must somehow be learned from training data. Different learning techniques could be used to obtain a set of atoms from available high-quality DEMs. We have tried 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. This is similar to the approach used in Yang et al. (2008) for forming the dictionary. Hence, for the construction of D_h we use a training set of high resolution DEMs of high quality (that 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 bi-cubic interpolation. The preparation of the dictionaries is off-line and needs to be done only once. Our empirical results in Section 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.
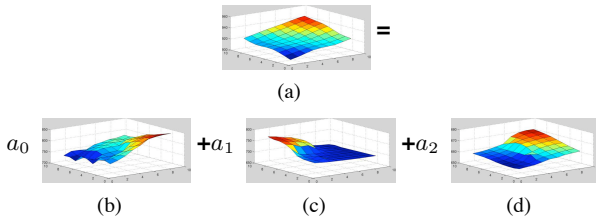

Figure 1: Reconstruction of a 9×9 patch (a) from three non-zero atoms (b)-(d), where $[a_0, a_1, a_2]$ is the sparse non-zero coefficient vector.

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 the 3D separations between a template (master) DEM and a second search (slave) DEM i.e. Gruen and Akca (2005). After co-registration the low resolution DEM is resampled to the nodes of the high resolution DEM by bicubic interpolation.

3.1 Fusion Weights

The fusion is accomplished with the support of weight maps, which reflect the estimated relative accuracy of the two DEMs at every single grid point. In some cases DEM providers deliver such error maps, which then can be directly used for fusion. However, in most cases these maps are not available or not reliable. Then, 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.

The slope is extracted using Horn's formula Horn (1981) to find the first order derivatives in x and y direction. Roughness refers to the height viariation within a local neighborhood, and can be measured in different ways, e.g. by the standard deviation or fractal dimension. We have experimented with several methods and have 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 3. After the mapping, we adjust the expected residuals by linearly scaling the values between 0 and 100. The result is an accuracy map. In the last step, we normalize the resulting accuracy maps of both input DEMs at each overlapping point. The inverse 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 Section 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 each time the produced results with available high quality reference data. Best results are achieved with the following set of parameters. The overlap parameter β is set to the interval $[0.5, 1.5]$. The number of the non-zero atoms used in OMP is set between 7 to 15. Under 7 the results are not reliable and over 15 the processing time increases while the results do not improve. 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' dictionary should contain atoms that describe every possible geomorphological structure, e.g. urban, forest, flat, mountainous areas, in order that OMP can find for every patch a well fitting combination of atoms.

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, over a test site located at Thun, Switzerland characterized by areas with different morphology and land cover. We used three DEMs produced with image matching and SAR interferometry. Figure 2 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[2] Reference 3D DEM (S): 30 m 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≤20%) is 10 m, for hilly terrain (20 % ≤ slope ≤ 40%) is 18 m and for the mountainous terrain (slope> 40%) is 30 m. The Reference 3D DEM is delivered with a High Resolution orthoimage (SPOT 5 sensor) with 5 m ground pixel size.

ALOS/PALSAR[3] DEM (A): 15 m 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 20 m and it has been estimated using the Lidar DEM.

ERS[4] DEM (E): 25 m grid spacing. Image acquisition date: master 22.10.1995 and slave 23.10.1995. C-Band (ca. 6 cm wavelength). It is produced by Sarmap SA. The overall accuracy is 29 m and it has been estimated using the Lidar DEM.

Lidar DEM (L): 2 m 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 0.5 m and 1.5 m 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: $\sigma = \text{AF} \cdot (1 - \text{coherence}^2)/(2 - \text{coherence}^2)$, where $\text{AF} = \text{R} \cdot (\sin(\theta)/(B_n \cdot 4\pi/\lambda))$, 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 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 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 com-

Master	Slave	σ_0	T_x	T_y	T_z
L	A	13.4	18.6	6.4	1.0
L	E	18.8	-6.7	20.0	1.7
L	S	8.9	16.6	3.2	2.9

Table 1: Co-registration results. σ_0 is the σ a posteriori, and T_x, T_y, and T_z are the three translations. All units are in meters.

puted 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]**S**ystème **P**our l'**Ob**servation de la **T**erre
[3]**A**dvanced **L**and **O**bserving **S**atellite/**P**hased **A**rray type **L**-band **S**ynthetic **A**perture **R**adar
[4]**E**uropean **R**emote-sensing **S**atellite

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 mean 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, roughness, and land cover. 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≤15°, 15°<Slope≤45°, Slope>45°. The three roughness classes are: Roughness≤10, 10<Roughness≤30, Roughness>30. The roughness is scaled in the interval [0, 100]. For the land cover analysis we selected different patches of (a) forest (15491 cells, 3.5 km^2), (b) urban (5015 cells, 1.1 km^2), and (c) flat fields (29507 cells, 6.6 km^2) areas using the SPOT orthoimage. The cells correspond to a grid of 15×15 m size.

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.
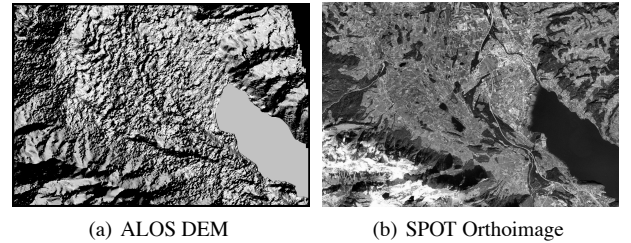

(a) ALOS DEM   (b) SPOT Orthoimage

Figure 2: Overlap area.

4.1 Fusion ALOS-SPOT

A grid was generated for the input DEMs at 15 m intervals according to the spatial resolution of the ALOS DEM. After the resampling, the input DEMs have a 800×1167 grid size. The processing time required for the fusion was 3.8 minutes. The weights for the ALOS DEM are calculated using the given accuracy map. In order to create a weight map for the SPOT DEM we used the information that is given from SpotImage relating the accuracy of the DEM with the slope (Figure3(a)).
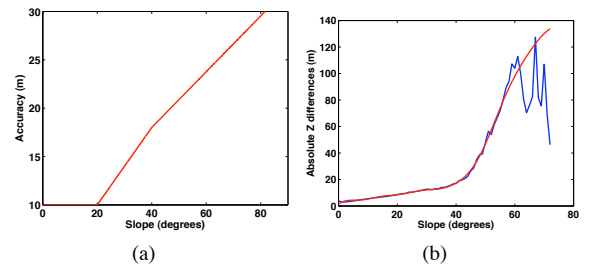

(a)   (b)

Figure 3: (a) Absolute elevation accuracy versus slope (mean every degree) for the SPOT DEM. (b) Absolute Z differences between the ERS and the SPOT DEM versus slope (mean every degree). The blue line are the original values. The red line is a process regression that we fit to the original values weighted with the corresponding roughness values.

In Table 2, we can see that compared to the ALOS DEM, the fusion achieved up to 35% 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 19%. Figure 4 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.

	ALOS-SPOT		
	MEAN	RMSE	MAD
L-A	-1.1	19.7	11.5
L-S	-1.6	15.7	7.9
L-F1	-1.2	12.8	7.3

Table 2: Statistical results of the ALOS-SPOT fusion for the complete area.
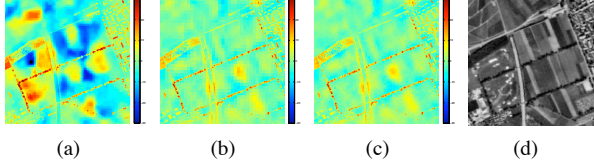

(a) (b) (c) (d)

Figure 4: 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.

The results of the slope, roughness and land cover assessment are presented in the Table 3. 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. Regarding dependence on the land cover we do not see an improvement after fusion for any of the three classes examined but these results are preliminary and non conclusive.

Slope	S≤15°, 45.3%			15°<S≤45°, 29.8%			S>45°, 24.9%		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-A	-1.8	10.8	6.1	0.1	20.4	13.3	-0.8	28.5	18.7
L-S	-0.8	6.2	3.8	-1.9	17.8	9.4	-2.7	23.6	13.5
L-F1	-0.8	5.5	3.7	-1.4	13.5	8.3	-1.8	19.6	12.7
Rough.	R≤10, 41.8%			10<R≤30, 36.4%			R>30, 21.8%		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-A	-1.7	10.4	6.0	0.5	18.2	11.7	-2.0	31.3	21.3
L-S	-0.7	6.2	3.7	-1.2	15.2	7.9	-4.1	26.3	15.9
L-F1	-0.6	5.0	3.5	-0.9	11.5	7.1	-3.0	21.9	14.9
Land Cover	Fields, 29507 cells			Forest, 15491 cells			Urban, 5015 cells		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-A	-2.1	14.9	8.4	-8.1	31.1	21.6	2.8	5.8	4.0
L-S	-0.8	3.4	2.4	-1.8	10.7	8.1	0.1	4.5	3.5
L-F1	-0.8	3.4	2.3	-2.3	11.2	8.4	0.1	4.4	3.5

Table 3: ALOS-SPOT fusion. Slope, roughness and land cover classes analysis.

4.2 Fusion ERS-SPOT

A grid was generated for the input DEMs at 25 m intervals according to the spatial resolution of ERS. After the resampling, the input DEMs have a 480×700 grid size. The processing time required for the fusion was 1.3 minutes. The weights for the ERS 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 3(b) the used weighting function is shown. It is a 6th degree polynomial function that we fit to the original values weighted with the corresponding roughness values. We used a different approach for the calculation of the weights than in ALOS-SPOT fusion because we found out that for the case of ERS-SPOT fusion the last approach gives slightly better results.

In Table 4, we can see that as compared to the ERS DEM, the technique achieved up to 52% 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 11%. Figure 5 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, roughness and land cover assessment are presented in the Table 5. The analysis of these results conforms

	ERS-SPOT		
	MEAN	RMSE	MAD
L-E	-0.8	29.3	13.3
L-S	-1.6	15.7	7.9
L-F2	-1.2	14.3	7.9

Table 4: Statistical results of the ERS-SPOT fusion for the complete area.
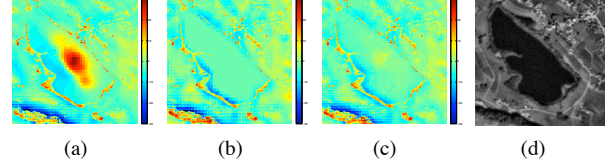

(a) (b) (c) (d)

Figure 5: 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.

to the analysis for ALOS-SPOT fusion. Here the imrpovement by fusion is larger for medium and high slopes and roughness than for low ones.

Slope	S≤15°, 45.3%			15°<S≤45°, 29.8%			S>45°, 24.9%		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-E	-2.3	12.0	4.2	-1.6	32.7	17.5	2.2	41.0	22.1
L-S	-0.8	6.2	3.8	-1.9	17.8	9.3	-2.7	23.6	13.5
L-F2	-1.1	5.8	3.5	-1.4	16.6	10.0	-1.0	20.7	13.2
Rough.	R≤10, 41.8%			10<R≤30, 36.4%			R>30, 21.8%		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-E	-2.2	11.4	4.0	-1.3	28.5	14.5	1.9	45.5	26.1
L-S	-0.7	6.2	3.7	-1.2	15.2	7.9	-4.1	26.3	15.9
L-F2	-0.9	5.5	3.4	-0.9	14.4	8.4	-2.1	23.1	15.5
Land Cover	Fields, 29507 cells			Forest, 15491 cells			Urban, 5015 cells		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-E	-0.9	4.9	3.1	0.3	13.3	10.3	-0.3	4.3	3.4
L-S	-0.8	3.4	2.4	-1.8	10.7	8.1	0.1	4.5	3.5
L-F2	-0.8	3.1	2.1	-1.6	10.8	8.3	-0.1	4.2	3.3

Table 5: ERS-SPOT fusion. Slope, roughness and land cover classes analysis.

4.3 Fusion ALOS-ERS

A grid was generated for the input DEMs at 15 m intervals according to the spatial resolution of the ALOS DEM. After the resampling, the input DEMs have a 800×1167 grid size. 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. The values of the accuracy maps are rescaled to the interval $[0, 1]$. In Table 6, we can see that as compared to the ALOS DEM, the technique achieved up to 4% 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 improving the RMSE by 36%. Figure 6 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, roughness and land cover assessment are presented in the Table 7. For ALOS and forest areas the large negative indicates that the ALOS DEM is systematically higher than the lidar one. Since L-band penetrates the tree canopy the mean should be large but positive. This cannot be explained except by destruction or cutting of trees between 2000 and 2006.

Here for medium and large slopes / roughness there is no imrpovement by fusion while for low values the improvement is significant. Regarding land cover we see a deterioration by fusion at forest areas. This is expected due to very different canopy penetration of the C- and L- bands.

	ALOS-ERS		
	MEAN	**RMSE**	**MAD**
L-A	-1.1	19.7	11.5
L-E	-0.8	29.3	13.3
L-F3	-0.7	18.9	10.1

Table 6: Statistical results of the ALOS-ERS fusion for the complete area.
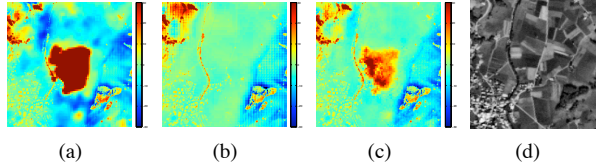

(a) (b) (c) (d)

Figure 6: 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.

Slope	$S \leq 15°$, 45.3%			$15° < S \leq 45°$, 29.8%			$S > 45°$, 24.9%		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-A	-1.8	10.8	6.1	0.1	20.4	13.3	-0.8	28.5	18.7
L-E	-2.3	12.0	4.2	-1.6	32.7	17.5	2.2	41.0	22.1
L-F3	-1.9	8.2	3.8	-0.2	21.0	12.9	0.9	28.0	17.4

Rough.	$R \leq 10$, 41.8%			$10 < R \leq 30$, 36.4%			$R > 30$, 21.8%		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-A	-1.7	10.4	6.0	0.5	18.2	11.7	-2.0	31.3	21.3
L-E	-2.2	11.4	4.0	-1.3	28.5	14.5	1.9	45.5	26.1
L-F3	-1.7	7.6	3.6	0.0	18.4	11.0	0.2	31.0	20.3

Land Cover	Fields, 29507 cells			Forest, 15491 cells			Urban, 5015 cells		
	MEAN	RMSE	MAD	MEAN	RMSE	MAD	MEAN	RMSE	MAD
L-A	-2.1	14.9	8.4	-8.1	31.1	21.6	2.8	5.8	4.0
L-E	-0.9	4.9	3.1	0.3	13.3	10.3	-0.3	4.3	3.4
L-F3	-1.0	5.1	3.5	-3.8	18.6	13.5	0.9	4.3	3.3

Table 7: ALOS-ERS fusion. Slope, roughness and land cover classes analysis.

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 ample 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 likeley to be realized in the near future for the calculation of the weight maps.

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, http://www.cs.technion.ac.il/ elad/software/.

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

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

Gruen, A. and Akca, D., 2005. Least squares 3D surface and curve matching. ISPRS Journal of Photogrammetry & Remote Sensing 59(3), pp. 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), pp. 83–89.

Horn, B., 1981. Hill shading and the reflectance map. Proceedings of the IEEE 69(1), pp. 14–47.

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), pp. 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), pp. 69–89.

Rao, Y. S., Rao, K. S., Venkataraman, G., Khare, M. and 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. and 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. M., Stolle, F. R. and Woo, D.-M., 1999. Error detection and DEM fusion using self-consistency. In: International Conference on Computer Vision.

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

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. Royal. Statist. Soc B. 58(1), pp. 267–288.

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

Research Reports