A COMPARISON OF REGISTRATION TECHNIQUES ON QUICKBIRD SATELLITE IMAGES

V. Arévalo*, J. González

Dept. of System Engineering and Automation, University of Málaga, Campus Teatinos–Complejo Tecnológico, 29071, Málaga (Spain), {varevalo,jgonzalez}@ctima.uma.es

KEY WORDS: Elastic image registration, non-rigid deformations, high resolution satellite imagery, QuickBird.

ABSTRACT:

In remote sensing, because of the wide diversity of image characteristics (size, spatial and radiometric resolution, terrain relief, observation poses, etc.), image registration methods that may work well on certain satellite images may not produce acceptable results for others, requiring more powerful techniques. A variety of registration techniques that account for images with non-rigid geometric deformations has been proposed, including piecewise (linear or cubic) functions, weighted mean functions, radial basis functions, B-spline functions, etc. This paper compares three of them: polynomial, piecewise linear and radial basis –thin plate spline– functions; and evaluates their accuracy according to two well-known metrics: root mean square error (RMSE) and circular error with 90% confidence (CE90). The comparison focuses on pan-sharpened QuickBird images (0.6 meters/pixel) acquired on different dates, from different observation attitudes, and sensing different land covers: urban area, high relief terrain and a combination of both. The experimental results show that local methods (as the radial basis functions) perform much better than global ones based on polynomial functions since they can exploit the geometrical information captured by a (desirable) large number of control points.

1. INTRODUCTION

Image registration is the process of spatially aligning two or more images of the same scene acquired on different dates (multitemporal analysis), from different viewpoints (multiview analysis) and/or using different sensors (multimodal analysis). In this process, one image remains without modification (the fixed image) whereas the other (the moving image) is spatially transformed until fitting with the fixed one. Image registration is a crucial step in those image analysis applications where the final result comes from the association of several data sources, for example, image fusion, change detection, 3D scene reconstruction, etc.

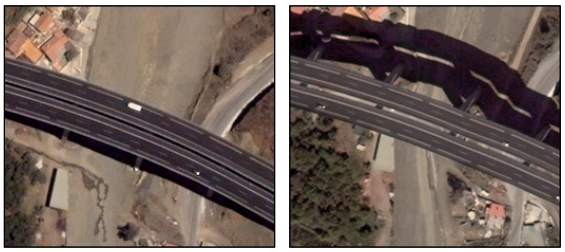

Figure 1: A pair of QuickBird satellite images of the same scene taken from different observation angle. Observe the geometric and radiometric differences between both images.

In remote sensing, because of the wide range of image characteristics (size, spatial and radiometric resolution, acquired scene, observation position, etc.), one registration method that may work well on certain satellite image will not produce acceptable results for others, requiring more powerful techniques. In concrete, global polynomial functions usually perform well with low-medium resolution images (Landsat, IRS, Spot, etc.), but may not be effective enough to register high resolution images such as QuickBird, Ikonos or Orbview. The main reasons for that limited performance include (see figure 1):

1. Larger image distortions because of the higher resolution. Distortion comes from different sources, but the most significant one is the off-nadir observation angle. For offering shorter revisit periods, these satellites can observe the scene from very different paths and angles, which gives rise to images with significant non-linear geometric differences (visit Space Imaging or Digital Globe web pages for sample imagery).
2. Changes in the scene (even small ones) appear now clearer in the images, hence it makes more difficult to successfully accomplish some stages of the registration process like detecting corresponding points or measuring the goodness of registration. Examples of these changes include temporal changes, cast shadows, different sides of a building, etc.

To address the registration of this kind of images some elastic fitting techniques have been proposed in the image processing literature, including: piecewise linear (or cubic) functions (Goshtasby, 1986), weighted mean functions (Goshtasby, 1988a), radial basis functions (Bookstein, 1989), B-spline functions (Kybic & Unser, 2003), etc. These techniques can be grouped into any of the following approaches:

1. *Intensity-based methods*, where registration is approached as an optimization process in which a cost function based on the radiometric similarity of both images is maximized.
2. *Landmark-based methods*, where registration is accomplished by a mapping function estimated from a set of representative pairs of control points (also called landmarks) identified in both images.

* Corresponding author.

Most of the elastic registration techniques applied within the remote sensing field follow this second approach since they are more efficient computationally. In spite of that, landmark-based methods may require to spend a considerable amount of time to identify precise corresponding points in both images, which may become a non-trivial problem in many practical situations where hundred of control points are required for capturing the relative geometric distortions. From a user point-of-view, it is clear that if a simple registration method (i.e. global polynomial function) achieves the accuracy required for a particular application, there is no need to waste time clicking extra corresponding points demanded by a more powerful technique.

Thus, we are interested in knowing the performance of these techniques on QuickBird imagery and in which cases (relative poses, terrain relief, etc.) more sophisticated registration procedures become necessary. To this aim, this paper compares three representative landmark-based registration methods[†]: polynomial, piecewise linear and thin plate spline functions. We evaluate their performance according to two metrics for the registration consistency: root mean square error (RMSE) and circular error with 90% confidence (CE90). The comparison focuses on QuickBird images (0.6 meters/pixel) acquired on different dates, from different viewpoints, and sensing different terrain profiles.

The remainder of this paper is organized as follows. In section 2 we review the non-rigid registration techniques considered in this study. In section 3, the datasets (images and sets of corresponding points) and metrics for measuring the registration consistency in the comparison are described. In section 4, we present and discuss the experimental results. Finally, some conclusions and future work are outlined.

2. IMAGE REGISTRATION METHODS

Landmark-based registration is usually carried out in three stages (Zitová & Flusser, 2003). In the first, the positions of a set of control points (CP) are accurately identified in the fixed and moving images; in the second one, this set of CP is used to estimate a geometric transformation function between both images; and finally, the moving image is spatially transformed to overlap the fixed one using the estimated mapping function and by applying some interpolation technique such as nearest neighbour, bilinear, bicubic or splines. For the registration to be successful, it is necessary that a) the correspondence pairs must be distributed on the images according to their geometric differences, and b) the applied transformation must be powerful enough to cope with the (possibly, non-linear) distortions.

This paper evaluates three well-known methods of image registration and analyzes their performance on QuickBird imagery. More precisely, we have compared a *global* procedure based on polynomial functions (of diverse orders), a *local* method based on piecewise linear functions, and a *hybrid* technique based on radial basis functions (concretely, thin plate splines). Formally, a pair of generic image mapping functions can be expressed as follow:

$$\begin{aligned} x &= f_x(x', y') \\ y &= f_y(x', y') \end{aligned} \tag{1}$$

where (x, y) = CP localization in the fixed image.
(x', y') = CP localization in the moving image.

In next sections, the functions analyzed in this work are described in detail.

2.1 Polynomial functions

Polynomial functions have been broadly used in remote sensing to register low-medium resolution images (Landsat, IRS, Spot, etc.), thematic data, etc. (Estrada et. *al*., 2001). The kind of geometric differences that these functions can manage depends on the polynomial order. Thus, we can use 1st order transformation to model translations, rotations and scale changes, that is, *rigid* distortions and 2nd and higher orders to model more complex distortions, named *non-rigid* or *elastic* distortions. A pair of polynomial functions of order t are defined as follow:

$$\begin{aligned} x = f_x(x', y') &= \sum_{i=0}^{t}\sum_{j=0}^{i} a_{ij} \times (x')^{i-j} \times (y')^{j} \\ y = f_y(x', y') &= \sum_{i=0}^{t}\sum_{j=0}^{i} b_{ij} \times (x')^{i-j} \times (y')^{j} \end{aligned} \tag{2}$$

where a_{ij} and b_{ij} are the polynomial coefficients.

The order of the polynomial also determines the minimum number of CP to be estimated. For example, to estimate a 1st order transformation, 3 non-collinear points are required. The following expression provides the minimum number N of CP required to estimate a polynomial function of order t:

$$N = \frac{(t+1)(t+2)}{2} \tag{3}$$

In practice, since the number of CPs is usually higher than N, the coefficients are computed by means of a least-squares fitting.

The limitation of this transformation comes from its global scope which permits us to cope with important image differences, but only if they are spread over the whole image (not locally).

2.2 Piecewise linear functions

Piecewise linear functions deal with the registration process by dividing the images in triangular elements (for example, by a Delaunay's triangulation method) which are then individually mapped using a linear transformation (Goshtasby, 1986). Although this approach guarantees the continuity of adjacent triangles, it does not produce smooth transitions between them, which causes an undesirable visual effect in the transformed image (i.e. line segments are not preserved). A pair of piecewise linear functions are defined as follow:

[†] All of them are included in most of the current commercial satellite image processing packages, such as ERDAS, ENVI and PCI.

$$x = f_x(x',y') = \begin{cases} a_{11} + a_{12}x' + a_{13}y' & if \quad (x',y') \in t_1 \\ \vdots \\ a_{n1} + a_{n2}x' + a_{n3}y' & if \quad (x',y') \in t_n \end{cases}$$
$$y = f_y(x',y') = \begin{cases} b_{11} + b_{12}x' + b_{13}y' & if \quad (x',y') \in t_1 \\ \vdots \\ b_{n1} + b_{n2}x' + b_{n3}y' & if \quad (x',y') \in t_n \end{cases} \quad (4)$$

where t_i = triangular element built upon three CPs (provided by the triangulation method).

a_{ij} and b_{ij} = polynomial coefficients of the linear functions corresponding to the triangle t_i.

n = number of triangular elements.

In general, the number of CPs required for a good registration will depend on the type of deformation to be modelled in the image, ranging from some dozens to thousands. Notice that, the mapping functions are only defined inside the convex hull of the control point set. Although extrapolation is possible outside this region, we do not apply it in this work because it introduces a specific error that could distort the registration result.

2.3 Radial basis functions (Thin plate splines functions)

Radial basis functions (RBF) are scattered data interpolation methods where the spatial transformation is a linear combination of radially symmetric basis functions, each of them centred on a particular CP. RBFs provide smooth deformations with easily controllable behaviour. In 2-dimensions, an RBF consists of two mapping functions that comprise a global component (typically, an affine transformation) and a local component. Given *n* corresponding CPs, we can define a pair of radial basis functions as follow:

$$x = f_x(x',y') = \sum_{i=0}^{t}\sum_{j=0}^{i} a_{ij} \times (x')^{i-j} \times (y')^{j} + \sum_{i=1}^{n} A_i g(r_i)$$
$$y = f_y(x',y') = \sum_{i=0}^{t}\sum_{j=0}^{i} b_{ij} \times (x')^{i-j} \times (y')^{j} + \sum_{i=1}^{n} B_i g(r_i) \quad (5)$$

where a_{ij} and b_{ij} = polynomial coefficients of the global component.

A_i and B_i = coefficients of the local component.

$g(r_i)$ = non-linear basis function where r_i denotes the Euclidean norm, that is $r_i = \|(x',y') - (x'_i, y'_i)\|$.

t = polynomial order.

These mapping functions are linear combinations of radially symmetric functions $g(r_i)$ and a low degree polynomial. The type of basis function determines the influence of each CPs on the RBF, that is, the CP scope. Some RBFs have a global behaviour (e.g. thin plate spline) whereas others have a more local influence (e.g. Gaussians). Table 1 shows some RBFs commonly used.

Basis function	$g(r_i)$	Parameters	Scope
Thin-Plate spline	$r_i^2 \log r_i^2$	-	global
Multi-quadric	$(r_i^2 + \delta)^{\pm\mu}$	$\delta > 0, \mu \neq 0$	local
Gaussian	$e^{(-r_i^2/\sigma)}$	$\sigma > 0$	local
Shifted-LOG	$\log(r_i^2 + \delta)^{\frac{3}{2}}$	$\delta \geq 0$	local
Cubic spline	$\|r_i\|^3$	-	global

Table 1: Some types of Radial Basis Functions (RFB).

In particular, in this paper we evaluate the behaviour of the thin plate spline functions (TPS), which is perhaps the transformation most widely employed for elastic registration. TPS was introduced by (Harder & Desmarais, 1972) and successfully applied to register Landsat (Goshtasby, 1988b) and medical images (Bookstein, 1989).

3. DATASETS AND METRICS

The goal of this work is to evaluate the performance of the elastic mapping functions mentioned above regarding the following two issues:

a. How they model the non-linear deformations that arise in a pair of QuickBird images because of the terrain relief and off-nadir observation angle.
b. How the number of CPs influences the registration accuracy.

For this purpose, we consider a variety of dataset (images and sets of CP) and two metrics of image registration consistency which are described in the following sections.

3.1 Datasets

In this section, we give a brief description of the test images as well as the CP sets used in this study.

3.1.1 Test images: We have considered three images of the city of Rincón de la Victoria (Málaga-Spain) acquired on different dates, from different poses and covering different terrain profiles: urban area (which is almost plain) and mountainous area (high relief). Figure 2 shows two of these images where the regions of interest are marked. The plot on top of the figure is a terrain profile where elevation levels range from 300 meters at the mountains to 0 meters at the coastline (information obtained from a DEM with a spatial resolution of 20×20 meters).

Angles	Image 1 (i1)	Image 2 (i2)	Image 3 (i3)
In track	15.9	-12.8	-1.1
Cross track	10.8	-2.3	11.0
Off nadir	19.3	12.8	10.9

Table 2: Satellite positioning data (in degrees) for the images used in the tests (please, refer to figure 3 for the meaning of these angles).

Table 2 shows the QuickBird viewing direction for the images used in the experiments. The pairs considered for registration are {<i1–i2>, <i1–i3>, <i2–i3>}. These image pairs have different relative observation angles which, in combination with the terrain relief (the same for all the images), give rise to diverse geometric distortions (see figure 3).

3.1.2 Control point sets: We have identified both control points (CP) to estimate the function coefficients and

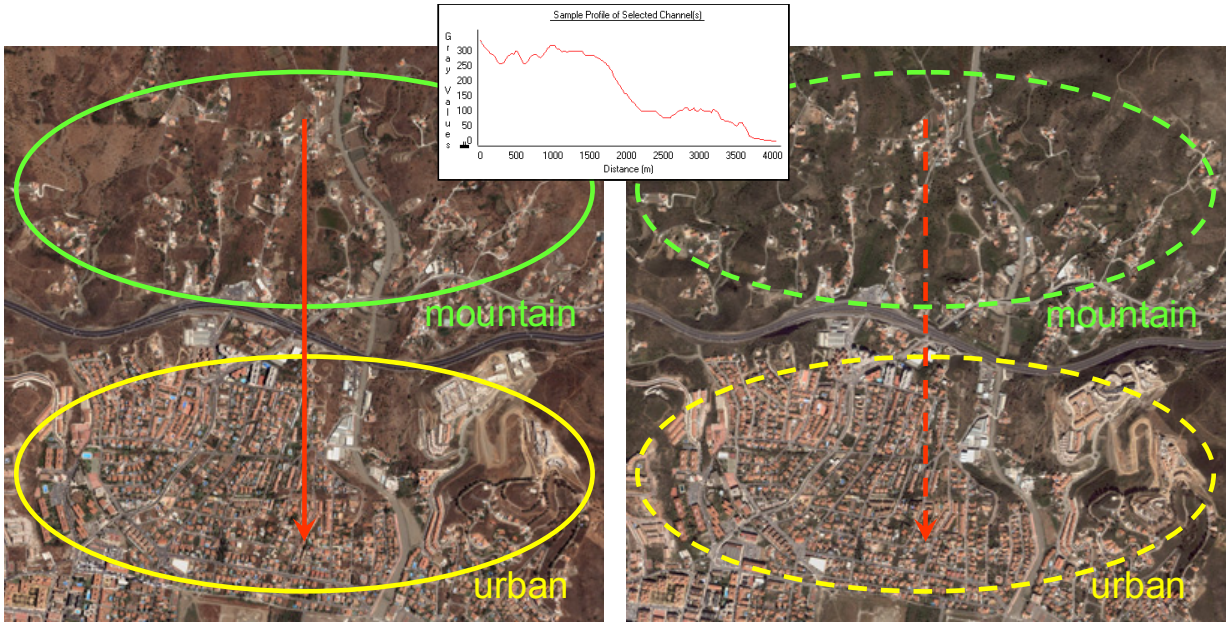

Figure 2: City of Rincón de la Victoria (Málaga-Spain). This image covers approximately 4 km^2 and contains all different zones of interest for the comparison tests (mountains, city, and coast).

independent checkpoints (ICP) to evaluate the accuracy of the estimated functions. To guarantee a uniform distribution of them over the image we pick a point from every cell of a rectangular grid. In order to have different densities of control points we define different cell sizes. In particular, the cell widths used in this work are 50 pixels for the ICP set and 100, 200 and 400 pixels for the CP sets. For QuickBird images these widths correspond to 30, 60, 120 and 240 meters, respectively (see figure 4). In other words, we dispose of three sets of 900, 225 and 64 CPs, and one set of 3600 ICPs for each pair of images. To accurately identify the pairs of CPs or ICPs, we have used an automatic procedure based on the following techniques:

- The Harris detector (Harris & Stephens, 1988) to identify distinctive feature points in the fixed image, and
- the Lucas-Kanade point tracker (Lucas & Kanade, 1981) to correspond fixed-image points with others in the moving one. This algorithm works well when both images are roughly aligned, which was done manually.

Next, the affine epipolar geometry of the two images is robustly estimated and the outliers (CPs not consistent with the estimated geometry) are removed from the initial set of point correspondences (Hartley & Zisserman, 2003). Finally, one single point is chosen for each cell of the grid. This process assures the uniform distribution of the control points throughout the image.

It is important to remark that though this automatic procedure allows us to manage a large number of consistently-matched point pairs (both CP and ICP), it does not guarantee that the corresponding points are the most suitable ones to capture the goodness of the image registration. For example, a perfectly-matched ICP pair at a flat roof of a very high building is not appropriate for measuring how good the ground or low-height buildings have been registered. This problem shows up in some of our experimental results in section 4.

3.2 Metrics for registration consistency

To evaluate the precision of each method, the following two metrics are utilized: circular error with 90% confidence (CE90) and root mean square error (RMSE). Both measures are applied to the registration errors of the ICP pairs, which are computed from the distances between corresponding ICPs of the fixed image and the registered one.

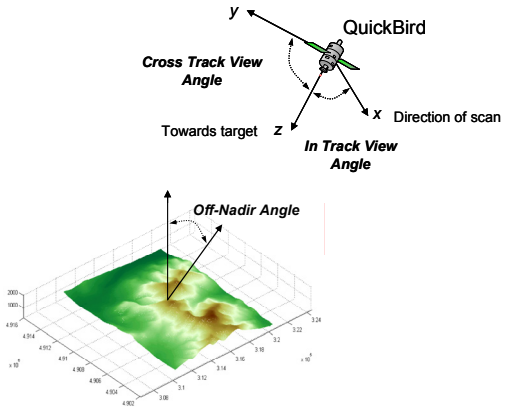

Figure 3: Satellite positioning system.

3.2.1 Circular error with 90% confidence: The CE90 is a measure to describe the accuracy in map or image products at the 90% confidence level, that is, the 90% of the ICP pairs must have distance errors within a circle of radius CE90. From a statistical point of view, this is stated in terms of the probability P such that:

$$P\left(\left\|(\hat{x},\hat{y})-(x,y)\right\| \le CE90\right) = 90\% \qquad (6)$$

where (x,y) = pixel coordinates of the a fixed image ICP.
$(\hat{x},\hat{y})$ = pixel coordinates of its corresponding ICP in the registered image.

3.2.2 Root mean square error: The RMS is also a measure to describe the accuracy of registration process. In particular, RMS measures the magnitude of the total longitudinal error. The name comes from the fact that it is the square root of the mean of the squares of the errors associated to each ICP pair. It is mathematically computed from:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left\| (\hat{x}, \hat{y}) - (x, y) \right\|^2} \qquad (7)$$

where n is the number of ICPs, and $(\hat{x}, \hat{y})$ and (x, y) are the ICP coordinates (as defined above).

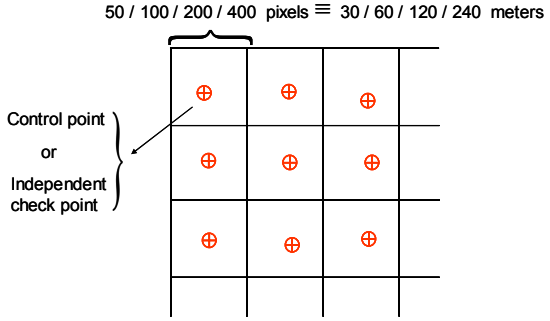

Figure 4: To spread both CPs and ICPs on the image, we divide it into a grid and choose one single point from each cell. The cell sizes used in this work are 50 pixels for the ICP set and 100, 200 and 400 pixels for the CP sets (for QuickBird images these widths correspond to 30, 60, 120 and 240 meters, respectively).

RMSE and CE90 are two complementary metrics commonly used in remote sensing. The main difference between them is that RMSE takes into account all the errors, including those that could be considered outliers. On the contrary, CE90 is only affected by the majority of them (the 90%), not capturing how bad the remaining 10% could be.

4. EXPERIMENTS AND RESULTS

In this section, the performance of the polynomial (from 2nd up to 6th order), piecewise linear and thin plate splines functions (with affine global components) are compared for different numbers of CPs, terrain profiles and observation angles. Since the effect of any of these parameters on the registration accuracy depends on the values of the others, an exhaustive analysis of them requires trials over all the possible combinations. We have carried out all these tests, though here we only show the comparative plots for those that we understand are of more significance. Regarding the implementation of the mapping functions, two points must be highlighted:

- We have experimentally verified that polynomials of order higher than 4th barely improve the registration although they require much more CPs to be estimated. Consequently, we only show results for the 4th order one, which requires a minimum of 15 CPs.
- As reported in the literature (Madych, 1992), when the number of control points is high (e.g. more than 1000) the equation system to be solved for the thin plate spline functions becomes ill-conditioned. To overcome this problem we have implemented the iterative method proposed in (Beatson et. *al*., 2001).

Figure 5.a shows the influence of the *number of control points* (i.e. their density) in each method. In these charts, we have fixed the angle of observation (Image pair <i1-i2> which has the largest difference angle) and the terrain profile (Mixed). The most interesting conclusion from it is that local methods (PWL and TPS) take advantage of the number of points, achieving very accurate results for the dense CP set (about 1 meter RMSE). On the contrary, global functions (POL4) do not improve the performance for higher number of points since nothing (except robustness) is gained when more than the required 15 CPs are employed. Notice also that, when the number of CPs decreases (sparse CP sets) local transformations become global (weak local) and can not adapt well to the local geometric differences between images.

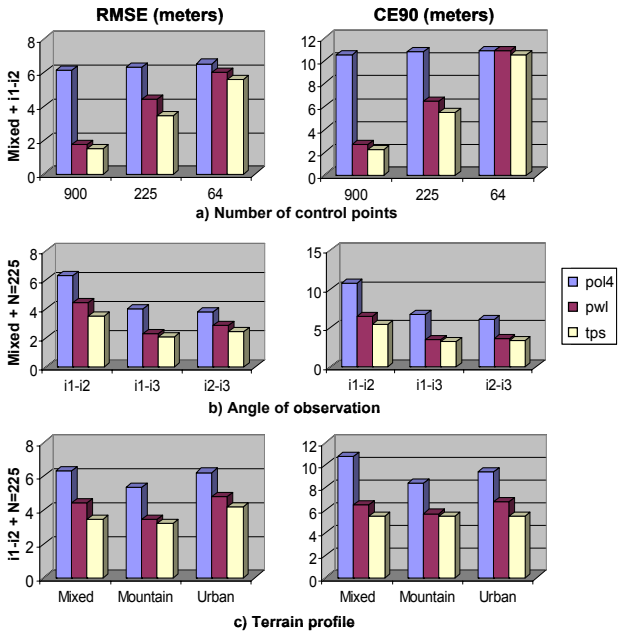

Figure 5: RMSE and CE90 values for the elastic functions analyzed in this work according to: a) number of CPs, b) angle of observation (image pairs), and c) terrain profile. On the left of each row, we indicate the fixed values of the other parameters.

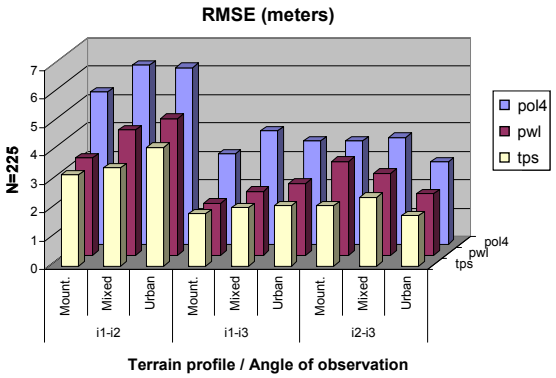

Figure 6: RMSE values for the considered functions grouped by the angle of observation and the terrain profile.

Figure 5.b shows the influence of the *angle of observation* for a (Mixed) terrain profile and 225 CPs (taken from cells of 120 m. width). From this comparative, we can observe that all methods

(local and global) are sensitive to the angle of observation although the influence is bigger for the POL4 global method. Although not shown here, for a denser set of CPs, POL4 performs similarly (practically not improvement of the results of figure 5.b), while the local methods achieve better results, especially for the pair <i1-i2>.

The effect of the *terrain relief* is displayed in figure 5.c, where we can highlight two things: a) local methods perform very similar in all profiles, b) the polynomial function (POL4) works worse on mixed terrain, which makes sense since it can not fit well different deformation models simultaneously. Figure 6 (which also contains figure 5.c) is a more comprehensive chart aimed to give us, at a single glance, the increasing effect of the observation angle for more irregular terrains (relative viewing angles of <i1-i3> and <i2-i3> are very similar). Please, notice that, maybe unexpectedly, the estimated registration error for the urban area is significantly bigger than that for the mountain. As commented earlier in section 3.1.2, we attribute this result to the existence of an important portion of ICPs that lie on top of the buildings.

From these experiments we can draw the following conclusions:

1. Local methods (particularly, thin plate splines) present better performance in all the experiments than the polynomial functions of 4^{th} order.
2. Global methods are suitable for high resolution image registration only in any of the following situations:
   a. Both images have been acquired from very close viewing angles, or
   b. the scene under observation is almost flat (neither mountains nor elements at different heights, e. g. buildings).
3. It is worthy to have as many points as possible when using a local method (not for global ones).

Thus, except for those cases where we certainly know that either the captured scene is practically plain, with no elements at different heights, or the two images have been taken with almost the same observation angle, we would suggest to employ local methods with as many CPs as possible (the larger the number of points, the better the registration accuracy). This result brings up the importance of developing an automatic and reliable procedure to find well-distributed pairs of CPs in the images.

5. CONCLUSIONS AND FUTURE WORK

High resolution satellite images such as QuickBird are expected to play an important role in many remote sensing applications. For achieving that, tools commonly used for lower resolution images may not be appropriated, as it is the case of traditional image registration methods. This paper has experimentally analyzed the performance of three well-known elastic registration techniques such as polynomial, piecewise linear and thin plate splines functions for diverse conditions: number of control points, terrain relief, and acquisition angles. We have evaluated their suitability and accuracy for registering QuickBird images according to two metrics: root mean square error (RMSE), circular error with 90% confidence (CE90).

From this analysis we have verified some of the intuitions that we had, but more importantly, this has allowed us to quantify the influence of the above factors in the performance of each registration method. Local methods (PWL or TPS) beat by far the global one (POL4) since they can exploit the information provided by many CPs. This indicates the importance in having tools for automatically detecting in the images as many CPs as possible. This is one of our concerns for the next future, as well to develop some kind on strategy for distributing the CPs in the image in a way that captures as better as possible the image relative deformations.

REFERENCES

Beatson, R.K., Light, W.A. and Billings, S.D. (2001). Domain decomposition methods for solution of the radial basis function interpolation equations. *SIAM Journal on Scientific Computing*, 22, pp. 1717-1740.

Bookstein, F. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. on Pattern Anaylsis and Machine Intelligence*, 11(6), pp. 567-585.

Estrada, M., Yamazaki, F. and Matsuoka M. (2000). Use of LANDSAT images for the identification of damage due to the 1999 Kocaeli, Turkey Earthquake. *In Proc. of The 21st Asian Conference on Remote Sensing*, Taipei, Taiwan.

Goshtasby, A. (1986). Piecewise linear mapping functions for image registration. *Pattern Recognition*, 19, pp. 459-466.

Goshtasby, A. (1988a). Image registration by local approximation methods. *Image and Vision Computing*, 6, pp. 255-261.

Goshtasby, A. (1988b). Registration of image with geometric distortion. *IEEE Transactions on Geosciences and Remote Sensing*, 26(1), pp. 60-64.

Harder, R.L. and Desmarais, R.N. (1972). Interpolation using surface splines. *Journal of Aircraft*, 9, pp. 189-191.

Harris, C.J. and Stephens, M. (1988). A combined corner and edge detector. *In Proc. of 4th Alvey Vision Conference*, pp. 147-151. Manchester.

Hartley, R.I. and Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge University Press, Cambridge, pp. 279-292.

Kybic, J. and Unser, M. (2003). Fast parametric elastic image registration. *IEEE Transactions on Image Processing*, 12(11), pp. 1427-1442.

Lucas B.D. and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. *IJCAI81*, pp. 674-679.

Madych, W. (1992). Miscellaneous error bounds for multi-quadric and related interpolators. *Computers Math. Applic.* 24, pp. 121-138.

Zitová, B. and Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21, pp. 977-1000.

ACKNOWLEDGEMENTS

The ©DigitalGlobe QuickBird imagery used in this study is distributed by Eurimage, SpA. (www.eurimage.com) and provided by Decasat Ingeniería S.L., Málaga, Spain. (www.decasat.com).