

# Vehicle positioning in urban areas using photogrammetry and digital maps

Barzaghi Riccardo, Carrion Daniela, Cazzaniga Noemi Emanuela, *DIAR, Politecnico di Milano*  
Forlani Gianfranco, *Dipartimento di Ingegneria Civile, Università di Parma*

## BIOGRAPHY

Riccardo Barzaghi. Born in 1958. Professor of Geodesy and Surveying at Politecnico di Milano. Main research interests: Physical Geodesy, Satellite Geodesy, GPS networks for deformation analysis.

Daniela Carrion. Born in 1976. PhD in Geodesy and Surveying in 2003, she is research fellow at Politecnico di Milano. Main research interests: Geodesy, in particular geoid computation, Cartography, local administration GIS design and historical applications of GIS, Remote Sensing.

Noemi E. Cazzaniga, born in 1976, is research fellow at Politecnico di Milano, Italy. She holds a Ph.D. in Geodesy and Geomatics and a degree in Environmental and Land Planning Engineering. Her main research interests are INS/GPS navigation, statistical analysis and satellite geodesy.

Gianfranco Forlani. Born in 1958. Professor of Surveying and Photogrammetry at the University of Parma. Main research interests: Mobile Mapping, surface reconstruction, automatic orientation of photogrammetric blocks, filtering of point clouds for terrain and building extraction.

## 1. INTRODUCTION

Medium precision positioning (below 50 cm) can be a valuable support in many surveying techniques as, e.g., Ground Penetrating Radar (GPR). GPR is an active instrument that uses the radar technology to penetrate the ground up to a few meters. Analysing the echo collected by the detector, a picture of the different kind of materials or structures underground can be produced. Among other applications, the detection of pipes in urban streets is of great importance, since in most cases there are no accurate maps of the underground network of cables and pipes laid by public utilities. This makes it difficult planning new installations and maintenance of the existing; when digging, damages and interference to

the service often occur. To draw a plan of the underground status quo, a GPR installed on a trailer is driven along the roads. At each pass, a strip about 2 m large can be surveyed; operating speed can be as high as 15-20 km/h, but the normal daily productivity is about 3 km. In order to provide a useful map to users, the georeferencing accuracy of the GPR data should be at maximum around 20 cm. This accuracy target might be achieved today by several positioning system that rely on GPS or GNSS as a key (or as the sole) system component. Indeed, real-time as well as post processed kinematic positioning within a network of permanent GPS stations (NRTK or PPK) can deliver this accuracy.

Nevertheless, the specific conditions of GPR urban surveys, make it impossible to rely on GPS alone for positioning. In front of large blocks, in narrow streets, or in tree-lined avenues it is well known that GPS positioning (even should the code solution be accurate enough) is not sufficient to guarantee uninterrupted coverage, because the number of visible satellites can often be less than 4. Besides the time to fix after a complete GPS outage can vary from 20 s to more than 200 s (Petovello et al., 2003), extending the duration of the GPS outages and sometimes forcing to stop the vehicle during the survey to recover the GPS solution. Usually, to overcome this limit, fusion of data collected from different sensors is operated. The most used integrated system combines GPS and inertial navigation system (INS) data performing the so called GPS-aided inertial navigation, adding to the accuracy of positioning the resistance to short term GPS outages (Zhang et al., 2005). By mean of the tightly coupled solution, the most recent systems allow a positioning precision of more than 1m after almost 1 min of complete absence of GPS signal. Besides, inertial aided navigation is currently still rather expensive.

Other cheap instruments used to aid the GPS vehicle positioning are odometers, but their precision is not sufficient for our purposes.

So our idea is to introduce photogrammetry to determine the positions of the vehicle, aided by GPS when available. In urban environment, the coordinates of some points located at every side of the surveying area are known with sufficient precision, thanks to the urban maps: this information can be used to georeference the photogrammetric strip.

### 1.1 A FEASIBILITY STUDY

The objective of this paper is to present the first results of a feasibility study of a system where photogrammetry and GPS are combined to provide data georeferencing in GPR surveys.

On the trailer carrying the GPR a pair of digital cameras and a GPS antenna will be mounted; the three sensor (GPS, GPR and cameras) will be synchronized to the same time base. The cameras will point leftwards and rightwards so as to look at the building façades on the sides with camera axis approximately perpendicular.

As far as the GPS solution is concerned, processing and data quality analysis of GPS data will provide time-tagged positions deemed reliable.

As far as the photogrammetric solution is concerned, two photogrammetric strips will be acquired by the cameras on the building façades of the street at each pass of the trail. Using robust Structure and Motion algorithms the inner geometry of each strip can be reconstructed. Ground control information, manually extracted from large scale maps, will be used to georeference each strip.

The position of the GPS antenna phase centre and the rigid transformation (rotations and shifts) from the camera systems to a vehicle-fixed reference system will be determined in a calibration stage; likewise, the transformation from the GPR to the vehicle-fixed reference system will be determined.

Since the orientation of images in object space will provide the transformation from object to camera system, using the calibration parameters both the photogrammetric and the GPS solution may be referred to the GPR location. A time interpolation will provide the location of the GPR scans.

Both positioning techniques will run in parallel; as it will be pointed out later, the assumption is that they should not fail at the same time, so a solution will always be available. In case of simultaneous failure to provide a solution with the standard procedure, additional field measurement will provide ground control for the photogrammetric solution, so there will always be a photogrammetric recovery of the trajectory, though at a larger cost.

How to choose the best solution (or to combine them when both are available with good quality

parameters) has not yet been investigated. Indeed, survey specifications are expected as a result of this study and work is still in progress.

In the following, the main focus will be on the feasibility of positioning using photogrammetric measurements and ground control taken from digital maps, considering the unfavourable case of complete absence of GPS solutions.

### 1.2 PREVIOUS WORK ON THE SUBJECT

Using images to locate the camera centre has long been a topic in robot navigation and techniques for image orientation have been proposed that can be successfully applied to the specific domain of urban scenes. A few authors have been working on Structure and Motion applied to Mobile Mapping sequences. To the best of our knowledge, none of the proposed methods reached an operational level. Automatic image sequence orientation to support IMU/GPS systems in overcoming GPS outages was proposed in (Chaplin and Chapman, 1998; Chaplin and Chapman, 2001; Tao et al., 1999); feature extraction was based on the Förstner interest operator and image correspondences found by adopting a parallax field and finding the best match through a least squares matching algorithm. Constraints were imposed on the relative orientation of synchronous pairs and the fundamental matrix was applied to estimate the epipolar geometry along the sequence. Although later the problem still finds some attention (Kim, 2004), not many figures are given on the success rate and the achieved accuracy.

In a previous paper (Forlani et al, 2005) it has been shown that automatic orientation of a sequence of front stereo images from a mobile mapping van over a path of about 300 m was possible, with projection centre errors in the order of 1 m. That result was achieved in a countryside environment, with medium resolution video-cameras (1.3 MB) and the very unfavourable imaging geometry of synchronous MMV front images. Using larger resolution digital cameras with appropriate focal length, the façade of a building block on the roadside provides a much better number and distribution of tie points, improving the performance of the matching algorithms and providing a better imaging geometry, especially thanks to the enforcement of the relative orientation constraint. Tracking points along the sequence does not have to accommodate for large variations in image scale or image rotation. This provide strength to the solution and makes it possible to track a tie point along the whole overlapping area of the images.

## 2. GEOREFERENCING GPR PASSES

### 2.1 HARDWARE AND DATA PROCESSING

As already mentioned, the trailer carrying the GPR will also host at least a pair of reflex digital medium resolution cameras, that will be placed on a rigid frame, high enough to ensure that the view towards the left and right side is not occluded, at least by parked cars (the most frequent obstruction). The cameras will be mounted looking leftwards and rightwards (i.e. in opposite directions) so as to look at the building faces with camera axis approximately perpendicular, as in the normal case of photogrammetry. In order to increase as much as possible the image coverage in narrow streets, the left camera should be pointed to the right and vice versa.

The data acquisition from the different vehicle's sensors should be synchronized, in order to allow the time interpolation of the trajectory; cameras will be triggered (as in MMVs) by an odometer. This avoids the recording of unnecessary images when the vehicle stops; it also maintains a stable overlap between consecutive images.

Each image sequence will be oriented in an arbitrary reference system by automatically extracting tie points with a Structure&Motion algorithm using fundamental matrix and trifocal tensor estimation to control matching errors. Manual measurement will provide the image coordinates of ground control points. After that, a bundle adjustment will be executed, using the relative orientation constraint and the available GPS positions (if any) to georeference the strip and to constrain the bundle solution so as to avoid the drifts.

### 2.2 FEATURE EXTRACTION

In order to speed up the orientation process of the image sequence, tie points can be extracted using an interest operator. The urban environment, especially in narrow streets, supplies an abundance of details that makes it easy to extract well defined points from windows, bricks and stones, shop fronts, etc. With cameras axes almost normal to the building façades, perspective distortions are small and the percentage of corresponding features extracted can be reasonably estimated in the range 40-60 %. Past experience with the Harris interest operator (Harris and Stephens, 1987) shows that typically from 500 to 800 points can easily be extracted in each image. The problem is to automatically find out which features are actual matches and to discard false correspondences. Here robust techniques are necessary, because many features are quite similar to each other, so mismatches may easily occur.

### 2.3 TIE POINTS MATCHING AND IMAGE SEQUENCE ORIENTATION

The input data to an image orientation algorithm are the list of interest points in each image and, if available, the inner orientation parameters of the camera. To run smoothly and to ensure reliability, the procedure normally assumes that images have been taken ensuring overlap between at least three consecutive images, as it would be along a single strip in aerial blocks. Images are then processed according to the shooting time. The right sequence and the left sequence are processed independently at this stage; they will be combined later in the bundle block adjustment.

A list of potential matches is established based on a disparity threshold and on the similarity of the grey values in the neighbourhoods; both information can be combined in a single score for correspondence through a version of the SVD match algorithm (Pilu, 1997).

Wrong correspondences in the preliminary list can be filtered in two steps. In the first, the epipolar geometry between consecutive images of the same camera is recovered with a robust algorithm (Fischler and Bolles, 1981), improved by adding more correspondences through a guided matching and finally by l.s. estimation of the fundamental matrix using all matches. In the second, to get rid of the remaining errors, the relative geometry of three views (i.e. the trifocal tensor) is computed, again by means of a robust algorithm (Hartley and Zisserman, 2000).

For images taken close to the normal case, the epipolar lines will normally be almost horizontal, at least along smooth road sections. The epipolar constraint therefore performs poorly in discriminating points selected on the same horizontal edge. Remaining mismatches will be dealt with by the trifocal tensor, combining a set of three consecutive images, after the epipolar check has been performed between the 1<sup>st</sup> and 2<sup>nd</sup> image and the 2<sup>nd</sup> and 3<sup>rd</sup> image.

Finally, starting from the first pair, points are labelled through the sequence, providing the input image coordinates for the bundle adjustment.

### 2.4 GROUND CONTROL

Thanks to the development of Digital Cartography and Geographic Information Systems most local administrations have produced digital base maps at large scale, like 1:1000 or 1:2000. Whenever large scale maps are available it can be useful to take advantage of their large amount of accurate informa-

tion. In fact map standards in Italy prescribe a 95% error tolerance in base maps of 40 cm for 1:1000 scale and of 80 cm for 1:2000 scale (Blachut et al., 1979). The surveying of urban areas can be aided by the coordinates obtained from the maps, where many elements can be located, like buildings corners, sidewalks or flower bed edges. Such points may as well be used as ground control points (GCP) to georeference a photogrammetric strip. As far as the (E, N) coordinates of GCP are concerned, building block edges can be easily recognized and measured on the images of the sequence. As far as height GCP are concerned, points on the ground will be used, mainly building corners at the ground level; eaves elevation is often available in modern digital maps and might be used as well. The visibility of the edges depends on the field of view of the camera lens and on the distance between the camera and the object. Slightly tilting upwards the cameras may help, although normally those points will only be available for low-rise buildings.

## 2.5 METRIC RECONSTRUCTION AND BUNDLE ADJUSTMENT

Using calibrated cameras, the orientation of each pair in an arbitrary metric frame can be recovered through the essential matrix; approximate points coordinates can be computed by forward intersection using a linear algorithm. By alternating intersection and resection the whole sequence will be oriented in an arbitrary frame. Using the coordinates of the ground control points, the transformation from the arbitrary reference to the map reference system will be computed.

Depending on the imaging geometry (and more specifically on the ratio of the object distance to the base) some points may have poor intersection geometry and therefore large uncertainties. Alternating resection and intersection may accumulate errors. The inner strength of the strip is anyway much better than that of Mobile mapping front stereo cameras, so there should be no particular problem in running the final bundle adjustment.

## 2.6 RELATIVE ORIENTATION CONSTRAINT

As well known and confirmed by preliminary simulations, the accuracy of the perspective centres decreases along the strip with the distance from the closest GCP. Although the mean accuracy of projection centres looks likely to remain within specification in most cases, the inner block strength can be increased combining the left and right strip. To this aim, the relative orientation (RO) as well as the distance between the left and right cameras, known by calibration, can be enforced in the adjustment, to constrain the solution.

By introducing two groups of observation equations for the projection centre and the attitude matrix of one camera, expressed as a function of the Exterior Orientation parameters of the other camera and of the asymmetric RO parameters, we have:

$$\left(R_o^1\right)^T b_1 = X_{0,2} - X_{0,1} \quad (1)$$

$$R_1^2 = R_o^2 \cdot \left(R_o^1\right)^T \quad (2)$$

where:

$X_{0,1}, X_{0,2}$  = projection centres of the cameras in the calibration reference system

$R_o^1, R_o^2$  = rotation matrices from object to camera systems

$b_1$  = baselength, at true scale, with respect to camera 1 frame, known from calibration

$R_1^2$  = rotation matrix of the asymmetric RO of the cameras, known from calibration

## 2.7 GPS AND ANCILLARY GCP SURVEY

As long as the camera records building façades, tie points will be extracted and matched seamlessly, because of the building (normally) rich texture and optimal viewing geometry. When crossing other roads, though, the perspective is much less favourable and features will therefore be difficult to find and trace along the sequence. Only a few good tie points will be available and instability of the strip orientation or even gaps in the strip may result. The same will happen when crossing squares, with gaps more likely to arise. Even finding GCP on the map or recognizing them in the images may not be easy. If the square width is large enough, maybe the GPS position becomes available; in road crossings, though, this is unlikely because the vehicle will take just a few seconds to cross it, not enough for the OTF to fix the ambiguity. If some perturbing of the traffic is acceptable during the surveys, a "Stop&Go" (Hofmann-Wellenhof et al, 2001) approach may be adopted. So when the crossroad is large enough to allow an adequate GPS satellite visibility, the vehicle will stop until the ambiguities are fixed, and then it will restart. In this case the photogrammetric positioning can be substituted by the GPS.

Due to the specific characteristics of Ground Penetrating Radar surveys, it can be reasonable to complement map information with additional GCP surveyed with traditional techniques. For example fast results can be achieved using total stations, surveying short traverses. Temporary markers can be put on the pavement to fill gaps between building blocks; well distinct points may be surveyed on building faces when blocks are so huge that the dis-

tance between GCP becomes too large (see later). This of course makes the procedure less automatic, but for such applications a ground survey can solve most of the problems.

### 3. SIMULATIONS

#### 3.1 TEST AREAS

In order to verify the accuracy of the positioning in different conditions, some simulations have been performed. The test area is Milano (Northern Italy) urban centre, where a recent city base map at 1:1000 scale is available (Bezoari et al., 2005). The elements to be tested, that influence the accuracy and the surveying characteristics, are: streets width, buildings height, building blocks length, crossroads width, obstacles (like trees or parked vehicles).

Different locations have been chosen in order to vary the critical parameters stated above.

The streets considered are:

- Site 1 – Viale Romagna (Figure 1): this road is very large, about 50 m. It consists of four lanes, two in the centre, and two on the sides; between the central and the side lanes there are flower beds with trees and sidewalks. In the section used in the tests building blocks are about 100 m long; there are two road crossing, with gaps in the building façades of 23 and 31 m. The trajectory has been simulated in the central lane.

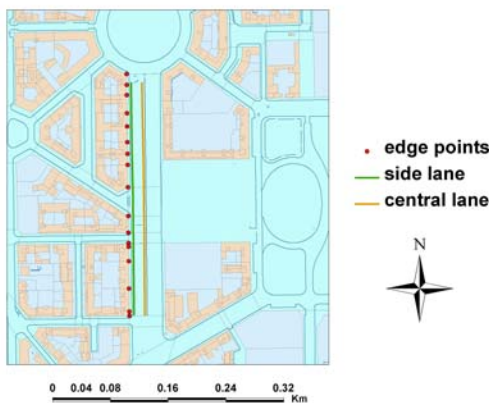


Figure 1 - Trajectories of the vehicle (on the central lane and on the side lane) and edge points for Site 1 (Viale Romagna).

- Site 2 – Via Cosimo del Fante: this road is 16 m wide. In this case there is a particularly long block, about 230 m, useful to test the drift of the photogrammetric solution.
- Site 3 – Via Felice Casati (Figure 2): this road presents two different widths, about 19 m and about 12 m, and a square recess of about 21 m facet.

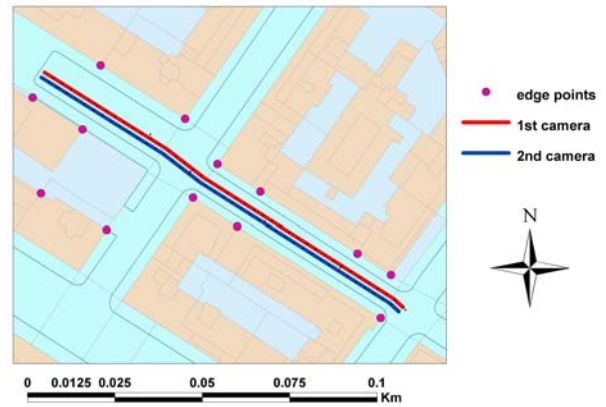


Figure 2 – Trajectories of the two cameras and edge points for Site 3 (Via Felice Casati). Notice the recess of one building on the left side of the map.

#### 3.2 SIMULATION PARAMETERS

##### 3.2.1 SCENE AND TRAJECTORY MODELING

The 3D scene of each simulated survey has been set up using the City Base Map. Tie points on buildings were represented by a grid of points with regular spacing of 4 m; some tie points on the ground, located on large crossroads, were also used.

A full GCP was established at every building corner that was recognizable on the map. Finally, the exterior orientation parameters of both cameras along the street pass have been simulated assuming the acquisition of a synchronous image pair every 3 m. The image coordinates of the tie point were generated by projecting the tie points grid on the images of the sequence.

##### 3.2.2 CAMERA, FOCAL LENGTH, IMAGE COORDINATES ACCURACY

The question of which image acquisition device (camcorder, digital reflex camera, video camera) is the best suited to this application and which set of camera parameters is optimal (if any) has not yet been addressed in detail and will therefore be considered out of the scope of this paper, whose aim is just a feasibility study.

To put some constraint on the project parameters, a fixed and relatively short base (3 m) was used (rather than, for instance, a fixed forward overlap). Therefore, for a given focal length, overlap percentage will vary with the object distance (although it must always be greater than 60% to allow tie point transfer). The short baselength is meant to assure that no significant position error will arise from the time interpolation of the position of the GPR measurements.

A second constraint, that has nevertheless been somehow released in one of the test cases, was the assumption that tie points would be selected from a fixed grid with 4 m spacing. The reason was to be somehow on the safe side as far as tie point measurement is concerned: although we are convinced, based on previous experience, that the rich texture of the building façades will easily provide several tens of tie points traced over at least three images, this rather large gap between grid points ensures that, even when texture lacks, at least some feature will reasonably be found in real images.

With those two constraints, the problem becomes getting enough points in each image; the two obvious recipes are, the closer the pass to the building, to decrease the focal length and/or to tilt the camera axis to get less ground and more building façade on the image.

Tilting the camera axis means getting tie points distributed over the whole camera format. For some range of inclinations this will improve accuracy with respect to a level axis. At highly tilted views, though, the decrease in average image scale and the increasing perspective deformation of the image will start to reverse the trend. An upper limit to the inclination is set by the need to measure all building corners: the image must always include the ground level at the bottom of the building blocks. Departing too much from the normal case may also adversely affect the image matching success ratio.

Since the GPR survey proceeds along the whole street width in several parallel passes, the distance to the façades varies. Should the image coverage be optimized, the camera inclination would have to change from pass to pass. This may be impractical and would also rule out using the relative orientation constraint, unless a stable mechanical device is used to precisely rotate the cameras at each pass.

Another way to adapt to varying distances from the façades is to use lenses with different focal length. Using too short focal length when driving in large avenues may decrease too much the resolution of the object, with a loss of accuracy and tie point density. Using a too long focal length will on the other hand decrease the accuracy of the projection centres. In the simulated tests, where the distance from the façades has been varied from 30 to 7 m we found that using a 18 mm and a 12 mm lens would be enough and that a maximum inclination of  $9^\circ$  is a good compromise.

The frame rate to sustain, assuming a maximum operating speed of 20 km per hour and a 3 m base-length, is not a problem for almost all cameras (including SLR cameras). Should the actual rate be larger (e.g. because it turns out that an optimal overlap percentage can be found) then this might rule

out some camera types or put constraints on the vehicle speed.

Since the goal is also to build a cheap system, high resolution ( $> 4$  Mpixel) digital video cameras will probably be too expensive (a limit on vehicle speed may be adopted), while the digital video cameras routinely used in Mobile Mapping Vehicles may not be the one suited for this application because of lack of resolution. For the time being, the simulations have been performed assuming a SLR 6 Mpixel camera with pixel size of 6 micrometers and a 18 mm or a 12 mm lens. Most of such cameras can now be computer controlled and can sustain a frame rate of 2 fps. Should the simultaneous acquisition of the left and right camera be required to enforce the relative orientation constraint, the cameras should be synchronized to within about 1 ms.

Exposure time (sensor integration time) must be adapted to the distance from the façades, to avoid image blurring due to motion. Indeed, should 1 pixel be the amount of theoretical image motion allowed, exposure times should be well below 1 ms for distances of 5-6 m at 20 km/h. This may also represent a limiting factor for the operating speed in narrow streets.

As far as the measurement of image coordinates is concerned, an accuracy of 1 pixel has been assumed. This is a safe figure, since image matching algorithms normally deliver accuracies of 0.5-0.7 pixels even with image textures less rich in good features than those in a urban environment.

### 3.2.3 GROUND CONTROL

As mentioned in the previous section, it is assumed that only building corners can be identified and used as 3D control points. On the other hand, points measured at different elevations along the same vertical building edge will have the same East and North coordinates of the corner at the ground. Since in most buildings the demarcation between floors is clearly distinguishable, it has been assumed that one horizontal point can be measured at every floor. Eaves can be used as height points; nevertheless, in none of the simulations this information could be exploited: due to the closeness of the vehicle trajectory to the buildings and building height, the eaves were always outside the image format. This fact has obvious consequences on the stability of the block: with a single strip the height points would all be located at ground level, in straight streets often very well aligned. If this is the case, the rotation of the whole strip along an axis parallel to camera trajectory is not sufficiently controlled by the GCP. This is reflected in large uncertainties for rotations and for the elevations of the camera centres. Running two opposite-looking strips and enforcing the rela-

tive orientation somehow helps to reduce the problem, but does not fully get rid of it: indeed, the condition number of the normal matrix of the blocks is always rather poor (better with 2 strips than with a single one). Besides, correlations between the elevations and the rotation might lead to estimation errors in the projection centres heights.

### 3.2.4 GENERATION OF SYNTHETIC IMAGE COORDINATES

In order to compute the accuracy of the camera projection centres an error propagation from the image coordinate errors to the camera projection centres coordinates has been performed. In other words, the covariance matrix of the bundle block adjustment of the strip has been computed from the known covariance matrix of the image coordinates and the jacobian of the collinearity equation system. The coordinates of the ground control points were introduced as additional pseudo-observation equations with given accuracy.

The actual computation was performed with a bundle adjustment program. Synthetic error-free observations were generated by another computer program from the tie point grid and the camera position and attitudes, computed from the vehicle trajectory.

## 4. RESULTS

### Site 1 – Viale Romagna

Both the left and right sequence have been generated; in the former the camera centres are at about 22 m from the buildings, in the latter at about 29 m. The camera principal distance is 18 mm. Two series of camera settings on the vehicle were simulated: with camera axis horizontal and with camera axis tilted above the horizontal by 9°.

The image sequence runs from South to North and is about 320 m long; Figure 3 shows the right hand side: blue dots depicts building façades; yellow dots the points actually imaged; red dots the control points; green dots the camera centres.

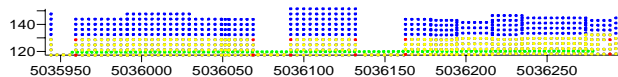


Figure 3 – Site 1: side view of the right strip: blue: building façades; yellow: image points; red: GCP; green: camera centres.

Each strip is made of 108 images; on average, in an image there are about 35 tie points; a point is measured in about 10 images. Although these figures seem to depict a strong block (the redundancy ratio is larger than 4) in fact the height ground control is

poor, unless also points on building eaves can be measured.

Five different block configurations have been simulated, varying the number of strips (2s: both left and right strips; 1s: only the right one) and the amount of ground control (A: 4 full GCP, two per strip, at ground level, at the strip ends; B: the same as A plus 4 horizontal GCP on the same vertical building edge of the points in A; C: at each building corner and road crossing one full GCP at ground level plus an horizontal GCP on the same vertical building edge). When both strips are used, the relative orientation (RO) from the left and right synchronous images is always enforced (this allows also to use the otherwise rank-deficient ground control of case A).

Table 1 summarizes the results of the 5 configurations: for each coordinate of the camera projection centres, the RMS value of the variances computed from error propagation are shown.

Block type	RMS( $\sigma^2$ )				
	2sA	2sB	2sC	1sB	1sC
(mm)	RO, 2 GCP /strip	RO, 2+2 GCP /strip	RO, all GCP	1 strip, 2+2 GCP	1 strip, all GCP
E	146	124	63	203	103
N	123	98	59	221	103
H	122	122	82	547	285

Table 1 – Site 1: average accuracies of the camera projection centre coordinates for 5 different block configurations and ground control; camera axes both horizontal.

It is apparent that, with the same ground control, combining the two strips increase the overall accuracy of the projection centres (from 50% to 100% in (E,N), from 300 to 400% in height).

Overall, horizontal accuracy requirements can be met with sufficiently dense ground control (i.e. if building blocks are not too long). The accuracy of the attitude angles (not shown in the table) is quite large for omega and kappa: more than 13° even in the best case. The reason is that a too little area of the image format is covered by tie points. By keeping the camera axis tilted above the horizon by 9° the results are significantly better in terms of attitude (0.2°) and in (E, N) while the elevation accuracy does not change very much, at least for the two strips (see Table 2).

RMS( $\sigma^2$ )					
Block type	2sA	2sB	2sC	1sB	1sC
(mm)	RO, 2 GCP /strip	RO, 2+2 GCP /strip	RO, all GCP	1 strip, 2+2 GCP	1 strip, all GCP
E	130	86	86	106	106
N	128	83	83	116	116
H	130	127	127	269	269

Table 2 – Site 1: average accuracies of the camera projection centre coordinates for 5 different block configurations and ground control; camera axes tilted by 9° above the horizon.

### Site 2 – Via Cosimo del Fante

Both the left and right sequence have been generated; in the former the camera centres are about 4.4 m from the buildings, in the latter about 12 m. There are just two blocks, each about 230 m long, on both road sides. This is a very demanding site, since ground control points are located only at the strip ends and the closeness of the trajectory to the roadside force to tilt the camera axis of the left camera (right strip) by 3° below the horizontal, in order to image the building foot (the other was kept at 9° above, instead). The block geometry is therefore rather asymmetric, because the left strip has only 28 images that could be added to the block (in the other there were too few tie points, so they were discarded) against the 68 images of the right strip. As before, five different block configurations have been simulated, varying the number of strips (2s or 1s) and the amount of ground control (A, B, C). In this case, though, ground control B and C actually coincide, since there are no other GCP except at the block ends. Results for the single strip refer to the right strip only.

Table 3 summarizes the results of the 5 configurations: for each coordinate of the camera projection centres, the RMS value of the variances computed from error propagation are shown.

RMS( $\sigma^2$ )					
Block type	2sA	2sB	2sC	1sB	1sC
(mm)	RO, 2 GCP /strip	RO, 2+2 GCP /strip	RO, all GCP	1 strip, 2+2 GCP	1 strip, all GCP
E	123	68	37	157	76
N	125	67	37	148	67
H	118	119	79	345	210

Table 3 – Site 2: average accuracies of the camera projection centre coordinates for 5 different block configurations and ground control; camera axes tilted by 9° above the horizon.

The accuracy figures with respect to the other sites are less good, because of the sparse ground control. It is interesting to notice that, even if one of the two strips is somehow fragmented in patches, it still helps to make the block more stable, especially in elevation but also in horizontal: this was not the case in the other sites.

### Site 3 – Via Casati

Both the left and right sequence have been generated; in both the minimum distance of the camera centres from the buildings is about 7 m; the maximum distance (left strip) is about 30 m. Strip length is just 120 m; there are 38 images/strip and 230 tie points. The camera focal length is 12 mm and the camera axes are tilted 9°. The ratio of equations to unknowns is 3.1; on average, there are 11 pts/photo and 3.6 rays/pt.

The same five different block configurations of previous cases have been simulated, varying the number of strips (2s or 1s) and the amount of ground control (A, B, C). When both strips are used, the relative orientation (RO) from the left and right synchronous images is always enforced.

Table 4 summarizes the results of the 5 configurations: for each coordinate of the camera projection centres, the RMS value of the variances computed from error propagation are shown.

RMS( $\sigma^2$ )					
Block type	2sA	2sB	2sC	1sB	1sC
(mm)	RO, 2 GCP /strip	RO, 2+2 GCP /strip	RO, all GCP	1 strip, 2+2 GCP	1 strip, all GCP
E	118	92	47	93	84
N	115	96	47	95	86
H	115	118	77	137	137

Table 4 – Site 4: average accuracies of the camera projection centre coordinates for 5 different block configurations and ground control; camera axes tilted by 9° above the horizon.

With respect to Site 1, the accuracy figures show that the shorter strip length (i.e. the shorter average distance of GCP) partly compensates for the smaller number of tie points per image and rays per point. A shorter focal of 12 mm is necessary, due to the closeness to the building façades, since the grid spacing of 4 m has been maintained. As in Site 1, with all GCP, the gain in accuracy when using both strips is about 100%.

## 5. CONCLUSIONS AND PERSPECTIVES

In this paper the first results of a feasibility study of a positioning system that employs photogrammetry and GPS are presented. A series of preliminary simulations have been carried out to assess the feasibility of using photogrammetry and ground control points from maps to locate a vehicle in a urban environment. Three test sites have been investigated: the results suggest that the objective can be achieved, provided that GCP can be recognized on images at intervals not much larger than 200-250 m, with precisions below 10 cm.

At the design level, optimization of several system components and surveying parameters is still needed. Much more still remain to be proven in terms of operational feasibility of the system, before moving to an implementation stage: many practical problems will arise from the great variety of situations in the urban environment. GPS surveying will be conducted to evaluate which is the minimum size of crossroads to permit the positioning in order to clarify the specifications of the procedure. Moreover, L1 receivers will be used to test if, in these applications, low cost receivers can give accuracies comparable with those attainable with L1/L2 instruments.

## ACKNOWLEDGEMENTS

The authors wish to thank the Milan Municipality and the Province of Milan (Sistema Informativo dei dati Ambientali) for the use of the 1:1000 City Base Map.

## REFERENCES

Bezoari G., Monti C., Selvini A. (2005), *La cartografia numerica della città di Milano: interventi per il collaudo*, Rivista dell'agenzia del Territorio, 2/2005, pp. 58-76.

Blachut T.J., Chrzanowski A., Saastamoinen J.H., (1979), *Urban Surveying and Mapping*, Springer-Verlag, New York.

Chaplin B.A., Chapman M.A. (1998), *A procedure for 3D motion estimation from stereo image sequences for a mobile mapping system*. International Archives of Photogrammetry and Remote Sensing, Vol. 32(3W1).

Chaplin B.A. Chapman, M.A. (2001), *Collection of roadway asset data along an obstructed urban corridor*. CD-ROM Proc. of the 3rd International Symposium on Mobile Mapping Technology, Cairo, Egypt. Session 2: 5 pages.

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., Vol. 24(3), pp. 81-95.

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.

Hartley R., Zisserman A. (2000), *Multiple View Geometry in computer vision*. Cambridge University Press.

Harris C., Stephens M. (1987), *A combined corner and edge detector*, Proc. Alvey Conference, pp. 189-192.

Hofmann-Wellenhof B., Lichtenegger H., Collins J. (2001), *GPS - Theory and Practice*, Springer, Wien.

Kim S., Choi K., Lee S., Choi J., Hwang T., Jang B., Lee J. (2004), *A Bimodal Approach for Land Vehicle Localization*, ETRI Journal, Vol. 26 (5), pp. 497-500.

Petovello, M.G., Cannon M.E., Lachapelle G., (2003), *Quantifying improvements from the integration of GPS and a tactical grade INS in high accuracy navigation applications*. Proc. of the ION NTM-2003, January 22-24, Anaheim, Canada, pp. 454-465.

Pilu M. (1997), *Uncalibrated stereo correspondences by SVD*, Proc. CVPR , pp. 261-266.

Tao C., Chapman M.A., El-Sheimy N., Chaplin B. (1999), *Towards automated processing of mobile mapping image sequences*. International Archives of Photogrammetry and Remote Sensing, Vol. 32(2W1).

Zhang H., Petovello M.G., Cannon M.E. (2005), *Performance Comparison of Kinematic GPS Integrated with Different Tactical Level IMUs*. CD-ROM Proc. of NTM 2005, San Diego, The Institute of Navigation, 12 pages.