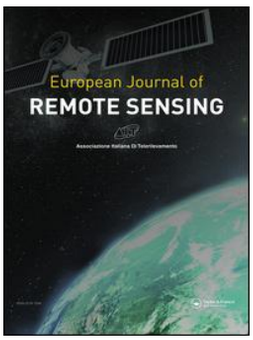

European Journal of Remote Sensing

ISSN: (Print) 2279-7254 (Online) Journal homepage: https://www.tandfonline.com/loi/tejr20

Automatic Co-registration of Satellite Time Series via Least Squares Adjustment

Luigi Barazzetti, Marco Scaioni & Marco Gianinetto

To cite this article: Luigi Barazzetti, Marco Scaioni & Marco Gianinetto (2014) Automatic Co-registration of Satellite Time Series via Least Squares Adjustment, European Journal of Remote Sensing, 47:1, 55-74, DOI: 10.5721/EuJRS20144705

To link to this article: https://doi.org/10.5721/EuJRS20144705

European Journal of Remote Sensing - 2014, 47: 55-74
doi: 10.5721/EuJRS20144705
Received 18/06/2013, accepted 24/01/2014

European Journal of Remote Sensing
An official journal of the Italian Society of Remote Sensing
www.aitjournal.com

Automatic Co-registration of Satellite Time Series via Least Squares Adjustment

Luigi Barazzetti[1]*, Marco Scaioni[2] and Marco Gianinetto[1]

[1]Politecnico di Milano, Department of Architecture,
Built Environment and Construction Engineering, Via Ponzio 31, 20133 Milano, Italy
[2]Tongji University, College of Surveying and Geo-Informatics, Center for Spatial Information Science and Sustainable Development Applications, 1239 Siping Road, 200092 Shanghai, P.R. China
*Corresponding author, e-mail address: luigi.barazzetti@polimi.it

Abstract
Image-to-image co-registration is a fundamental task for data processing of satellite time series. This paper presents a new multi-image co-registration algorithm that simultaneously uses multi-image corresponding points in the whole multi-temporal sequence. Image co-registration parameters are then computed on the basis of a global adjustment. The implemented algorithm provides sub-pixel accuracy similar to that achievable with interactive measurements, but it is also able to register images which do not directly share corresponding features with the reference. Results for a (i) synthetic dataset and a (ii) real complex multi-temporal series made up of 13 Landsat images collected over a period of 30 years are illustrated and discussed. The obtained results showed that the implemented algorithm is atmospheric resistant and quite robust against land cover changes, cloud cover, and snow.

Keywords: image processing, least squares adjustment, matching, medium-resolution, co-registration, satellite images, time series.

Introduction
Image-registration can be defined as 'the process of overlaying images (two or more) of the same scene taken at different times, from different viewpoints, and/or by different sensors' [Brown, 1992]. It plays an essential role in remote sensing as it is a prerequisite for the analysis of multi-date image sequences [Atzberger and Rembold, 2012; Gianinetto, 2012; Maselli, 2012; Balenzano et al., 2013; Schucknecht et al., 2013].

Images are co-registered when the pixel-to-pixel alignment is provided with at least pixel (or sub-pixel) precision. Image co-registration issues might seriously influence the final quality of the remote sensing analysis and products because small misalignments in the input data could give large errors in the final outputs [Townshend et al., 1992], e.g. for change detection applications [Gianinetto and Villa, 2011]. The co-registration process is

also termed *image-to-image* approach, because images are directly co-registered without external data. An alternative approach for co-registration is based on the *image-to-map* techniques, where a map (or a set of ground control points, e.g. from measurements obtained with Global Navigation Satellite System (GNSS) techniques) is adopted as reference for precise geo-correction.

Image-to-image methods have two important advantages. The first one concerns the analyzed data: many applications require the comparison of multi-temporal and multi-source data that should be precisely aligned. The relative co-registration of different images is therefore more important than their absolute geo-referencing. Secondly, image-to-image co-registration methods are based on the analysis of dataset of the same type. This leads to a general simplification of the co-registration process (as better clarified in the next sections), especially in the case automatic co-registration techniques are used. For these reasons this paper will deal with image-to-image methods.

The basic co-registration process works on two images showing a partial or complete area with corresponding points. One image is selected as reference (or 'master') and the other ('slave') is transformed to match the 'master.' Different geometric transformations can be used for mapping the 'slave' onto the 'master', whereas the multi-spectral content of the images is usually preserved in order to use the original information for classification purposes. The estimation of transformation parameters is carried out with a set of corresponding features (usually some control points - CPs) which should be homogeneously distributed on the area. In the case several images have to be co-registered, the 'master' is manually selected and the others will be co-registered by using pairwise matching techniques (*one-to-one* approach). An alternative strategy is based on a preliminary reordering of the images in a sequence, for example according to data acquisition time (*sequential* approach). Then co-registration is computed with single image pairs, from the beginning of the sequence up to the end. Each image plays as 'slave' when is co-registered to the previous one and as 'master' when it is used for the co-registration of the next one. The advantage of this strategy is that every image can be directly co-registered to the previous one, thus the differences in the image content are expected to be smaller. On the other hand, the problem of this approach is that co-registration errors propagate and accumulate along the whole sequence.

The reader is referred to Goshtasby [2005] and Le Moigne et al. [2011] for a comprehensive review on satellite image co-registration methods. However, the problem is still open and no standard efficient and robust solution is available in the field of satellite remote sensing, especially when data processing at a very high level of automation (without interactive measurements) is needed. Indeed, in the case when few images have to be processed, this task is usually carried out manually and the selection of a set of CPs between the 'master' and several other 'slaves' is carried out interactively. This operation is not trivial and expert operators have to measure precise homologous features with a homogenous geometric distribution in the images. Obviously, if a huge volume of data needs to be co-registered, this operation becomes laborious and time consuming.

When a large dataset is processed [Chirici et al., 2004], image matching algorithms [Gruen, 2012; Lemmens, 1988] can be used for automatically detecting corresponding CPs. Traditionally, the basic co-registration process works on pairs of images [Gianinetto and

Scaioni, 2008]. Consequently, if the data set is composed of many images, each one has to be independently co-registered to another (one-to-one or sequential approach) 'master' image (Fig. 1 - left). This is the approach implemented in several commercial software packages. Moreover, automatic algorithms should be able to deal with different radiometric variations (e.g. illumination, shadows, etc.) and 'disturbing elements' (e.g. clouds, snow, vegetation growth, etc.). In some cases, this step is quite complex and manual measurements become the only solution to extract a set of corresponding features. A typical workflow includes the extraction of elements like corners, edges, line intersections, centroids of specific regions, etc. These should be visible in different images of the time series, meaning that stability in time and invariance to both geometric and radiometric changes are needed.
Image matching can be performed with different methods that exploit the intensity value of adjacent pixels, their spatial distribution or description by means of mathematical vectors. The scientific literature is vast and some of the most used techniques are correlation-like methods [Pratt, 1991], mutual information [Pluim et al., 2001], Fourier methods [Castro and Morandi, 1987], relaxation methods [Price, 1985], and spatial relations [Goshtasby et al., 1986].

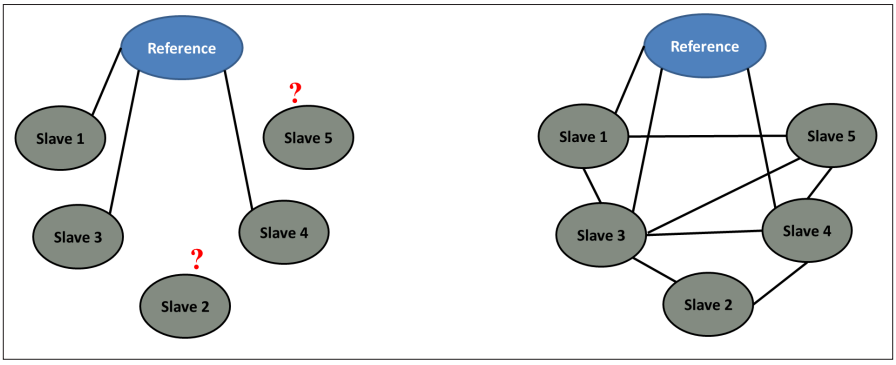

Figure 1 - Left: standard image co-registration scheme where an image is set as 'master' and the co-registration is carried out independently for the remaining 'slave' images ('one-to-one' approach). Images ID 2 and 5 do not share tie points with the 'master'. Right: the implemented matching strategy based on multiple image connections followed by a global LS adjustment for simultaneous estimation of all the transformation parameters ('one-to-many' approach).

The extracted CPs (either in manual or automatic way) are then used to estimate a mapping function between the different images. Overall, the problem consists in (i) the preliminary selection of an appropriate mathematical model for image co-registration and (ii) the estimation of its unknown parameters. The number of CPs (and their distribution) should be sufficient to obtain a reliable estimation of these parameters. Unfortunately, automated methods may produce gross errors (outliers) and robust estimation techniques are recommended for the model estimate. Once blunder rejection process is completed, a final transformation can be estimated for the whole image by means of Least Squares techniques. Remotely sensed data are often considered connected by similarity, affine,

homography, second or third-order polynomial functions. Much more involved Rational Function Models [Poli and Toutin, 2012] are used especially for high and very-high resolution satellite imagery.

This paper describes the first results of a new automatic multi-image co-registration technique where all the images of a time series are used in a global adjustment, therefore integrating the one-to-one and the sequential approach (Fig. 1 - right). This approach provides a global mapping function and an adjustment process that simultaneously includes all the images of the data set (*one-to-many* approach). Image matching is performed between all the image pair combinations which share common features extracted by using the SURF operator [Bay et al., 2008]. Features are then clustered and reordered to generate a regular structure of image coordinates for the following global Least Squares adjustment. The implemented algorithm extracts points often visible in more than two images and the global Least Squares adjustment provides the simultaneous estimation of all the parameters for different images. One image is selected as 'master' for the whole block, but its aim is only to define the datum for the co-registration of the other images (still called 'slaves'). Consequently, the design matrix of the system of equations will include the coefficients of both (i) 'master'-to-'slave' and (ii) 'slave'-to-'slave' equations for the co-registration of images without a direct connection to the 'master' (see next sections for more details). A connection graph showing multiple links among the images is used to initialize a co-registration algorithm that maps every 'slave' image to the 'master'.

The method allows the co-registration of images without a direct visibility with the 'master' and avoids error propagation in the case of sequential co-registration where errors can be accumulated. The chosen mathematical model is a 2D similarity transformation including four parameters [Hartley and Zisserman, 2004].

This method was used for a synthetic dataset, consisting in a Landsat-4/TM image that was artificially deformed to create different 'slave' images. This experiment was carried out in order to check the correctness of the procedure and its implementation in a scientific package. Then, the algorithm was used for a real satellite time series made up of 13 Landsat-4/TM and Landsat-5/TM images collected over a period of 30 years, in different seasons and with different snow, cloud and vegetation cover, along with land changes.

In the following sections the procedures for multi-image matching and co-registration are described. Then, some results for both simulated and real data are illustrated in order to prove the correctness of the implemented algorithm and the feasibility of the method. Finally, some conclusions and improvements for future development are illustrated.

Automated measurement of corresponding features

As mentioned in the introduction, the multi-image technique can extract corresponding image features among all the image combinations of the time series. Image matching is carried out between image pairs and the different image pair combinations are progressively analyzed. The same label is assigned to points that are matched in more than two images, so that they can be tracked along the whole dataset and show better inner reliability [Kraus et al., 1997] in the final global block adjustment.

In the case of N images, CPU time depends on N^2, notwithstanding the process can be easily parallelized by using multi-core CPUs because the matching procedure of a specific combination does not influence any other. The current implementation runs in MATLAB®

environment where the parallelization process is automatically managed. Points are then clustered and reordered at the end of the point extraction stage.
The operator for image matching was implemented in order to take into consideration the following issues:

- correspondences have to be automatically extracted and gross errors should be rejected without user's interaction;
- multiple satellite images are used together and the same feature should be matched in several images (even all images);
- sub-pixel precision is needed;
- data have to be processed at their original resolution; and
- the method should be sufficiently robust against radiometric and geometric differences.

As the method was developed for fully automated image processing, preliminary data processing must be avoided (such as the extraction of subsets - spatial or spectral domains - or atmospheric correction, where the method proved to be quite atmospheric resistant).
The multi-image matching starts by extracting features with the Speeded up Robust Features (SURF) operator [Bay et al., 2008], that has a detector capable of finding interest points in the images and a descriptor to associate a vector of information to each single detected point for further matching purposes.
In this work the Landsat TM spectral band 3 (0.63 - 0.69 μm) was used in data processing but further investigation will be carried out in future work, where multiple bands or their combined use could provide more information. The SURF operator is popular in applications where terrestrial images are used, like close-range photogrammetry and computer vision. Here, sub-pixel precision during the triangulation phase was reached [Vergauwen and Van Gool, 2006; Juan and Gwun, 2010; Barazzetti et al., 2010, 2011].
As this descriptor is quite innovative (2008), it has been scarcely applied to remote sensing images. On the other hand, the technical literature reports some scientific contributions [Teke and Temizel, 2010; Bouchiha and Besbes, 2013] where the method was tested on image pairs and compared to other similar operators (SIFT [Lowe, 2004], PCA-SIFT [Ke and Sukthankar, 2004], GLOH [Mikolajczyk and Schmid, 2005], see Mikolajczyk et al. [2005] for a detailed review of image matching results), obtaining sub-pixel precision and robustness against scale variation, translation, rotation and changes in brightness values. SURF relies on a Hessian matrix-based measure for the detector and the distribution of the first-order Haar wavelet responses for the descriptor [Bay et al., 2008].
Corresponding points can be found by simply comparing the descriptors of points extracted from two images, without any preliminary information such as seed points or any manual measurements. In this study, a SURF implementation derived from the original one (available at www.vision.ee.ethz.ch/~surf/) was adapted to handle Landsat images at their original geometric resolution (approximately the size is 60,000,000 pixel). A descriptor of 128 elements is used in this implementation, although the variant U-SURF (rotation invariance is not considered) has a descriptor of only 64 elements. This choice is motivated by the expected use of the algorithm: although Landsat images usually do not present geometric deformations including scale variations and large rotations, future applications with different data would require a more flexible and comprehensive

matching method.
In addition, initial geo-referencing parameters can be derived from the metadata associated to the GeoTIFF format which are distributed by USGS.
Given a generic image pair (I, J), features were detected by using the SURF algorithm. Descriptor vectors DI_m and DJ_n were computed per each generic feature m on image I and on image J, respectively. The matching procedure exhaustively compares the descriptors of features found on image I with all the ones on image J by using as a measure of the difference the Euclidean distance $d_{mn} = || DI_m - DJ_n ||$ between them.
All computed Euclidean distances dmn are ranked from the largest to the smaller one in the vector $d = [d(1)_{mn}, d(2)_{mn}, \ldots, d(p)_{mn}]$, where p is the total number of combinations. Starting from the best distance $d(1)_{mn}$, the corresponding features are selected as candidates to be corresponding CPs. Moreover, a constraint between the first - and the second-best candidates ('ratio-test') is added to be more distinctive [Snavely et al., 2007]:

$$d(1)_{mn} < 0.75 d(2)_{mn} \quad [1]$$

The selection process finishes when all features are assigned or there are no matched features passing the 'ratio-test'.
This automated strategy for the comparison of the feature descriptors retrieves a sufficient number of image correspondences but some mismatches are often still present. The procedure integrates the robust estimation of a similarity transformation between each pair of images to remove these errors. The proposed method is based on the analysis of several sets of image coordinates randomly extracted from the whole dataset following the approach of the popular high-breakdown point estimator RANSAC [Fischler and Bolles, 1981].
After image matching and outlier rejection for all the image pair combinations, a connection graph similar to that shown in Figure 1 (right) is created to obtain a structure of features that completes the image matching phase and initialize the global image co-registration algorithm.

Estimation of the multi-image co-registration parameters

The *connection graph* (Fig. 1 - right) allows one to analyze all the image pair combinations in order to extend the pairwise matching phase towards a global adjustment. Pairwise features are organized into tracks and the comparison of the numerical values (2D pixel coordinates) of image points gives the set of image correspondences for the whole dataset. The output is an ordered structure where an "ID" is added to every point of every image, in a certain way similar to the input of a bundle adjustment in the case of standard photogrammetric projects [Granshaw, 1980]. Here, this principle is replicated but a different mathematical model is employed. This can be intended as a global and simultaneous adjustment where a similarity transformation is employed. Observations are image point measurements along with their geometric multiplicity (visibility in two or more images).
The mathematical model for image co-registration is based on the following relationships (similarity transformation):

$$x'_{iJ} = s_J x_{iJ} cos\alpha_J - s_J y_{iJ} sin\alpha_J + t_{xJ}$$
$$y'_{iJ} = s_J x_{iJ} sin\alpha_J + s_J y_{iJ} cos\alpha_J + t_{yJ} \quad [2]$$

where i is the feature index and J the image index, meaning that, in general, each image may be shifted (t_{xJ}, t_{yJ}), rotated (α_J) and scaled (s_J) with respect to the 'master'.
Eq.s (2) are non-linear and cannot be directly implemented into the Least Squares adjustment. This drawback can be overcome by operating a simple substitution ($a = s\cos(\alpha)$ and $b = s\sin(\alpha)$) that provides a system of linear equations:

$$x'_{iJ} = a_J x_{iJ} - b_J y_{iJ} + t_{xJ}$$
$$y'_{iJ} = b_J x_{iJ} + a_J y_{iJ} + t_{yJ} \quad [3]$$

Two different categories of equations can be formulated depending on the nature of the matched features:

a) If corresponding features are found between a generic image and the 'master' (or reference, index R), the unknowns are only the four transformation parameters underlined in the Eq. 4:

$$x'_{iR} = \underline{a}_J x_{iJ} - \underline{b}_J y_{iJ} + \underline{t}_{xJ}$$
$$y'_{iR} = \underline{b}_J x_{iJ} - \underline{a}_J y_{iJ} + \underline{t}_{yJ} \quad [4]$$

b) If corresponding features are found between two generic images not including the 'master', they become a sort of tie points and give rise to a pair of equations where point coordinates projected onto the reference image are also estimated (unknowns are underlined in next Eq. 5):

$$\underline{a}_J x_{iJ} - \underline{b}_J y_{iJ} + \underline{t}_{xJ} - x'_{iJ} = 0$$
$$\underline{b}_J x_{iJ} + \underline{a}_J y_{iJ} + \underline{t}_{yJ} - y'_{iJ} = 0 \quad [5]$$

It is clear that this mathematical formulation does not include pseudo-observations and two different sets of Eq.s 4 and 5 are employed. The solution here proposed is the same obtainable with a weighted adjustment based on pseudo-observations, assuming that points of the master image remain fixed (σ = 0 pixel) and those of 'slave'-to-'slave' connections have the same precision. This choice is supported by the following considerations:

i. eq. 4 ('master'-to-'slaves'), with master points assumed as fixed known quantities, allows the creation of a stable master that is the reference for data processing. The parameters of 'master' image after Least Squares adjustment are forced to become $[a_R\ b_R\ t_{xR}\ t_{yR}]^T = [1\ 0\ 0\ 0]^T$. In the case of pseudo-observations with precision $\sigma>0$

the adjustment will modify these transformation parameters. This would be a problem for new co-registration processes where the same image is assumed as reference, because the master could be resampled in a different way according to the 'slave' images adopted;

ii. all the observations (even the points selected as reference) are image coordinates measured with the same automatic operator and therefore the same precision can be assumed for all (unary weights in this case); and

iii. it is quite evident the analogy of this methodology with a photogrammetric independent model block adjustment (2D case – [Kraus, 2007]), where features of the reference image play as ground control points. Other matches are tie points and allow one to include images without a direct connection to the reference in the global adjustment.

The system of Eq.s 4 and 5 can be written in matrix form as follows:

$$\begin{bmatrix} \boldsymbol{y}_1 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{1\text{ms}} & 0 & 0 \\ 0 & \mathbf{A}_{1\text{ss}} & \mathbf{A}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{1\text{ms}} \\ \boldsymbol{x}_{1\text{ss}} \\ \boldsymbol{x}_2 \end{bmatrix} = \mathbf{A}\boldsymbol{x} \quad [6]$$

where sub-design matrix A_{1ms} contains the coefficients of co-registration parameters from 'master'-to-'slave' (eq. 4) which are stored in vector x_{1ms}, in similar way. Matrix A_{1ss} contains the coefficients of co-registration parameters from 'slave'-to-'slave' (eq. 5) in vector x_{1ss}. Sub-design matrix A_2 contain the coefficients of the image coordinates of tie points (in vector x_2) re-projected onto the reference image. Vector y_1 contains the coordinate of tie points on the 'master' image.

The system of normal equations has the following structure:

$$\mathbf{N}\boldsymbol{x} = \mathbf{A}^{\mathrm{T}}\mathbf{A}\boldsymbol{x} = \mathbf{A}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{y}_1 \\ 0 \end{bmatrix} = \boldsymbol{b} \quad [7]$$

and the Least Squares solution and the estimated variance-covariance matrix C_{xx} are computed as follows [Mikhail et al., 2001]:

$$\begin{aligned} \hat{x} &= \boldsymbol{N}^{-1}\boldsymbol{b} \\ \sigma_0^2 &= \frac{\left(\boldsymbol{N}\hat{x} - \boldsymbol{b}\right)^{\boldsymbol{T}} \left(\boldsymbol{N}\hat{x} - \boldsymbol{b}\right)}{r} \\ \boldsymbol{C}_{\hat{x}\hat{x}} &= \sigma_0^2 \boldsymbol{N}^{-1} \end{aligned} \quad [8]$$

where r is the global redundancy: a point correspondence provides two equations whereas unknowns include both transformation parameters for all the images and point coordinates re-projected onto the reference.

Since a robust outlier rejection technique is applied after pairwise matching between all image combinations, the final dataset should have a limited number of blunders. This allows on to apply standard techniques for outlier rejection which are used in Least Squares theory (data snooping) on the basis of the statistical distribution of standardized residuals of observation equations.

Results: processing of synthetic data

A Landsat image (image ID 1 - 8,101×7,210 pixels, ground resolution 30 m) was cropped, rotated and scaled to simulate different similarity transformations. In all, four new images with the following 'deformations' were created:

- image ID 2: rotation of 180°;
- image ID 3: cropped from x = 250 pix and y = 422 pix: this means that a rectangular sub portion of the original data was extracted by using the point (x, y) = (250, 422) as new (false) origin. The new image has therefore a resolution of 7,851x6788 pixels;
- image ID 4: scaled of a factor 2;
- image ID 5: cropped from x = 250 pix and y = 422 pix (see above), scaled of a factor 4 and rotated of 90°.

This small dataset includes all possible single deformations and their combinations, from the basic translation (or rotation or scale variation) up to the full similarity transformation. Similar datasets (with more images) were created and processed but the authors decided to report only this examples because results were always similar and well-representative to understand the output of the method.

Matching was carried out by using all images without any preliminary image compression in order to improve precision, obtaining 37,372 image correspondences. Figure 2 (left) shows the extracted points and their distribution. As can be seen, the visualization is similar to that of photogrammetric projects, where points are matched in more than two images. In this case the distribution of feature points is rather homogenous in the images.

A visualization of the connection graph for this simulated block is shown in Figure 2 (right). Images are represented by the nodes of the graph and matching results (i.e., the number of points) are displayed along the lines (the number of extracted features is reported on the line). As can be seen, images 1, 2 and 3 are well-connected as the geometric deformations are rotations and translations (the operator is completely invariant against these effects). The number - 8,999 - is not trivial since it represents the maximum number of points that the algorithm can extract between a generic pair. This threshold was included in the implementation in order to reduce the number of points and simplify the following adjustment phase (in terms of CPU time). This could be intended as a limitation for this specific dataset, as several good points could be automatically removed although they were probably visible in more than two images. On the other hand, this impressive number of points is due to the good quality of this dataset: real data (see next section) usually give less points.

In addition, the connection with image ID 4 is very good notwithstanding a scale variation of 50%. The last image has instead less points since there is a strong scale factor (the new image is about 25%). On the other hand the number of points is enough to estimate all transformation parameters for the whole dataset.

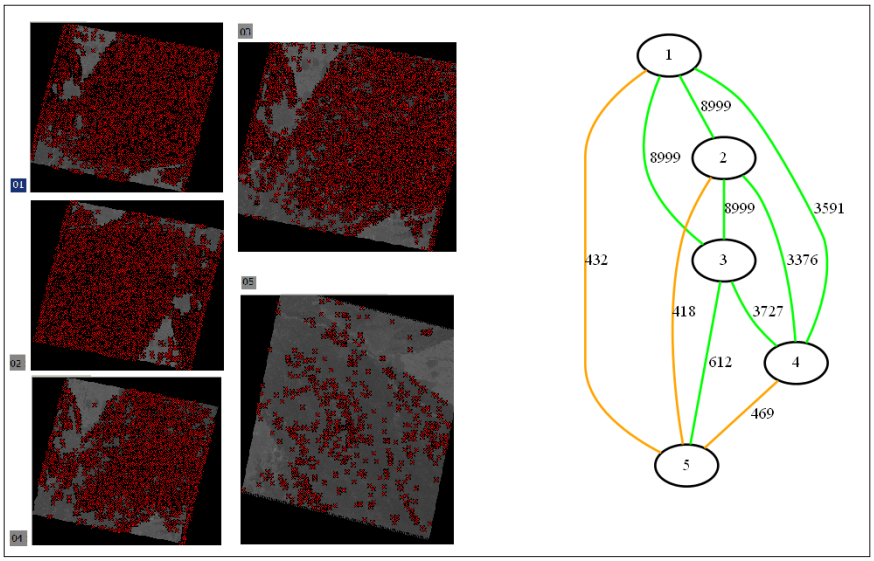

Figure 2 - Left: the distribution of the extracted points in the images. It seems quite homogenous and sufficient to estimate reliable parameters for the similarity transformation. Right: the connection graph for the synthetic dataset along with the number of matches. Most image pairs share corresponding points and no image is excluded from data processing, meaning that the LS co-registration will provide parameters for all images.

Synthetic data are particularly suitable to check both matching and adjustment procedures as the deformations are known. The estimated parameters after Least Squares co-registration are shown in Table 1. Sigma-naught was ±0.35 pixels (global redundancy is given by 74,744 equations and 13,338 unknowns) and demonstrates that for this dataset sub-pixel precision was reached. It is simple to carry out a comparison between simulated and estimated parameters.

The estimated translation parameters for image ID 2 are exactly the simulated ones because the original image was only cropped. For the remaining images translation parameters slightly differ from the original ones since the scale factor of the simulation process makes image resampling necessary.

Finally, we analyzed the estimated transformation parameters and compared them to the corresponding simulated values that in this case can be assumed as 'ground-truth.' The RMSE of the results for the rotation angles and the isotropic scale factor are quite good, meaning that these parameters could be well estimated also in case of very large values. The differences between translation parameters resulted larger ($RMSE_{tx}$=0.78 pixels and $RMSE_{ty}$=0.39 pixels, respectively). The main contribution to this result was due to the co-registration of images with large simulated scale factors, i.e. ID 4 (s=2) and ID 5 (s=4). This means that large scales do not influence their estimates, but they affect the estimate of shifts. On the other hand, the last simulated value is really large and was used to try out the algorithm in extreme conditions. The achieved results can considered sufficient to demonstrate the correctness of both matching and adjustment methodology.

Table 1 - Comparison between simulated and estimated parameters for the synthetic dataset in the case of a roto-translation with scale variation as mathematical model.

image		*α (deg)*	*s*	$t_x(pix)$	$t_y(pix)$
1	*simulated*	0.0	1.0	0.0	0.0
	estimated	0.0	1.0	0.0	0.0
	difference	**0.0**	**0.0**	**0.0**	**0.0**
2	*simulated*	180.0	1.0	8100.0	7199.0
	estimated	180.0	1.0	8100	7199.0
	difference	**0.0**	**0.0**	**0.0**	**0.0**
3	*simulated*	0.0	1.0	250.0	422.0
	estimated	0.0001	1.0	249.996	422.012
	difference	**-0.0001**	**0.0**	**0.0040**	**-0.0120**
4	*simulated*	0.0	2.0	0.0	0.0
	estimated	0.0006	2.0	0.4499	0.5078
	difference	**-0.0006**	**0.0**	**-0.4499**	**-0.5078**
5	*simulated*	90.0	4.0	250.0	7199.0
	estimated	90.0002	4.0	251.542	7198.4
	difference	**-0.0002**	**0.0**	**-1.5420**	**0.6000**
	average	**-0.0002**	**0.0**	**-0.3976**	**0.0160**
	std.dev	**0.0002**	**0.0**	**0.6689**	**0.3926**
	RMSE	**0.0003**	**0.0**	**0.7781**	**0.3929**

Results: co-registration of a complex case study

The second dataset is made up of a multi-temporal time series (see Tab. 2 for more details) including 13 Landsat-4/TM and Landsat-5/TM images (path 197, row 28) acquired over a mountain area in Valtellina (Province of Sondrio, Northern Italy) from 1984 to 2011 and made available through the USGS Earth Resources Observation and Science (EROS) Center (http://glovis.usgs.gov/). The location of the image dataset is displayed in Figure 3. All the Landsat/TM images (185 km x 185 km) were processed at their original resolution (30 m for VNIR-SWIR).

The images of 1987 have a significant cloud cover that made their co-registration really challenging. During this year the bad weather affected the region. A flood emergency started in the second half of July and lasted until the beginning of September. On Tuesday 28 July 1987 a landslide occurred in Val Pola (see Fig. 4), which is located in the study area. More than 40 million m^3 of rock debris came off from the flank of Monte Zandila moving downslope, reaching the bottom of the valley and resulting in an artificial dam. Some villages were destroyed and there were many causalities [Crosta et al., 2004]. This artificial dam stopped the flow of the River Adda and created a lake with 6 million m^3 of water.

The remaining images of the sequence, collected over a period of 30 odd years, were selected according to a specific season (August). In some cases different months (July and September) were selected because the study area was completely cloudy.

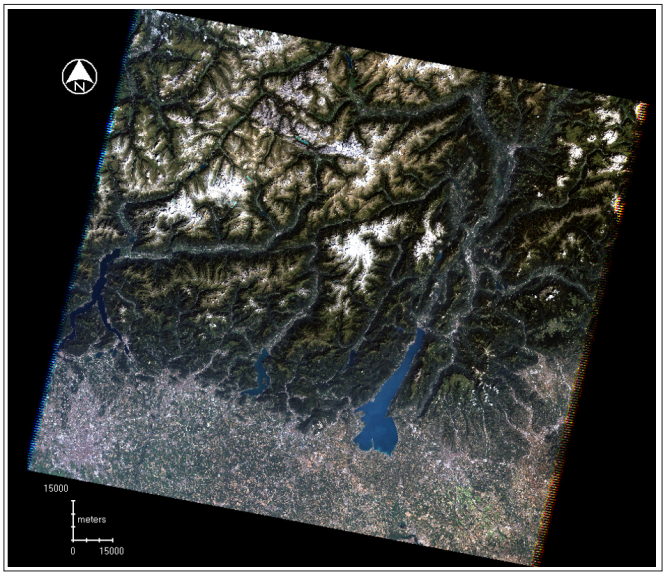

Figure 3 - A global overview of the area under investigation (path 197, row 28) seen from Landsat-5/TM in August 21, 2011.

As stated before, this is a challenging dataset for an automatic algorithm because the time series is made up of images collected in different seasons, years, and images are characterized by different cloud and snow cover. Moreover, important land cover changes are also visible in the study area, since the heavy rainfall and the melting of glaciers resulted in the flooding of various rivers and torrents which in turn brought about much destruction in many villages and towns (for more details the reader is referred to Azzoni et al. [1992]). The process of formation of the lake, and the successive works for recovering the regular hydraulic functions of the valley gave an impressive modification of the topography and land use. For these reasons, the traditional image co-registration algorithms based on the standard 'one-to-one' approach could fail for some images of the sequence. In particular, tests carried out with ERDAS AutoSynch® confirmed that standard techniques were not able to manage this dataset, whereas the new simultaneous multi-image ('one-to-many') approach here described allowed one to process all the images of the time series.
Matching with the proposed algorithm was simultaneously run for all image combinations (in this case N=13 images provided $N(N\text{-}1)/2$=78 combinations).
As previously introduced, the Landsat time series includes several 'tricky' images (with clouds, snow, etc.) and was a good test for the method. It is important to note that different images had a different number of corresponding features with a different geometric distribution. For instance, the images acquired in summer (in Fig. 5 a global overview of matching results is reported) have many corresponding features whereas those acquired during the year of the flood have less features because of the heavy cloud coverage. From this point of view, the algorithm was also able to highlight the presence of clouds or other problems, i.e. outcomes that could be also exploited in the future for other purposes. On the other hand, we decided to include all the images in the adjustment because matching

results suggested that this operation was feasible. In addition, the algorithm automatically removed all the image pairs with less than 40 corresponding features as a reliable estimate is not feasible (when few points are matched - say less than 20/30 - the percentage of outliers increases).

Table 2 - Landsat/TM time series over Valtellina (Northern Italy) used for the study (in red the year of the flood).

ID	*Image source*	*Acquisition date*	*Satellite/sensor*
1	LT51930281984207AAA04	July 25, 1984	Landsat-5/TM
2	LT51930281987007XXX01	January 7, 1987	Landsat-5/TM
3	LT51930281987023XXX01	January 23, 1987	Landsat-5/TM
4	LT51930281987039XXX01	February 8, 1987	Landsat-5/TM
5	LT51930281987119XXX02	April 29, 1987	Landsat-5/TM
6	LT51930281987183XXX02	July 2, 1987	Landsat-5/TM
7	LT51930281987215XXX02	August 3, 1987	Landsat-5/TM
8	LT51930281987263XXX02	September 20, 1987	Landsat-5/TM
9	LT41930281988226XXX05	August 13, 1988	Landsat-4/TM
10	LT41930281990215XXX03	August 3, 1990	Landsat-4/TM
11	LT51930282003259MTI01	September 16, 2003	Landsat-5/TM
12	LT51930282007238MOR00	August 26, 2007	Landsat-5/TM
13	LT51930282011233MOR00	August 21, 2011	Landsat-5/TM

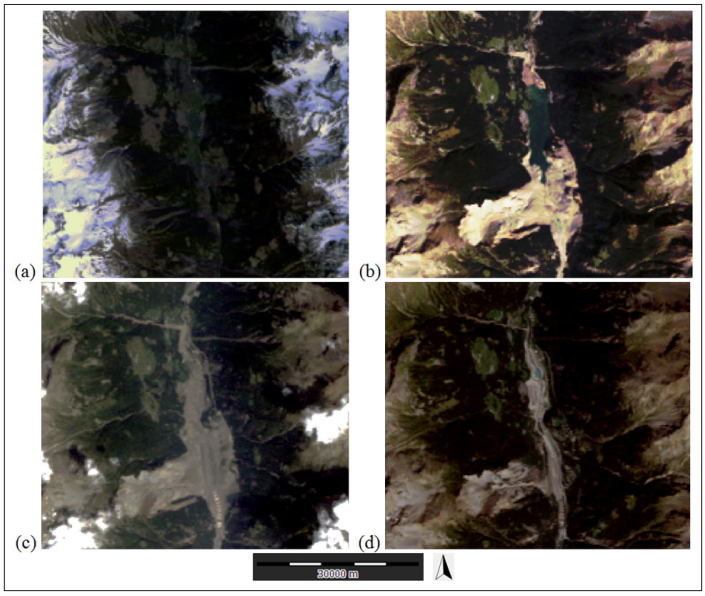

Figure 4 - (a) Landsat images before the landslide in Val Pola (April 29, 1987); (b) and (c) after the landslide with the "new" lake (September 20, 1987) and its modification after three years (August 3, 1990); and (d) a recent view (August 21, 2011) where significant land cover changes are visible.

After completing the pairwise matching phase, a connection graph (Fig. 6) was created with

the Neato tool (http://www.graphviz.org/). The nodes of the graph are the satellite images and the lines their connections (meaning that homologous points were found). In our case study it is interesting to note that the graph 'clusters' the data set into two main blocks: roughly speaking data collected in the year of the flood and the remaining ones. Image ID 1, acquired on July, has several connections with images ID 6-13 because the season does not vary significantly. Images ID 2-5 were instead collected between January and April and were clustered together, probably on the basis of a different land-cover. In any case, all the groups were connected and thus a global adjustment of the whole time series was feasible.

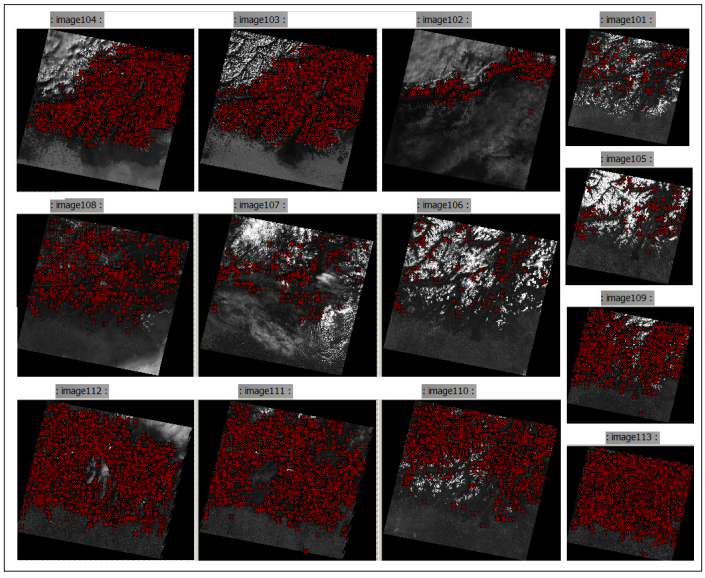

Figure 5 - Landsat/TM images with o the extracted corresponding features. It is clear that images with a high snow or cloud cover have bad point distributions (e.g. images 102-105-106-107). 100 is here added to each image ID for implementation purposes: a the present the algorithm is able to process up to 898 image simultaneously, e.g. from ID 101 up to ID 999.

Figure 7 provides an alternative visualization of the connection graph by means of a double-entry matrix, where some black dots show the 'visibility' between the images. Here, the term 'visibility' expresses the presence of a consistent set of point correspondences between different image pair combinations. This visualization provides a better overview of the matching issues during the year of the flood.

The first image of the time series (ID1) was then set as reference in this experiment. A simple inspection of the connection graph confirmed that the parameters of all the images could be simultaneously estimated by using a system of equations that combines both categories of Eq.s 4 and 5.

Finally, Least Squares adjustment was carried out within the implementation described in subsection 2.2. A set of multi-image correspondences was manually extracted by

using interactive measurements to verify the correctness of the adopted methodology (as mentioned both image matching and Least Squares adjustment were implemented in an in-house scientific software under MATLAB® environment).
The first significant difference between manual and automated measurements concerned the number of tie points extracted: a human operator matched few tens of points whereas the automated operator provided more than 25,000 image correspondences. Results are shown in Table 3, where a balance equations vs unknowns is reported along with the sigma-naught of Least Squares adjustment.

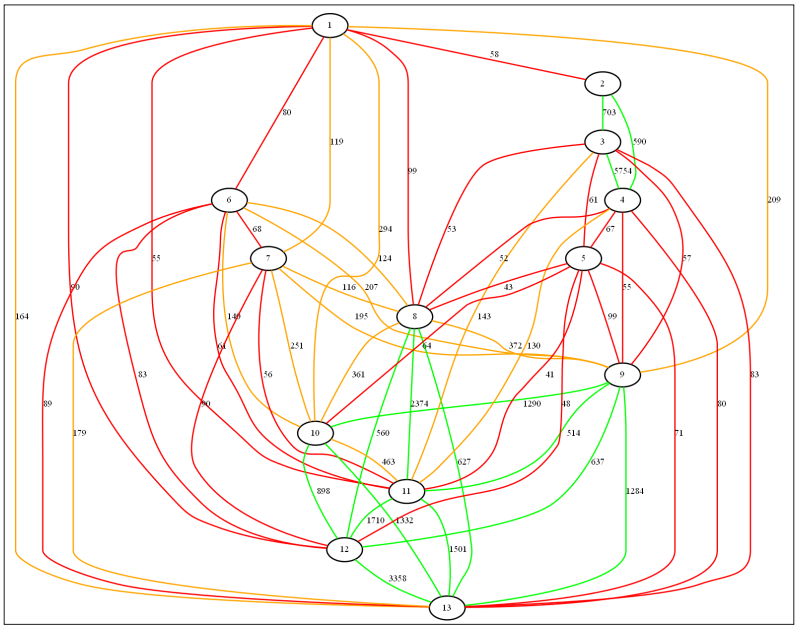

Figure 6 - Connection graph of the time series of Landsat images over Valtellina showing how data are related and clustered.

As can be seen in Table 3, average sub-pixel precision in the whole sequence of images was reached for both manual (±0.57 pixel) and automated (±0.73 pixel) measurements, as indicated by the estimated sigma naught of Least Squares adjustment. Manual measures provided a slightly better precision.

Table 3 - Least Squares adjustment information for the multi-image similarity-based adjustment.

	Automated measurement	Manual measurement
# Equations	59, 720	922
# Unknowns	25, 968	78
Sigma-naught (pixels)	±0. 73	±0. 57

Another important result concerns the estimated scale factors and rotation angles: an almost unary scale factor *s* and a null rotation α were found for all the images, meaning that all the

images were essentially affected by a rigid translation. Moreover, results are comparable for both manual and automated measurements. Table 4 shows the estimated statistics (average and standard deviation) of these parameters and confirms the previous statement.
On the other hand, the results for the estimated translation vectors were significantly different for every image of the time series and required a more exhaustive investigation. For this reason, the differences between the translation values for manually extracted ('man') and automatically extracted ('auto') correspondences ($\Delta t_x = t_{x\_man} - t_{x\_auto}$ and $\Delta t_y = t_{y\_man} - t_{y\_auto}$) were estimated and are shown in Figure 8. The average and standard deviation values of the translation parameter differences resulted in -0.05±0.36 pixels (for Δt_x) and -0.10±0.35 pixels (for Δt_y). These sub-pixel results confirm the consistence of automated measurements with manual data, assumed as ground truth in this work.

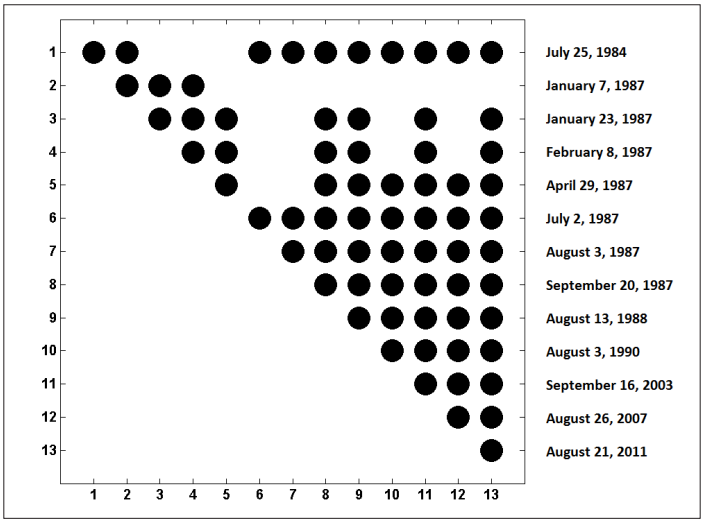

Figure 7 - An alternative visualization of the connection graph between Landsat images over Valtellina by means of a double-entry matrix. Black dots indicate that the algorithm was able to find corresponding features for the specific image pair combination.

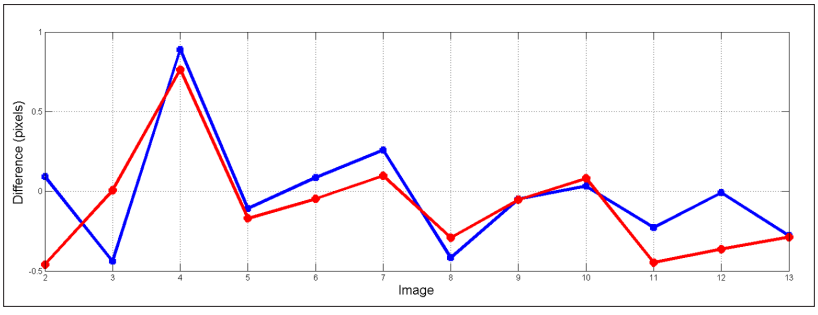

Figure 8 - Comparison between the estimate translation values for automated and manual measurements. The difference of the translation parameters are given in pixels (blue and red for x and y components, respectively).

Table 4 - Summaries of results for the dataset acquired over Valtellina: average scale factors, rotation angles and their RMS for both manual and automated measurements. As can be seen, an almost unary scale factor and null rotation was found for this real dataset.

	Automated measurement	*Manual measurement*
$\overline{s}$	1.000225	1.000193
RMS_s	0.000069	0.000027
$\overline{\alpha}$(deg)	0.00192	0.00089
RMS_α(deg)	0.00147	0.00062

Conclusions and future developments

This paper illustrated the first results of a new multi-image matching and adjustment methodology for the co-registration of satellite time series. The main advantage concerns the simultaneous estimation of the transformation parameters along with a Least Squares adjustment that assumes a multi-image similarity-based transformation as geometric model.
This method avoids standard matching between several images and only a reference ('master'). An image without corresponding points with the 'master' can be co-registered in a way similar to an aerial (or close-range) bundle block adjustment, where images without ground control points can be oriented by means of tie points.
Two sets of equations can be written: those including points measured in the 'master' image and those for tie points (i.e. points matched between images without a direct connection with the 'master'). The Least Squares problem provides the unknowns with a global adjustment where transformation parameters are simultaneously estimated.
The method here proposed is fully automated and very robust against outliers, as techniques for gross error detection were included. The method demonstrated to be atmospheric resistant as atmospheric correction was not carried out beforehand.
The first experimental results with simulated and real medium-resolution satellite data proved that manual measurements were not needed, notwithstanding the complexity of land-cover changes (images collected in different seasons and years), different snow and cloud cover and changes due to several land changes in the study area. However, two main computational problems were found: (i) the matching time that depends on the number of images squared N^2 and the (ii) size of the normal matrix in the Least Squares problem. The first problem could be partially overcome with speeded-up matching strategies for descriptor comparison [Brown and Lowe, 2007]. The second drawback can be coped by reducing the set of image points to a smaller one, which would include less observations with a good multiplicity and spatial distribution in the images [Barazzetti et al., 2010]. A similar strategy, called multi spanning tree (MST) has been recently introduced in the processing of large datasets of InSAR images [Refice et at., 2006]. In this case a graph connecting different images is constructed to help the selection of the best 'master' and to avoid the matching between all image combinations, but limiting this task to those which can give better results.
As things stand at the moment, the first image of the time series provides the reference features and removes the rank deficiency of Least Squares problem. This means that image

ID1 fixes the datum of the combined registration process. Moreover, other geometric models for image co-registration (i.e. affine, homography, or bilinear interpolation) will be included to take into considerations other deformations. Future work will also investigate the multispectral content during image matching.

Acknowledgement

This research has been supported by the Italian Ministry of Education, University and Research (MIUR) within the grant FIRB 2010 entitled: 'Subpixel techniques for matching, image co-registration and change detection with applications to civil and environmental engineering'. Acknowledgements also go to the National Basic Research Programs of China (No. 2012AA121302 and No. 2013CB733204).

References

Atzberger C., Rembold F. (2012) - *Portability of neural nets modelling regional winter crop acreages using AVHRR time series*. European Journal of Remote Sensing, 45: 371-392. doi: http://dx.doi.org/10.5721/EuJRS20124532.

Azzoni A., Chiesa S., Frassoni A., Govi M. (1992) - *The Valpola Landslide.* Engineering Geology, 33: 59-70. doi: http://dx.doi.org/10.1016/0013-7952(92)90035-W.

Bay H., Ess A., Tuytelaars T., Van Gool L. (2008) - *SURF: Speeded Up Robust Features*. Computer Vision and Image Understanding, 110 (3): 346-359. doi: http://dx.doi.org/10.1016/j.cviu.2007.09.014.

Balenzano A., Satalino G., Lovergine F., Rinaldi M., Iacobellis V., Mastronardi N., Mattia F. (2013) - *On the use of temporal series of L- and X-band SAR data for soil moisture retrieval. Capitanata plain case study*. European Journal of Remote Sensing, 46: 721-737. doi: http://dx.doi.org/10.5721/EuJRS20134643.

Barazzetti L., Remondino F., Scaioni, M. (2010) - *Orientation and 3D modelling from markerless terrestrial images: combining accuracy with automation*. The Photogrammetric Record, 25 (132): 356-381. doi: http://dx.doi.org/10.1111/j.1477-9730.2010.00599.x.

Barazzetti L., Forlani G., Remondino F., Roncella R., Scaioni M. (2011) - *Experiences and achievements in automated image sequence orientation for close-range photogrammetric projects*. Proceedings of SPIE Optics+Photonics, 8085, 23-26 May, Munich, Germany.

Bouchiha R., Besbes K. (2013) - *Automatic Remote-sensing Image Co-registration Using SURF*. International Journal of Computer Theory and Engineering, 5 (1): 88-92. doi: http://dx.doi.org/10.7763/IJCTE.2013.V5.653.

Brown G.L. (1992) - *A survey of image Co-registration Techniques*. ACM Computing surveys, 24: 325-376. doi: http://dx.doi.org/10.1145/146370.146374.

Brown M., Lowe D.G. (2007) - *Automatic panoramic image stitching using invariant features.* International Journal of Computer Vision, 74 (1): 59-73. doi: http://dx.doi.org/10.1007/s11263-006-0002-3.

Castro E.D., Morandi C. (1987) - *Co-registration of translated and rotated images using finite Fourier transform*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 9: 700-703. doi: http://dx.doi.org/10.1109/TPAMI.1987.4767966.

Chirici G., Gianinetto M., Scaioni M. (2004) - *Experiences in Upgrading of Large Databases of Satellite Images*. In: Procedings of XX ISPRS Congress, Istanbul (Turkey), 12-23

July, IAPRS&SIS, Vol. XXXV, Part B/III, pp. 382-387.

Crosta G.B., Chen H., Lee C.F. (2004) - *Replay of the 1987 Val Pola Landslide, Italian Alps*. Geomorphology, 69: 127-146. doi: http://dx.doi.org/10.1016/j.geomorph.2003.07.015.

Fischler M.A., Bolles R.C. (1981) - *Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*. Communications of the ACM, 24: 381-395. doi: http://dx.doi.org/10.1145/358669.358692.

Gianinetto M., Scaioni M. (2008) - *Automated Geometric Correction of High-Resolution Pushbroom Satellite Data*. Photogrammetric Engineering & Remote Sensing, 74 (1): 107-116. doi: http://dx.doi.org/10.14358/PERS.74.1.107.

Gianinetto M., Villa P. (2011) - *Mapping hurricane Katrina's widespread destruction in New Orleans using multi-sensor data and the normalized difference change detection technique (NDCD)*. International Journal of Remote Sensing, 32 (7): 1961-1982. doi: http://dx.doi.org/10.1080/01431161003645808.

Gianinetto M. (2012) - *Automatic Co-Co-registration of Satellite Time Series*. The Photogrammetric Record, 27 (140): 462-470. doi: http://dx.doi.org/10.1111/j.1477-9730.2012.00689.x.

Goshtasby A., Stockman G.C., Page C.V. (1986) - *A region-based approach to digital image co-registration with subpixel accuracy*. IEEE Transactions on Geoscience and Remote Sensing, 24: 390-399. doi: http://dx.doi.org/10.1109/TGRS.1986.289597.

Gosthtasby A. (2005) - *2-D and 3-D image co-registration*. John Wiley & Sons, Hoboken, NJ, U.S.A., 258 pp.

Granshaw S.I. (1980) - *Bundle adjustment methods in engineering photogrammetry*. The Photogrammetric Record, 10 (56): 181-207. doi: http://dx.doi.org/10.1111/j.1477-9730.1980.tb00020.x.

Gruen A. (2012) - *Development and Status of Image Matching in Photogrammetry*. The Photogrammetric Record, 27 (137): 36-57. http://dx.doi.org/10.1111/j.1477-9730.2011.00671.x.

Hartley R., Zisserman A. (2004) - *Multiple View Geometry in Computer Vision*. Cambridge University Press. doi: http://dx.doi.org/10.1017/CBO9780511811685.

Juan L., Gwun O. (2010) - *SURF applied in panorama image stitching*. 2nd International Conference on Image Processing Theory, Tools and Applications, pp. 495-499.

Ke Y., Sukthankar R. (2004) - *PCA-SIFT: A More Distinctive Representation for Local Image Descriptors*. Computer Vision and Pattern Recognition.

Kraus K., Jansa J., Kager H. (1997) - *Photogrammetry. Advanced Methods and Applications*. Vol. 2, 4th edition, Dümmler, Bonn, Translated by Peter Stewardson. ISBN 3-427-78694-3, 466 pp.

Kraus K. (2007) - *Photogrammetry: Geometry from Images and Laser Scans*. Walter de Gruyter, 459 pp. doi: http://dx.doi.org/10.1515/9783110892871.

Lemmens M. (1988) - *A Survey on Stereo Matching Techniques*. International Archives of Photogrammetry and Remote Sensing, 27 (B8): 11-23.

Le Moigne J., Netanyahu N.S., Eastman R.D. (2011) - *Image Co-registration for Remote Sensing*. Cambridge University Press, U.K., 484 pp. doi: http://dx.doi.org/10.1017/CBO9780511777684.

Lowe D. (2004) - *Distinctive Image Features from Scale-Invariant Keypoints*. International Journal of Computer Vision, 60 (2): 91-110. doi: http://dx.doi.org/10.1023/B:VISI.000

0029664.99615.94.

Maselli F. (2012) - *A method to improve the spatial features of NDVI data series*, European Journal of Remote Sensing, 45: 407-420. doi: http://dx.doi.org/10.5721/EuJRS20124534.

Mikhail E.M., Bethel J.S., McGlone J.C. (2001) - *Introduction to Modern Photogrammetry.* John Wiley & Sons, Inc., New York, USA, 479 pp.

Mikolajczyk K., Tuytelaars T., Schmid C., Zisserman A., Matas J., Schaffalitzky F., Kadir T., Van Gool L. (2005) - *A Comparison of Affine Region Detectors* . International Journal of Computer Vision, 65 (1/2): 43-72. doi: http://dx.doi.org/10.1007/s11263-005-3848-x.

Mikolajczyk K., Schmid C. (2005) - *A performance evaluation of local descriptors*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 10 (27): 1615-1630. doi: http://dx.doi.org/10.1109/TPAMI.2005.188.

Poli D., Toutin T. (2012) - *Review of developments in geometric modelling for high resolution satellite pushbroom sensors*. The Photogrammetric Record, 27 (137): 58-73. doi: http://dx.doi.org/10.1111/j.1477-9730.2011.00665.x.

Pluim J.P.W., Maintz J.B.A., Viergever M.A. (2001) - *Mutual information matching in multiresolution contexts.* Image and Vision Computing, 19: 45-52. doi: http://dx.doi.org/10.1016/S0262-8856(00)00054-8.

Pratt W.K. (1991) - *Digital Image Processing*. John Wiley & Sons Inc. (2nd Ed.), New York, NY, USA.

Price K.E. (1985) - *Relaxation matching techniques - a comparison*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 7:617-623. doi: http://dx.doi.org/10.1109/TPAMI.1985.4767709.

Refice A., Bonvenga F., Nutricato R. (2006) - *MST-based stepwise connection strategies for multipass radar data, with application to coco-registration and equalization*. IEEE Transactions on Geosciences and Remote Sensing, 44 (8): 2029-2041. doi: http://dx.doi.org/10.1109/TGRS.2006.872907.

Schucknecht A., Erasmi S., Niemeyer I., Matschullat J. (2013) - *Assessing vegetation variability and trends in north-eastern Brazil using AVHRR and MODIS NDVI time series*. European Journal of Remote Sensing, 46: 40-59. doi: http://dx.doi.org/10.5721/EuJRS20134603.

Snavely N., Seitz S.M., Szeliski R. (2007) - *Modeling the World from Internet Photo Collections*. International Journal of Computer Vision, 80 (2): 189-210. doi: http://dx.doi.org/10.1007/s11263-007-0107-3.

Teke M., Temizel A. (2010) - *Multi-spectral Satellite Image Co-registration Using Scale-Restricted SURF*. 20th International Conference on Pattern Recognition, Istanbul, 2310-2313.

Townshend J.R., Justice C.O., Gurney C., McManus J. (1992) - *The impact of misco-registration on change detection*. IEEE Transactions on Geosciences and Remote Sensing, 30 (5): 1054-1060. doi: http://dx.doi.org/10.1109/36.175340.

Vergauwen M., Van Gool L. (2006) - *Web-Based 3D Reconstruction Service*. Machine Vision Applications, 17: 411-426. doi: http://dx.doi.org/10.1007/s00138-006-0027-1.