

A METHOD FOR ROBUST EXTRACTION OF CONTROL POINTS ON HIGH-RESOLUTION SATELLITE IMAGES

J. González, V. Arévalo*, and C. Galindo

Dept. of System Engineering and Automation
University of Málaga
Campus Teatinos, 29071 Málaga, Spain
*varevalo@ctima.uma.es
<http://babel.isa.uma.es/mapir>

Commission III/3

KEY WORDS: Digital Terrain Model, Control-Point Extraction, Elastic Registration, High-Resolution Imagery, Quadtree Decomposition, Affine Epipolar Geometry

ABSTRACT:

This paper presents a procedure to robustly distribute control point (CP) pairs in high-resolution satellite images as a preliminary step for accurate image registration. The proper distribution of the CPs is achieved by means of a quadtree decomposition of a coarse digital terrain model (DTM) of the sensed region. This technique parcels up the image according to its relief variance yielding almost planar pieces of land. A corner detector is then employed to identify key points in the *reference* image and an affinity-based feature tracker that searches for their corresponding corner in the *target* one. This search is executed in every parcel, selecting (at-least) one CP, ensuring thus denser distributions in rugged regions than in flat ones. Additionally, robustness to mismatches is attained by exploiting the intrinsic affine epipolar geometry of the two images. The proposed method has been successfully tested with a broad variety of panchromatic high-resolution images of the city of the Rincón de la Victoria (Málaga, Spain).

1 INTRODUCTION

Image registration is the process of spatially fitting two images of the same scene acquired on different dates, from different viewpoints, and/or using different sensors. Image registration is required in a variety of applications, like, image fusion, 3D scene reconstruction, and multi-temporal analysis (i. e. natural disaster monitoring, urban change detection, etc.). See (Schowengerdt, 2007) for a comprehensive survey.

Image registration is typically accomplished by (automatically or manually) identifying common features, called control points (CP) pairs, in the involved images. Through such CPs it is possible to estimate the underlying geometrical transformation between the considered images, which is used to spatially transform (register) the *target* image. The accuracy of this process is, then, strongly tied to:

1. the type of geometrical transformation considered for the registration (affine, projective, piecewise linear, thin-plate-spline, etc.), which should account for the relative geometric differences between the images, and
2. the distribution of CPs over the images, which should take into account the nature of their differences.

A correct combination of both aspects is crucial to guarantee the accuracy of the registration: while only two pairs of CPs suffice to perfectly overlap images of a flat terrain (since they may only differ in shift, scale and rotation), a large number of them will be necessary to capture the geometric difference between images of high-relief surfaces acquired from different viewing angles, requiring, also, complex (so-called elastic) transformations. While elastic transformations have been broadly studied in the remote sensing field (see (Arévalo and Gonzalez, 2008), for example),

the proper distribution of the CP pairs has not been addressed indeed. This paper focuses on this issue.

In the absence of information about the relief of the imaged surface, the more effective (but surely not more efficient¹) approach is the straightforward solution of distributing regularly as many CPs as possible all over the images (Arévalo and Gonzalez, 2008). However, when some information about the terrain profile is available, a more elaborated algorithm can help us to decide their appropriate distribution on the images.

This paper presents an automatic method to distribute CPs for the accurate registration of high-resolution satellite images. Exploiting the terrain profile information provided by a coarse digital terrain model (DTM) of the imaged scene, our approach generates a minimal distribution of CPs, achieving significant speedup in the CPs extraction, without jeopardizing accuracy in the registration.

Our method is intended to be applied to *basic* high-resolution satellite imagery, that is, products that are only featured with corrections for radiometric distortions and adjustments for internal sensor geometry, optical and sensor distortions. As the effect of the terrain is not compensated, two images of a rugged region acquired from different viewpoints may present severe local geometric differences. Main providers, as it is the case of GeoEye or DigitalGlobe², distribute several of these products, as the Ikonos Ortho Kit, QuickBird Orthoready, etc., which are significantly cheaper than geometrically corrected ones.

The rest of this paper is organized as follows. In section 2, we describe in detail the proposed method. In section 3, some experimental results are presented. Finally, some conclusions and future work are outlined.

¹The reader can refer to (Fonseca and Kenney, 1999) for an interesting control-point assessment for image registration.

²The reader can refer to the websites <http://www.geoeye.com/> and <http://www.digitalglobe.com> for further information.

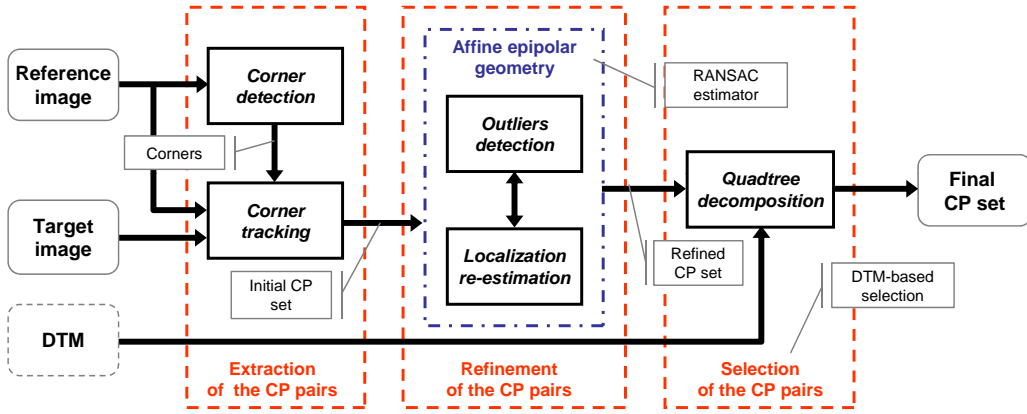


Figure 1: Scheme of the proposed method. Please refer to the section 2 for a detailed description.

2 DESCRIPTION OF THE PROPOSED METHOD

The proposed method combines techniques adapted from the computer vision field to divide the images according to their estimated relative distortions and to robustly distribute CPs pairs following the obtained partitioning. Figure 1 schematically shows an overview of our approach.

In a nutshell, the proposed method consists of three steps: 1) the extraction of the CP candidates, 2) the detection of mismatches (we take advantage of the affine epipolar geometry of the images to robustly deal with this stage), and finally 3) the selection of the CPs according to the distribution obtained from the DTM.

2.1 Extraction and Matching of the CPs

This stage is accomplished in two steps: first, a corner detector (Harris and Stephens, 1988) is applied to identify distinctive points in the reference image and then, a feature tracker searches for their correspondences in the target image, assuming local affinity deformations (Lucas and Kanade, 1981). Our implementation of the feature tracker relies on a variant of the sum of square differences (SSD) which provides robustness to image brightness differences through a local linear radiometric correction (Fonseca and Kenney, 1999).

We also achieve image scale invariance by means of a Gaussian pyramid decomposition (Burt and Adelson, 1983) of the image pair. This technique, broadly used in image processing, consists of creating a series of images down-scaled by convolving the image with a Gaussian kernel (a low-pass filtering). Thus, a stack of successively smaller images is created, where each pixel contains the local gaussian-weighted average that corresponds to a pixel neighbourhood on a higher level of the pyramid.

More formally, let the 2D original image be denoted by $I(x, y)$. The Gaussian pyramid decomposition of $I(x, y)$ can be recursively defined as follow

$$G_l(x, y) = \begin{cases} I(x, y) & l = 0 \\ \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) G_{l-1}(2x + m, 2y + n) & l > 0 \end{cases} \quad (1)$$

where $w(m, n)$ is the gaussian kernel (identical at all levels).

This technique allows the feature tracker to cope with large displacements between corresponding corners.

2.2 Detecting Mismatches and Refining the Coordinates of the CPs

The intrinsic affine epipolar geometry of two views is exploited for attaining robustness to mismatches (the so-called *outliers*) and, collaterally, refining the coordinates of the extracted CPs (Torr, 2002). To this aim we employ the RANdom SAMple Consensus algorithm (Fischler and Bolles, 1981), a robust estimator which exploits the redundancy of samples to provide a robust estimate of the parameters of a model which fits to the majority of them.

The final step of the RANSAC consists of re-estimating the model but only considering the *inliers*. In our case, this step is accomplished by minimizing the symmetric epipolar error from which we derive the Maximum Likelihood (ML) estimate of the affine epipolar matrix, F_A , and refine the coordinates of the CPs. In this process, we assume that the image point localizations are affected by Gaussian noise.

The ML estimate is obtained by minimizing³ the following cost function based on geometric image distances:

$$\min_{\{F_A, \hat{x}_i, \hat{x}'_i\}} \sum_{i=1}^n d(x_i, \hat{x}'_i)^2 + d(x_i, \hat{x}_i)^2 \quad (2)$$

where as usual $x_i \longleftrightarrow x'_i$ are the measured correspondences, and \hat{x}_i and \hat{x}'_i are the estimated “true” correspondences that satisfy $\hat{x}_i^T F_A \hat{x}'_i = 0$ exactly for the estimated affine epipolar matrix.

2.3 Distribution of the CPs According to the Image Distortions

The proper distribution of the CPs is regulated by means of a quadtree decomposition of a medium-resolution DTM of the sensed scene. This decomposition uses the relief variance to parcels up the image in almost planar plots of land. The algorithm 1 depicts this process.

An illustrative example of a coastal city surrounded by mountains is shown in figure 2. Upon a DTM of 20 m. of spatial resolution of this region provided by the “Consejería de Medio Ambiente” of the “Junta de Andalucía” (Spain) (figure 2-b), our method generates a quadtree decomposition according to the relief of the different parts: a region is divided in 4 equal pieces, when elevation

³Minimizing expression (2) is equivalent to fitting the hyperplane f to the set of points $X_i = (x_i, y_i, x'_i, y'_i)^T$ in \mathbb{R}^4 . The refined points $\hat{X}_i = (\hat{x}_i, \hat{y}_i, \hat{x}'_i, \hat{y}'_i)^T$ satisfy the equation $\hat{x}_i^T F_A \hat{x}'_i = 0$ which may be expressed as $(\hat{X}_i^T, 1)f = 0$ (i.e. the equation of a point in \mathbb{R}^4 on the plane f) where $f = (a, b, c, d, e)^T$ (Hartley and Zisserman, 2004).

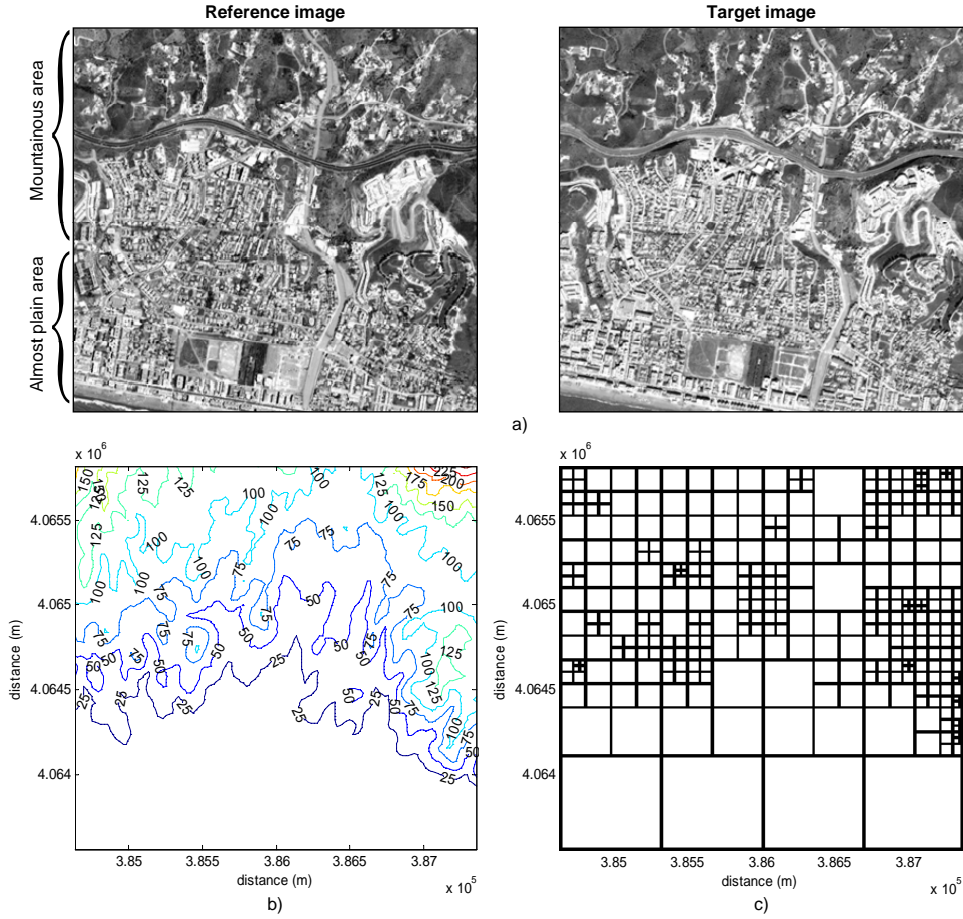


Figure 2: a) Two images of the coastal city of the Rincón de la Victoria (Málaga-Spain). b) Digital terrain model (DTM) of the region of interest. c) Quadtree decomposition of the DTM. Each parcel will contain a CP for posterior image registration.

differences are above 10 m. (i.e. $t = 10$). By doing so, high-relief areas, which provoke large image distortions, will be more intensively decomposed (figure 2-c). If a subdivision operation gives rise to four regions whose size is less than a square of 25 pixels of side (i.e. $s = 25$), it is rejected. This means that the smallest cell size, for this example, will be bigger than 25 pixels and smaller than 50.

Algorithm 1 Quadtree decomposition of the DTM.

```

1: //  $R_0$  contains the coordinates of the regions to be analyzed
2:  $R_0 \leftarrow \{\text{coord}(\text{DTM})\}$  //  $R_0$  is initialized with the
3: // coordinates of the DTM
4: //  $R$  will contain the coordinates of the final regions
5:  $R \leftarrow \emptyset$ 
6: for all  $r \in R_0$  do
7:    $v \leftarrow \text{DTM}(r)$  //  $r = \{x, y, \text{width}, \text{height}\}$ 
8:   if  $\text{size}(r) > s$  and  $(\max(v) - \min(v)) > t$  then
9:     // quad divides  $r$  into 4 equal pieces and returns
10:    // their coordinates
11:     $R_0 \leftarrow \{R_0 \cup \text{quad}(r)\}$ 
12:   else
13:     //  $r$  is not divided and it is stored in  $R$ 
14:      $R \leftarrow \{R \cup r\}$ 
15:   end if
16:   //  $r$  is removed from  $R_0$ 
17:    $R_0 \leftarrow \{R_0 - r\}$ 
18: end for

```

Finally, the selection of the final CP set is accomplished as fol-

lows: for each parcel of the decomposition, we check the number of detected CP pairs and, if this number is greater than one, we select the CP pair that exhibits the best score in the matching process, that is, the CP pair with the minor SSD value.

3 EXPERIMENTAL RESULTS

The benefits of the proposed method has been successfully verified by elastically registering a number of panchromatic (Orthoready) QuickBird image pairs (0.6 m./pixel), as the one shown in figure 2-a. The multitemporal series considered in our tests present significant relative geometric distortions induced by the off-nadir observation of no-planar regions as well as radiometric changes. The reader can found more details on satellite positioning data and the acquisition dates in (Arévalo and Gonzalez, 2008).

The registration process is accomplished by means of radial basis functions (RBF). Radial basis functions are scattered data interpolation methods where the spatial transformation is a linear combination of radially symmetric basis functions (second term of (3)), each of them centered on a particular CP, typically combined with a global affine transformation (first term of (3)). Mathematically

$$\begin{aligned}
 x &= \sum_{j=0}^m \sum_{k=0}^j a_{jk} (x')^{j-k} (y')^k + \sum_{j=1}^n A_j g(r_j) \\
 y &= \sum_{j=0}^m \sum_{k=0}^j b_{jk} (x')^{j-k} (y')^k + \sum_{j=1}^n B_j g(r_j)
 \end{aligned} \tag{3}$$

where

$$r_j = \|(x', y') - (\hat{x}'_j, \hat{y}'_j)\| \quad (4)$$

being $\hat{x}_j \longleftrightarrow \hat{x}'_j$ the refined CPs.

The type of basis function, g , determines the influence of each CP on the RBF, that is, the CP scope. So, the accuracy of the registration depends extremely on the distribution of CPs on the image. In this work we employ the thin plate spline (TPS) function $g(r_j) = r_j^2 \log r_j^2$ (Bookstein, 1989), which is perhaps the RBF most widely employed for elastic registration.

To evaluate the method performance, we have compared the registration accuracy obtained using the resultant CP set with respect to uniform and random CP distributions. The uniform distribution is obtained by selecting CPs according to a regular grid of squared cells of 50 pixels of side, while the random distribution is obtained by arbitrarily selecting the same number of CPs than the uniform one.

The results of the comparison, displayed in figure 3, show how the proposed method yields better results in terms of accuracy. The accuracy of the registration process has been assessed comparing the geometric errors (RMSE and CE90%) of a set of independent control points (ICPs) manually identified, achieving on average RMS errors under 1.4 m. with CPs distributed according to the DTM information.

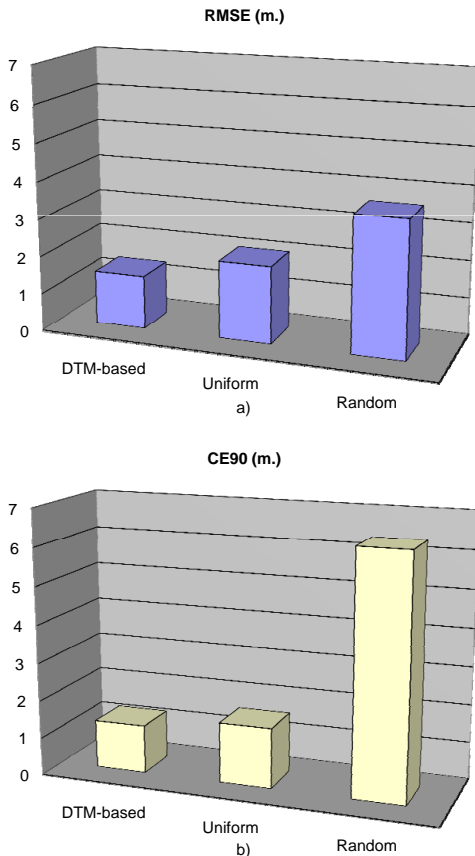


Figure 3: Accuracy of the proposed method compared to uniform and random distribution of CP considering a) RMSE and b) CE90%.

Observe that the results of the uniform distribution and our approach are similar, since the smallest squared cell generated by our approach has the same size that the one considered in the uniform distribution. The number of CP required in our approach, however, is, on average, around 37% lower. The benefits of our

approach are clear, specially, when the CP extraction must be manually performed.

4 CONCLUSIONS

This paper presents a technique to distribute the CP pairs according the relative image distortions, more severe in rugged terrains, and proposes an automatic procedure to robustly extract CPs in two images by applying computer vision techniques. The experimental results reveal the advantage of employing our method, in comparison with other two strategies (uniform and random distributions) implemented in most of popular commercial packages of remote sensing like ERDAS, ENVI and PCI.

ACKNOWLEDGEMENTS

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

This work has been partly supported by the Spanish Government under research contract CICYT DPI-2008-03527.

REFERENCES

- Arévalo, V. and Gonzalez, J., 2008. An Experimental Evaluation of Non-Rigid Registration Techniques on QuickBird Satellite Imagery. *International Journal of Remote Sensing* 29(2), pp. 513–527.
- Bookstein, F., 1989. Principal Warps: Thin-Plate-Splines and the Decomposition of Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(6), pp. 567–585.
- Burt, P. and Adelson, E., 1983. The Laplacian Pyramid as a Compact Image Code. *Communications, IEEE Transactions on [legacy, pre-1988]* 31(4), pp. 532–540.
- Fischler, M. and Bolles, R., 1981. Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography. *Communications of the ACM* 24, pp. 381–395.
- Fonseca, L. M. G. and Kenney, C. S., 1999. Control Point Assessment for Image Registration. In: *Proc. XII Brazilian Symposium on Computer Graphics and Image Processing*, pp. 125–132.
- Harris, C. and Stephens, M., 1988. A Combined Corner and Edge Detector. In: *Alvey Vision Conference*, Vol. 4, Manchester, UK, pp. 147–151.
- Hartley, R. and Zisserman, A., 2004. *Multiple View Geometry in Computer Vision*. Second edn, Cambridge University Press, Cambridge, UK.
- Lucas, B. and Kanade, T., 1981. An Iterative Image Registration Technique with an Application to Stereo Vision. In: *International Joint Conference on Artificial Intelligence*, pp. 674–679.
- Schowengerdt, R., 2007. *Remote sensing: Models and Methods for Image Processing*. Academic Press.
- Torr, P., 2002. *A Structure and Motion Toolkit in Matlab: Interactive Adventures in S and M*. Technical Report MSR-TR-2002-56, Microsoft Research.