ROAD SURVEY BY KALMAN FILTER RECTIFICATION OF IMAGE SEQUENCES ACQUIRED WITH A MONOSCOPIC LOW-COST MMS

Domenico Visintini

Department of Georesources & Territory, University of Udine, via Cotonificio, 114 I-33100 Udine, Italy
domenico.visintini@uniud.it

KEY WORDS: Sequences, Dynamic, Orientation, Matching, Rectification.

ABSTRACT
This paper proposes a method for a mosaic digital representation of a road by means of a Kalman filter based rectification of a monoscopic image sequence acquired with a low-cost MMS without INS sensors. The frontal images of the road acquired by a pointing ahead-down camera are rectified onto the road surface by means of a homographic transformation among image and ground planes. As basic idea, the "external/homographic" parameters constitute the unknown "state vector" of a Kalman filter pseudo-dynamic model, describing its evolution with respect to the covered distance, with observation equations given from image coplanarity and collinearity conditions and GPS kinematic measures. To make easy the relative orientation among images, an original Kalman filter matching is exploited: once selected some lane tract points in the "first" image, the homologous points in the "second" image are so automatically detected. Exploiting only near points in the lower part of the images, a high-scale optical model with a strong geometrical consistency is achieved: therefore, a reliable rectification of the image sequence can be expected.
This method has been implemented in a Matlab language program making possible the rectification of the image sequence by a digital resampling with automatic mosaicking onto finite planes modeling the road surface.
First numerical experiments with real data of our MMS prototype have given satisfactory results here described and analyzed.

1. INTRODUCTION

In these lasts years, the photogrammetric survey technology by means of a Mobile Mapping System (MMS), that is terrestrial vehicles equipped with GPS, INS, CCD on other integrated sensors, presents an increasing development and interesting applications, mainly to 3D-survey the geometry and the pavement characteristics of roadways.
This advanced technique, involving completely different measuring sensors (satellite, inertial, odometer, imaging, etc), is characterized by lots of analytical, methodological, and technological sub-aspects. Obviously, this paper only briefly hints to some of them: the interested reader can investigate in depth within the wide available literature. One can start just from the proceedings of the previous International Symposiums on Mobile Mapping Technology (Ohio State University, 1995; Li and Murai, 1999; El-Sheimy, 2001; Tao, 2004) and from the recent volume of Tao and Li (2007).
As general expectation on the technology, knowing the high level of technological instruments and analytical models involved in acquiring and processing MMS data, a high efficiency survey in terms of correctness, accuracy, reliability, completeness, and productivity is realistically expected.

This work of this paper fits in the philosophy of our researches (1996-2003) devoted to envelop/asses a simplified MMS equipped with only one CCD camera, a GPS receiver, an odometer, but without INS sensors (section 2).
The so-called "coplanarity condition", linking homologous points of successive images, is the essential analytical tool to actually achieve a survey with such low-cost MMS (section 3)
The "external/homographic" parameters constitute the unknown "state vector" of a Kalman filter pseudo-dynamic model, describing its geometric evolution in the sequence (section 4).
For the orientation problem (subsection 4.1), the filter "state equations" are obtained by means of cubic spline functions. The "updating observations" are mainly provided from image coordinates of homologous road lane points in successive images. Other two kinds of observations are used: kinematic GPS measures acquired in the shot instants and, if reported in a digital map, road edge corner submitted to the collinearity condition among image and cartographic coordinates.
Also a matching Kalman filter based is applied (subsection 4.2): once selected some lane points in the "first" image, the homologous points in the "second" one are automatically found. Thus for each couple of images, the orientation and the matching problems are recursively and dynamically solved (section 5). Using only points in the lower part of the images, far to the vanishing point, a high-scale optical model with strong geometrical consistency is achieved.
The frontal images of the road acquired by a pointing ahead-down camera are rectified onto the road surface by a homographic transformation among image and ground planes (section 6). Thanks to the implementation in Matlab language, the image lower parts are rectified onto a horizontal plane by a digital resampling with automatic mosaicking.
Two real CCD image sequence have been processed and rectified with encouraging numerical results (section 7).

2. SURVEYING WITH A SIMPLIFIED MMS

As well-known, the 3D surveying of every point placed "within" a road can be done by means of the MMS sensors thanks to the below three steps of coordinate transformations:

1. Rotation and translation from a given mapping frame to the GPS/INS frame, by means of the position and attitude measured by GPS/INS post-processing techniques.
2. Rotation and translation from the GPS/INS-frame to each CCD frame, by means of the position and attitude parameters suitably measured in the MMS calibration.
3. Reverse projection (space resection) from, at least, two different CCD frames to the point position, by means of digital photogrammetric techniques.

Inside of this technology of surveying, our research group has developed an original analytical model to determine the

external orientation of the CCD frames, i.e. to simultaneously solve the steps 1. and 2.. It extends the usual "direct method" (e.g. El-Sheimy, 1996) GPS/INS measurements based, since the CCD position and attitude is also achievable by an "indirect method". In fact, this traditional photogrammetric process makes use of an image sequence also for the orientation and not only for the 3D-surveying, well displayed from the aerial triangulation procedure. Moreover this MMS configuration has low costs since, in principle, GPS/INS sensors are not required.
In order to fix the ideas, in our researches we have tried to envelop and assess a simplified low-cost monoscopic MMS with only one CCD camera, a GPS receiver, an odometer, but without INS sensors. This choice has some operative drawbacks to overcome by proper analytical techniques. In this sense:

- the requirement of stacks of control points along the route can be avoided if 3D digital map points are available;
- the influence of the precision of such points on the image orientation accuracy can be handled with a "mixed model of estimation and prediction" (Dermanis, 1990);
- the not optimal estimation by classical (static) analytical models can be replaced by a Kalman filter "pseudo-dynamic" model for a computationally efficient process.

3. COPLANARITY CONDITION FOR A SIMPLIFIED MMS

Throughout this section, it is explained how the so-called "coplanarity condition", involving homologous points of two consecutive images, is extremely precious to actually carry out the surveying with a monoscopic low-cost MMS without INS.

Figure 1. First and second image of a monoscopic sequence.

Let consider a "first" image acquired at the instant *t* and a "second" one acquired at the *(t+1)* one, as those frontally acquired with a pointing ahead-down camera in our experiments and reported in Figure 1. The following equation states the coplanarity condition between the two 3D segments joining each shot point with the corresponding image point (see e.g. Crosilla and Visintini (1998) for the analytical proof):

$$\chi_{t+1}\mathbf{q}_{t+1} = 0 \qquad (1)$$

Equation (1) gathers together, in its two terms, the image coordinates with the orientation parameters, in fact:

- $\chi_{t+1} = \begin{bmatrix} x_{t+1}x_t & x_{t+1}y_t & -x_{t+1} & y_{t+1}x_t & y_{t+1}y_t & -y_{t+1} & -x_t & -y_t & 1 \end{bmatrix}$ is one row (for each couple of homologous points) of a nine column matrix, with $x_t, y_t, x_{t+1}, y_{t+1}$ normalized image coordinates onto the CCD consecutive frames;
- $\mathbf{q}_{t+1}$ is a nine row vector obtained by stacking the rows (and transposing) of the matrix $\mathbf{Q}_{t+1}$ involving the external orientation image parameters as:

$$\mathbf{Q}_{t+1} = \mathbf{R}_{t+1}\begin{bmatrix} 0 & H_{t+1}-H_t & -N_{t+1}+N_t \\ -H_{t+1}+H_t & 0 & E_{t+1}-E_t \\ N_{t+1}-N_t & -E_{t+1}+E_t & 0 \end{bmatrix}\mathbf{R}_t^{\,T}$$

where $\mathbf{R}_t$ and $\mathbf{R}_{t+1}$ are the CCD rotation matrices, while $\mathbf{r}_t = \begin{bmatrix} E_t & N_t & H_t \end{bmatrix}^T$ and $\mathbf{r}_{t+1} = \begin{bmatrix} E_{t+1} & N_{t+1} & H_{t+1} \end{bmatrix}^T$ are the coordinates of the CCD positions, each one in the two respective shot instants.

It should be now clear as equation (1) is fundamental for a monoscopic MMS without INS: it contains either image coordinates or orientation parameters of successive frames of an image sequence. By means of (1), it is indeed possible to solve:

- "External orientation problem": if $\mathbf{r}_t, \mathbf{R}_t$ are known (that is the first image is oriented) and $x_t, y_t, x_{t+1}, y_{t+1}$ are measured, the external orientation $\mathbf{r}_{t+1}, \mathbf{R}_{t+1}$ of the second image can be estimated, simply by considering enough homologous points in the equation (1).
- "Epipolar matching problem": if instead $\mathbf{r}_t, \mathbf{R}_t, \mathbf{r}_{t+1}, \mathbf{R}_{t+1}$ are known (that is both images are oriented) and x_t, y_t are chosen, a relationship (epipolar line) joining together the image coordinates x_{t+1}, y_{t+1} of the homologous point can be achieved starting again from equation (1) and suitably exploiting it to automatically find such a point.

In more detail, to estimate the external orientation parameters $E_{t+1}, N_{t+1}, H_{t+1}, \omega_{t+1}, \phi_{t+1}, \kappa_{t+1}$ of the second image, it is not difficult to retrieve some homologous points to directly apply the equation (1), may be rather a problem to automatically find them! From the geometrical point of view, the coplanarity condition allows the relative orientation with respect to the already externally oriented first image. This is exactly the same contribution given from the INS gyros and accelerometers, if used without GPS geometrical constraints. Also for our orientation, the accuracy decreases for the "drift effect" as long as we use images without any control point constraint.

For the "epipolar matching problem" instead, equation (1) has to be rewritten for a couple of homologous points using pixel values relating to the template and patch window centres (**T** and **P**) involved in such a problem. In this way, it becomes a straight-line equation of (epipolar) constraint on $\mathbf{P} \equiv \begin{bmatrix} k_P, h_P \end{bmatrix}$:

$$h_P = \frac{D}{N}k_P + \frac{1}{N\lambda}\left(Nx_c - Dy_c + C\right) \qquad (2)$$

where:

- $N = x_t(\mathbf{q}_{t+1})_1 + y_t(\mathbf{q}_{t+1})_2 - (\mathbf{q}_{t+1})_3$
- $D = x_t(\mathbf{q}_{t+1})_4 + y_t(\mathbf{q}_{t+1})_5 - (\mathbf{q}_{t+1})_6$
- $C = x_t(\mathbf{q}_{t+1})_7 + y_t(\mathbf{q}_{t+1})_8 - (\mathbf{q}_{t+1})_9$
- $x_c = x_0/c$ normalized coordinates of the pixel frame origin along x-direction;
- $y_c = y_0/c$ normalized coordinates of the pixel frame origin along y-direction;
- $\lambda = 25{,}4/c \cdot r$ image scale-factor [1/pixel], where r is the image sensor equivalent resolution [dpi].

It must be stressed that, working with a simplified MMS, a reliable area based image matching procedure is practically mandatory. In fact, a lot of homologous points have to be detected on successive images and used for "photogrammetric and not inertial relative (external) orientation". Once the user has chosen a point **T** in the first image, the automatic point detection of the homologous point **P** in the second image is then essential to actually achieve this alternative orientation, without increase to much the user job for the point collimation.

4. DYNAMIC ANALYSIS OF AN IMAGE SEQUENCE BY KALMAN FILTERS MODELS

Explained in the previous section the basic geometry of a monoscopic image sequence, here the algorithmic solutions of the orientation and matching problems are described. As main concept, the unknowns of these problems constitute "state vectors" of specific Kalman filter models (Visintini, 2001c), describing their geometrical evolution instead of the temporal one, as usually done in processing dynamic phenomena.
In addition, since the orientation of each image practically solves the problem of its rectification by a homography, as will be explained in section 6, the Kalman filter solution of the orientation parameters implicitly regards the homographic ones.

4.1 Kalman filter orientation of an image sequence

To obtain the external orientation parameters of the "second" CCD image acquired at *(t+1)* time:

$$\mathbf{x}_{t+1} = [E_{t+1} \quad N_{t+1} \quad H_{t+1} \quad \omega_{t+1} \quad \varphi_{t+1} \quad \kappa_{t+1}]^T$$

considered as a state vector, a Kalman filter model has been applied (Visintini, 2001c). Since the MMS motion has a "regular trajectory (in local sense) with an irregular course (in global sense)", a transition matrix $\mathbf{\Phi}_t$ describing approximately the evolution of $\mathbf{x}_t$ can be suitably obtained. The filter "state equations" are in fact achieved by applying cubic spline functions on the East, North and Height coordinate of points in a cartography approximately describing the MMS route (Crosilla and Visintini, 1998).
The linearized form of the state equation is then:

$$\delta\mathbf{x}_{t+1} = \mathbf{\Phi}_t \delta\mathbf{x}_t + \boldsymbol{\mu}_{t+1} \qquad \text{with } \boldsymbol{\mu}_{t+1} \sim N(0, \mathbf{\Theta}_{t+1}) \quad (3)$$

Three kinds of linearized observations are exploited:

1. Image coordinates $\mathbf{l_1}$ of couples of homologous points (e.g. the road lane corners) visible in successive images submitted to the coplanarity condition (1), here rewritten as:

$$\mathbf{h}\{\bar{\mathbf{x}}_{t+1}, \mathbf{l_1}\} + \mathbf{C_1}\delta\mathbf{x}_{t+1} + \mathbf{D_1}\mathbf{v_1} = 0 \quad (4)$$

2. Image $\mathbf{y_2}$ and absolute $\mathbf{X}$ coordinates of visible 3D map points submitted to the well-known collinearity condition:

$$\mathbf{g}\{\bar{\mathbf{x}}_{t+1}, \mathbf{X}\} + \mathbf{C_2}\delta\mathbf{x}_{t+1} + \mathbf{E_2}\mathbf{s} + \mathbf{v_2} \quad (5)$$

3. CCD frame coordinates $\mathbf{m}$ obtained by kinematic GPS measures and taking into account known eccentricity $\mathbf{a}_{ccd}^{gps}$ and rotation $\mathbf{R}_{ccd}^{gps}$ between CCD and GPS frames:

$$\mathbf{r}_{gps} \cong \mathbf{m}\{\bar{\mathbf{x}}_{t+1}\} + \mathbf{C_3}\delta\mathbf{x}_{t+1} + \mathbf{v_3} \quad (6)$$

where:

- $\mathbf{h}\{\bar{\mathbf{x}}_{t+1}, \mathbf{l_1}\}, \mathbf{g}\{\bar{\mathbf{x}}_{t+1}, \mathbf{X}\}, \mathbf{m}\{\bar{\mathbf{x}}_{t+1}\}$ equations values computed with an approximated value $\bar{\mathbf{x}}_{t+1}$;
- $\mathbf{C_1}, \mathbf{D_1}, \mathbf{C_2}, \mathbf{E_2}, \mathbf{C_3}$ partial derivatives matrices of the equations with respect to the unknowns $\mathbf{x}_{t+1}$ and $\mathbf{s}$;
- $\mathbf{s}$ predicted coordinate increments of 3D-map points, starting from the stochastic values $\mathbf{X}$ of the digital mapping.

Gathering equations (4), (5) and (6) together, the following observation system can be written:

$$\mathbf{b}_{t+1} = \mathbf{C}\delta\mathbf{x}_{t+1} + \mathbf{E}\mathbf{s} + \mathbf{D}\mathbf{v} \quad (7)$$

where:

$$\mathbf{b}_{t+1} = \begin{bmatrix} \mathbf{h}\{\bar{\mathbf{x}}_{t+1}, \mathbf{l_1}\} \\ \mathbf{y_2} - \mathbf{g}\{\bar{\mathbf{x}}_{t+1}, \mathbf{X}\} \\ \mathbf{r}_{gps} - \mathbf{m}\{\bar{\mathbf{x}}_{t+1}\} \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} -\mathbf{C_1} \\ \mathbf{C_2} \\ \mathbf{C_3} \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} 0 \\ \mathbf{E_2} \\ 0 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} -\mathbf{D_1} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \quad \mathbf{v} = \begin{bmatrix} \mathbf{v_1} \\ \mathbf{v_2} \\ \mathbf{v_3} \end{bmatrix}$$

Finally, the Kalman filter iterative solution of state (3) and observation (7) equations gives out the prediction of the image external orientation parameters as:

$$\tilde{\mathbf{x}}_{t+1} = \mathbf{\Phi}_t \tilde{\mathbf{x}}_t + \mathbf{K}_{t+1}\mathbf{b}_{t+1} \quad (8)$$

where:

$$\mathbf{Q}_{\delta\tilde{\mathbf{x}}_{t+1}} = \mathbf{\Theta}_{t+1} + \mathbf{\Phi}_t \mathbf{Q}_{\mathbf{x}_t} \mathbf{\Phi}_t^T = \mathbf{Q}_{\tilde{\mathbf{x}}_{t+1}}$$

$$\mathbf{Q}_{\mathbf{b}_{t+1}} = \mathbf{D}\mathbf{Q}_{\mathbf{v}}\mathbf{D}^T + \mathbf{E}\mathbf{Q}_{\mathbf{XX}}\mathbf{E}^T + \mathbf{C}\mathbf{Q}_{\delta\tilde{\mathbf{x}}_{t+1}}\mathbf{C}^T$$

$$\mathbf{K}_{t+1} = \mathbf{Q}_{\delta\tilde{\mathbf{x}}_{t+1}}\mathbf{C}^T\mathbf{Q}_{\mathbf{b}_{t+1}}^{-1} \quad \text{Kalman gain matrix}$$

Discerning this analytical orientation model, it can be defined as an "indirect & direct method" since it exploits for the goal either indirectly image coordinates or directly GPS orientation parameters. From the photogrammetric point of view, the available 3D map points submitted to the collinearity condition (5) only opportunely fix the datum of the "optical model" constituted by the image sequence. This last is mainly built thanks to coplanarity condition (4) and, using a lot of homologous (whatever) points, the resulting optical model has a strong geometrical internal (relative) auto-consistency. Furthermore, the exploiting of road points near to the MMS and far from the vanishing point warrants a very high 3D-accuracy.
Concerning instead absolute accuracy, a fine benefit is assured from the "mixed model" (Dermanis, 1990) applied for the unknowns in the equations (5) and (7). It allows simultaneously the orientation estimation and the prediction of 3D-coordinates starting from stochastic values stored in map. In such a way, also for external orientation, the high scale of MMS images is well exploited, enforcing the optical model onto more precise 3D point positions, keeping into account the GPS constraint (6).

4.2 Kalman filter matching in an image sequence

To exploit the coplanarity-epipolarity condition for the matching problem also if the external orientation parameters are "not available" in equation (2), a Kalman filter approach has been again pursued (Visintini, 2001a). In this way, we take advantage of the epipolarity contribution in whatever condition, even to automatically find homologous points to externally orient the second (not oriented) image! That is mathematically possible introducing in equation (2) the approximated orientation values computed by the cubic spline functions, as just described in subsection 4.2.
Anyway, for the pseudo-dynamic solution of the matching problem, the incremental state unknowns are alternatively:

$$\kappa = \delta\mathbf{k} = \mathbf{k} - \bar{\mathbf{k}} \qquad \text{with } \kappa \sim N(0, \Sigma_{\mathbf{k}})$$

$$\eta = \delta\mathbf{h} = \mathbf{h} - \bar{\mathbf{h}} \qquad \text{with } \eta \sim N(0, \Sigma_{\mathbf{h}})$$

where:

$$\mathbf{k} = [a \quad b \quad c]^T \qquad \mathbf{h} = [d \quad e \quad f]^T$$

with a, b, c, d, e, f, affine transformation parameters from the template to the patch window centered on the homologous points **T** and **P**. The incremental state unknown is then a three row vector $\boldsymbol{\kappa}$ switching in $\boldsymbol{\eta}$ (and again in $\boldsymbol{\kappa}$) during the iterations and containing three of the six coefficients of the affine geometric window transformation.
The maximum likelihood condition on every pixel grey-values of the template and patch windows yields to the following linearized alternate systems of observation equations:

$$\mathbf{b} = \mathbf{A}\pi + \mathbf{E}\eta + \mathbf{v} \qquad (9.1)$$

$$\mathbf{b} = \mathbf{A}\pi + \mathbf{F}\kappa + \mathbf{v} \qquad (9.2)$$

where:

$$\mathbf{b} = -\mathbf{f}(i,j) + \left[r_0 + r_1 \mathbf{g}(ai+bj+c, di+ej+f)\right]_{\bar{\mathbf{p}},\bar{\mathbf{k}},\bar{\mathbf{h}}}$$

$$\pi = \delta\mathbf{p} = \mathbf{p} - \bar{\mathbf{p}} \qquad \text{with } \pi \sim N(0, \Sigma_{\mathbf{p}})$$

with:

$$\mathbf{p} = \begin{bmatrix} r_0 & r_1 \end{bmatrix}^T$$

- r_0, r_1 : radiometric transformation parameters on grey-values from the template to the patch window;
- **A**, **E**, **F**: partial derivative matrices of the grey disparity respect to the **p**, **h** and **k** matching unknowns.

As mentioned before, the epipolar line obtained from equation (2) can be more or less correct depending from the orientation values accuracy of $\mathbf{r}_{t+1}, \mathbf{R}_{t+1}$, and also, within the same image couple, can be truthful for a point and untruthful for another one. For such reason, the equation (2) has to be handled in stochastic way, exploiting it as a state equation, so to take into account possible errors in the straight-line coefficients. The following alternate state equations in the $\boldsymbol{\eta}$,$\boldsymbol{\kappa}$ vector space have been suitably obtained, with very simple transition matrices:

$$\eta = \frac{D}{N}\mathbf{I}\kappa + \nu_\eta \qquad \text{with } \nu_\eta \sim N(0, \Sigma_{\nu_\eta}) \quad (10.1)$$

$$\kappa = \frac{N}{D}\mathbf{I}\eta + \nu_\kappa \qquad \text{with } \nu_\kappa \sim N(0, \Sigma_{\nu_\kappa}) \quad (10.2)$$

The Kalman filter iterative solution of the state (10.1) and observation (9.1) equations gives now:

$$\hat{\mathbf{p}} = \bar{\mathbf{p}} + \hat{\pi} = \bar{\mathbf{p}} + \left[\mathbf{A}^T\mathbf{Q}_{\mathbf{b}}^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^T\mathbf{Q}_{\mathbf{b}}^{-1}\mathbf{b} \qquad (11.1)$$

$$\tilde{\mathbf{h}} = \bar{\mathbf{h}} + \tilde{\eta} = \bar{\mathbf{h}} + \mathbf{G}_\eta\left[\mathbf{b} - \mathbf{A}\hat{\pi}\right] \qquad (11.2)$$

where:

$$\mathbf{Q}_{\bar{\eta}} = \mathbf{Q}_{\nu\eta} + \frac{D^2}{N^2}\mathbf{Q}_\kappa = \mathbf{Q}_{\bar{\mathbf{h}}}$$

$$\mathbf{Q}_{\mathbf{b}} = \mathbf{Q}_{\mathbf{v}} + \mathbf{E}\mathbf{Q}_{\bar{\eta}}\mathbf{E}^T$$

$$\mathbf{G}_\eta = \mathbf{Q}_{\bar{\eta}}\mathbf{E}^T\mathbf{Q}_{\mathbf{b}}^{-1} \text{ Kalman gain}$$

Every step done for vector $\boldsymbol{\eta}$ can be exactly redone to predict vector $\boldsymbol{\kappa}$ and estimate again vector $\boldsymbol{\pi}$ by exploiting now equations (10.2) and (9.2) with solutions analogous to (11). Once this new prediction is carried out, an updated value of the first unknown becomes available for a second prediction $\tilde{\tilde{\eta}}$ of $\boldsymbol{\eta}$ (or $\tilde{\tilde{\kappa}}$ of $\boldsymbol{\kappa}$). Such recursive predictions will be thus repeated as long as the solution convergence is satisfied. The patch centre **P** hence alternatively translates along rows or columns only (as a staircase series), but staying in the epipolar line surroundings, until the final position is reached. The satisfying numerical results in matching CCD image sequences acquired by our simplified MMS are described in Visintini (2001b).

5. FLOW CHART OF THE DYNAMIC ANALYSIS OF AN IMAGE SEQUENCE

In the flow chart of the dynamic analysis (DA) of a sequence of *n* CCD images, let consider now having to orient the *(s+1)*-th image, operation possible once the previous *s*-th image has been already oriented. The spline orientation values of the transition matrix in (3) are now better of those computed via a mapping, since the first *s* shot points have been exactly evaluated yet. In this way, the cubic spline interpolation will better interpolate the remaining *n-(s+1)* points the more *(s+1)* leans to *n*.
Following the flow chart reported in Figure 2, once a first point (template center) has been manually selected in the *s*-th image:

1. The Kalman filter matching computes by (11) the **h**, **k** and **p** unknowns, namely automatically extracts the homologous point (patch center) in *(s+1)*-th image.
2. The Kalman filter orientation uses the so found couple of homologous points as coplanarity observation and computes by (8) the $\mathbf{x}_{t+1}$ unknown external orientation.

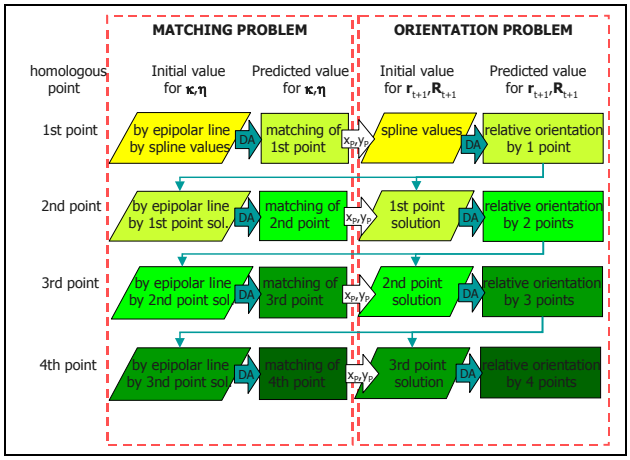

Figure 2. Flow chart for the matching and orientation problems.

Steps 1. and 2. are repeated for a second, a third, and so on point, but each time with better improved orientation approximated values $\bar{\mathbf{x}}_{t+1}$ as well as a new point is introduced.
In this way, each time the Kalman filter estimations and predictions either for the matching problem or orientation one will be numerically more stable (for this reason the green coloring grows in Figure 2). Moreover, introducing a new point, the solutions will require each time a smallest number of iterations, since the stochastic contribution of the state equations becomes biggest with respect to the observation ones.
Of course, when the search of homologous points among *s*-th and *(s+1)*-th images and the external orientation of the *(s+1)*-th image have been solved, the dynamic analysis involves the following *(s+2)*-th image and its homologous points, and so on until to orient the last *n*-th image.
Summarizing, the matching and orientation pseudo-dynamic problems are strictly correlated together and can be here preferred the comprehensive expression of "matching for the orientation" as the global aim of the dynamic analysis. Remembering that image orientation means, de facto, image

homographic rectification, the global dynamic analysis here described involves all the sequence issues.
In spite of this, the external orientation of the whole MMS image sequence makes clearly possible to compute the $\mathbf{r}_i$ position of each point of our interest. Using precise known map points and/or GPS data ensure best accurate external orientation $\mathbf{r}_t, \mathbf{R}_t$ and off course best accurate coordinate $\mathbf{r}_i$, given by:

$$\mathbf{r}_i = \mathbf{r}_t + S_i \mathbf{R}_t \mathbf{r}_i^{pix} \tag{12}$$

Equation (12) considers a reverse projection from, at least, two different images, but in this form it considers pixel coordinates $\mathbf{r}_i^{pix}$ of one image and a point-varying image scale value S_i: in truth, this value is only implicitly computed during the photogrammetric space resection computations.

6. RECTIFICATION OF IMAGE SEQUENCES

Throughout the previous sections the concept of the Kalman filter orientation of monoscopic images has been shown: now the paper philosophy of the digital rectification of such frontal images is finally delineated. As first general consideration, this 2D or 2,5D raster representation is characterized by a fully automation in its production and such a reason makes this kind of output more popular then a 3D vector representation.
To obtain a photo-mosaic of images ortho-projected or rectified, one has to know (for each image):

- the image external orientation, attained as shown up to now;
- the surface reconstruction, i.e. the Digital Surface Model (DSM) for complex geometry requiring differential rectification (orthophoto) or just the plane displacement for simplest case (rectification).

In any case, geometrical information about the surface where remap (texturizing) the images has to be known and these data have to be acquired in some way by means of, in order of complexity, topographic, photogrammetric or (today greatly emerging) laser scanning surveying.
A part methodological and practical reasoning about integration of survey techniques, here the "alone" MMS photogrammetric technology and the image projection onto the road surface are only considered, sending for example to Visintini (2003) for the projection over the façades of the buildings facing the road.
In this context, the geometry of the road surface can be modelled with DSMs of different levels of detail and accuracy, anyway keeping in mind that such a DSM has to be defined in the same (mapping) reference frame of the MMS image sequence. In other words, the relative congruence among CCD image and road surface has to be fixed, that is a sort of "relative orientation" between the image plane and the road surface.
Equation (12) can be applied to measure a certain number of 3D points allowing building the road DSM in form of a Triangular Irregular Network (TIN) more or less detailed with vertexes in such a points. However, the qualitative shape of a DSM of a road is generally a quasi horizontal plane. In any case, the projection over this DSM is formally a mosaicked orthophoto.
Some simplifying hypothesis can be however done, as that the road surface is composed by two planes slanting of some ‰ from the axis towards the borders, so defining a triangular section according to the constructive requirements. This is the shape for rectilinear road tract, varying in the clothoid up to a unique transversal inclination in a bend tract. In this way, the image projection is simplified as a rectification onto these planes. In Roncella and Forlani (2006) the further simplification of a road surface as a unique plane is not recommended since this can introduce up to decimetric errors.

As definitive working hypothesis, in this paper we consider to rectify each monoscopic image of the sequence over a "small" plane constituted from the part of the road acquired just in front of the MMS, being or not this plane transversally slanting. The road is then a sort of "mosaic of finite planes". Nevertheless, it must be recalled as the raster representation of a road survey is anyway done onto a horizontal plane. This means, as ultimate working assumption, that it is sufficient to share each image projection over different planes, e.g. one for the right track and one for the left one.
To finally perform such projection-rectification a homographic transformation is applied between two planes (Hartley and Zisserman, 2000): in robotic vision literature it is known as "inverse perspective mapping" (Mallot et al., 1991) and it is applied for both tracks since the goal is not a precise surveying but the obstacle detection instead (e.g. Bertozzi and Broggi, 1998). Our frontal images of the road acquired by a pointing ahead-down camera are then resampled according to:

$$E = \frac{a_{t1}x + b_{t2}y + c_{t3}}{u_t x + v_t y + 1} \qquad N = \frac{a_{t2}x + b_{t2}y + c_{t2}}{u_t x + v_t y + 1} \tag{13}$$

where the eight parameters of the homography are expressed as combinations of the six parameters of the image external orientation as follows:

$$a_{t1} = \frac{1}{R_{t33}c}\left(R_{t11}H_t - R_{t13}E_t\right)$$

$$b_{t1} = \frac{1}{R_{t33}c}\left(R_{t21}H_t - R_{t23}E_t\right)$$

$$c_{t1} = E_t - \frac{R_{t31}}{R_{t33}}H_t$$

$$a_{t2} = \frac{1}{R_{t33}c}\left(R_{t12}H_t - R_{t13}N_t\right)$$

$$b_{t2} = \frac{1}{R_{t33}c}\left(R_{t22}H_t - R_{t23}N_t\right)$$

$$c_{t2} = N_t - \frac{R_{t32}}{R_{t33}}H_t$$

$$u_t = -\frac{R_{t13}}{R_{t33}c}$$

$$v_t = -\frac{R_{t23}}{R_{t33}c}$$

with:

$$R_{t11} = \cos\phi_t \cos\kappa_t$$
$$R_{t12} = \cos\omega_t \sin\kappa_t + \sin\omega_t \sin\phi_t \cos\kappa_t$$
$$R_{t13} = \sin\omega_t \sin\kappa_t - \cos\omega_t \sin\phi_t \cos\kappa_t$$
$$R_{t21} = -\cos\phi_t \sin\kappa_t$$
$$R_{t22} = \cos\omega_t \cos\kappa_t - \sin\omega_t \sin\phi_t \sin\kappa_t$$
$$R_{t23} = \sin\omega_t \cos\kappa_t + \cos\omega_t \sin\phi_t \sin\kappa_t$$
$$R_{t31} = \sin\phi_t$$
$$R_{t32} = -\sin\omega_t \cos\phi_t$$
$$R_{t33} = \cos\omega_t \cos\phi_t$$

Thanks to these relationships, each externally oriented MMS images can be directly rectified onto "horizontal" planes composing the road surface, with the H_t value defined in truth as the elevation of the CCD with respect to the road plane, moreover practically constant during the MMS surveying.
This method has been implemented in a Matlab language program, exploiting in particular the functions *maketform* and *imtransform* of the Image Processing Toolbox. Once solved the analytical problem by equation (13), the rectification of the image lower parts is so accomplished by a digital resampling with automatic mosaicking. From the numerical point of view, the analytical parameters of equation (13) are suitably converted in other ones for the resampling, since the i,j pixel coordinate system is a clockwise and pointing down frame.

7. NUMERICAL APPLICATION

Here are reported first numerical applications of the method proposed in this paper, considering image sequences dynamically acquired with a Fuji FinePix F700 CCD camera (2.832x2.128 pixel, 35 mm equivalent focal length) in the parking area of the Stadio Friuli in Udine (Figures 3). Such wide area is composed from grids of opposite parking seats with significant contrary slopes, visible in the lanes in Figure 3 at left. In this way, such area well simulates a rectilinear way if routed orthogonally to the parking seats, where the parking splitting lane stands for the central (axial) lane of a certain road.

Figure 3. Left, the area of the test, only roughly horizontal; right, an irregularity of the asphalting in elevation and in slope.

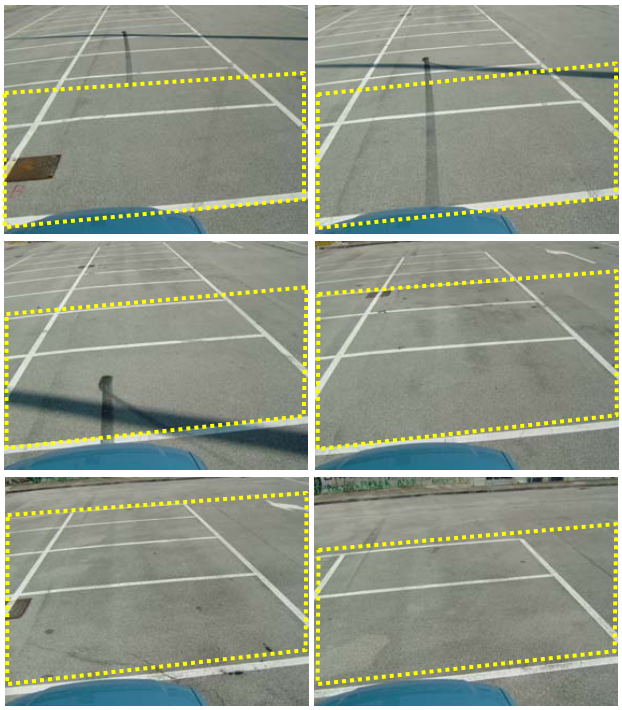

Figure 4. From the third to the eighth image of a monoscopic sequence acquired by a rectilinear trajectory.

Unfortunately, the road pavement presents an imperfection due to a step edge of some millimetres in the asphalting in the "right" tract (Figure 3 at right). The levelling apparatus shows how the pavement does not slant towards the right external border but in the opposite left direction. This means an unconventional transversal section just in the track routed by our simplified MMS, whose consequence in the image rectification will be shown afterwards.
Anyway, this free area allowed to completing without restraint also a second curved MMS trajectory in the same framework.
The first experiment involves eight frontal pointing down images acquired with a mean interval of 5 m during a rectilinear surveying route from a height of about 2,35 m with respect to the road surface. Images from the third to the eight compose Figure 4, while the first two images are displayed in Figure 1.
In Table 5 the six external orientation parameters are reported with E_t, N_t, H_t locally referred, while Table 6 reports the eight homographic parameters resulting form equations (13).

image	E_t [m]	N_t [m]	H_t [m]	ω_t [g]	ϕ_t [g]	κ_t [g]
R1	6,986	-3,053	2,380	72,772	1,800	-4,791
R2	6,838	1,811	2,361	72,356	-0,145	-4,362
R3	6,776	6,767	2,396	73,283	2,534	-5,734
R4	6,828	11,793	2,374	71,168	2,910	-6,400
R5	6,814	16,615	2,362	72,788	3,109	-5,877
R6	6,818	21,651	2,358	72,510	2,717	-5,538
R7	6,906	26,670	2,345	70,850	3,696	-6,059
R8	6,921	31,741	2,357	71,561	3,153	-5,921

Table 5. External orientation parameters of the image sequence.

image	a_{t1} [m/pix]	b_{t1} [m/pix]	c_{t1} [m]	a_{t2} [m/pix]	b_{t2} [m/pix]	c_{t2} [m]	u_t [1/pix]	v_t [1/pix]
R1	2,39e-03	-5,03e-03	6,82	-2,11e-04	3,07e-03	-3,22	6,54e-05	-7,40e-04
R2	2,23e-03	-4,85e-03	6,85	3,05e-05	-5,22e-04	1,82	4,92e-05	-7,29e-04
R3	2,54e-03	-4,94e-03	6,54	5,52e-04	-4,30e-03	6,53	8,18e-05	-7,55e-04
R4	2,41e-03	-4,54e-03	6,58	1,00e-03	-7,35e-03	11,5	8,53e-05	-6,92e-04
R5	2,50e-03	-4,86e-03	6,54	1,43e-03	-1,15e-02	16,3	8,51e-05	-7,40e-04
R6	2,43e-03	-4,81e-03	6,58	1,70e-03	-1,50e-02	21,4	7,82e-05	-7,30e-04
R7	2,38e-03	-4,55e-03	6,60	2,28e-03	-1,74e-02	26,4	8,50e-05	-6,83e-04
R8	2,41e-03	-4,69e-03	6,65	2,62e-04	-2,15e-03	31,5	8,14e-05	-7,03e-04

Table 6. Homographic parameters of the same image sequence.

It can be numerically noticed how, for their definition and dimensionality, c_{t1}, c_{t2} coefficients have a magnitude order many times higher that the others; furthermore all the c_{t1} values are practically constant as E_t are, while c_{t2} coefficients grow according to increasing values of N_t.
The images have been therefore rectified onto two planes, one for each "tract", assuming a 2,5‰ transversal slope towards the external. In truth, the MMS trajectory was practically centred on the right portion of the parking area, nevertheless a small part of then depicts also the left side.
By means of the inverse perspective mapping the rectangular images become trapeze shapes on the road, recovering the true surface geometry, e.g. the orthogonality among the lanes (Figure 7): the proposed method correctness is hence qualitatively proved. If we observe again Figure 7, it can be seen also the good congruence of the resulting automatic mosaicking, of course better for the part near to the MMS and worse for the far ones involving the vanishing point. The mistakes in the digital reconstruction near to the central lane are due to the step edge irregularity previously remarked: this is essentially the well-known planimetric effect of a height error in the modeling of the surface to be remap.

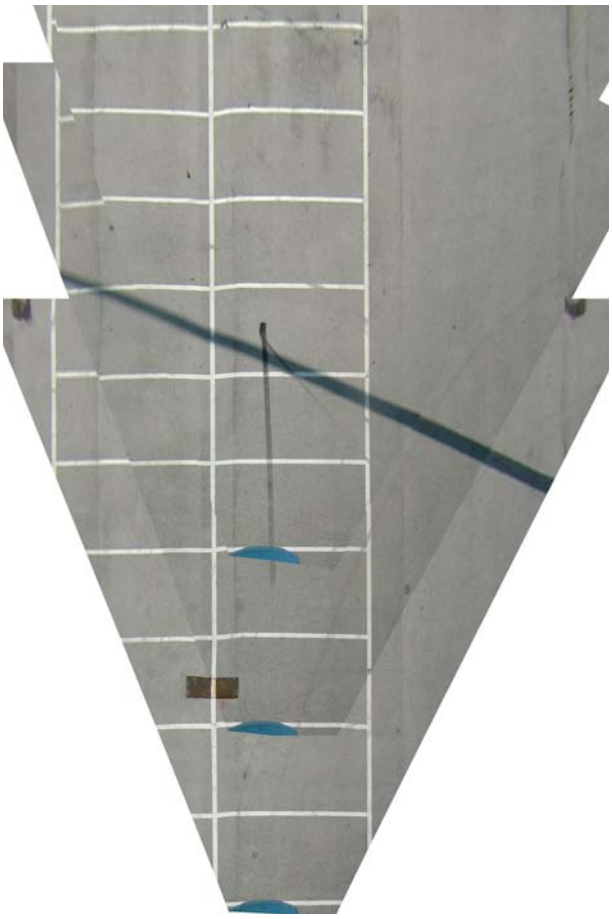

Figure 7. Mosaic of the second, third and fourth rectified image.

The final global mosaic is reported in Figure 8, with images lowest part depicting the car covered from the image before: the rectified areas are highlighted in Figure 6 by yellow dot lines.

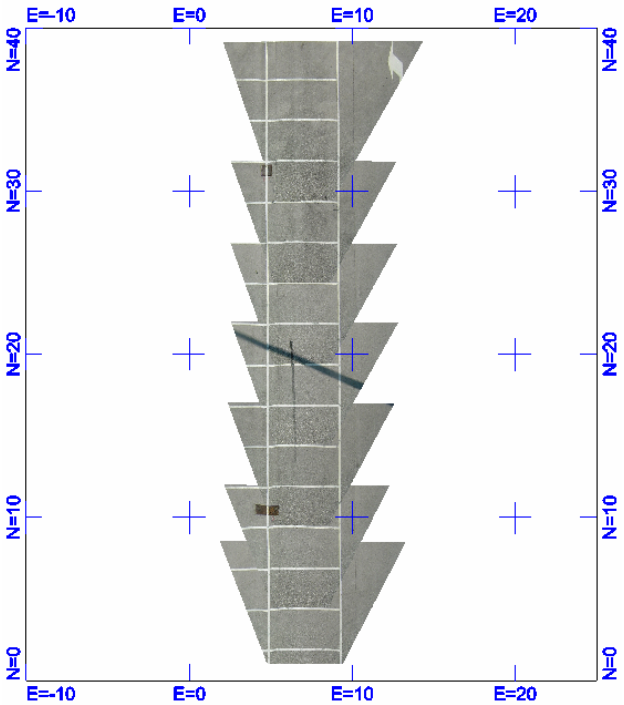

Figure 8. Mosaic of eight rectified images.

The second experiment regards instead a curved trajectory, having routed a right-left bend over the same area (Figure 9).

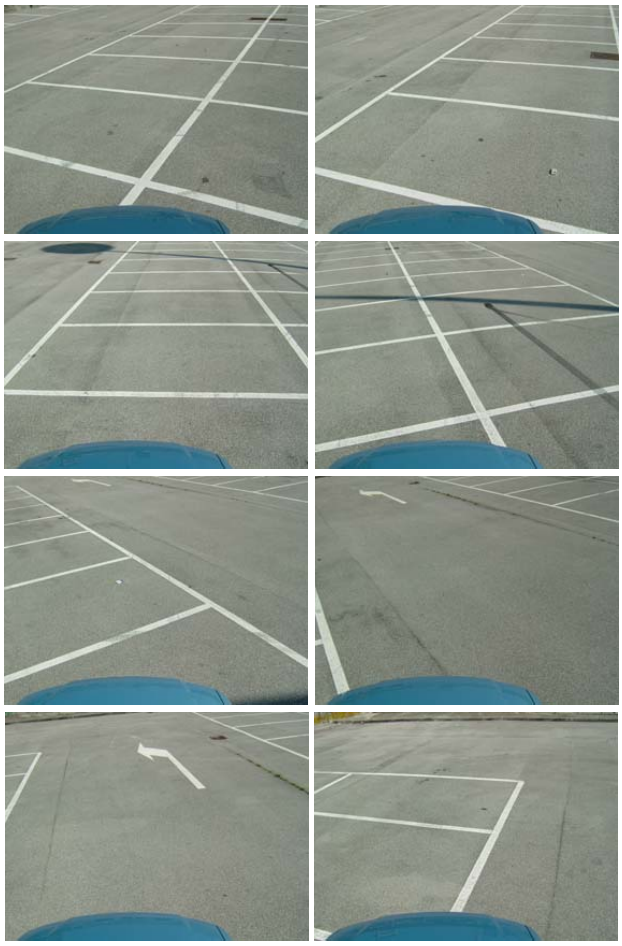

Figure 9. From the first to the eighth image of a monoscopic sequence acquired by a curved trajectory.

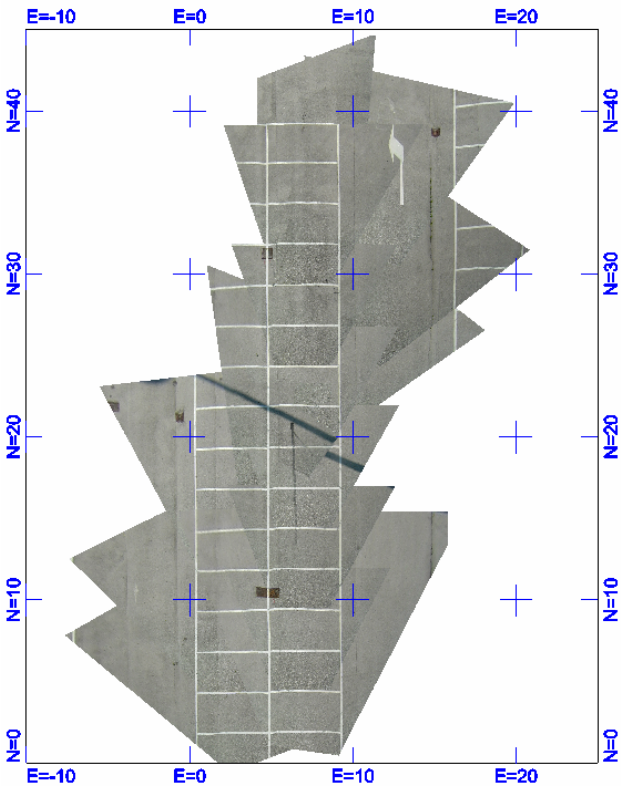

Figure 10. Mosaic of sixteen rectified images.

Figure 10 is the altogether mosaic of the sixteen rectified images, showing again a good geometrical congruence (a part the shadow of the illumination pole constituting a meridian angularly reporting the time between the acquisition of the first image sequence and the second one!).
Summarizing the road surveying obtained in these first numerical experimentations, we can state its general likelihood, postponing to next future the necessary statistical evaluations about the achieved values for accuracy and reliability, sharing the effects resulting from the homographic parameters to those due to the road surface modeling.

8. CONCLUSIONS

This paper proposes an original method for a mosaic digital representation of a road by a Kalman filter based rectification of monoscopic images acquired with a pointing ahead-down camera with a low-cost MMS.
To this goal, the image sequence is oriented mainly exploiting image coordinates submitted to classical photogrammetric equations as the coplanarity and collinearity conditions. In addition, a matching procedure applying another Kalman filter is exploited to improve the automation of the relative orientation among the successive images. Exploiting only points in the lower part of the images, i.e. near to the MMS, a high-scale optical model is achieved.
The image rectification requires either image orientation parameters or geometrical information about the surface to be projected: for the second aim, it is sufficient to share the road pavement in finite planes, one for the right tract and one forth the left tract for each image. Regarding instead the image orientation, the six external parameters univocally determine the corresponding eight homographic ones.
The numerical experiments in the processing of real CCD image sequences with a Matlab routine implementing this method are qualitatively satisfying, although quantitative analyses have to be carried out as soon as possible.
As possible development of this research, the Kalman filter state vector could be directly constitutes from the eight homographic parameters, but the derivation of the corresponding state equations from the image shot coordinates and attitude angles, not trivial from the analytical point of view. In particular geometrical conditions, e.g. a perfect plane surface where the shot direction is more constrained or can vary only in azimuthal sense, as happens for industrial applications and in robotic vision. In this case, some homographic parameters remain constants or it is easier to analytically derive the prediction of their evolution.

REFERENCES

Bertozzi, M., Broggi, A., 1998. GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Transactions on Image Processing*, 7 (1), pp. 62-81.

Crosilla, F., Visintini, D., 1998. External orientation of a mobile sensor system for cartographic updating via dynamic vision of digital map points. *Bollettino di Geodesia e Scienze Affini*, 1, Istituto Geografico Militare, Firenze, pp. 41-60.

Dermanis, A., 1990. Modeling and solution alternatives in photogrammetry. In: *Proceedings of the ISPRS ICWGIII/VI Tutorial "Mathematical aspects of data analysis"*, Rhodes, Greece, pp. 77-121.

El-Sheimy, N., 1996. *The development of VISAT - A Mobile Survey System for GIS applications*. PhD thesis, University of Calgary, Dept. of Geomatics Engineering, Canada, 179 pp.

El-Sheimy, N., (ed.), 2001. *Proceedings of the 3rd International Workshop on "Mobile Mapping Technology"*, Cairo, Egypt (on CD).

Hartley, R., Zisserman, A., 2000. *Multiple view geometry in computer vision*. Cambridge University Press, Cambridge, 496 pp.

Li, R., Murai, S., (eds.), 1999. *Proceedings of the 2nd International Workshop on "Mobile Mapping Technology"*, Bangkok, Asian Institute of Technology, Pathumthani. Thailand.

Mallot, H.A., Bülthoff, H. H., Little, J.J., Bohrer, S., 1991. Inverse perspective mapping simplifies optical flow computation and obstacle detection. *Biological Cybernetics*, 64, pp. 177-185.

Ohio State University (ed.), 1995. *"1995 Mobile Mapping" Symposium*. Proceedings of the 1st International Workshop on "Mobile Mapping Technology", Columbus, Ohio, American Society of Photogrammetry & Remote Sensing, 274 pp.

Roncella, R., Forlani, G., 2006. Automatic lane parameters extraction in Mobile Mapping sequences. In: *Proceedings of the ISPRS Commission V Symposium*, Dresden, Germany, IAPRS&SIS, vol. XXXVI, 5, 6 pp. (on CD).

Tao, V., (ed.), 2004. *Proceedings of the 4th International Symposium on "Mobile Mapping Technology"*, Kunming, China, (on CD).

Tao, V., Li, J., 2007. *Advances in Mobile Mapping Technology*. ISPRS Book Series, vol. 4, Taylor & Francis, London, 191 pp.

Visintini, D., 2001a. A Kalman filter based matching for a MMS image sequence. In: *Proc. of the 3rd International Symposium on "Mobile Mapping Technology"*, Cairo, Egypt, 14 pages (on CD).

Visintini, D., 2001b. Testing a dynamic algorithm for image matching in a CCD sequence acquired by a low cost MMS. In: *Proceedings of the Italy-Canada Workshop on "3D digital imaging and modeling applications"*, Padua, Italy, 8 pp. (on CD).

Visintini, D., 2001c. A simplified mobile mapping system for 3D-measurements via image sequence dynamic analysis. In: *Proceedings of the 5th "Optical 3-D measurement techniques" Conference*, Vienna, Austria, pp. 137-144.

Visintini, D., 2003. High-efficiency building surveys by dynamic and static digital images oriented with "mixed models". In: *Proc. of the ISPRS WG V/4 and IC WG III Workshop on "Vision techniques for digital architectural and archaeological archives"*, Ancona, Italy, IAPRS&SIS, vol. XXXIV, 5/W12, pp. 335-340.

ACKNOWLEDGEMENTS

This work was carried out within the research activities supported by the INTERREG IIIA Italy-Slovenia 2003-2006 project "Cadastral map updating and regional technical map integration for the GIS of the regional agencies by testing advanced and innovative survey techniques".