2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops

Efficient and automated multimodal satellite data registration through MRFs and linear programming

Konstantinos Karantzalos
Remote Sensing Laboratory
National Technical University of Athens, Greece
karank@central.ntua.gr

Aristeidis Sotiras
Section of Biomedical Image Analysis
Center for Biomedical Image Computing & Analytics
University of Pennsylvania, Philadelphia, USA
aristeidis.sotiras@uphs.upenn.edu

Nikos Paragios
Center for Visual Computing
Ecole Centrale de Paris
France
nikos.paragios@ecp.fr

***Abstract*—The accurate and automated registration of multimodal remote sensing data is of fundamental importance for numerous emerging geospatial environmental and engineering applications. However, the registration of very large multimodal, multitemporal, with different spatial resolutions data is, still, an open matter. To this end, we propose a generic and automated registration framework based on Markov Random Fields (MRFs) and efficient linear programming. The discrete optimization setting along with the introduced data-specific energy terms form a modular approach with respect to the similarity criterion allowing to fully exploit the spectral properties of multimodal remote sensing datasets. The proposed approach was validated both qualitatively and quantitatively demonstrating its potentials on very large (more than 100M pixels) multitemporal remote sensing datasets. In particular, in terms of spatial accuracy the geometry of the optical and radar data has been recovered with displacement errors of less than 2 and 3 pixels, respectively. In terms of computational efficiency the optical data term can converge after 7-8 minutes, while the radar data term after less than 15 minutes.**

***Keywords*-Remote Sensing; Multisensor; Multitemporal; Markov Random Fields; Image; Radar; Alignment;**

I. Introduction

The significant progress during the last decade in optics, photonics and remote sensing platform and sensor technology has led to an unprecedented volume of earth observation data. Remote sensors, network of sensors, location sensing devices and the generation of dynamic, and geographically distributed spatiotemporal data has exploded. With trends such as the quantified-self, the Internet of Things, both US and EU Open Data initiatives and applications like Google Earth Engine, the amount of geospatial data will continue to grow exponentially in the coming years and therefore, the need to exploit effectively these massive volumes of geospatial big data will be of fundamental importance.

However, the key and necessary condition of exploiting these multitemporal geospatial datasets is to operationally and accurately manage to register them in a common georeference system [1], [2], [3]. Although, most acquired 2D and 3D geospatial data are *a priori* linked to a certain global, world geodetic system, when one seeks for accurate spatial positioning, local reference ellipsoids must be employed. This registration/ ortho-rectification process is not trivial [4], [5], [6] requiring, usually, detailed Digital Elevation Models (DEMs), labor-intensive and time-consuming procedures, especially, for very and ultra high resolution imaging data or data which are not linked a priori with a reference system.

Despite the numerous research efforts and developed algorithms [7], [8], [5], [9] there are, still, important challenges regarding the automated and accurate registration of very large images, multivariate and multimodal data. Processing data with hundreds of millions of pixels require optimized algorithms with a relative low complexity when near real time performance is desired. Powerful feature descriptors do not perform with the same robustness in multispectral, hyperspectral or multimodal data (*e.g.,* optical, radar, Lidar data) [10], [11], [12], [13], [14]. Therefore, there is a current need for algorithms that will be generic, automated, and able to process and register multimodal data without sacrificing on spatial accuracy. However, processing multimodal data is not straightforward [15], [16], [17], [11], [18] and requires novel, sophisticated algorithms that can accept as an input multiple data from different sensors, data with different dimensions, data with different geometric, spatial and spectral properties and can automatically register and process them.

While, in medical imaging, multimodal data registration through mutual information similarity measures has become a standard reference [19], [20], [9], for very large optical remote sensing datasets with significant rotation and/or scaling differences, these approaches may fail or become extremely time consuming. For optical data simpler, more spectral-based similarity measures are, in the general case, more robust and may be combined with entropy correlation coefficients when data of another modality (*e.g.*, SAR radar or Lidar data) are included in the dataset [10].

A. Contribution

Towards this end, we propose a generic registration framework for very large, multimodal, multitemporal remote sensing data. The proposed approach is based on recent robust formulations incorporating Markov random fields and

978-1-4799-4308-1/14 $31.00 © 2014 IEEE
DOI 10.1109/CVPRW.2014.57

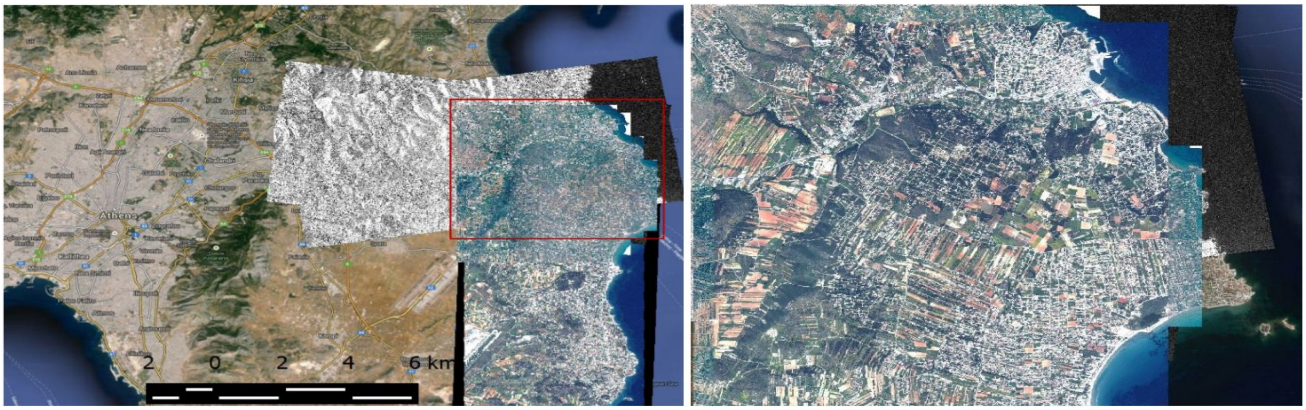

Figure 1. The multimodal dataset is covering a 25km^2 region in the East Prefecture of Attica (Greece), denoted with a red rectangular on the left side. The dataset includes multitemporal, multispectral optical and radar satellite data over a complex terrain with urban, peri-urban, agricultural, coastal and forest regions.

powerful discrete optimization [21], [22]. In a similar way, we investigate the application of minimal cost graph formulation in the case of (deformable) registration of remote sensing data, where nodes correspond to the deformation grid, a node's connectivity corresponds to regularization constraints, and labels correspond to 2D deformations.

The discrete optimization setting along with the introduced data-specific energy terms form a modular approach with respect to the similarity criterion allowing to fully exploit the spectral properties of multimodal remote sensing datasets. The employed efficient inference algorithm boosts the computational efficiency of the proposed framework enabling the automated registration of very large (more than 100M pixels) remote sensing data in less than 10 minutes for the optical and less than 15 minutes for the other modalities. Furthermore, the iterative local evaluation of deformed grid's control points allows us to better handle the significant changes of the multitemporal remote sensing datasets which are well-localized in space. Last but not least, the proposed generic and automated registration framework was validated both qualitatively and quantitatively on a very large multimodal, multitemporal remote sensing dataset covering a 25km^2 region in the East Prefecture of Attica in Greece (Figure 1, Table I).

II. Methodology

Let us consider a target image $I_T : \Omega \mapsto R^d$, where Ω is the target image domain and d is the number of bands of the multispectral data, and a source image I_S. We aim at computing a transformation $\mathcal{T} : \Omega \mapsto \mathbb{R}^2$ such that the two images get aligned under its influence.

We employ a grid based deformation model that can provide for one-to-one and invertible transformations. The basic idea of the deformation model is that by superimposing a uniform deformation grid $G : [1, K] \times [1, L]$ (K and L are significantly smaller than the image dimensions; and the distance between the control points is δ), one can deform the underlying image by controlling the nodes of the grid and using an interpolation strategy to calculate their influence in the rest of the image domain.

$$\mathcal{T}(\mathbf{x}) = \mathbf{x} + \sum_{i=1}^{K} \sum_{j=1}^{L} \eta_{ij}(\mathbf{x}) \mathbf{d}_{ij}, \tag{1}$$

where $\mathbf{d}$ denotes the displacement of one control point and bold is used to denote coordinate vectors. η corresponds to an interpolation or weighting function that determines the influence of a control point ij to the image point x; the closer the image point, the higher the influence of the control point. The control points are evenly placed at a distance δ, forming a uniform grid.

The optimal transformation is typically estimated through the minimization of an objective function that comprises two terms: i) a data term that quantifies the level of alignment between the two images by employing a similarity (or dissimilarity) criterion, and ii) a regularization term that aims to account for the ill-posedness of the problem and introduce user knowledge regarding the desired properties of the solution.

In this work, we formulate the registration problem by using discrete Markov Random Field theory [23]. The discrete setting for image registration has some interesting properties that make it ideal for remote sensing applications: First, this approach is modular with respect to the similarity criterion. Because of the discrete nature of the proposed approach, one may seamlessly use different metrics under the same framework. This allows us to fully exploit the spectral properties of the dataset at hand. Second, the availability of efficient inference algorithms boosts the computational efficiency of the discrete registration framework. The importance of this property is further underlined by the large data size of the remote sensing data. Moreover, the global search over the set of candidate solutions renders the registration method less prone to get stuck to local minima.

An MRF is a probabilistic model that can be represented

Multimodal Satellite Data	**Date of Acquisition**	**Spectral Bands** (number)	**Spatial Resolution** (meters)	**Dimensions** (pixels)	**Size** (GBs)
Pleiades PanSharpened	May 2013	4	0.5m	10915 x 14045	1.17
Worldview-2 Panchromatic	April 2011	1	0.5m	10940 x 14033	0.31
Worldview-2 Multispectral	April 2011	8	2.0m	2735 x 3508	0.15
Worldview-2 Panchromatic	April 2010	1	0.5m	10940 x 14033	0.31
Worldview-2 Multispectral	April 2010	8	2.0m	2735 x 3508	0.15
TerraSAR-X Radar Data	Jan. 2013 (asc.)	1	1.0m	5476 x 7012	0.02
TerraSAR-X Radar Data	Jan. 2013 (desc.)	1	1.0m	5476 x 7012	0.02
Aerial Ortho-Mosaic	2010	3	0.4m	13024 x 16706	0.65

Table I
THE MULTIMODAL DATASET INCLUDES AERIAL AND SATELLITE OPTICAL MULTISPECTRAL IMAGES AND SAR RADAR SATELLITE DATA WITH DIFFERENT SPATIAL RESOLUTIONS AND ACQUISITION DATES.

by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ and $\mathcal{E}$ denote the vertices and the edges of the graph, respectively. The vertices encode the random variables that can take values from a discrete label set $\mathcal{L}$, while the edges encode the interactions between the variables. In the specific case of image registration, the random variables of the discrete MRF model correspond to the displacements of the control points. Moreover, the discrete label set corresponds to a quantized version of the solution space, or in other words, it consists of a set of n labels $\mathcal{L} = \{l_1, \cdots, l_n\}$ that index the allowed displacements, $l \equiv \mathbf{d}$.

The goal is to infer the optimal labeling $\mathbf{l}^\star$ that assigns a label l to every node so that the following energy is minimized:

$$E_{MRF}(\mathbf{l}) = \sum_{p \in \mathcal{V}} U_p(l_p) + \lambda \sum_{(pq) \in \mathcal{E}} P_{pq}(l_p, l_q). \qquad (2)$$

The first term encodes a data term that measures the data likelihood of applying all allowed displacements to each random variable through the use of unary potentials U. The second term of the energy encodes a regularization term that penalizes non-desirable interactions between the random variables through the use of pairwise potentials P. The intuition behind the second term is that tightly related variables p and q should have similar labels assigned to them. λ is a scalar value that weights the influence of the regularization term.

As far as the unary potentials are concerned, we employ a block-matching similarity criterion that evaluates a similarity metric over a patch centered around each control point:

$$U_p(l_p) = \int_\Omega \hat{\eta}(\|\mathbf{x} - \mathbf{p}\|) \rho(I_S \circ \mathbf{d}^{l_p}, I_T) d\mathbf{x}, \qquad (3)$$

where $\hat{\eta}$ denotes the function that defines the block around the control point p and $\|\cdot\|$ denotes the l^2 euclidean norm‘. The similarity criterion ρ can be used either as a point-wise intensity difference one, or a statistical criterion (*e.g.*, normalized cross-correlation or mutual information). In the first case, $\hat{\eta}$ can be used to weight more pixels that are spatially closer to the control point and thus are more influenced by it, and can be defined as $\hat{\eta} = (\eta(\|\mathbf{x}-\mathbf{p}\|))/(\int_\Omega \eta(\|\mathbf{y}-\mathbf{p}\|) d\mathbf{y})$. In the second case, $\hat{\eta}$ is defined as an identity function that is equal to 1 only for the pixels that belong to the block centered at p.

Moreover, the block-matching strategy is adapted for multimodal remote sensing datasets because it allows us to evaluate locally the difference between the data to be registered. Furthermore, the local evaluation allows us to better handle temporal changes since changes are usually well-localized in space and thus, their contribution to the registration energy should be also spatially limited.

Lastly, as far as the regularization term of the discrete registration energy is concerned, we employ a simple strategy that is based on the vector differences between the candidate displacements normalized by the grid distance δ:

$$P_{pq}(l_p, l_q) = \frac{\|l_p - l_q\|}{\delta}. \qquad (4)$$

Regarding the edge system, we adopt a 4-neighborhood system and take into account the interactions between each control point and its closest neighbors in the deformation grid. Let us note that, depending on the interpolation strategy that is used for the deformation model, the employed neighborhood system constitutes a simplification with respect to whole interactions between the deformation nodes. Nonetheless, it preserves the most important ones while allowing us extreme computational efficiency.

A. Implementation

The algorithm was implemented in a multi-resolution function where a pyramidal representation of the images was coupled with a multi-scale approach for the deformation model. The Gaussian pyramid for the images allows us to reduce the computational cost. The multi-scale approach for the deformation model consists of employing deformation grids of increased resolution by halving the control point distance. In that way, it is possible to recover first the coarser displacements and then, gradually refine the result. Cubic B-splines were used as the interpolation strategy of the grid-based deformation model.

For every grid resolution level, an iterative scheme was used in order to enhance the efficiency of the discrete scheme. Instead of using a large set of discrete labels which would result in high computational burden, we keep the cardinality of the solution space reasonable and refine the solution at each iteration. At each iteration we apply the result of the previous iteration refining the set of solutions to capture different smaller displacements.

Throughout our experiments the following configuration was used: a Gaussian pyramid of three levels, a deformation grid of three levels with initial control point distance of 120 pixels. The distances of the other two levels are set to 60 and 30 pixels, respectively.The distances of the other two levels are set to 60 and 30 pixels, respectively. For the iterative scheme 5 iterations were used. The cardinality of the label set was 41. The label set was constructed by evenly sampling 10 labels along the x-, y-axis and the diagonals. The no-displacement (zero pixels displacement) was also part of the solution set. In the first iteration, the maximum displacement that was sampled was equal to $0.4 \times \delta$ thus guaranteeing the preservation of topology [24]. During the following iterations, the maximum sampled displacement corresponded to 0.67 of the maximum displacement of the previous iteration. λ was empirically chosen, however, results are robust to choices ranging from 6 to 10.

The patch size was equal to $2\delta \times 2\delta$. Depending on data modality two energy terms were introduced, *i.e.*, one to address the registration of optical data and one the other modalities like radar data. Normalized Cross Correlation (NCC) and Normalized Mutual Information (NMI) were used as similarity metrics. In the case of optical data, one expects temporal changes (*e.g.*, new buildings, vegetation, *etc.*) to locally change the intensities in a linear fashion. Therefore, we used NCC to register optical data because NCC is able to optimally account for such linear intensity relations. In the case of other modalities *e.g.,* radar data where such a well-founded assumption regarding the intensity relations does not exist, we employed NMI. The reason behind this choice is that NMI is able to account for unknown statistical relations between the intensities of the two images. Last but not least, discrete inference was performed by using the fast and efficient Primal-Dual scheme [21].

Multimodal Satellite Data	**dx** (pixels)	**dy** (pixels)	**D** (pixels)
Pleiades PanSharpened	1.188	0.972	1.535
Worldview-2 Panchromatic	0.632	1.098	1.267
Worldview-2 Multispectral	2.279	1.547	2.754
Worldview-2 Panchromatic	0.958	0.965	1.359
Worldview-2 Multispectral	1.158	1.971	2.285
mean errors (optical data)	1.243	1.310	1.806
TerraSAR-X (2013a)	2.185	1.811	2.838
TerraSAR-X (2013b)	1.790	2.349	2.953
mean errors (radar data)	1.988	2.080	2.877

Table II
QUANTITATIVE EVALUATION RESULTS AFTER THE APPLICATION OF THE PROPOSED MULTIMODAL REGISTRATION FRAMEWORK. ERRORS ARE IN PIXELS AND HAVE BEEN CALCULATED BASED ON MANUALLY DENOTED GCPS. IN ALL CASES, THE THE AERIAL ORTHOMOSAIC WAS THE REFERENCE AND ALL MEAN ERRORS WERE LOWER THAN 3 PIXELS.

III. EXPERIMENTAL RESULTS AND EVALUATION

The proposed algorithm was applied to a multimodal dataset containing a number of high resolution satellite multispectral images and radar data (Table I). In particular, the dataset consisted of two Worldview-2 images acquired in 2010 and 2011 with eight multispectral bands and one panchromatic, one Pleiades pansharpened image acquired in 2013, one aerial orthomosaic with 40cm spatial resolution and two TerraSar-X images acquired in January 2013 one descending and another ascending. For all the optical data we did experiments both with the raw and the pan-sharpened images. All data were fully covering a 25km^2 region in the East Prefecture of Attica, in Greece (Figure 1).

The evaluation of the developed registration algorithm was performed both qualitatively and quantitatively. For the qualitatively evaluation various checkerboard and blend visualization figures were closely reviewed along with animated images showing the unregistered and registered data. For the quantitative evaluation an expert after careful photo-interpretation manually denoted the same Ground Control Points (GCPs) in all registered datasets and reference image. The calculated registration errors (the dx, dy displacements along the axis and the distance D, in pixels) when the multimodal dataset was registered to the aerial orthomosaic are given in (Table II).

Experimental results after the application of the developed algorithm are demonstrated in Figure 2 (optical data) and Figure 4 (radar data). For the qualitative validation the

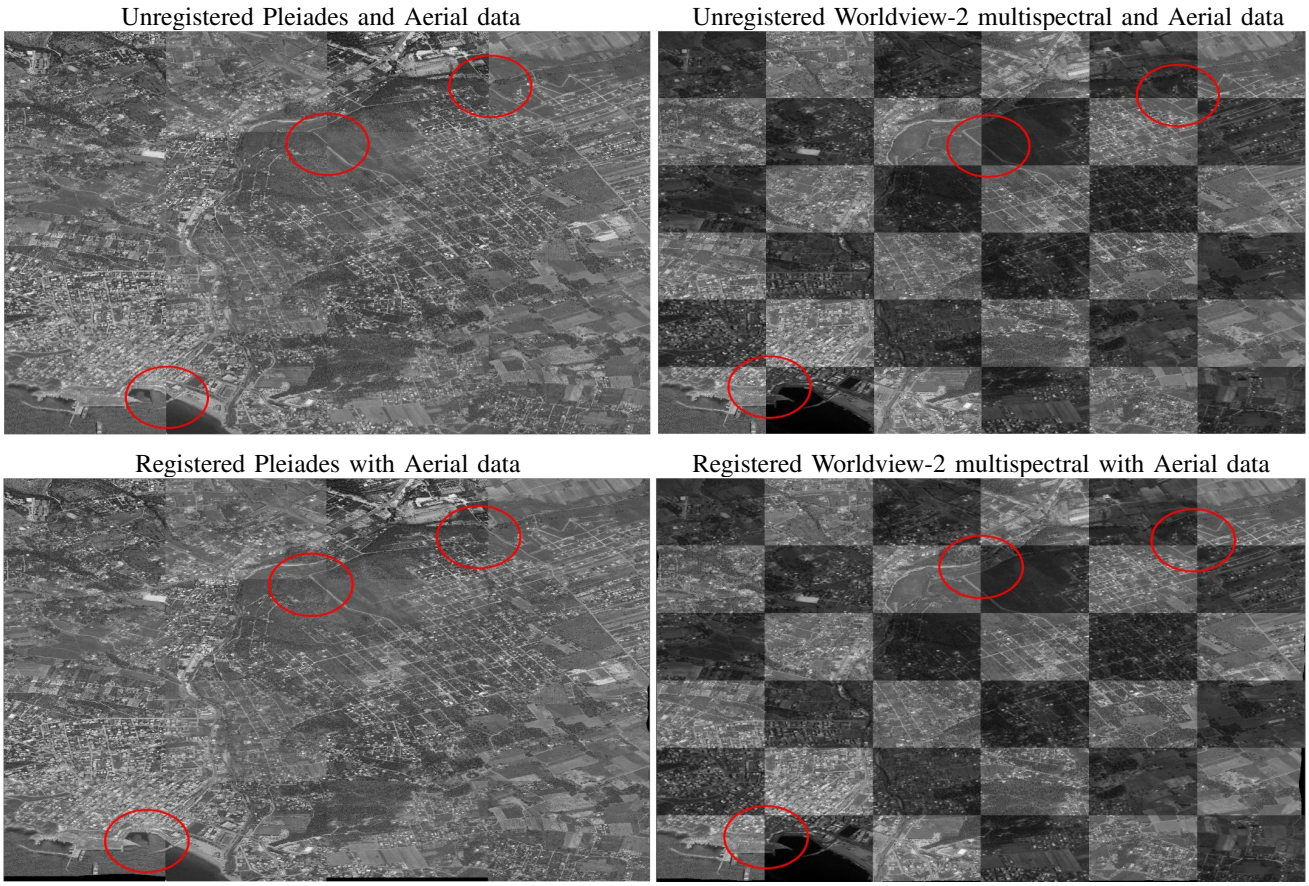

Figure 2. Checkerboard visualization for the qualitative evaluation of the proposed multimodal registration framework. The upper row illustrates the raw unregistered optical data including Pleiades, Worldview-2 multispectral and the aerial orthomosaic optical data. The bottom row demonstrates the registration result and the effectiveness of the proposed algorithm which managed to recover the geometry of the multispectral data.

checkerboard visualization of unregistered and registered optical data are shown in Figure 2. In particular, the upper row illustrates the raw unregistered optical data including Pleiades, Worldview-2 multispectral and the aerial orthomosaic as a reference. The bottom row demonstrates the registration result and the effectiveness of the proposed algorithm which managed to recover the geometry of the multispectral (*i.e.*, 8 or 4 spectral bands), multitemporal data with different spatial resolutions (*i.e.*, 0.4m, 0.5m and 2.0m). Due to the very large spatial size of all data certain regions are denoted with red circles indicating the initial and final stage after algorithms convergence. It should be noted that based on metadata we known the initial data position in relation usually with a global reference system.

In Figure 3 a smaller region of the dataset is shown along with a checkerboard visualization. The upper row illustrates the raw unregistered optical data and the bottom row the registration result. After close inspection, one can observe the important non-uniform deformation required to recover the geometry of radar data. Regarding the computational efficiency of the developed algorithm, the registration functional for the optical data required for its convergence (432 seconds) 7.0 minutes employing the normalized correlation coefficient similarity measure with, in all cases, three image pyramid levels, three control grid scales and with λ, the parameter which controls the influence of the regularization term, set to 4. Moreover, regarding the quantitative evaluation, the optical data were registered with a mean displacement error lower than 2 pixels (Table II).

Experimental results demonstrating the efficiency of the developed algorithm for multimodal data registration are shown in Figure 4 and Figure 5. In particular, in the upper row of Figure 4 the raw unregistered Worldview-2 multispectral data (left) and the Pleiades (right) are shown along with the TerraSAR-X (left, ascending) and TerraSAR-X (right, descending) with a checkerboard visualization. The bottom row demonstrates the registration result and the effectiveness of the proposed algorithm which managed to recover the geometry of the TerraSAR-X radar data.

In Figure 5 a smaller region of the dataset is shown with a checkerboard visualization. Regarding the computational efficiency of the developed algorithm, the multimodal registration functional required (793 seconds) 13.2 minutes with the normalized mutual information similarity measure

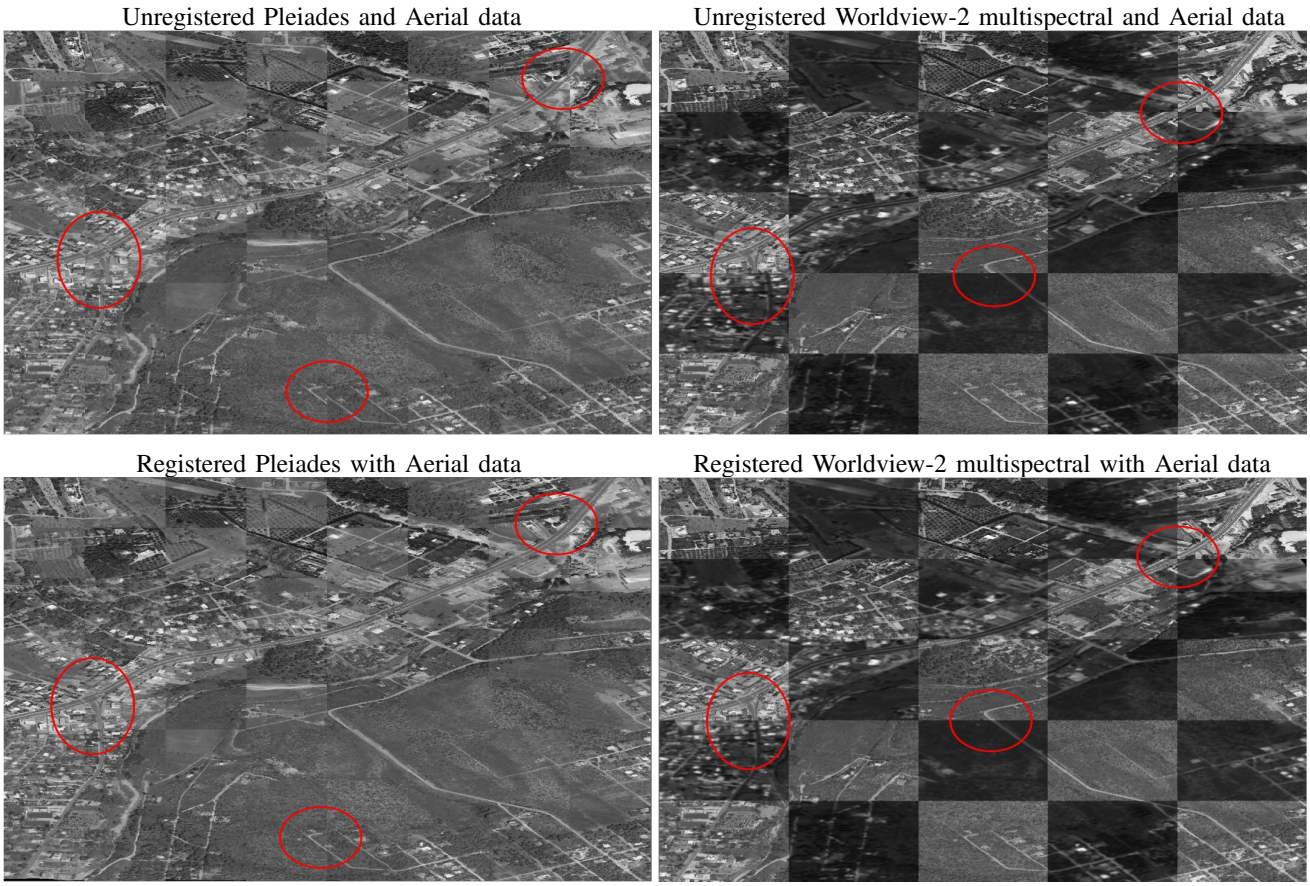

Figure 3. Checkerboard visualization on a smaller region of Figure 2 for the qualitative evaluation of the proposed multimodal registration framework. The upper row illustrates the raw unregistered optical data and the bottom row the registration result.

with, in all cases, three data pyramid levels, three control grid scales and with λ, the parameter which controls the influence of the regularization term, set to 4. Moreover, regarding the quantitative evaluation, the multimodal data were registered with a mean displacement error lower than 3 pixels (Table II).

IV. Conclusions and Future Perspectives

We studied the application of MRF image registration in remote sensing by appropriately adapting it to account for the multi-modality nature of remote sensing data. The MRF registration framework possess certain advantages due to its modularity w.r.t. similarity criterion that allows us to address multiple registration scenarios in a unified way, efficiency and robustness to initialization. We have experimentally demonstrated that graph-based deformable registration can exploit the spectral variation of multitemporal satellite data. The quantitative validation demonstrated the potentials of this approach on very large (more than 100M pixels) multimodal, multitemporal remote sensing datasets. In particular, in terms of spatial accuracy the geometry of the optical and radar data has been recovered with displacement errors less than 2 and 3 pixels, respectively. In terms of computational efficiency the optical data term can converge after 7-8 minutes, while the radar data term after less than 15 minutes. In particular, the main bottleneck of the computational complexity of the proposed approach is the calculation of the unary terms. Nonetheless, this calculation is highly parallelizable. Thus, we plan to explore GPU implementations towards real time performances. Moreover, we plan to validate the proposed approach on other modalities like Lidar, DEMS, hyperspectral data *etc*.

Acknowledgment

This research has been co-financed by the European Union (European Social Fund-ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF)-Research Funding Program THALES: Reinforcement of the interdisciplinary and/or inter-institutional research and innovation.

References

[1] D. Barber, J. Redding, T. McLain, R. Beard, and C. Taylor, "Vision-based target geo-location using a fixed-wing miniature air vehicle," Journal of Intelligent and Robotic Systems, vol. 47, no. 4, pp. 361–382, 2006. 1

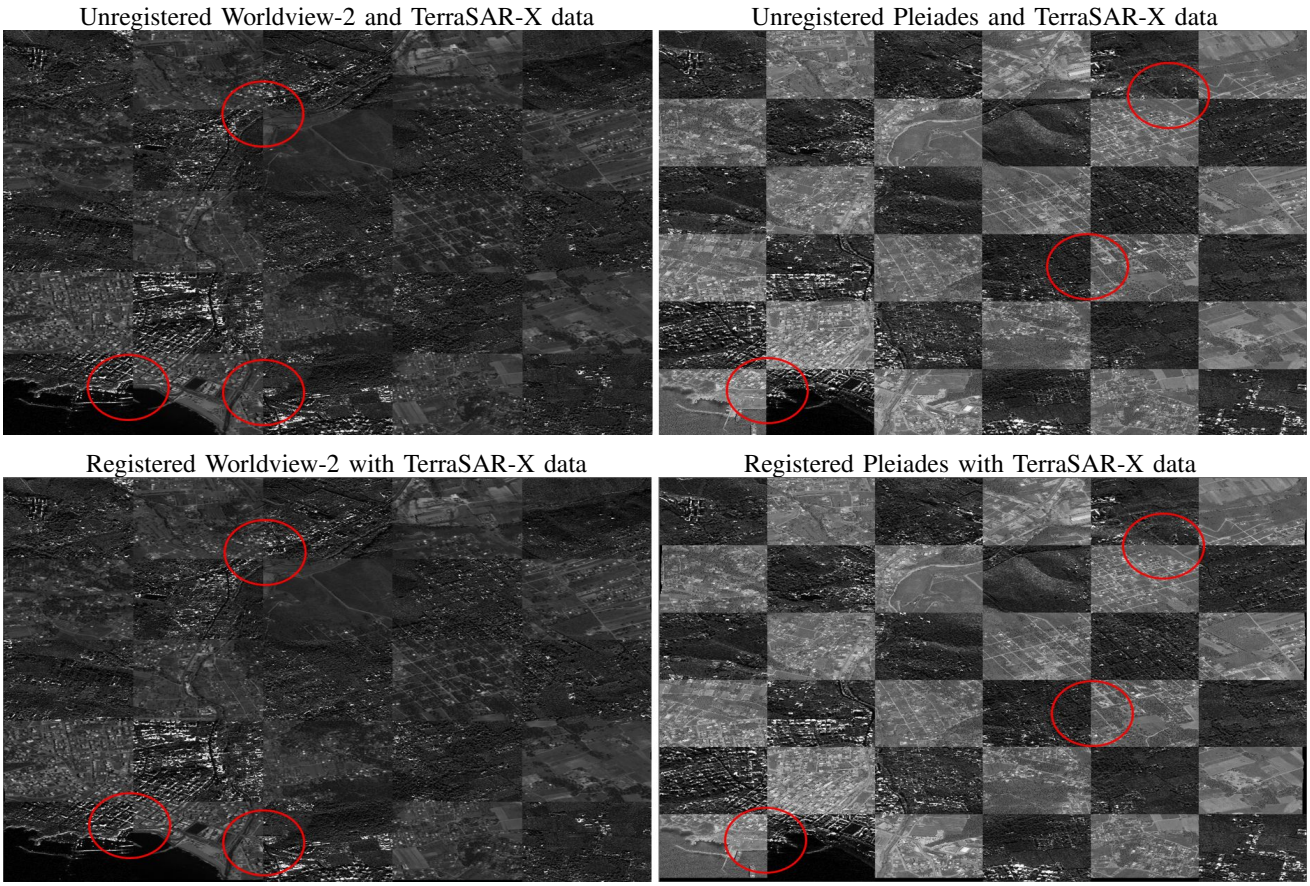

Figure 4. Checkerboard visualization for the qualitative evaluation of the proposed multimodal registration framework. The upper row illustrates the raw unregistered optical (Pleiades and Worldview-2 multispectral) and radar (TerraSAR-X) satellite data. The bottom row demonstrates the registration result and the effectiveness of the proposed algorithm which managed to recover the geometry of the radar data.

[2] J.-M. Frahm, J. Heinly, E. Zheng, E. Dunn, P. Fite-Georgel, and M. Pollefeys, "Geo-registered 3D models from crowd-sourced image collections," Geo-spatial Information Science, vol. 16, no. 1, pp. 55–60, 2013. 1

[3] S. Freire, T. Santos, A. Navarro, F. Soares, J. Silva, N. Afonso, A. Fonseca, and J. Tenedrio, "Introducing mapping standards in the quality assessment of buildings extracted from very high resolution satellite imagery," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 90, pp. 1–9, 2014. 1

[4] C. Cariou and K. Chehdi, "Automatic georeferencing of airborne pushbroom scanner images with missing ancillary data using mutual information," Geoscience and Remote Sensing, IEEE Transactions on, vol. 46, no. 5, pp. 1290–1300, 2008. 1

[5] J. Le Moigne, N. S. Netanyahu, and R. D. Eastman, Image Registration for Remote Sensing. Cambridge University Press, 2011. 1

[6] M. A. Aguilar, M. d. M. Saldaa, and F. J. Aguilar, "Assessing geometric accuracy of the orthorectification process from GeoEye-1 and WorldView-2 panchromatic images," International Journal of Applied Earth Observation and Geoinformation, vol. 21, pp. 427–435, 2013. 1

[7] A. Goshtasby, "Registration of images with geometric distortions," Geoscience and Remote Sensing, IEEE Transactions on, vol. 26, no. 1, pp. 60–64, 1988. 1

[8] H. Erives, S. Teare, and G. J. Fitzgerald, "An automated nonrigid registration for a tunable hyperspectral imaging system," in In: Fourier Transform Spectroscopy/ Hyperspectral Imaging and Sounding of the Environment. Optical Society of America, 2007, p. JWA20. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=FTS-2007-JWA20 1

[9] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," IEEE Transactions on Medical Imaging, vol. 32, no. 7, pp. 1153–1190, 2013. 1

[10] M. Hasan, M. Pickering, and J. Xiuping, "Robust automatic registration of multimodal satellite images using CCRE with partial volume interpolation," Geoscience and Remote Sensing, IEEE Transactions on, vol. 50, no. 10, pp. 4050–4061, 2012. 1

[11] C. Berger, M. Voltersen, R. Eckardt, J. Eberle, T. Heyer, N. Salepci, S. Hese, C. Schmullius, J. Tao, S. Auer, R. Bamler, K. Ewald, M. Gartley, J. Jacobson, A. Buswell, Q. Du, and F. Pacifici, "Multi-modal and multi-temporal data fusion: Outcome of the 2012 GRSS Data Fusion Contest," Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, vol. 6, no. 3, pp. 1324–1340, 2013. 1

[12] B. Fan, C. Huo, C. Pan, and Q. Kong, "Registration of Optical and SAR satellite images by exploring the spatial relationship of the improved sift," Geoscience and Remote Sensing Letters, IEEE, vol. 10, no. 4, pp. 657–661, July 2013. 1

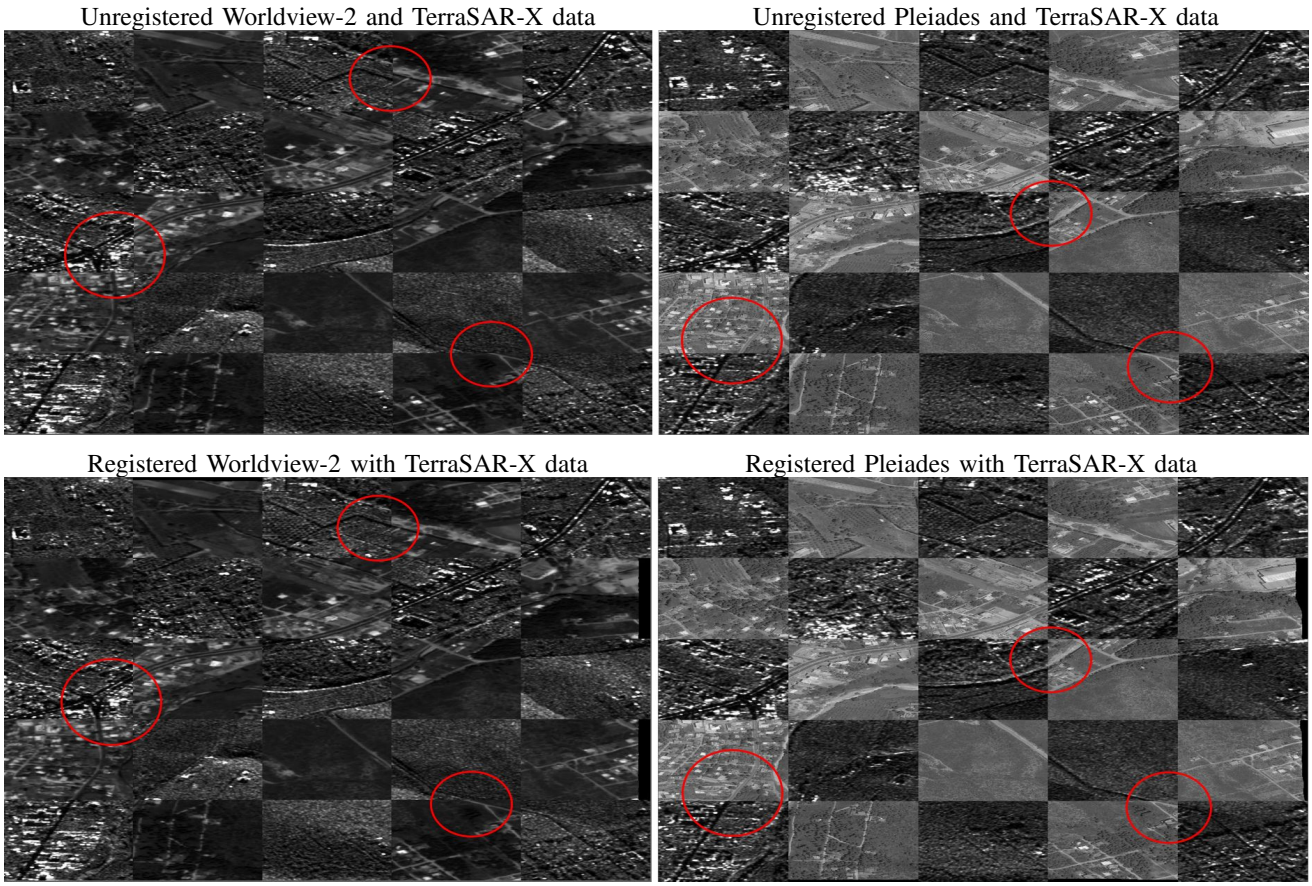

Figure 5. Checkerboard visualization on a smaller region of Figure 4 for the qualitative evaluation of the proposed multimodal registration framework. The upper row illustrates the raw unregistered optical and radar data and the bottom row shows the deformed after the application of the proposed registration functional.

[13] T.-A. Teo and S.-H. Huang, "Automatic co-registration of optical satellite images and airborne lidar data using relative and absolute orientations," Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, vol. 6, no. 5, pp. 2229–2237, 2013. 1

[14] M. Vakalopoulou and K. Karantzalos, "Automatic descriptor-based co-registration of frame hyperspectral data," Remote Sensing, vol. 6, no. 5, pp. 2229–2237, 2014. 1

[15] D. Holtkamp and A. Goshtasby, "Precision registration and mosaicking of multicamera images," Geoscience and Remote Sensing, IEEE Transactions on, vol. 47, no. 10, pp. 3446–3455, 2009. 1

[16] J. Le Moigne, A. A. Cole-Rhodes, R. D. Eastman, N. S. Netanyahu, H. S. Stone, I. Zavorin, and J. T. Morisette, "Multitemporal and multisensor image registration," in Image Registration for Remote Sensing. Cambridge University Press, 2011, pp. 293–338. 1

[17] N. Longbotham, F. Pacifici, T. Glenn, A. Zare, M. Volpi, D. Tuia, E. Christophe, J. Michel, J. Inglada, J. Chanussot, and D. Qian, "Multi-modal change detection, application to the detection of flooded areas: Outcome of the 2009-2010 data fusion contest," Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, vol. 5, pp. 331–342, 2012. 1

[18] C. Cariou and K. Chehdi, "Dense registration of CHRIS-PROBA and IKONOS images using multi-dimensional mutual information maximization," in In: Proceedings of SPIE, Image and Signal Processing for Remote Sensing, 2013. 1

[19] D. De Nigris, D. Collins, and T. Arbel, "Multi-modal image registration based on gradient orientations of minimal uncertainty," IEEE Transactions on Medical Imaging, vol. 31, no. 12, pp. 2343–2354, 2012. 1

[20] T. Hopp, M. Dietzel, P. Baltzer, P. Kreisel, W. Kaiser, H. Gemmeke, and N. Ruiter, "Automatic multimodal 2D/3D breast image registration using biomechanical FEM models and intensity-based optimization," Medical Image Analysis, vol. 17, pp. 209–218, 2013. 1

[21] N. Komodakis, G. Tziritas, and N. Paragios, "Fast, approximately optimal solutions for single and dynamic MRFs," in Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on, June 2007, pp. 1–8. 2, 4

[22] B. Glocker, A. Sotiras, N. Komodakis, and N. Paragios, "Deformable medical image registration: Setting the state of the art with discrete methods," Annual Review of Biomedical Engineering, vol. 13, pp. 219–244, 2011. 2

[23] S. Z. Li, Markov Random Field Modeling in Image Analysis, 3rd ed. Springer-Verlag London, 2009. 2

[24] D. Rueckert, P. Aljabar, R. A. Heckemann, J. V. Hajnal, and A. Hammers, "Diffeomorphic registration using b-splines," in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006. Springer, 2006, pp. 702–709. 4