Terrestrial photogrammetry without ground control points

G. Forlani · L. Pinto · R. Roncella · D. Pagliari

Received: 7 March 2013 / Accepted: 9 July 2013
Published online: 06 August 2013

G. Forlani · R. Roncella
Department of Civil Engineering, Parma University, Parma, Italy

G. Forlani
e-mail: gianfranco.forlani@unipr.it

R. Roncella
e-mail: riccardo.roncella@unipr.it

L. Pinto · D. Pagliari (✉)
D.I.C.A., Politecnico di Milano, Milan, Italy
e-mail: diana.pagliari@polimi.it

L. Pinto
e-mail: livio.pinto@polimi.it

Introduction

Motivations

Terrestrial photogrammetry has many inherent strengths: besides the wealth of information content and archiving capabilities of the images, the flexibility of the network design can provide both accurate and reliable results in most cases. Despite this, it is often disregarded in several surveying tasks in favour of total station surveys or, as far as architectural surveys or surveys of historical buildings are concerned, it has to compete with the more fashionable laser scanner.

Networks of GPS Permanent Stations (NPS) offer 3 to 5 cm accuracy in real-time with the NRTK (Network Real Time Kinematic) service to GPS users, freeing them from the need to set up and maintain a reference station, cutting fix times for the integer ambiguity to around 60 s almost irrespectively of the distance to the nearest station (sometimes tens of km away).

In our view, this opens the possibility to effectively combine GPS and photogrammetry in terrestrial applications. Two major benefits may be gained: accurate georeferencing of the survey with GPS is straightforward and does not require ground control points; besides, GPS fixes may add strength and reliability to the photogrammetric block.

Keeping up-to-date the information stored in a topographic Data Base (DB) or a GIS in urban areas demands for procedures and techniques to detect changes, acquire new data and update the DB. In medium sized or small towns, the technical staff of the town council often either does not include surveyors at all or just high school surveyors, mostly with little GPS and/or photogrammetric expertise.

Map or GIS updating is performed today with GPS receivers where collected data are stored in a local DB or even directly uploaded on the main remote DB. Adding georeferenced image information with photogrammetric quality enhances the span of objects and information that can be retrieved.

The information to retrieve from surveys is getting more and more rich in content (i.e. the attribute information is as important as or more important than geometry); a topographic survey (although modern reflectorless total stations speed up operations in the survey of buildings) is still time consuming and inappropriate for the task (it provides too much accuracy on a too few points, and no attribute information at all). A terrestrial photogrammetric survey not only makes a better match with the accuracy required for the basic spatial data (ground plan and elevation), but gives the possibility to add many more building details (e.g. shape, materials, colour, etc.).

This need for fast and rich data acquisition for map updating in urban context might be answered by the integration of photogrammetry and GPS in a NPS.

GPS-assisted photogrammetry

Since the early days of GPS, the idea of measuring the position of the projection centres to support aerial triangulation (AT) (Ackermann 1984) has been proposed; by integration of GPS and INS, direct georeferencing was also foreseen (Schwarz et al. 1984). After the first successful aerial test Flevoland (Van der Vegt 1989) and the completion of the GPS constellation, GPS-assisted aerial triangulation (AT) became for some years the most effective (though not widely used) technique for image orientation in aerial photogrammetry. Since early 2000, the progress in INS/GPS integration made it look obsolete, direct georeferencing for mapping purposes looking feasible (Forlani and Pinto 2002; Heipke et al. 2002). To our opinion, however, the same trend does not apply to terrestrial photogrammetric surveys: in other words, direct georeferencing is not yet feasible to reasonable costs and accuracy for hand held systems.

The simple version of terrestrial GPS-assisted photogrammetry that we propose is made of a GPS receiver tightly tied to a photogrammetric camera. The GPS provides the coordinates of the projection centre of each camera station, taking into account the eccentricity between the camera perspective centre and the antenna phase centre. The camera attitude parameters are estimated using tie points in a bundle block adjustment. Therefore there is no need for ground control points. The measurement of tie points, once a time-consuming task, can be obtained by Structure and Motion techniques, which benefit either from the richness of details of the urban environment as well as from the a priori information on the camera station position provided by the GPS.

Related work

Although much of the research and development on mobile mapping focused on aerial and vehicle application, some attention has been paid in early 2000 also to the so-called portable mobile mapping systems. i.e. lightweight position and orientation systems that can be carried by pedestrians and are typically equipped with a camera and perhaps a laser finder. In this field, three major developments can be highlighted. The first system was developed at ETH Lausanne in the context of the study of artificially released avalanches (Vallet et al. 2000; Vallet 2001) with the objective to measure the volume of snow involved; being impractical using control points in such environment, the images are taken from helicopter, with a large format camera with a tactical grade INS tied to its back and a GPS antenna connected to the camera by a mast. The system has been later upgraded to accommodate also a laser scanner (Buckley et al. 2008) indeed definitely becoming not a truly portable system, due to weight and power requirements.

A second prototype system from the University of Calgary was presented in two different versions (Ellum and El-Sheimy 2000, 2001); in both cases the objective was to achieve direct georeferencing. After some test with different sensors, in the final version the position information is derived from GPS and the attitude information from a Leica Digital Magnetic Compass (DMC); the camera, the GPS antenna and the DMC are all mounted on a lightweight bar and their relative position and attitude calibrated; results with an off-the-shelf Kodak DM260 varied from meter level accuracies to about 25 cm, depending also on the number of images used.

A last example (Gillet et al. 2001) is the Applanix Backpack. It is composed by a tactical grade Inertial Measurement Unit (IMU) and a GPS receiver aimed to provide meter level 3D accuracy on environments with very poor GPS signal quality, such as forests, for seismic surveys as well as tree surveys. Position is obtained by aided inertial navigation, either with GPS or with Zero velocity UpdaTes (ZUPT) necessary to improve the calibration of velocity error sources when GPS outages last too long. Although variants of the above described systems are commercially available, no references have been found for a photogrammetric use of such systems.

To our opinion, all the above mentioned products have limitations of some sorts, which prevented their diffusion. As far as inertial aided navigation systems are concerned, weight and costs are still an issue if GPS outages may last long. As far as using GPS to provide positions and a DMC to provide attitude, the attainable accuracy with direct georeferencing is lower or comparable with that of the method we propose. Whenever the survey requires the measurement of several points, most perhaps not even directly accessible by the operator, photogrammetry is a very flexible tool and with our proposal can be implemented simply acquiring some more-than-strictly-necessary images. With calibrated digital cameras and the progress made by Structure and Motion (SfM) algorithms (Hartley and Zisserman 2000; Furukawa and Ponce 2010; Agarwal et al. 2011) such additional cost is limited. The gain is dispensing with ground control points and adding the accuracy and reliability provided to the survey by a bundle adjustment (Kraus 1997).

System concept and implementation

As mentioned in Section 1.2, the proposed system is made of a GPS receiver rigidly tied to a calibrated photogrammetric camera. A block of images (the minimum required is three from non-aligned camera stations) is acquired, registering the GPS position at each station where the signal is available. Tie points are measured in the images to establish correspondences manually or preferably by SfM algorithms. The block is then oriented and georeferenced introducing the GPS data as auxiliary information in the bundle adjustment. The GPS receiver can be operated in kinematic mode or in Stop&Go; RTK positioning is not required but is recommended since it ensures that the GPS position of the camera station is successfully measured with a known accuracy (only FIX solution may be accepted). The GPS position in Stop&Go can be registered holding the camera and shooting at the same time or the shooting time can be recorded in the GPS receiver via a cable connection from the camera flash pins (see Section 2.4).

To include the GPS information in the bundle block adjustment the mathematical model of the collinearity equations must be extended with an ad-hoc observation equation.

The mathematical model

The GPS position refers to the antenna phase center or to the antenna mount point, depending on the settings of the GPS data processing. Since the camera is fixed with respect to the antenna, the offset between camera projection center and GPS position is constant. The vector components of this offset can be measured in the image space with a calibration procedure (see Section 2.5).

The camera interior orientation parameters are assumed to be known by a previous calibration; indeed, the accuracy (and validity) of such parameters is more important here than with conventional surveys with ground control, since it is well known that correlations between interior and exterior orientation parameters can compensate at least part of such errors (Jacobsen 2000). Interior orientation and distortion parameters are modelled with the model proposed by (Fraser 1997); in principle the bundle adjustment with GPS observations can also incorporate self-calibration parameters; however, this is sound only when the overall conditions (number and distribution of camera stations, object shape, site constraints) allows for reliable results.

The mathematical model of the bundle adjustment therefore includes the collinearity equations (Kraus 1997):

$$\begin{aligned} \xi &= \xi_0 - c\frac{r_{11}(X-X_0)+r_{12}(Y-Y_0)+r_{13}(Z-Z_0)}{r_{31}(X-X_0)+r_{32}(Y-Y_0)+r_{33}(Z-Z_0)} \\ \eta &= \eta_0 - c\frac{r_{21}(X-X_0)+r_{22}(Y-Y_0)+r_{23}(Z-Z_0)}{r_{31}(X-X_0)+r_{32}(Y-Y_0)+r_{33}(Z-Z_0)} \end{aligned} \tag{2.1}$$

where

ξ,η	image coordinates
ξ_0,η_0, c	interior orientation parameters
X,Y,Z	ground coordinates of a point
X_0, Y_0, Z_0,	perspective centre coordinates
R	attitude matrix (rotation matrix from object system to camera system with elements r_{ij})

and the observation equation relating the antenna phase centre position to the camera centre, which is basically the same used in aerial triangulation (Forlani and Pinto 1994).

$$X_a = X_0 + R^t e \tag{2.2}$$

where

X_a	GPS antenna position at exposure time in a cartesian reference frame (often a local level frame)
X_0	perspective centre coordinates
e	offset perspective center – antenna phase center in image space

Drift parameters of the aerial triangulation, which are supposed to take care of systematic discrepancies between the GPS solution and the photogrammetric solutions on a strip-by-strip basis or on a block basis, have not been included. Indeed, use of such parameters are not justified in terrestrial blocks due to different block structure and GPS conditions (PDOP, number of satellites in sight), therefore there is no point in modelling “systematic” discrepancies between GPS and photogrammetry. Besides, using drift would require measuring GCPs: to avoid this is precisely the idea behind developing this system.

Only the perspective centre is considered as unknown in the equation: the eccentricity vector is considered a known value determined by calibration (see Section 2.5); the attitude matrix is determined by the photogrammetric observations and updated during iterations for non linearity of the bundle adjustment. The observation accuracy of (2.2) is taken from the GPS data processing results, normally rather optimistic.

The reference system for the adjustment should be a local level frame in the gravity centre of the area; using mapping coordinates is also possible, but results in a scale error in Z; however, this is normally within the error bounds of the method.

Block georeferencing and block control

Since the GPS stations (i.e. the camera stations where a FIX solution has been obtained from GPS data processing) should act as control points, they also establish the reference system for the adjustment. Unlike traditional ground control points (GCPs), which are often introduced as error-free quantities because their accuracy surpasses that of photogrammetry, GPS positions from a kinematic survey have about the same accuracy (and sometimes a lower accuracy) than photogrammetry. Therefore, they are introduced with the accuracy provided by the GPS data processing or at least with the expected accuracy of such survey in the actual conditions. Rather than control points, they should be seen as additional information, in principle error prone as the photogrammetric observations; indeed, errors in GPS positions can be expected because the operating environment (an urban area) is unfavourable, due to sudden and frequent changes in satellite configuration in sight.

Ill-conditioning due to inadequate block geometry

Using pseudo-observations removes the rank deficiency of the block adjustment; there are cases, though, where near-singularities may arise because of the unfavourable location of the GPS stations in the block geometry. For instance, in a single strip the projection centers are aligned and therefore the ω angle of the whole strip cannot be determined (the strip can rotate around its axis).

To get around such cases and to address the problem in general cases, simulations and experience suggest two ways:

a) work on the relative positions of GPS and camera stations to get a well georeferenced block;

b) exploit object characteristics to constrain the block adjustment.

A stable georeferencing is obtained when the convex hull obtained by the projection on the XY plane the camera stations where the GPS signal is available should not be elongated, but of rather compact shape. With objects with height, width and length of similar size, the best is obviously to have such stations all around the object. If the object has two dimensions larger than the remaining one (such as a building facade or a wall) then two strips should be taken, either at a significantly different elevation (but this would imply using an elevated platform, which is not convenient) or rather take the second strip from a larger distance from the wall (maybe changing the focal length to get an image scale about the same as the first one). Notice that, as far as georeferencing is concerned, it is not necessary that the GPS signal is received at all camera stations; of course this is desirable, since it may increase block control or the reliability of the GPS positions, since it allows a mutual check of photogrammetry and GPS.

In order measure the effect on the overall accuracy of the block as a function of the GPS stations positions, some simulations have been performed.

Figure 1 shows one of the block configurations, taken all around a rectangular building 10 m high where tie points are marked in red.

Twelve images have been taken at 2 m above ground (blue), 8 at 1.5 m (green). Three configurations have been compared: all 20 images, 12 images (3 on each building side) and 8 (2 on each side): Table 1 shows the results. As expected, the accuracy decreases with the number of images in the block, but the configuration "all around" is always stable.

Surveying just a building facade with 5 images from the same height, the condition number of the normal matrix drops by three orders of magnitude and the RMS of the tie points accuracy gets worse in width and in elevation (but

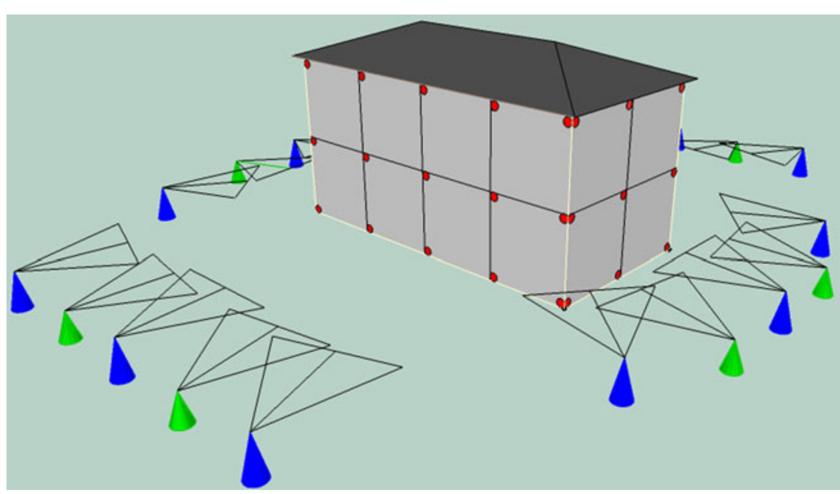


Fig. 1 Simulation of GPS camera stations around a building; tie points are marked in red; camera&GPS stations at different elevations are in *green* and *blue*

much less in depth, which is almost not affected by indeterminability around the straight line joining the projection centers); using just 3 images of this strip and two from a second strip, taken at twice the object distance (see Fig. 2), the RMS improves significantly in both depth and height, obviously without going back to the values of the stronger "all round" configuration.

The effect of additional constraints on the object has also been studied, namely constraining pairs of points to be on a horizontal line or on a vertical line; besides, the effect of constraining also just the elevation or the 3D coordinates of a single object point has been verified.

Overall, it has been found that vertical constraints are effective only when the height difference between the two points is larger than the camera-to-object distance; likewise, the horizontal constraint is effective the more the two points are aligned in a direction perpendicular to the wavelength and the farther they are from each other. This means that tie points should also be measured on a plane orthogonal to the facade plane.

System hardware and software

The basic system hardware is made of a digital camera and a GPS receiver, with the antenna and the camera tied together (see Fig. 3). In our prototypal versions we used a relatively expensive sort of cameras: a Nikon D100 and a Nikon D70s equipped with a 18 mm and 20 mm lens respectively, both calibrated. Their resolution (6 Mpixels) is today matched by many less expensive cameras, although the quality of the optics and the possibility to interchange calibrated lenses ensure accuracy and flexibility. In order to achieve subdecimeter (or better) accuracies, a geodetic GPS receiver should be used, which is indeed rather expensive. In the various test we used a Leica 1230 GG with AT 1202, a Leica 530 with AT502 and a Trimble 5700 with a Zephir antenna; recent tests point to the possibility of using also a single frequency.

Since the receiver must operate in differential mode to provide the above mentioned accuracy level, a reference station is necessary. Due to the frequency of the loss of lock in urban environment, short reacquisition times and fast and reliable ambiguity fixing are necessary for the system to be of practical use. As already underlined, the availability of a network of GPS permanent stations with NRTK service improves dramatically the performance of the system.

Table 1 Theoretical accuracy of tie points with different configurations of GPS stations (Figs. 1 and 2)

Block	RMS (std. dev.)		
	X(mm)	Y(mm)	Z(mm)
12_2+8_1.5	8	8	8
12_2	9	10	10
8_2	12	12	12
5 (façade)	41	288	999
3+2 (façade)	42	53	120

Indeed, the results obtained using the NRTK service within the IREALP (Biagi et al. 2006) GPS permanent network (with the closest station at about 15 km away, compare very well with others cases where positioning was obtained with respect to a single reference station just a few hundred meters away (see Section 4).

Positions were obtained mainly with NRTK; in two cases only post-processing was used.

As far as connection between antenna and camera is concerned, the antenna is mounted on a 2 m high pole with the camera fixed to the GPS terminal carrier; in this way the measurement of object points can be performed, should this be necessary; the position at shooting time can be recorded in Stop&Go, just holding firm the pole and pressing shutter release button. Otherwise it can be recorded by sending to the receiver input port a signal (e.g. the TTL signal of the camera flash head). Some receiver (e.g. the Leica 1230 GG) allows for the interpolation of the position of every input event recorded.

System calibration

The eccentricity vector components with respect to the camera frame can be recovered easily, in many ways. The simplest but less accurate is taking images of a testfield with known points in WGS84 while also measuring with the receiver. By space resection the camera exterior orientation parameters X_0 and R can be recovered, while processing the GPS data provides the antenna phase center position X_a. The eccentricity vector from camera center to antenna is then computed as:

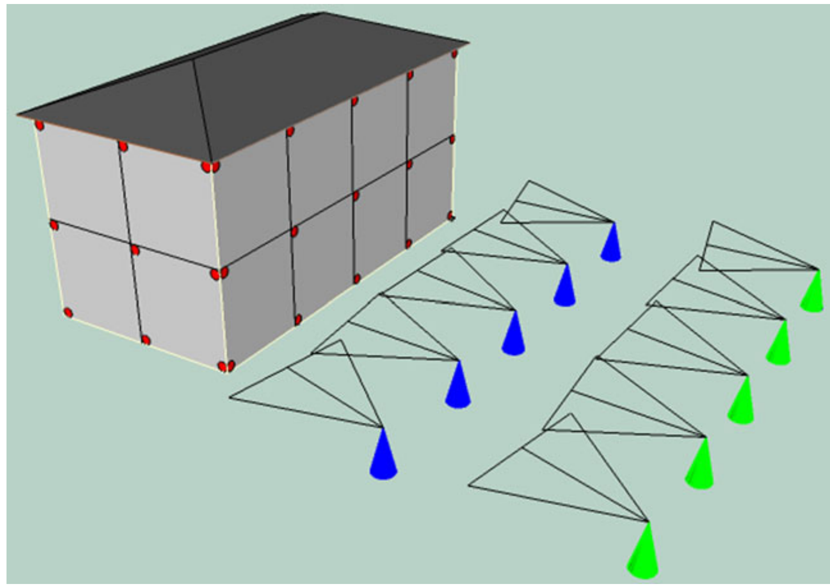
$$e = R(X_a - X_0) \tag{2.3}$$

With this method, the standard deviation of the vector magnitude with 5–6 images ranges between 1 and 2 cm, i.e. the accuracy of GPS.

A more accurate determination is performed using a total station and a retroreflective prism mounted on the pole, kept vertical, in place of the antenna. The camera stations are determined by space resection, while the total station provides the position of the prism (see Fig. 4). By accounting for the height difference with respect to the pole screw of the prism and of the antenna phase centre, the accuracy of the vector magnitude ranges between 2 and 3 mm.

Indeed, the accuracy of the eccentricity vector need not to be better than that of the GPS; the vertical component, since the shootings are mostly taken with camera axis nearly

Fig. 2 Simulation of two strips at different object distance from a building facade



horizontal, is the most important: its effect is systematic. Errors in the horizontal components tend to cancel out in "all round" blocks.

Automatic tie point extraction

Though in most of the tests tie points were measured manually, in fact tie points can be extracted largely automatically. In our processing pipeline, images are first processed with the program EyeDEA (Roncella et al. 2011) that implements a SfM (Structure from Motion) strategy based on SURF (Bay et al. 2008) or Harris (Harris and Stephens 1988) interest operators, SURF descriptors and the kd-tree search (Beis and Lowe 1997) to find correspondences. Elimination of outliers

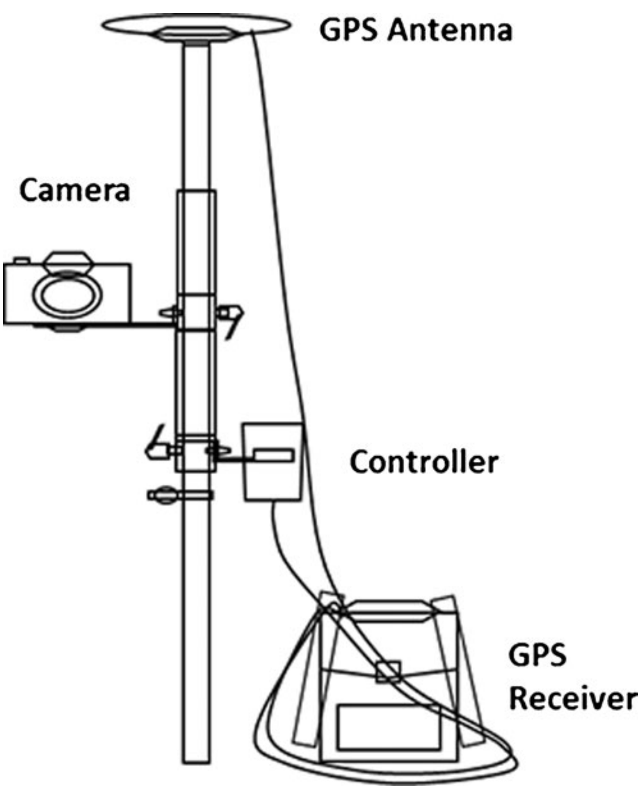


Fig. 3 System hardware

Fig. 4 System calibration

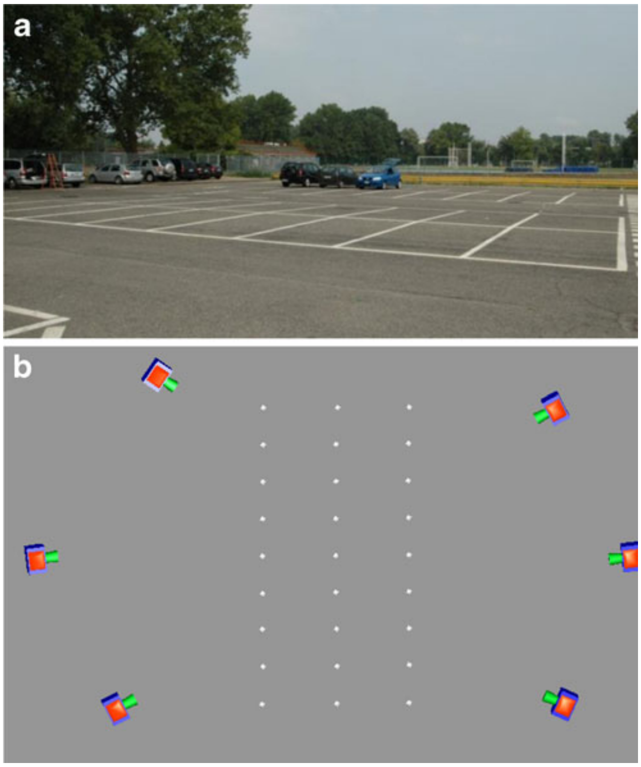


Fig. 5 Test 1 - Car parking: **a** image; **b** camera stations and tie points

in the correspondences is performed with epipolar and trifocal filtering combined with RANSAC (Fischler and Bolles 1981), a robust estimation technique. The epipolar constraint can be implemented either with the fundamental or the essential matrix (Nister 2004); the trifocal constraint can be implemented with the trifocal tensor or by a simpler method, again using RANSAC (see Roncella et al. 2011 for more details).

We have also recently made some trials with SIFT (Lowe 2004), on other image sequences. However, there seem to be no remarkable difference of performance, while SURF is faster.

Additional processing steps can be executed to improve the number of rays per point or to reduce the number of tie points to save computing time by preserving multiplicity and optimizing the distribution on images. After this step, a preliminary bundle block adjustment (BBA) is executed with Photomodeler™ in an arbitrary reference system, to check for inconsistencies by also taking advantage of its graphic and display tools to measure check points; then the photogrammetric observations are exported and reformatted for processing in the BBA program CALGE (Forlani and Pinto 2002). Besides handling the traditional block control using GCP, CALGE implements also the pseudo-observation Eq. (2.2). The antenna coordinates are provided by the GPS data processing, while the shooting time is recorded in the GPS receiver via the Event Input port connected to the camera flash pins. To get the position of the GPS antenna at the shooting time, interpolation may or may not be required, depending on whether kinematic positioning or Stop&Go is used.

The camera-to-antenna eccentricity vector, expressed in the camera system and determined by a previous calibration, is also input to the program. Processing the GPS and photogrammetric observations the main problem is to assign the correct weights to the GPS positions with respect to the photogrammetric observations. It is well known that GPS data processing deliver quite optimistic (unrealistic) accuracy estimates. Weight estimation techniques or simple rescaling to realistic values can be used; the major difficulty is to cope with the possibility of outliers in the GPS antenna trajectory (e.g. sudden "jumps" due to satellite changes or systematic deviations from true trajectory).

System tests

Several tests have been performed. Two buildings were surveyed with kinematic post-processed GPS data; a car parking, two more buildings and a city block were measured by NRTK in VRS mode and communication link by the NTRIP protocol.

In all cases but one, check points were measured on the object by total station and GPS; where no check points were available, the survey has been repeated twice, with a different camera and a different satellite configuration or points from maps have been identified in the images. The photogrammetric measurements (tie points selection and collimation) were performed manually with Photomodeler™ or by using EyeDEA; the image measurements and the GPS antenna positions were adjusted with CALGE.

A short description of each test site and of the photogrammetric block follows; the overall results of the tests are summarized in Table 2.

Test 1 Car parking

The corners of the markings of a parking (Fig. 5a) were measured by GPS and total station to an accuracy of about 2 cm. A block of 6 images was taken (Fig. 5b) with the GPS receiver working in NRTK. The test on standardized residuals led to the identification of an erroneous GPS observation (an elevation error of about 15 cm).

Test 2 Dept. of Mathematics, Parma University

Twenty-two images (Fig. 6b) were taken around the Department of Mathematics (Fig. 6a) with GPS measuring in Stop&Go. Post-processed GPS data were used. At some stations the GPS had loss of lock due to shadowing by trees.

Test 3 Building, Parma University Campus

Eleven images (Fig. 7b) were taken around the Campus Technical Unit building (Fig. 7a) with GPS

Table 2 Theoretical accuracy of tie points and empirical accuracy at check points

	# chk pts	σ chk pts [mm]	# eq., # unkn.	RMS(σ) X, Y, Z [mm]	RMS(chk pts) X, Y, Z [mm]
1 - Car parking	27	20	304, 117	14, 15, 16	38, 35, 6
2 - Dept. of Math.	16	20	1160, 489	22, 23, 26	53, 31, 61
3 - Campus Building	19	10	789, 387	27, 25, 31	29, 57, 33
4 -Building	n/a	n/a	1500, 822	36, 41,47, 18, 21, 21 (‘)	40, 35, 49 (*)
5 - City block	52	300 (°)	112726, 56385	90,100,38	300, 300, 230

(‘) first row: D70s block; second row: D100 block

(*) RMS refers here to the coordinate differences of 238 tie points in the two different surveys

(°) nominal accuracy of 1:1.000 maps

working in Stop&Go. Post-processed GPS data were used.

Test 4 Building

The building shown in Fig. 8a has been surveyed twice, the first time with a D70s Nikon kept in landscape position: 18 images, were taken, but for only 8 (along the shortest sides) a NRTK fixed solution could be measured with respect to the GPS permanent network; the second survey was executed one week later with a D100 Nikon kept in portrait position: 18 images, all with NRTK fixed solution (Fig. 8b). In this case no check points were available; the comparison between the two surveys is therefore meant to give an indication on the stability of block georeferencing by GPS.

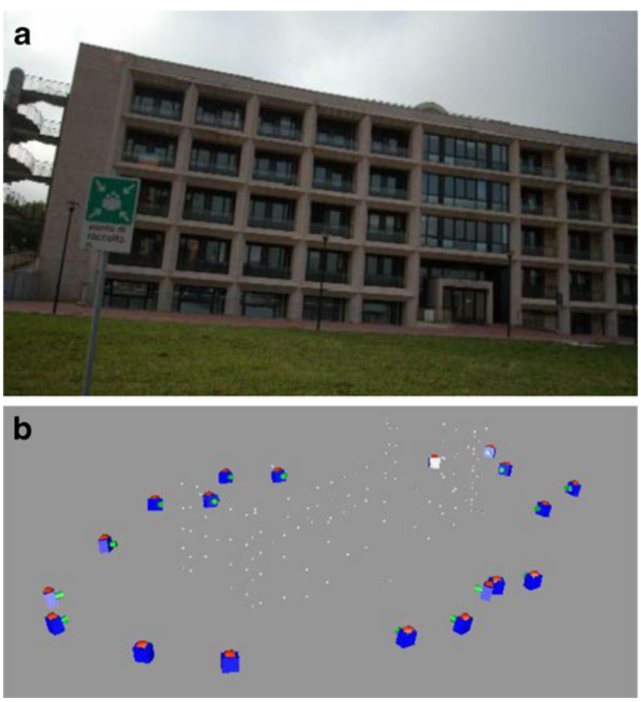


Fig. 6 Test 2: **a** image; **b** camera stations and tie points

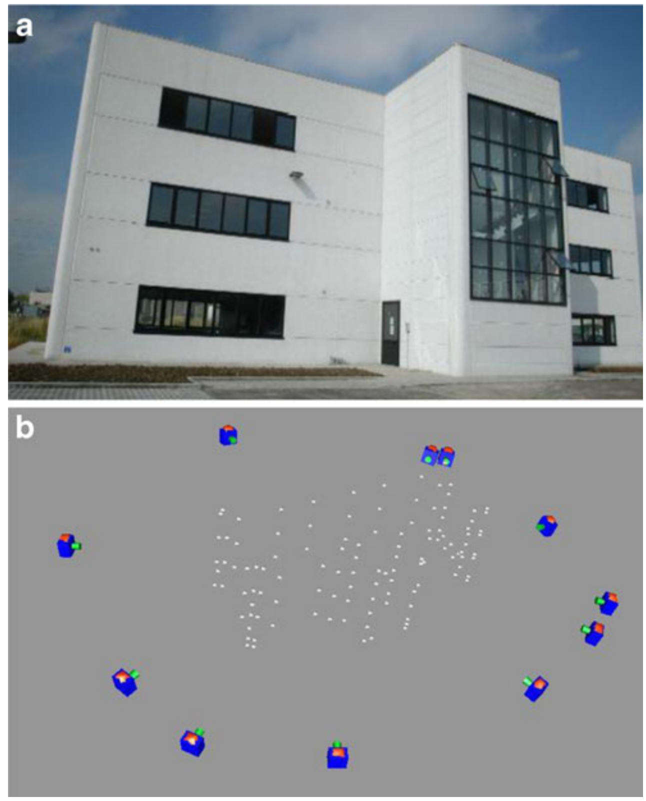


Fig. 7 Test 3 - Building at the Campus of Parma University **a** image; **b** camera stations and tie points

The photogrammetric observations were given an accuracy of 8 µm, i.e. 1 pixel. The GPS observations were introduced with an accuracy of 2 cm in X,Y and 3 cm in Z; in all cases, several standardized residuals of the antenna pseudo-observation equation are out of tolerance (larger than 4); this marks a certain disagreement between GPS and photogrammetry.

Test 5 City block

In this last test a mobile mapping version of the system has been adopted to survey a city block of detached houses in Cremona; the goal of this experiment is mainly to show that with SfM techniques even in a challenging urban environment a great number of images can be automatically oriented and that even long image sequences can be georeferenced with a rather small number of GPS camera stations.

Two Nikon D70s synchronized cameras were mounted with a base length of 1 m on the left side of a van top, together with a GPS receiver; the camera axes are directed normal to the buildings front. A

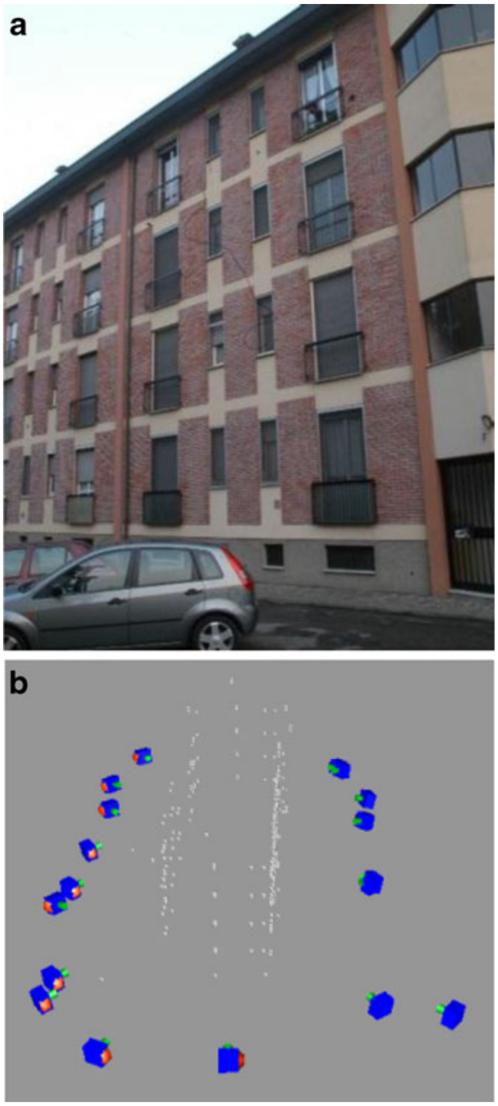


Fig. 8 Test 4: building

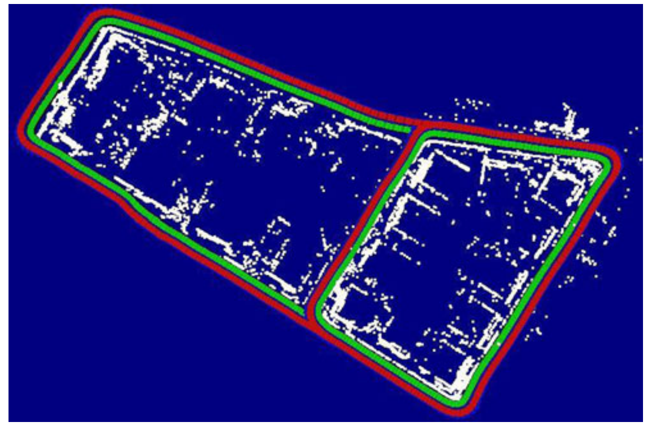


Fig. 9 Test 5: City block: camera stations and tie points

sequence of about 600 images (300 image pairs) was taken with a very short spacing (moving about 2 m from a station to the next) along the city block, completing two closed loops (see Fig. 9); the total length of the loops is approximately 600 m. Tie points were automatically extracted and matched with EyeDEA; due to the large number and uneven distribution of features detected, a filtering step was executed to optimize the extraction (see Fig. 10) preserving image coverage and ray multiplicity. To do so, each image is divided in cells (whose size is user-defined): within each cell the point with the higher multiplicity is selected and the corresponding observations in the other images are stored. All the cells of each image are processed and whenever a cell contains already a point it proceeds further until every image is completely analysed. We verified that the use of such a strategy for the homologous point selection reduce the block adjustment processing time without loss of accuracy with respect to using all the observation.

In the end about 18000 tie points were selected with overall 60000 image points measured. As can be seen from Fig. 10, feature matching is applied to a demanding environment, due to a large scene depth with objects very close as well as far away (large perspective differences even with the small base used, occlusions); moreover bushes and grass do not offer reliable features and there is plenty of repetitive patterns (bricks).

On average, there are 106 points per image; the average image coverage is about 79 %; the average number of rays per image is 3 (with a maximum of 22); the average intersection angle between homologous rays is 17°.

The GPS positions were computed from the NRTK fixed solution and were available for all image stations, thanks to the low height of the houses and lack of trees along the road.

The sequences along the two loops have been adjusted with CALGE adding as a constraint the relative orientation parameters and the distance between the two cameras (Forlani et al. 2005); in such a way, for each image pair only the exterior parameters of the left image are free unknown.

In order to study the dependence of the tie points coordinates and their precision from the amount of GPS stations in the adjustment, several adjustment were executed, including each time a smaller number of GPS stations. Moving from all 300 stations down to just eight at the corners of the main loop the RMS of the corrections to the coordinates and the theoretical accuracy change barely by 2 cm. In addition to the changes to the tie points coordinates, also 52

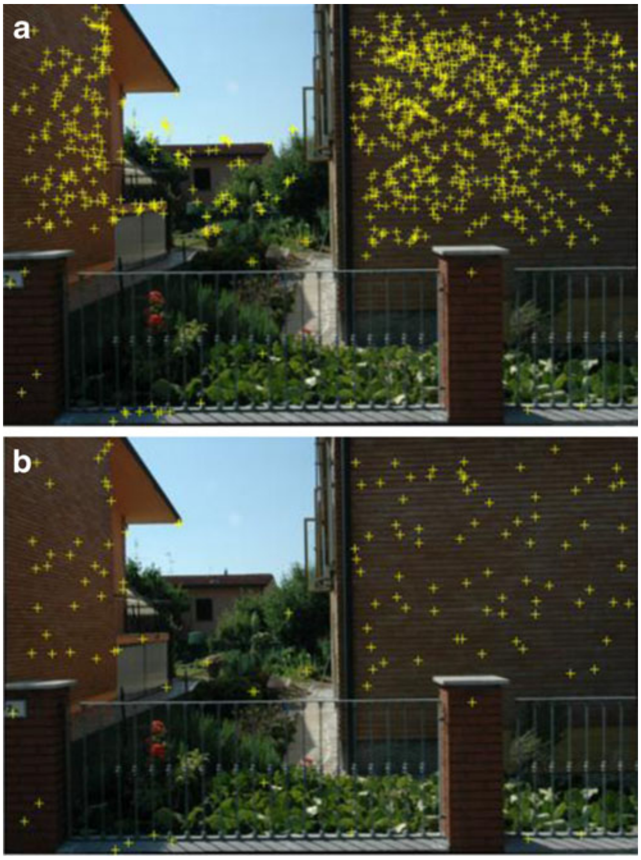


Fig. 10 Test 5: Tie points extracted and filtered

check points were identified on a 1:1000 digital map (see Fig. 11) on the buildings corners to check independently the accuracy of georeferencing. Most likely due to the peculiar image sequence characteristics, reducing the number of GPS stations does not change significantly the RMSE on the check points, even using as few as 8 GPS positions. The local redundancy of the GPS observations, that is more the 0.8 with all

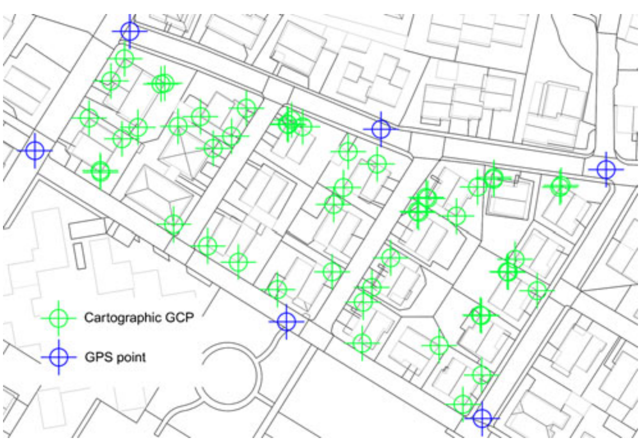


Fig. 11 Test 5: Tie points extracted and filtered

GPS stations, obviously drops to unacceptable values (less than 0.1) with just 8 GPS stations.

Table 2 summarizes the results of the block adjustments on the five tests: the fifth and the sixth column show respectively the theoretical accuracy from the block adjustment for all tie points of the block and the RMS of the discrepancies at the check points. The empirical accuracy is about the same or two times worse than the theoretical accuracy (computed with the estimated sigma naught). As far as the 4th test is concerned, the theoretical accuracy of the second block is about two times better than the first. This may be due to a better image quality, to the camera pose (portrait rather than landscape) and finally to a larger and well distributed number of GPS positions.

In Test 5 the RMSE is consistent with the quality of the check points while the theoretical accuracy is much better. Overall, therefore, the results confirm that the accuracy offered by the system is better than that required in urban maps updating.

Conclusions and prospects

A simple autonomous surveying system made of a photogrammetric camera and a GPS receiver has been presented, designed for use in the photogrammetric survey of sites where GCP are not available nor easy to set up, such as buildings in urban areas. To provide the object coordinates, tie points must be measured and a GPS-assisted aerial triangulation performed: the system cannot therefore be classified as a Mobile Mapping system, since camera orientation is not provided by navigation sensors.

Due to the difficult GPS signal and satellite configuration where the system is meant to operate, it is most conveniently used within the NRTK service of a network of GPS permanent stations, where an accuracy of a few cm may be reached and time to fix the integer ambiguity is in the order of a few tens of seconds.

From the number of test cases investigated so far, an accuracy in the range 3–7 cm (relative accuracy 1/1000-1/2000) have been demonstrated on the restitution of buildings by untrained personnel. Repeatability of the GPS georeferencing and of the object restitution was proved to be in about the same range, using different cameras and different satellite configurations. Although the relative accuracy figures are rather low, the absolute accuracy is more than enough for GIS data acquisition in urban areas.

Obstacles to GPS signal is the main limiting factor for the system, which may lead to ill-conditioning of the normal equation system in block adjustment. Some ways to get around these problems have been suggested, but there may nonetheless be cases where the system won't work.

The question of whether the GPS can also add strength to the block, rather than just providing georeferencing has been also addressed; simulations show that there are benefits in terms of the theoretical accuracy of object points, but the real improvement depends on the inner block strength and on the actual accuracy of GPS observations. Indeed, a balance between GPS and photogrammetry may help to increase block reliability and provide a check for the GPS observations.

References

Ackermann F (1984) Utilization of navigation data for aerial triangulation. International Archives of Photogrammetry and Remote Sensing, ISPRS Commission III, Vol. 25, Part A3a, Rio de Janeiro, pp. 1–9.

Agarwal S, Furukawa Y, Snavely N, Simon I, Curless B, Seitz SM, Szeliski R (2011) Building Rome in a day. Communications of the ACM 54(10):105–112

Bay H, Ess A, Tuytelaars T, Van Gool L (2008) SURF: Speeded up robust features. Computer Vision and Image Understanding (CVIU) 110(3):346–359

Beis JS, Lowe DG, (1997) Shape indexing using approximate nearest-neighbour search in high-dimensional spaces. In: Proc.of CVPR 1997, pp. 1000–1006.

Biagi et al. (2006) *Il Servizio di Posizionamento in Regione Lombardia e la prima sperimentazione sui servizi di rete in tempo reale*. Bollettino della SIFET n° 1, 2006

Buckley SJ, Vallet J, Wheeler W, Braathen A (2008) *Oblique helicopter-based Laser scanning for digital terrain modeling and visualization of geological outcrops*, The Internal Archives of the Photogrammetry, remote sensing and Spatial Information sciences, Beijing, Commission IV/4, pp. 493–498.

Ellum C, El-Sheimy N (2000) The development of a backpack mobile mapping systems, International archives of the Photogrammetry, Remote sensing and spatial information sciences, Vol. XXXIII Part B2, (CD), Amsterdam, The Netherlands, pp. 184–191.

Ellum C, El-Sheimy N (2001) A mobile mapping system for the survey community, Proc. of The 3rd Int. Symp. on Mobile Mapping Technology, Cairo, Egypt, January 3–5, 2001, (CD).

Fischler M, Bolles R (1981) Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography. Commun Assoc Comp Mach 24(3):81–95

Forlani G, Pinto L, (1994) Experiences of combined block adjustment with GPS data. International Archives of Photogrammetry and Remote Sensing, ISPRS Commission III, Vol. 30 Part 3/1, Muenchen, pp. 219–226.

Forlani G, Pinto L (2002) *Integrated INS/DGPS systems: Calibration and combined block adjustment*, *Proceedings*, OEEPE Workshop "Integrated Sensor Orientation", Hannover, Sept. 17–18, 2001, OEEPE Official Publication N°. 43, 2002, pp. 85–96.

Forlani G, Roncella R, Remondino F (2005) Structure and motion reconstruction of short mobile mapping image sequences, Proc. of the 7th Conf. On Optical 3D measurement techniques, Vienna, 3–5 Oct. 2005, Vichman Verlag, Vol I, pp. 265–274

Fraser CS (1997) Digital camera self-calibration. ISPRS Journal of Photogrammetry and Remote Sensing 52:149–159

Furukawa Y, Ponce J (2010) Accurate, dense, and robust multi-view Stereopsis. IEEE Transactions on PAMI, Vol. 32, Issue 8, August 2010, pp 1362–1376.

Gillet J, McCuiag R, Scherzinger B, Lithopoulos E (2001) Tightly coupled inertial/GPS system for precision forestry surveys under canopy: Test results. First International Precision Forestry Symposium, University of Washington, College of Forest Resources, Seattle, pp 131–138

Harris C, Stephens M (1988) A combined corner and edge detector. Alvey Vision Conference, pp. 147–151

Hartley R, Zisserman A (2000) Multiple view geometry in computer vision. Cambridge University Press, Cambridge, pp 1–496

Heipke C, Jacobsen K, Wegmann H (2002) Analysis of the Results of the OEEPE Test Integrated Sensor Orientation. OEEPE Official Publication, No 43:31–49

Jacobsen K (2000) Potential and limitation of direct sensor orientation, Int. Arch. of Photogrammetry and Remote Sensing, Amsterdam, The Netherlands, Vol. 33, Part B3/1, pp. 429–435

Kraus K (1997) Photogrammetry, vol 2. Dümmler, Bonn, 466 pp

Lowe D (2004) Distinctive image feature from scale-invariant keypoints. International Journal of Computer Vision 60(2):91–110

Nister D (2004) An efficient solution to the five-point relative pose problem. IEEE T, Pattern Anal 26(6):756–770

Roncella R, Re C, Forlani G (2011) Comparison of two structure and motion strategies. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume 38-5/W16, pp. 1–8, CD-ROM

Schwarz KP, Fraser CS, Gustafson PC (1984) Aerotriangulation without ground control. International Archives of Photogrammetry and Remote Sensing, ISPRS Commission I, Vol. 25, Part A1, Rio de Janeiro, pp. 237–250.

Vallet J (2001) Design of a helicopter based portable handheld mobile system for avalanche mapping. Proc. of The 3rd Int. Symp. on Mobile Mapping Technology, Cairo, Egypt, January 3–5, 2001, (CD).

Vallet J, Skaloud J, Koelbl O, Merminod B (2000) Development of a Helicopter-based integrated system for avalanche and hazard management. International. Archives of Photogrammetry and Remote Sensing. Vol. XXXIII part B2, (CD), Amsterdam, The Netherlands, pp. 565–572

Van der Vegt HJW (1989) GPS test flight Flevoland. Schriftenreihe des Instituts für Photogrammetrie, Universität Stuttgart, Vol 13, pp. 285–298