Automated Geometric Correction of High-resolution Pushbroom Satellite Data

Marco Gianinetto and Marco Scaioni

Abstract

In this article, we present the use of the Automatic Ground control points Extraction technique (AGE) for increasing the automation in the geometric correction of high-resolution satellite imagery. The method is based on an image-to-image matching between the satellite data and an already geocoded image (i.e., a digital orthophoto). By using an adaptive least squares matching algorithm which implements a very robust outlier rejection technique, AGE can automatically measure many hundreds of topographic features (TFs) on the images, whose cartographic coordinates are derived from the geocoded image and elevations are extracted from an associated digital elevation model (DEM). The AGE technique has been tested for different high-resolution data: (a) 0.62-meter QuickBird panchromatic data (basic imagery processing level), (b) 2.5-meter SPOT-5/HRG panchromatic supermode data (standard 1B processing level), and (c) 1-meter Ikonos panchromatic data (standard Geo product processing level) collected in the Northern of Italy, both in flat (Torino Caselle test site) and mountain areas (Lecco test site). Regardless the relative image resolution between the satellite and the aerial data (1-meter) and regardless the processing level of the original satellite data, a similar TFs density has been obtained for both the QuickBird and the SPOT-5/HRG data (4.4 GCPs/km^2 and 4.1 GCPs/km^2) respectively, with a geometric accuracy for the GCPs extracted of 0.90 m for QuickBird and 3.90 m for SPOT-5/HRG. For the Ikonos imagery, AGE extracted a more dense set of GCPs (8.7 GCPs/km^2) but with a lower accuracy (3.19 m). The TFs identified with AGE can be used as GCPs for the rational polynomial coefficients (RPCs) computation and, therefore, for implementing a full automatic orthoimage generation procedure. By using the commercial off-the-shelf software PCI Geomatica® v.9.1, orthoimages have been generated for all datasets. The geometric accuracy was verified on a set of 30 manually measured independent check points (CPs) and assessed a precision of 4.99 m RMSE for QuickBird, 5.99 m RMSE for SPOT-5/HRG, and 8.65 m RMSE for Ikonos. The use of a non-conventional image orthorectification technique implementing a neural network GCPs regularization, tested for the SPOT-5/HRG data, showed the full potential of the AGE method, allowing to obtain a 3.83 m RMSE orthoprojection in a fully automated way.

Introduction

Geocoding is one of the basic image processing operations in remote sensing, and it is involved in nearly all applications (i.e., multi-temporal change detection, multi-sensor data fusion, cartographic updating). Unfortunately, this task is largely time expensive, requiring an expert operator for measuring on the images a large number of ground control points (GCPs) whose coordinates are known in the cartographic reference system from topographic field measurements (i.e., by GPS) or are derived from existing maps (see Smith and Atkinson, 2001 for a comparison between both approaches). Due to the increasing number of imaging sensors and to the great amount of data collected every day, the development of automatic techniques for the extraction and measurement of topographic features has become a very important issue in the remote sensing research community (Inglada and Giros, 2004).

Broadly speaking, methods for the automatic georeferencing can be classified into two main categories: (a) image-to-image, and (b) image-to-map. In the first case, the un-geocoded image is co-registered to an already georeferenced image by means of image matching algorithms, which are used to extract from the image pair a set of homologous GCPs with known cartographic coordinates (derived from the already geocoded data).

In the second case, the strategy is quite similar, but homologous GCPs must be extracted either from an image and a digital map (raster or vector), requiring matching algorithms which are able to run at a higher level of abstraction (i.e., by relational matching techniques (Vosselmann, 1992).

This kind of problem is already well known in digital photogrammetry, because in the last two decades image orientation procedures have been implemented in automatic way by means of matching algorithms (see Heipke, 1997 for a review about this subject). However, when looking for correspondences in satellite imagery, some differences occur with respect to the same task carried out considering aerial imagery. These are mainly due to the potential large differences in the image content introduced by: (a) the long time elapsed between the acquisitions, where changes in the landscape might easily happen (i.e., anthropic activities, vegetation changes, building construction), (b) the geometric deformations introduced in the data by the different acquisition mode (i.e., whiskbroom, pushbroom, asynchronous pushbroom) and by the viewing geometry (i.e., non-nadir), and, (c) the differences in the geometric resolution, scale,

Marco Gianinetto is with the Remote Sensing Laboratory, Department of Hydraulics, Environment Engineering, Infrastructures and Surveying (DIIAR), Politecnico di Milano University, Milano, Italy (marco.gianinetto@polimi.it).

Marco Scaioni is with the Department of Hydraulics, Environment Engineering, Infrastructures and Surveying (DIIAR), Politecnico di Milano University, Lecco, Italy (marco.scaioni@polimi.it).

Photogrammetric Engineering & Remote Sensing
Vol. 74, No. 1, January 2008, pp. 107–116.

0099-1112/08/7401–0107/$3.00/0

and radiometry (spectral bands and bandwidth used) when multi-sensor data are used.

On the contrary, photogrammetric data are captured during a limited number of missions carried out in a short period (for medium and small projects this is frequently restricted to 1 day) with a nadir-viewing geometry and without relevant changes in image scale and radiometry for adjacent frames.

Besides, the co-registration of satellite data to aerial images (scanned films or digital orthophotos) must take into account all the aspects before described (this is the most general case).

All these considerations lead to the conclusion that, unlike aerial imagery, the automatic image registration of satellite data has to cope with large variations both in the geometric properties (shifts, rotations, and scales) and in the image quality (radiometry and texture).

Background

Substantially, two different approaches exist to manage this task: (a) a combination of Feature Based Matching techniques (FBM) and Area Based Matching techniques (ABM), and (b) the use of a multi-resolution strategy, also referred to as image pyramids. In the first approach, initial values for the image registration are computed from a set of common features (points, lines, and patches) extracted from both the images with one or more FBM algorithms. Then, approximations are used to initialize the search for homologous GCPs by means of an ABM algorithm. Finally, the transformation from an image to another is estimated on the basis of the extracted GCPs.

Most of the algorithms described in literature rely on the use of a single FBM algorithm (Fonseca and Manjunath, 1996). In Dare and Dowman (2001) a method incorporating several algorithms is proposed, resulting in a significant growing of GCPs number that can be extracted from a multi-sensor image pair (tests have been carried out considering SPOT-5/HRG and ERS imagery).

The multi-resolution approach is often used in the search space and generally applies a coarse-to-fine approach in order to reduce the search area and to limit blunders (Corvi and Nicchiotti, 1995). Using this strategy, the image registration is first performed at the lower resolution level (small image size and large pixel size), and then is extended to the higher resolution levels (large image size and small pixel size), where only refinements of the previous steps are computed. Multi-resolution also enables to cope with images having large differences between them which may easily lead to mismatching.

Both FBM and ABM techniques may be integrated in a multi-resolution approach, where the latter play the prevalent role, especially at the final stage.

The most used ABM algorithm is the well-known normalized cross-correlation, which gives the measure of the similarity between a search window in the reference (master) image and its homologous position in the second (slave) image (see Kraus, 1997). Unfortunately in cross-correlation, only shifts might be estimated, resulting in the impossibility of computing the image registration if other geometric deformations are significantly present. An evolution of this algorithm is the Adaptive Least Squares Matching (ALSM) proposed by Grün (1985), which currently is the widely adopted ABM method in digital photogrammetry. ALSM allows to compensate for both geometric and radiometric transformations between the images, whose parameters are computed by a Least Squares (LS) approach (usually this algorithm is referred to as Least Squares Matching, LSM).

A more efficient and less computational expensive class of methods are the sequential similarity detection algorithms, which are based on an L1 norm parameter estimation between two images (Barnera and Silverman, 1972). Phase correlation techniques rely on the translation property of the Fourier transform (shift theorem), being able to compute translations and rotations between an image pair in the frequency domain (De Castro and Morandi, 1987). However, their use is not suitable for multi-sensor registration.

Finally, a quick look should be given to the methods used for the geometric registration. In Hanley and Fraser (2001) can be found a general overview about these techniques. The global point-mapping technique is the approach often used for the registration of images with unknown misalignments. It makes use of GCPs to compute the parameters of a polynomial transformation between the image reference systems through an approximation or interpolation (Ghoshtasby, 1988).

Local point-mapping methods can handle more distortions than global point-mapping techniques, because they can take into account also local distortions by piecewise interpolation, and they can compute a mapping transformation for each coordinate value (Ghoshtasby, 1986).

Non-linear registration algorithms can be used for iterative warping, using local finite element transformations which provide a non-linear global warp. This technique has been successfully applied to images with severe local distortions and different radiometry, where the cross-correlation approach would fail (Valdivieso-Casique and Arridge, 1999).

Research Objective

Although many methods have been developed to perform the automatic image registration, this task is still scarcely implemented into commercial software, and only recently, automatic image-to-image matching is being introduced in some remote sensing packages (i.e., RSI ENVI® v.4.2, PCI Geomatica® v.10.0, ERDAS Imagine® v.9.0).

In this article is presented a method for the automatic image matching of multi-sensor imagery which is based on the Automatic Ground control points Extraction (AGE) technique developed during the last years by the authors (Scaioni and Gianinetto, 2003; Gianinetto *et al.*, 2004; Gianinetto, 2005). Some basic algorithms widely used in the automatic procedures of digital photogrammetry (especially for relative orientation and aerial triangulation) have been properly optimized and adapted to be used also with satellite data and have been implemented in a user friendly software (GEOREF v.1.0) running under Microsoft® Windows® platform (Chirici *et al.*, 2004). A detailed description of the algorithms is reported in the next section.

One of the applications thought for the AGE technique is the automatic orthorectification of high-resolution satellite imagery by means the rational function model (RFM) transformation (Dowman and Tao, 2002; Fraser *et al.*, 2002; Tao *et al.*, 2004; Toutin, 2004), where the rational polynomial coefficients (RPCs) are generated using a set of GCPs automatically extracted both from the satellite image and from an aerial digital orthophoto (Scaioni and Gianinetto, 2003; Gianinetto *et al.*, 2004).

As described in previous works (Tao *et al.*, 2004; Chmiel *et al.*, 2004), many tests have shown that, if the RPCs have been properly computed, the RFM geometric model can approximate the rigorous sensor model without any distinguishable loss of accuracy. Besides, the most critical element in the RPCs computation is the strong dependence of the solution from the number and the geometric distribution of GCPs; moreover, when using higher-degree RFMs (typically third-order), over-parameterization may cause instability in the LS solution with the consequence of undesired local image deformations (Gianinetto *et al.*, 2004). Moreover,

when using RFM for orthoimage generation, there is nearly a linear relationship between the Root Mean Square Error (RMSE) of the orthorectified images and the number of GCPs used for computing the RFM (Shi and Shaker, 2003).

Because the RFM can be used as a terrain-independent solution, making it a safe replacement to the physical sensor model, all the satellite image vendors normally provide pre-determined rapid positioning coefficients (RPCs) when an image is purchased. Unfortunately, these pre-computed RPCs are not suitable for precise orthorectification, because they are derived from orientation data originating from the satellite ephemeris and star tracker observations without reference to GCPs (Poon *et al.*, 2005).

To overcome this problem, there are two methods that are widely used: (a) RPCs refinement, and (b) RPCs completely self-computed from measured GCPs. The first approach tries to refine the pre-determined RFM solution through the use of few GCPs (Tao *et al.*, 2004), but again results may be poor in upland areas and subsets or pre-processed images can't be treated.

The second approach is based on a self-computation of the RPCs through the use of many scattered measured GCPs. Unfortunately, unless a large number of GCPs with a regular planimetric and altimetric distribution is used, this approach may lead to an inaccurate solution (Tao *et al.*, 2004), while poorly designed GCPs configurations may easily lead to heavy image distortions (asymptotes) in the orthoimages (Gianinetto *et al.*, 2004), as previously recalled.

In the last several years, many authors have also proposed methodologies for deriving the RPCs directly from the satellite's orbital parameters and attitude data (i.e., Hanley and Fraser, 2001; Dial and Grodecki, 2002; Jacobsen and Passini, 2003), thus limiting the demand of GCPs needed for the orthorectification. But again, the main disadvantage of these approaches lies in the need of the satellite's precise ephemeredes and ancillary data.

The main advantages of the methodology proposed can be summarized in: (a) all types of data can be processed (including all product levels and also image subsets), (b) no pre-determined RPCs are needed, and also, data without RPCs can be processed and, (c) no metadata are needed (i.e., satellite's ephemeredes and sensor attitude). In addition, this methodology proposes a new use for the (old) aerial imagery widely available in all developed countries.

In the Case Studies section, some experimental tests are presented. Using as reference data 1-meter resolution digital orthophotos and gridded digital elevation models (DEMs) with a 50-meter sampling distance, three different high-resolution satellite images have been processed (QuickBird panchromatic, Ikonos panchromatic, and SPOT-5/HRG panchromatic supermode). Hundreds of topographic features (TFs) have been automatically determined with the AGE technique, RPCs have been computed, and the corresponding orthoimages have been automatically generated. Finally, results have been assessed on the basis of a manual check.

Concept and Implementation of the AGE Strategy

The AGE algorithm is based on three main computational steps: (a) extraction of interest points (IPs) by means of an interest operator, (b) measurement of corresponding IPs by means of LSM, and (c) robust estimation of GCP dataset by means of a Least Median Squares outlier rejection technique implementing an affine model to map points from an image to another.

Before applying the AGE procedure, a data pre-processing stage is needed. When using panchromatic data, the only task to perform is a radiometric resampling to 8-bit per pixel resolution. This is the radiometric resolution usually adopted in applications of digital photogrammetry, and a higher bit per pixel resolution would not improve the final results.

When processing color images (multispectral or hyperspectral data), a single spectral band or a combination of more bands should be selected in order to derive unique 8-bit per pixel images. The criteria to be adopted is to select spectral bands similar to each others for all the images.

Interest Point Extraction

Interest Points (IPs) are features with special characteristics that can be easily recognized, even in an automatic way. In practice, they are defined as geometrically stable points that appear when high intensity differences occur in more than one direction (Stylianidis, 2003).

Among all interest operators, the Harris (Harris and Stephens, 1988) and the Förstner (1985) operators are the most well known and widely used in operational packages. The Förstner operator is based on the evaluation of the quality of corner points, by analyzing the shape and the size of the error ellipse describing the covariance matrix associated to the corner point location. Reliable corner points should have a near circular error ellipse with a small size (Habib *et al.*, 2003). This is the strategy adopted in the AGE technique, motivated by the possibility of selecting only IPs with the best characteristics among all those extracted to reduce the risk of mismatching due to the use of poorly defined points.

The theory of Förstner operator is currently a basic knowledge in the photogrammetric community, so it is omitted here. On the other hand, some addresses on its use deserve to be reported. The operator is applied to all the images to be co-registered by computing two parameters for every squared window of fixed side, l_{FW}, through the image. Due to the fact that the selection of l_{FW} will directly affect the final number of extracted IPs, according to the image texture, an auto-adaptive procedure has been implemented. A first attempt is tried with a large l_{FW}, and the extracted point density is compared to a minimum threshold. If this condition is not satisfied, then a further attempt with a smaller l_{FW} is carried out as far as the threshold is satisfied.

IP discrimination is performed on the basis of the Förstner operator parameters: roundness (q) and interest value (w). For all IPs, q and w are computed and: (a) the roundness is verified to be higher than 0.95, meaning that the shape of the error ellipse is quite circular and contrasted features are not unidirectional and, (b) the interest parameter is analyzed, considering that its value depends on the image texture, and no criteria exist for its normalization. However, the higher the value of w the better its quality, so all IPs are ranked on the basis of the w value, and only the first n of them are kept (Forlani *et al.*, 1996).

It is well known from Förstner (1985), that once a window has been selected as containing an IP, its center is not the most representative point of the TF. This is usually computed as the gravity center of window's digital numbers (DNs) resulting in the possibility that two adjacent windows could yield a pair of IPs too close to each other. To avoid this problem, a minimum distance between IPs is to be defined and checked.

Least Squares Matching

Least Squares Matching (LSM) is an image correlation algorithm which is able to map a starting window (template) extracted around a point in the master image to a search window extracted around the approximate position of the homologous point in the slave image (see Baltsavias, 1991 and Grün, 1996 for further theoretical and implementation details). To compensate for differences between them, several geometric and radiometric functions may be implemented, whose parameters are estimated by an iterative LS approach.

Geometric Model

The geometric model implemented in the LSM performed by AGE is an affine transformation, where any of the six parameters may be constrained adjusting the model to a simpler one based on the image content (i.e., shifts only, 2D conformal transformation, etc.). The first three iterations are performed by using only shifts, in order to get a rough localization of the shapes to match and to avoid the initial convergence to a false position. When the affine model turns over the shifts model, an automatic testing of the parameters' significance is performed in the iteration process to update the model.

To stop iterations, the relative change of sigma naught, of the correlation coefficient and of the shifts are used. Furthermore, an upper bound for the number of iterations is adopted for limiting the computation time.

Concerning the technique for constraining the shaping parameters, AGE uses a combination of two tests that are applied just after the normal matrix **N** of LS has been built.

The first test gives information about the determinability of shifts, by evaluating the shape of their error ellipse. The classic statistical test used to check the well determinability of parameters in LS is based on the evaluation of relative size and ratio of their estimated standard deviations, requiring the inversion of the normal matrix **N**. Baltsavias (1991) suggest a fast approximated approach that can be obtained by using the following simpler quantities:

$$ratio = \sqrt{\frac{max(n_{rr}, n_{cc})}{min(n_{rr}, n_{cc})}} \qquad \begin{cases} G_r = \sqrt{\frac{n_{rr}}{m}} \\ G_c = \sqrt{\frac{n_{cc}}{m}} \end{cases} \tag{1}$$

where n_{rr} and n_{cc} are the diagonal elements of the normal matrix **N** pertaining to shifts (in row and column direction, respectively), and m is the pixel number of the window. Two thresholds are needed: (a) the upper limit for the ratio (t_1), and (b) the lower limit for G_r and G_c (t_3). The logical scheme of the test is the following:

```
IF (ratio > t1) THEN
  IF (nrr > ncc) THEN
    constrain column shaping parameters
  ELSE
    constrain row shaping parameters
  ENDIF
ENDIF
IF (min(Gr,Gc) ≤ 0.5*t3) THEN
  constrain all the shaping parameters
ENDIF.
```

After several empirical trials, optimum thresholds have been found as $t_1 = 1.2$ and $t_3 = 8.0$. The second test deals with the analysis of correlations between shaping parameters. High correlations between two parameters mean their non-determinability. Of great importance are the correlations of each of the shaping parameters with the shifts, between similar shaping parameters (i.e., two scales, two shears), and between shaping parameters in the same direction. If correlation exceeds an *a priori* fixed threshold (i.e., 0.8), one of the correlating parameters must be excluded, and the test repeated until all correlations are below the threshold. Obviously, shifts can never be excluded. In case a parameter has a high correlation with a shift, it must be checked if this happens because that parameter has a high correlation with another parameter which is also highly correlated with the shift itself.

Radiometric Model

To improve the fitting between the template and the slave windows, these are radiometrically equalized previous the computation of the geometric matching. A Wallis filter (Wallis, 1976) is used to force the equalization of the mean value and the standard deviation of the DNs by applying a locally-adaptive contrast enhancement to a grayscale raster image. This kind of filter is just designed for images in which there are significant areas of bright and dark tones, for which a typical global contrast enhancement cannot simultaneously produce good local contrast at both ends of the brightness range. On the contrary, the Wallis filter adjusts brightness values in local areas so that the first-order (local mean) and the second-order (standard deviation) moments match specified values.

The enhanced image is increased in amplitude with respect to the original at pixels that deviate significantly from their neighbors and decreased in relative amplitude elsewhere. The parameters of the radiometric transformation are not included in the LS adjustment, but they are computed prior the matching to avoid the possibility of overparameterizing the model itself.

Strategy for Point Transfer

The Förstner operator extracts two sets of IPs from both the master and slave images. Before applying the LSM algorithm for detecting the homologous points, a preliminary search for possible candidates is carried out. This is done by computing a rough approximate position of the corresponding point on the slave image for each IP found on the master image. Around this position a squared search window is drawn, and all IPs falling inside are tried for matching. The point showing the higher correlation after LSM is selected as homologous, according however to a minimum threshold to reduce possible mismatching (usually fixed at 0.7).

An affine transformation to define the rough approximate point is implemented, whose parameters are computed by manually selecting at least three corresponding points in the image pair. An alternative is to use a multi-resolution strategy (image pyramids), which gives good results especially in case of mountain areas, but however would require the measurement of some corresponding points. In flat areas, the use of manually measured starting points reduces the total computation time and increases the number of homologous points. Moreover, the implementation of this task in the GEOREF software has resulted as highly operational, even more than the use of image pyramids.

Typically, the LSM algorithm yields a large number of outliers in the set of corresponding extracted TFs, which may overcome the 20 to 30 percent of the total amount. This fraction would prevent a correct estimation of the final transformation by a classical robust method (i.e., L1 or *M* estimators), so a very high breakdown-point estimator (Least Median Squares: LMEDS) is applied in order to make a preliminary cleaning of large outliers from the dataset to be further processed by a L1 estimator.

Robust Estimation of the Geometric Transformation Mapping Images

The basic concept of the LMEDS estimator is opposite to the conventional smoothing techniques (Rousseeuw and Leroy, 1987). Rather than using as much data as possible to obtain an initial solution and then attempting to identify outliers, a small subset made up of the minimum number of data necessary to compute the parameters is used. This process is repeated on different subsets to ensure that there is a fixed chance that one of the subsets will contain only good data points. The best solution is one that minimizes (or maximizes) a given criteria.

Theoretically, every possible sub-sample of data could be considered, but this is usually computationally infeasible (except in case of simple regressions). Therefore, the number of sub-samples m_s is chosen sufficiently high to give a probability P (i.e., $P = 0.95$) that at least a good sub-sample is selected. In the AGE procedure, it has been adopted the formulation proposed by Rousseeuw and Leroy (1987) for computing m_s:

$$m_s > \frac{log(1 - P)}{log(1 - (1 - \varepsilon)^s)} \tag{2}$$

where ε is the fraction of outliers (to be guessed), and S is the number of data in each sub-sample.

The solution obtained is used to reject large outliers, requiring also a robust estimation of the standard deviation of errors (σ_0). This is achieved by starting from a first value computed from the median squared error of the chosen parameter fit, multiplied by a finite-sample correction factor for the case of normal errors:

$$\sigma_0 = 1.4826\left(1 + \frac{5}{n - p}\right)\sqrt{med\ d_i^2} \tag{3}$$

where n and p are respectively the number of data points and parameters, and d_i is the distance of the estimated position of the i^{th} point from its observed position.

This preliminary estimation of σ_0 is then used to determine a weight w_i for the i^{th} observation:

$$w_i = \begin{cases} 1 & if\ |d_i| \le 2.5\sigma_0 \\ 0 & otherwise \end{cases} \tag{4}$$

that is used for the final computation σ_0^*, which is no more influenced by outliers:

$$\sigma_0^* = \sqrt{\frac{\left(\sum_{i=1}^{n} w_i d_i^2\right)}{\left(\sum_{i=1}^{n} w_i - p\right)}}. \tag{5}$$

Finally, outlier rejection can be carried out discarding all those points whose distance d_i is more than k times σ_0^* (usually $k = 1.96$).

When estimating the solution for each sub-sample, this might be made up of points that are very close to each other, resulting in an instable configuration. For this reason, AGE adopts the "bucketing" technique proposed by Zhang *et al.* (1995). After calculating the minimum and maximum of the point coordinates in the first image, the rectangular region inside these values is evenly divided into $n_b = a \cdot b$ zones called "buckets", each of them collecting all points falling in it. The size of the parameters a and b is set so that all the interest area is split into 20 to 30 buckets. To generate a sub-sample, each element is randomly selected from a different bucket among those with at least one point inside.

Case Studies

Tests involving 0.62-meter panchromatic QuickBird data, 1-meter panchromatic Ikonos data, and 2.5-meter panchromatic supermode SPOT-5/HRG data have been performed over two different sites in the Northern of Italy: (a) QuickBird and SPOT-5/HRG images have been collected over Caselle Torinese, Piemonte (dataset 1), and (b) the Ikonos image has been taken over Lecco, Lombardia (dataset 2). Both datasets derived from a national research projects (COFIN, 2001) aimed to study the use of high-resolution satellite data for mapping and land management.

Dataset Description

The test site 1 is located in the area of the Caselle Torinese International Airport. The following data were collected:

1. digital aerial orthoimages with 1 m ground resolution;
2. one 16 km × 17 km QuickBird panchromatic image with 0.62 m ground resolution, collected on 22 February 2002, 1032 GMT (processing level: basic imagery);
3. one 68 km × 60 km SPOT-5/HRG panchromatic supermode image with 2.5 m ground resolution, collected on 01 October 2002, 1020 GMT (processing level: standard 1B);
4. a 50 m grid DEM arranged in the Roma40 Italian geodetic datum and expressed in the Gauss-Boaga cartographic coordinate system.

The test site 2 is located in the surroundings of the city of Lecco. The following data were collected:

1. digital aerial orthoimages with 1 m ground resolution;
2. one 7 km × 24 km Ikonos panchromatic image with 1 m ground resolution, collected on 06 September 2001, 1030 GMT (processing level: standard Geo product);
3. a 50 m grid DEM arranged in the ED50 geodetic datum and expressed in the UTM cartographic coordinate system.

Table 1 shows a summary of the dataset characteristics.

Data Processing

For the dataset 1 (Caselle Torinese Airport), an area overlapping the aerial images has been selected from both the SPOT-5/HRG (166 km^2) and from the QuickBird data (178 km^2). The test site is relatively flat with an elevation ranging from 200 m to 400 m, and presents a mixture of urban areas and rural fields with a regular spatial distribution.

A smaller sample (22 km^2) has been selected from the dataset 2 (near the city of Lecco). In this case, the area is mostly mountain, with an elevation ranging from about 200 m up to 1,300 m in the northeast of the image.

As described in the first section, the 1 meter aerial orthophotos have been used as reference (geocoded) data, while onto the satellites data a set of homologous GCPs has been extracted by means of the AGE technique. Prior the extraction, for each image pair (aerial versus QuickBird, aerial versus SPOT-5/HRG, and aerial versus Ikonos), three homologous GCPs have been interactively selected onto the images for the computation of the initial approximate

TABLE 1. SUMMARY OF THE DATA CHARACTERISTICS

Test Site	Satellite	Pixel Resolution (m)	In-track View Angle (deg)	Cross-track View Angle (deg)	Processing Level
Caselle Torinese (dataset 1)	QuickBird	0.62	−5.47318	−9.18762	Basic imagery
	SPOT-5/HRG	2.50	n.a.	−6.37640	Standard 1B
Lecco test (dataset 2)	Ikonos	1.00	80.75540(1)	272.66750(2)	Standard Geo product

(1) Nominal Collection Elevation
(2) Nominal Collection Azimuth

transformation between the image reference systems (see the Strategy for Point Transfer section).

For maximizing the number of GCPs collected, the AGE parameters have been tuned using as test data the lower resolution SPOT-5/HRG imagery (Table 2). The strategy adopted has been to identify as much as possible IPs in high gradient areas, which mainly correspond to urban areas (i.e., buildings or crossroads). This task has been obtained by using a threshold for the circularity index ($q > 0.90$) and by selecting only the IPs featuring the highest interest value. Regarding the LSM between the satellite and the aerial data, the best results have been achieved by using a searching window size of 7 pixels × 7 pixels, a matching window size of 5 pixels × 5 pixels, and a threshold for the correlation of 80 percent.

For each identified GCP, the northing and easting coordinates have been automatically derived from the corresponding homologous points onto the orthophotos, while the elevations have been automatically extracted from the associated DEM.

Finally, after assessing their precision by means of a manual measurement of 30 Independent Check Points (ICPs), RPCs, and orthoimages have been generated.

TABLE 2. SUMMARY OF THE AGE PARAMETERS USED DURING THE TESTS

Operation	Parameter	Value
Förstner operator	Window size (pixel)	5 × 5
	circularity index	0.9
	Minimum point density (IPs/pixels^2)	2
LSM	Searching window size (pixel)	7 × 7
	Matching window size (pixel)	5 × 5
	Correlation coefficient	>0.80
	Maximum iterations (nr.)	<50
	Maximum homologous candidates for each IP (nr.)	<8
Outlier rejection	Geometric model	Affine

Analysis of Results

Topographic Features Extraction

On all datasets, AGE extracted a large number of GCPs: (a) 775 GCPs for the QuickBird image (Figure 1), (b) 677 GCPs for the SPOT-5/HRG image (Figure 2), and (c) 191 GCPs for the Ikonos image (Figure 3).

In order to evaluate their geometric accuracy, 30 of the GCPs have been randomly selected and manually measured. For the dataset 1 the manual check assessed: (a) a total RMSE of 0.90 m (1.45 pixel) with a mean RMSE of 1.27 m (2.05 pixel) and a standard deviation of 0.69 m (1.12 pixel) for the QuickBird data, and (b) a total RMSE of 3.90 m (1.56 pixel) with a mean RMSE of 5.85 m (2.34 pixel) and a standard deviation of 3.08 m (1.23 pixel) for the SPOT-5/HRG data. For the dataset 2, the manual check assessed a total RMSE of 3.19 m (3.19 pixel) with a mean RMSE of 2.37 m (2.37 pixel) and a standard deviation of 2.09 m (2.09 pixel) for the Ikonos data. A summary of results is reported in Table 3.

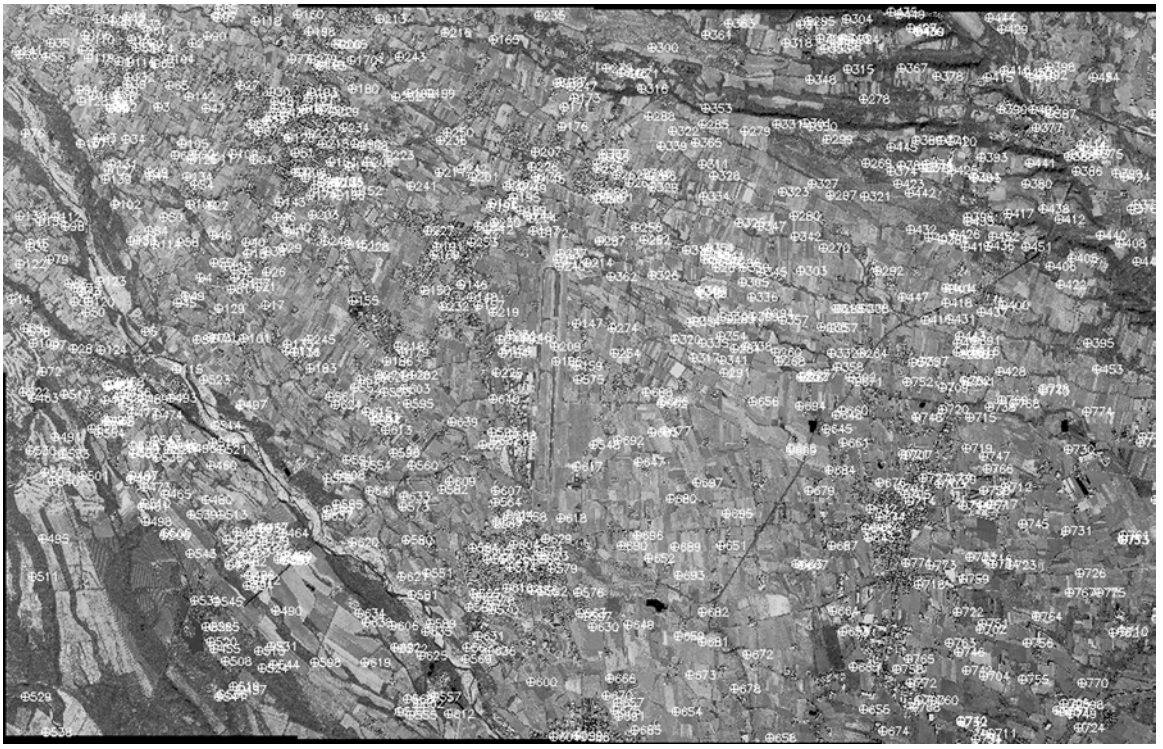

Figure 1. QuickBird image (0.62-meters ground resolution) acquired over the Caselle Torinese study area (dataset 1) with superimposed the 775 GCPs extracted with AGE.

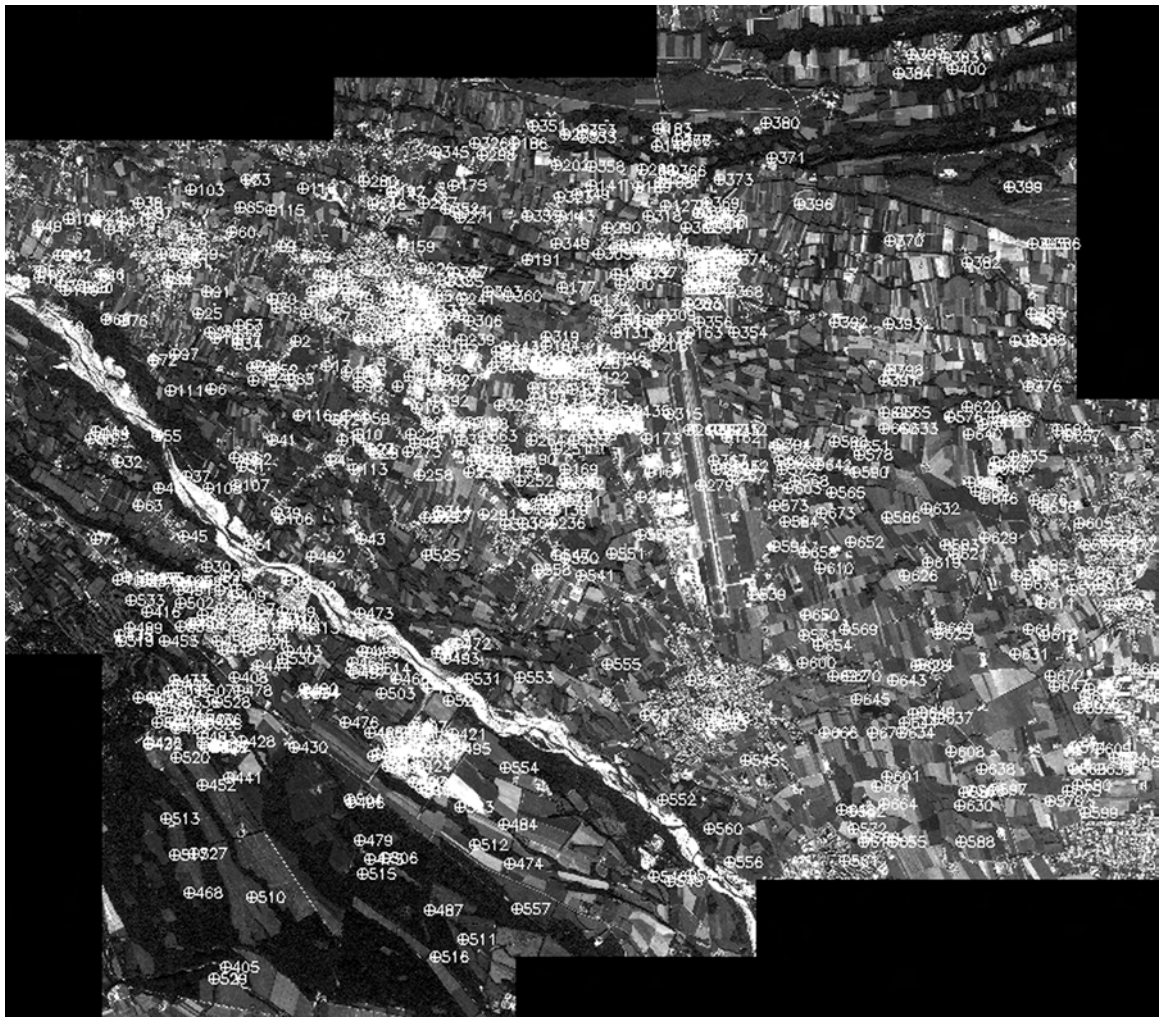

Figure 2. SPOT-5/HRG supermode image (2.5-meters ground resolution) acquired over the Caselle Torinese study area (dataset 1) with superimposed the 677 GCPs extracted with AGE.

TABLE 3. ACCURACY ASSESSMENT OF THE TOPOGRAPHIC FEATURES EXTRACTED WITH AGE BY MEANS OF MANUAL MEASUREMENT OF 30 INDEPENDENT CHECKPOINTS

Test Site	Satellite	GCPs (nr.)	Total RMSE (pixel)	Total RMSE (m)
Caselle Torinese (dataset 1)	QuickBird	775	1.45	0.90
	SPOT-5/HRG	667	1.56	3.90
Lecco test (dataset 2)	Ikonos	191	3.19	3.19

Standard Orthorectification

Standard image orthorectification has been performed by a commercial off-the-shelf software (COTS). Using the RFM algorithm implemented into PCI Geomatica® v.9.1 and computing the RPCs on the basis of the GCPs collected with AGE, orthoimages have been generated for all dataset. A set of 30 ICPs has been located on easily recognizable and well-contrasted image features, such as crossroads or building corners, and have been manually measured. Moreover, to get a realistic metric valuation, ICPs have been selected with a regular geometrical disposition in order to assess the quality of the whole images.

Regarding dataset 1, the QuickBird and SPOT-5/HRG orthoimages showed a similar total RMSE of 4.99 m and 5.99 m, respectively; while for dataset 2, the Ikonos orthoimage showed a greater total RMSE of 8.65 m. A summary of results is shown in Table 4.

Non-conventional Orthorectification

For the SPOT-5/HRG, a non-conventional image orthorectification procedure recently developed by the authors has been applied. In this case, all GCPs coming from the AGE stage (having a random geometric distribution) have been additionally processed. Instead of directly using them for RPCs estimation, they have been used to train a Multi-Layer Perceptron Neural Network (MLP-NN) whose goal is the computation of a new set of GCPs regularly gridded over the whole image. Using these gridded GCPs, RPCs were computed with a non-commercial RFM algorithm developed by the authors. Past experience demonstrate that this approach produces better results in comparison with standard RFM image orthorectification (for further details, see Gianinetto *et al.*, 2004).

This processing technique produced an accuracy increase of 36 percent in the SPOT-5/HRG orthoprojection, with a total RMSE of 3.90 m (1.56 pixel) on GCPs and 3.83 m (1.53 pixel) on ICPs.

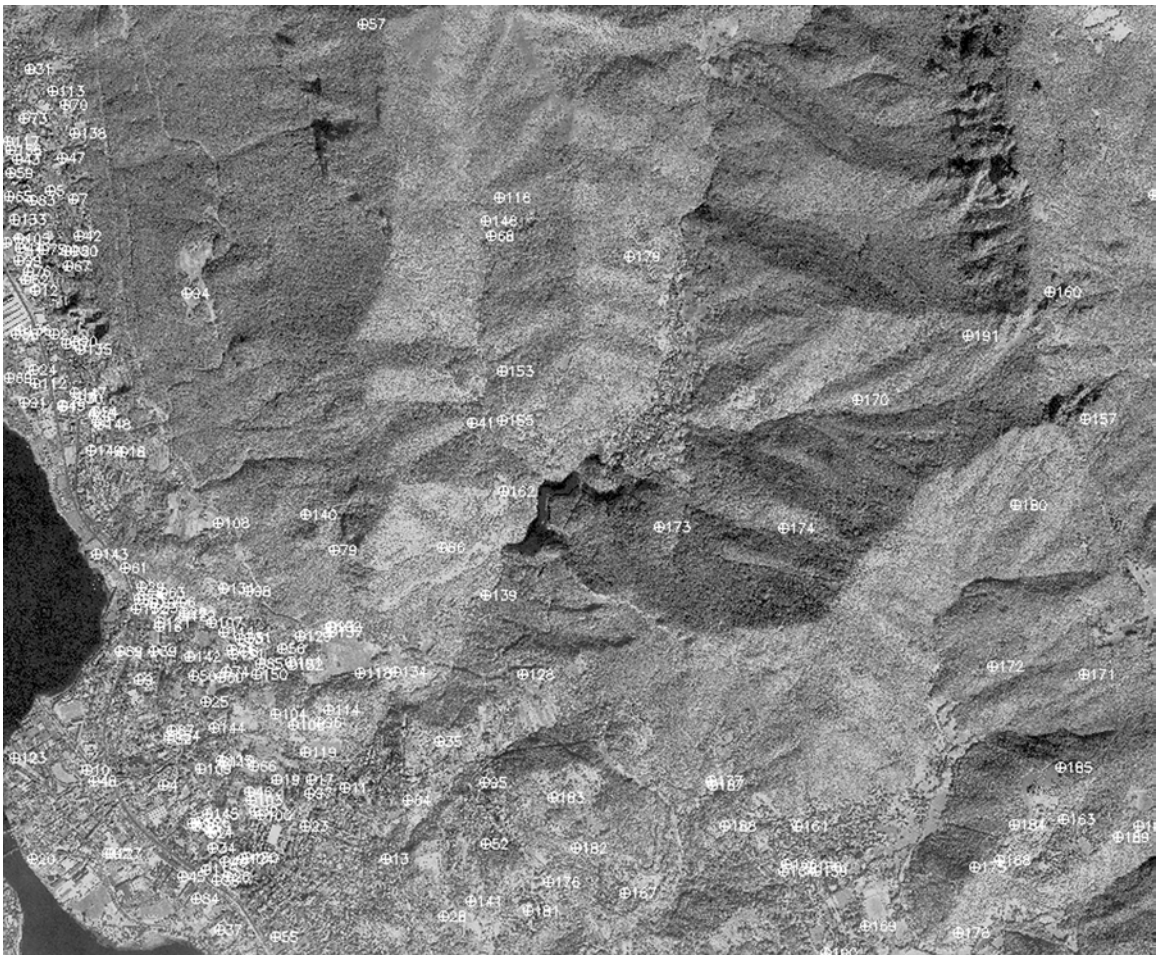

Figure 3. Ikonos image (1-meter ground resolution) acquired over the Lecco study area (dataset 2) with superimposed the 191 GCPs extracted with AGE.

TABLE 4. ACCURACY ASSESSMENT OF THE ORTHOIMAGES GENERATED WITH THE COMMERCIAL SOFTWARE PCI GEOMATICA BY MEANS OF MANUAL MEASUREMENT OF 30 INDEPENDENT CHECK POINTS

Test Site	Satellite	GCPs Density (GCPs/km^2)	Elevation Range (m)	Orthorectification Processing	Mean ΔE (m)	Mean ΔN (m)	RMSE (m)
Caselle Torinese (dataset 1)	QuickBird	4.4	200–400	RFM, PCI Geomatica v.91	+2.81	−2.80	4.99
	SPOT-5/HRG	4.1	200–400	RFM, PCI Geomatica v.91	+0.43	−4.01	5.99
	SPOT-5/HRG	4.1	200–400	Non-conventional (MLP-NN regularization)	+1.51	−1.01	3.83
Lecco (dataset 2)	Ikonos	8.7	200–1,300	RFM, PCI Geomatica v.91	+2.73	−2.50	8.65

Discussion and Conclusions

This study analyzed the accuracy of automatic GCPs extraction performed by means of the AGE (Automatic Ground control points Extraction technique) algorithm developed and implemented by authors. Some tests have been carried out for QuickBird, Ikonos, and SPOT-5/HRG data, either in flat or mountain areas. As reference data, 1-meter digital orthoimages have been used for the planimetric coordinate estimation, while 50-meter digital elevation models for the elevation.

Using the extracted GCPs, image orthorectification have been performed for all datasets with the commercial off-the-shelf software PCI Geomatica® v. 9.1 (RFM model). In addition, for the SPOT-5/HRG a non-conventional image orthorectification method recently developed by authors has been also tested.

Results show that without any knowledge of the camera model, without ephemeredes data, without using pre-computed RPCs, and without man-measured GCPs, the

AGE technique is able to extract a large number of GCPs for all dataset. Regardless, the relative image resolution between the satellite and the aerial data and regardless of the processing level of the original satellite data (basic imagery for QuickBird and standard 1B for SPOT-5/HRG), a similar GCP density has been obtained for both the QuickBird (4.4 GCPs/km^2) and the SPOT-5/HRG (4.1 GCPs/km^2) images, whereas for the Ikonos data (standard Geo product) AGE extracted a more dense set of GCPs (8.7 GCPs/km^2) but with a lower accuracy.

With reference to the GCP accuracy, the extraction process has been influenced by the relative viewing angles of the aerial and the satellite images. In spite of its lower geometric resolution, the smaller tilt angle of the SPOT-5/HRG configuration allowed to obtain interesting results in terms of image pixels ($RMSE_{SPOT\text{-}5/HRG}$ = 1.56 pixel), similar to those obtained with the higher-resolution QuickBird data ($RMSE_{QuickBird}$ = 1.45 pixel).

The impact of topography (elevations ranging from 200 m to 1,300 m) and the low image quality of the aerial data taken over the Lecco test site are mainly responsible for the lower quality of the Ikonos GCPs ($RMSE_{IKONOS}$ = 3.19 pixel).

Regarding the automatic image rectification, both the QuickBird and the SPOT-5/HRG standard orthorectified images show similar residuals ($RMSE_{QuickBird}$ = 4.99 m, $RMSE_{SPOT\text{-}5/HRG}$ = 5.99 m), while the Ikonos standard orthorectified image is afflicted by larger and widely spread errors ($RMSE_{IKONOS}$ = 8.65 m), both due to the wider elevation range of the study area, which reflects in a greater planimetric displacement (valuable in about 2 m) and to the non-uniform geometric distribution of GCPs (no GCPs were identified inside the wood in the northeast portion of the image).

This last aspect is the most critical element of the RFM method, because its solution has a strong dependence from the number and distribution of GCPs. Bad configurations, such as irregularly distributed GCPs on the whole image can easily lead to significant errors in the corrected image. This problem can be overcome by using non-conventional orthorectification techniques implementing GCP regularization (Gianinetto *et al.*, 2004). For the SPOT-5/HRG data, the use of a non-conventional orthorectification procedure (MLP-NN GCPs regularization followed by RPCs computation) showed an increased accuracy of image orthorectification from 5.99 m to 3.83 m RMSE (2.99 m in East and 2.39 m in North), corresponding to 1.53 pixel, on ICPs (Table 4).

To evaluate the potential of the method described, results are compared to those obtained by other authors. Regarding Ikonos orthorectification, for flat areas (Ontario, Canada) and using stereo images, Tao *et al.* (2004) assessed a RMSE of 7.54 m using RFM with pre-determined RPCs and a RMSE of 3.62 m using refined RPCs. For hilly areas in China, Wang and Ellis (2005) found a planimetric RMSE in the range of 0.90 m to 2.59 m using polynomial models and height compensation. For QuickBird stereo pairs, Noguchi and Fraser (2004) assessed an orthorectification accuracy in the range of 0.92 m (0.46 m in East and 0.80 m in North) to 1.12 m (0.77 m in East and 0.82 m in North) for RPC bundle adjustment model with bias compensation and 4.27 m (4.0 m in East and 1.5 m in North) using the affine bundle adjustment model.

Regarding SPOT-5/HRG, using stereo images and the Toutin's satellite orbital model (Toutin, 2004), Buyuksalih *et al.* (2005) found RMSE values in the orthoimages in the range of 3.74 m to 4.26 m (2.53 m to 3.23 m in East and 2.75 m to 2.77 m in North). Using automated orientation techniques and a rigorous sensor model developed for a SPOT-5/HRG stereo pair, in the framework of the HRS Scientific Assessment Program (HRS-SAP) organized by CNES and ISPRS, Poli *et al.* (2004) assessed a RMSE of 7.07 m (5.9 m RMSE in east, and 3.9 m RMSE in north). An alternative test based on the RPC model approach (Zhang *et al.*, 2004) assessed a RMSE of 6.88 m (5.6 m RMSE in east and 4.0 m RMSE in north).

The AGE technique here described proved to be an interesting method for the automatic GCPs extraction from aerial versus high-resolution satellite images and for implementing a complete automated procedure of image orthorectification when digital orthoimages are already available. In particular, the use of image orthorectification implementing a neural network GCPs regularization, tested for the SPOT-5/HRG supermode data, showed the capability of full image exploitation, allowing to obtain precise planimetric geopositioning in a full automated way, even better than those obtained with rigorous sensor models.

Acknowledgments

This work has been partially funded by the Italian Ministry for University and Research, contract title: "Nuovi metodi per l'analisi di deformazioni, tecniche innovative per la georeferenziazione di immagini ad alta risoluzione e uso di immagini satellitari multisensore e multirisoluzione per il controllo dei danni causati dalle alluvioni." QuickBird, SPOT-5/HRG, and Ikonos data have been provided under the research framework founded by the Italian Ministry for University and Research, contract title: "L'uso delle immagini satellitari ad alta risoluzione per le analisi territoriali." Aerial data have been provided by Politecnico di Torino University (Italy).

We gratefully acknowledge the help of Enrico Borgogno Mondino and Fabio Giulio Tonolo of the Politecnico di Torino University (Italy) in the SPOT-5/HRG data processing.

References

Baltsavias, E.P., 1991. *Geometrically Constrained Multiphoto Matching*, Mitteilungen 49, Institute of Geodesy and Photogrammetry, ETH, Zurich, Switzerland.

Barnera, D.I., and H.F. Silverman, 1972. A class of algorithms for fast digital image registration, *IEEE Transactions on Computers*, 2(C-21):179–186.

Buyuksalih, G., G. Kocak, H. Topan, M. Oruç, and A. Marangoz, 2005. Spot revisited: Accuracy assessment, DEM generation and validation from stereo SPOT 5 HRG images, *The Photogrammetric Record*, 20(110):130–146.

Chmiel, J., S. Kay, and P. Spruyt, 2004. Orthorectification and geometric quality assessment of very high spatial resolution satellite imagery for common agricultural policy purposes, *The International Archive of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, Turkey, 35(B4):1019–1024.

Chirici, G., M. Gianinetto, and M. Scaioni, 2004. Experiences in upgrading of large databases of satellite images, *The International Archive of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, Turkey, 35(B4):382–387.

Corvi, M., and G. Nicchiotti, 1995. Multiresolution image registration, *Proceedings of the IEEE International Conference on Image Processing*, Washington D.C., 3:3224–3234.

Dare, P., and I. Dowman, 2001. An improved method for automatic feature-based registration of SAR and SPOT images, *ISPRS Journal of Photogrammetry and Remote Sensing*, 56:13–28.

De Castro, E., and C. Morandi, 1987. Registration of translated and rotated images using finite Fourier transforms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(5):700–703.

Dial, G., and J. Grodecki, 2002. IKONOS accuracy without ground control, *Proceedings of ISPRS Commission I Mid-Term Symposium*, Denver, Colorado, pp. 10–15.

Dowman, I., and V. Tao, 2002. An update on the use of rational functions for photogrammetric restitution, *ISPRS Journal of Photogrammetry and Remote Sensing*, 7(3):26–29.

Fonseca, L., and B. Manjunath, 1996. Registration techniques for multisensor remotely sensed imagery, *Photogrammetric Engineering & Remote Sensing*, 62(9):1049–1056.

Forlani, G., A. Giussani, M. Scaioni, and G. Vassena, 1996. Target detection and epipolar geometry for image orientation in close-range photogrammetry, *The International Archive of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vienna, Austria, 31(B5):518–523.

Förstner, W., 1985. A feature based correspondence algorithm for image matching, *The International Archive of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Rovaniemi, Finland, 26(3/3):150–166.

Fraser, C.S., E. Baltsavias, and A. Gruen, 2002. Processing of IKONOS imagery for submetre 3D positioning and building extraction, *ISPRS Journal of Photogrammetry and Remote Sensing*, 56(3):177–194.

Ghoshtasby, A., 1986. Piecewise linear mapping functions for image registration, *Pattern Recognition*, 19(6):459–466.

Ghoshtasby, A., 1988. Image registration by local approximation, *Image Vision Computing*, 6(4):255–261.

Gianinetto, M., M. Scaioni, E. Borgogno Mondino, and F. Giulio Tonolo, 2004. Satellite images geometric correction based on non-parametric algorithms and self-extracted GCPs, *2004 IEEE International Geoscience and Remote Sensing Symposium*, Anchorage, Alaska, 6:3755–3758.

Gianinetto, M., 2005. Automatic topographic features extraction from Pléiades HR and OrbView-5 simulated data, *The International Archive of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, 36(2/W25):75–80.

Grün, A., 1985. Adaptive least squares correlation: A powerful image matching technique, *South Africa Journal of Photogrammetry*, 14(3):175–187.

Grün, A., 1996. Least squares matching: A fundamental measurement algorithm, *Close Range Photogrammetry and Machine Vision* (K.B. Atkinson, editor), Whittles, Caithness, Scotland, pp. 217–255.

Habib, A., Y. Lee, and M. Morgan, 2003. Automatic matching and three-dimensional reconstruction of free-form linear features from stereo images, *Photogrammetric Engineering & Remote Sensing*, 69(2):189–197.

Hanley, H.B., and C.S. Fraser, 2001. Geopositioning accuracy of Ikonos imagery: indications from two dimensional transformations, *The Photogrammetric Record*, 17(98):317–329.

Harris, C., and M. Stephens, 1988. A combined corner and edge detector, *Proceedings of the 4th Alvey Vision Conference*, Manchester, U.K., pp. 147–151.

Heipke, C 1997. Automation of inner, relative and absolute orientation, *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(1):1–19.

Jacobsen, K., and R. Passini, 2003. Accuracy of digital orthophotos from high resolution space imagery, *Proceedings of the Joint Workshop High Resolution Mapping from Space*, Hannover, Germany, unpaginated CD-ROM.

Kraus, K., 1997. *Photogrammetry*, Vol. 2., Dümmler Verlag, Bonn, Germany.

Inglada, J., and A. Giros, 2004. On the possibility of automatic multisensor image registration, *IEEE Transactions on Geoscience and Remote Sensing*, 42(10):2104–2120.

Noguchi, M., and C.S. Fraser, 2004. Accuracy assessment of QuickBird stereo imagery, *The Photogrammetric Record*, 19(106):128–137.

Poli, D., Z. Li, and A. Gruen, 2004. Orientation and automated DSM generation from SPOT-5/HRS stereo images, *Proceedings of the 25th ACRS Conference*, Chiang Mai, Thailand, 1:190–195.

Poon, J., C.S. Fraser, Z. Chunsun, Z. Li, and A. Gruen, 2005. Quality assessment of digital surface models generated from IKONOS imagery, *The Photogrammetric Record*, 20(110):162–171.

Rousseeuw, P.J., and A.M. Leroy, 1987. *Robust Regression and Outliers Detection*, John Wiley & Sons, New York.

Scaioni, M., and M. Gianinetto, 2003. Fusion of aerial and satellite imagery over the city of Venezia, *Proceedings of the 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas (URBAN 2003)*, Berlin, Germany, pp. 216–219.

Shi, W., and A. Shaker, 2003. Analysis of terrain elevation effects on Ikonos imagery rectification accuracy by using non-rigorous models, *Photogrammetric Engineering & Remote Sensing*, 69(2):1359–1366.

Smith, D.P., and S.F. Atkinson, 2001. Accuracy of rectification using topographic map versus ground control points, *Photogrammetric Engineering & Remote Sensing*, 67(5):565–570.

Stylianidis, E., 2003. A new digital interest point operator for close-range photogrammetry, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Corfu, Greece, 34(5/W12):319–324.

Tao, C.V., Y. Hu, and W. Jiang, 2004. Photogrammetric explotation of Ikonos imagery for mapping applications, *International Journal of Remote Sensing*, 25(14):2833–2853.

Toutin, T., 2004. Geometric processing of remote sensing images: models, algorithms and methods, *International Journal of Remote Sensing*, 25(10):1893–1924.

Valdivieso-Casique, M., and S. Arridge, 1999. Non-linear registration using a finite element method, *Proceedings of the 10th British Machine Vision Conference*, Nottingham, U.K., pp. 214–223.

Vosselmann, G., 1992. *Relational Matching*, Lectures Notes in Computer Sciences, Springler Verlag, 628, Berlin, Germany.

Wallis, R., 1976. An approach to the space variant restoration and enhancement of images, *Proceedings of the Symposium on Current Mathematical Problems in Image Science*, Naval Postgraduate School, Monterey, California.

Wang, H., and C. Hellis, 2005. Spatial accuracy of orthorectified IKONOS imagery and historical aerial photographs across five sites in China, *International Journal of Remote Sensing*, 26(9):1893–1911.

Zhang, Z., R. Deriche, O. Faugeras, and G.T. Luong, 1995. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry, *Artificial Intelligence*, 78(1–2):87–119.

Zhang, L., and A. Gruen, 2004. Automatic DSM generation from linear array imagery data, *International Archives of Photogrammetry and Remote Sensing*, 35(B3):128–133.

(Received 24 March 2006; accepted 09 June 2006; revised 17 June 2006)