

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



One Approach to the Fusion of Inertial Navigation and Dynamic Vision

Stevica Graovac

*School of Electrical Engineering, University of Belgrade
Serbia*

1. Introduction

The integration of different types of navigation systems is frequently used in the automatic motion control systems due to the fact that particular errors existing in anyone of them are usually of different physical natures. The autonomous navigation systems are always preferred from many of reasons and the inertial navigation systems (INS) are traditionally considered as the main representative of this class. The integration of such systems based on the inertial sensors (rate gyros and linear accelerometers) and other navigation systems is very popular nowadays, especially as a combination with global positioning systems [Farrel & Barth, 1999], [Grewal et al., 2001]. The vision based navigation systems (VNS) are also of autonomous type and there is a reasonable intention to make the fusion of these two systems in some type of integrated INS/VNS system. This paper is oriented toward the possibility of fusion of data originated from a strap-down INS on one side, and from a dynamic vision based navigation system (DVNS), on the other. Such an approach offers the wide area of potential applications including the mobile robots and a number of automatically controlled ground, submarine, and aerial vehicles.

The most usual approach in navigation systems integration is of "optimal filter" type (typical INS/VNS examples are given in [Kaminer et al., 1999] and [Roumeliotis et al., 2002]) In such an approach one of the systems is considered as the main one and the other supplies less frequently made measurements (corrupted by the noise, but still considered as the more precise) used in order to estimate in an optimal fashion the navigation states as well as the error parameters of the main system's sensors.

The approach adopted in this paper considers both systems in an equal way. It is based on the weighted averaging of their outputs, allowing some degrees of freedom in this procedure regarding to the actually estimated likelihood of their data. These estimates are based on reasoning related to the physical nature of system errors. The errors characterizing one typical strap-down INS are of slowly varying oscillatory nature and induced by the inaccuracies of inertial sensors. On the other hand, the errors in any VNS are mainly due to a finite resolution of a TV camera, but there is a significant influence of the actual scene structure and visibility conditions, also. In other words, it could be said that the accuracy of an INS is gradually decreasing in time while it is not affected by the fact where the moving object actually is. The accuracy of a DVNS is generally better in all situations where the

recognizable referent landmarks are inside the camera's field of view, occupying larger extent of it. Because the DVNS is based on processing of a sequence of images, the larger relative motion of the landmarks in two consecutive frames is preferable too.

Having in minds these basic features of INS and DVNS, their integration could be considered in two basic ways:

1. An INS is a kind of "master" navigation system while the corrections produced by DVNS are made in time when a moving object is passing the landmarks located around the trajectory. This approach is typically applicable in case of the flight control of remotely piloted vehicles and in the similar "out-door" applications;
2. A VNS is a basic navigation system assuming that the reference scene objects always exist in the scene, while an INS provides the required data related to the absolute motion of an object during the interval between two frames. This approach is oriented toward mobile robot "in-door" applications as well as in case of automatic motion control of the road vehicles.

The next chapter of paper introduces the fundamentals of INS and VNS in the extent required to understand their integration. In Chapter 3. the general case of suggested fusion procedure is presented. A number of particular implementation schemes including reduced set of sensors and/or reduced amount of calculations could be specified based on this general case. The next two chapters consist of the illustrative examples of application: A vision aided INS in the simulated case of remotely piloted vehicle's flight (Chapter 4), [Graovac, 2004]; and a VNS assisted by the acceleration measurements provided by an INS, for the robot control applications (Chapter 5), [Graovac, 2002].

The results related to "out-door" applications are obtained using the full 6-DOF simulation of object's motion and the model of the INS work. The computer-generated images of terrain and ground landmarks have been used during the tests of a DVNS algorithm. These images have been additionally corrupted by noise and textured. The "in-door" applications are illustrated using the laboratory experiments with an educational robot equipped with a TV camera.

2. Basic Concepts of INS and VNS

2.1 Fundamentals of an INS

Estimation of a position of moving object, $\vec{R}_I = [x_I \ y_I \ z_I]^T$, relative to an *inertial coordinate frame* (ICF) could be done according to the basic navigation equations as

$$\begin{aligned} \begin{bmatrix} \dot{V}_{XI} \\ \dot{V}_{YI} \\ \dot{V}_{ZI} \end{bmatrix} &= \begin{bmatrix} A_{XI} \\ A_{YI} \\ A_{ZI} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix}, \quad \begin{bmatrix} V_{XI}(0) \\ V_{YI}(0) \\ V_{ZI}(0) \end{bmatrix} = \vec{V}_I(0) \\ \begin{bmatrix} \dot{x}_I \\ \dot{y}_I \\ \dot{z}_I \end{bmatrix} &= \begin{bmatrix} V_{XI} \\ V_{YI} \\ V_{ZI} \end{bmatrix} = \vec{V}_I, \quad \begin{bmatrix} x_I(0) \\ y_I(0) \\ z_I(0) \end{bmatrix} = \vec{R}_I(0) \end{aligned} \quad (1)$$

Acceleration vector in ICF \vec{A}_I , on the right hand side, is obtained by transformation of the acceleration measurement vector \vec{A}_B . These measurements are made by a triad of linear accelerometers rigidly fixed to the body of moving object and they are referenced to the *body fixed coordinate frame* (BCF):

$$\vec{A}_I = \begin{bmatrix} A_{XI} \\ A_{YI} \\ A_{ZI} \end{bmatrix} = T_{I/B} \vec{A}_B = T_{I/B} \begin{bmatrix} A_{XB} \\ A_{YB} \\ A_{ZB} \end{bmatrix}. \quad (2)$$

Matrix transform $T_{I/B}$ is defined via Euler angles of pitch, yaw, and roll (ϑ, ψ, ϕ) as

$$T_{I/B} = T_3^T(\psi) T_2^T(\vartheta) T_1^T(\phi). \quad (3)$$

where T_1, T_2, T_3 , represent the elementary matrix transformations due to rotation around coordinate axes. Actual values of Euler angles are obtained by numerical integration of a set of differential equations:

$$\begin{bmatrix} \dot{\phi} \\ \dot{\vartheta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} \omega_{VX} + \tan \vartheta (\omega_{VY} \sin \phi + \omega_{VZ} \cos \phi) \\ \omega_{VY} \cos \phi - \omega_{VZ} \sin \phi \\ \sec \vartheta (\omega_{VY} \sin \phi + \omega_{VZ} \cos \phi) \end{bmatrix}. \quad (4)$$

where $\omega_{VX}, \omega_{VY}, \omega_{VZ}$, represent the projections of the angular velocity of a moving object in ICF onto the axes of BCF. These are measured by the set of rate gyros rigidly fixed to the body of the moving object.

The measurements of linear accelerations and angular velocities in BCF are inaccurate due to slowly varying bias introduced by a number of physical phenomena inside the inertial instruments. These are results of the complex motion of an object (with six degrees of freedom) as well as of sensor imperfections. Sensor signals are additionally corrupted by high frequency measurement noise caused by internal imperfections and by external influences due to the air turbulence, vibrations of vehicle, etc. A specific type of error associated to this particular mechanization (known as a *strap-down inertial navigation system* - SDINS) in case of flying object is a result of rectification introduced by multiplication shown in (2). The elements of matrix $T_{I/B}$ as well as of vector \vec{A}_B include the oscillatory components on natural frequency of body oscillations.

The inertial instruments analyzed here are of medium quality (typically used for the flight stabilization purposes). The numerical data illustrating their accuracy are:

- Rate gyros: Bandwidth - 80 Hz;
 Bias - 10^0 /hour; G-sensitive drift - 10^0 /hour/g;
 Scale factor error - 1%;
 Measurement noise: white, Gaussian, zero mean value, $\sigma = 1^0$ /s;
 Accelerometers: Bandwidth - 150 Hz;
 Bias - 0.1 m/s²; Resolution - 0.05 m/s²;
 Scale factor error - 1%;
 Measurement noise: white, Gaussian, zero mean value, $\sigma = 0.1$ m/s²;

The accuracy of an INS was analyzed using the complete 6-DOF horizontal flight simulation. As a way of on-line accuracy improvement, the Kalman filter was applied in order to make the filtration of rate gyro signals. This one was based on the linearized dynamic models in pitch and yaw channels. The results of the Kalman filter application in the estimation of pitch rate and pitch angle during the interval of ten seconds of horizontal flight are illustrated in Figure 1.

2.2 Fundamentals of a Dynamic Vision Based Navigation

The linear position of a moving object carrying a TV camera on-board relative to the environmental elements can be reconstructed either from one frame or from a sequence of frames. In the first case, a number of characteristic scene objects' features should be

extracted. The other approach, known as a *dynamic vision*, generally allows usage of a lower number of extracted and tracked features. If some additional information about linear and angular velocities or about angular orientation are known, the task can be radically simplified, allowing the tracking of just one reference object's feature [Frezza et al., 1994], [Menon et al., 1993]. In both cases, if the absolute position of a reference object in ICF is a priori known, the whole method can be interpreted as a reconstruction of the absolute position of a moving object - *visual navigation*.

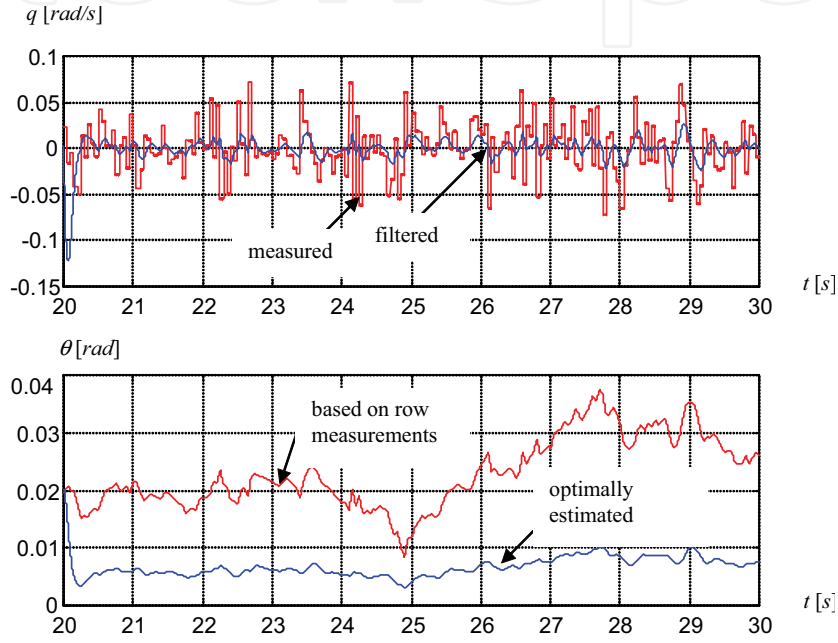


Fig. 1. The effects of application of Kalman filter in the estimation of pitch rate and pitch angle.

The dynamic vision method has been applied in this paper. Supposing that external information about linear and angular velocities of a moving object exists (supplied by an INS), the number of tracked features is reduced to one. In the case of an autonomously guided flying vehicle, typical ground reference objects could be bridges, airport runways, cross-roads, distinguishable buildings, or other large stationary landmarks located at known absolute positions. The task of a VNS consists in extracting the image of landmark itself and after that, calculating the position of some easily recognizable characteristic point (e.g., a corner). If the image of a whole landmark occupies just a small portion of the complete image, it is more reasonable to calculate the position of its centroid instead.

Primary detection (recognition) of a landmark is the most critical step. It is supposed that this task is accomplished using a bank of reference landmark images made separately in advance, under different aspects, from different distances, and under different visibility conditions. Once primary automatic detection has been done, the subsequent recognition is highly simplified. The recognized pattern from the actual image becomes the reference one for the next one, and so on.

In an ideal case, the only required information is regarding the shift of the characteristic point between two consecutive images. The existence of image noise and different other reasons may cause an erroneous calculation of a characteristic point location inside the picture. In order to minimize these effects, a greater number of characteristic points and/or consecutive frames should be analyzed.

Dynamic equations describing the stated approach are the following:

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{x_2}{x_1} \\ \frac{x_3}{x_1} \end{bmatrix} \quad x_1 \neq 0 \quad (5)$$

State vector $\vec{x} = [x_1 \ x_2 \ x_3]^T$ represents the position of a reference point with respect to viewer frame, while coefficient vectors $\vec{v} = [v_1 \ v_2 \ v_3]^T$ and $\vec{\omega} = [\omega_1 \ \omega_2 \ \omega_3]^T$ represent relative linear and angular velocities, also expressed in the coordinate frame fixed to the moving object. Measured outputs of this nonlinear dynamic system consist of two projections of the reference point onto the image plane (*picture coordinate frame* - PCF) which is perpendicular to the x_1 axis, at a conventional distance $f = 1$ from the origin. If the relative positions are known, the task consists of motion parameter estimation (coefficient vectors identification). If the motion parameters are known, the task is of state estimation nature (structure reconstruction). The second case is considered here.

If in some following moment of time (e.g., in the next frame) the state vector has the value $[x_1 + \Delta x_1 \ x_2 + \Delta x_2 \ x_3 + \Delta x_3]^T$, there would exist the shift of an image of reference point in PCF given as

$$\Delta y_1 = f \frac{\Delta x_2}{x_1 + \Delta x_1} - y_1 \frac{\Delta x_1}{x_1 + \Delta x_1} \quad \Delta y_2 = f \frac{\Delta x_3}{x_1 + \Delta x_1} - y_2 \frac{\Delta x_1}{x_1 + \Delta x_1} \quad (6)$$

The variation of position $\Delta \vec{x} = [\Delta x_1 \ \Delta x_2 \ \Delta x_3]$ is produced by the linear motion of a moving object $\Delta \vec{x}_I$ as well as by its change of angular orientation, defined now by matrix transformation $T_{I/B} + \Delta T_{I/B}$ instead of the previous $T_{I/B}$. After appropriate geometrical recalculations it could be shown that the variation of the relative linear position is represented as

$$\Delta \vec{x} = T_C [\Delta T_{I/B} T_{I/B}^T (T_C^T \vec{x} + \vec{l}) - (T_{I/B} + \Delta T_{I/B}) \Delta \vec{x}_I] \quad (7)$$

where the linear position of the camera relative to the center of gravity of a moving object is denoted as \vec{l} , while the angular orientation of a camera relative to the BCF axes is represented via transformation matrix T_C . Both these parameters are known because they are either constant ones or can be measured easily.

After division of both sides of (7), one obtains

$$\begin{bmatrix} \Delta x_1/x_1 \\ \Delta x_2/x_1 \\ \Delta x_3/x_1 \end{bmatrix} = T_C \Delta T_{I/B} T_{I/B}^T T_C^T \begin{bmatrix} 1 \\ y_1/f \\ y_2/f \end{bmatrix} + T_C \Delta T_{I/B} T_{I/B}^T \begin{bmatrix} l_1/x_1 \\ l_2/x_1 \\ l_3/x_1 \end{bmatrix} - T_C (T_{I/B} + \Delta T_{I/B}) \begin{bmatrix} \Delta x_I/x_1 \\ \Delta y_I/x_1 \\ \Delta z_I/x_1 \end{bmatrix} \quad (8)$$

Supposing that T_C , f , and \vec{l} are known and that $T_{I/B}$ and $T_{I/B} + \Delta T_{I/B}$ are supplied externally as well as $\Delta \vec{x}_I$ (e.g., by an INS), the task of VNS consists in determining the pair of coordinates in PCF (y_1, y_2) and at least one of the displacement components (6). Combining three scalar equations (8) with the proper one in (6), it is now possible to determine four unknown variables $(\Delta x_1, \Delta x_2, \Delta x_3, x_1)$. Once x_1 is calculated, one can reconstruct the remaining

two components of the relative position vector (x_2 and x_3) from the output part of (5). The knowledge of relative position vector \vec{x} and of the absolute position of the characteristic point in ICF is sufficient to reconstruct the absolute position of a moving object.

The crucial problem from an image processing point of view is how to determine the locations of characteristic points in PCF as accurately as possible. There exist a number of methods of distinguishing objects of interest inside the image. Practically all of them are application dependent. Various sequences of image enhancement/digital filtration procedures, segmentation approaches using multilevel gray or binarized picture, morphological filtration algorithms, etc. making these procedures, must be carefully chosen according to the actual scene contents.

Computer generated images of ground landmarks are used throughout this work. A relatively simple correlation technique consisting of matching of actual image contents and a reference pattern has appeared as the most robust one. It is based on calculation of a sum of absolute differences of light intensities inside the window scanning across the image. The feature has been defined as the light intensity distribution inside the rectangular window of $n_T = n_X n_Y$ pixels around the characteristic point. The displacement of characteristic point $(\Delta y_1, \Delta y_2)$ is calculated by maximizing the similarity of the actual image and the previous one, i.e., minimizing the criterion given as a sum of absolute values of differences (MAD) of light intensity I_N :

$$L = \frac{1}{n_T} \sum |I_N(y_1 + \Delta y_1, y_2 + \Delta y_2) - I_{N-1}(y_1, y_2)|. \quad (9)$$

The efficiency of the stated algorithm will be illustrated using the sequence of textured images of a bridge. The nearest holder has been used as a reference object while its crossing with the left edge of a runway was selected as a characteristic point (Figure 2.) Figure 3. illustrates matching results obtained for the multiple level gray and binarized pictures. The brightest points inside the black window are pointing to the locations of maximal similarity. The reference window was of dimensions 25 X 25 pixels. Higher sharpness of candidate area in case (a) suggests that one could expect better results if the multiple level gray pictures were used.

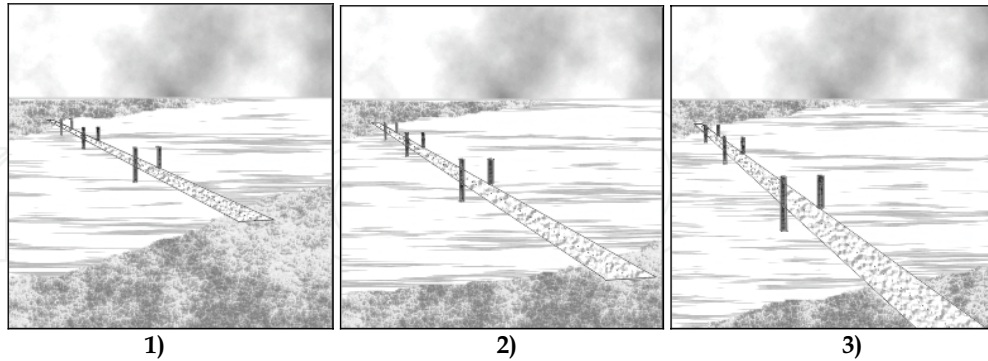


Fig. 2. Sequence of textured images of a bridge used as a landmark.

When the sequence of frames shown in Figure 2. was used for navigation purposes, the results given in Table 1. have been obtained. It is supposed that the angular position of the camera is constant during the observed time period (yaw angle, 12°, pitch angle, -5°, roll

angle, 0°). The linear velocities of the moving object are exactly known and constant also ($V_X = 500$ m/s, $V_Y = 100$ m/s, and $V_Z = 0$ m/s). Position estimates given in Table 1. were obtained using the pairs of frames 1-2, 1-3, and 2-3.

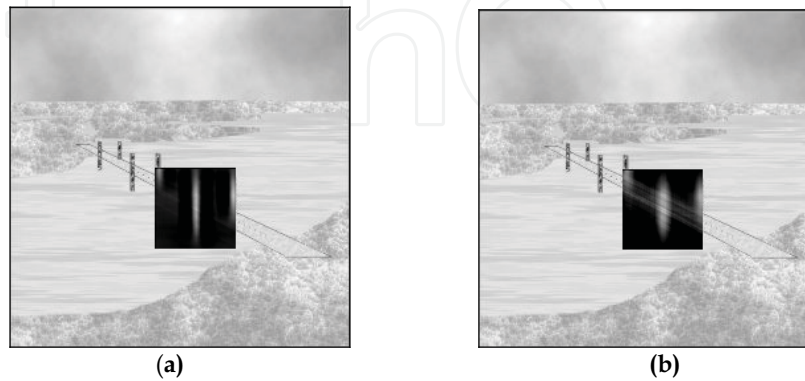


Fig. 3. Extraction of a characteristic point inside: (a) multiple level gray and (b) binarized picture.

	Position in frame 1.			Position in frame 2.	
	Exact	Estimate		Exact	Estimate
		Based on pair1-2	Based on pair1-3		Based on pair2-3
X (m)	600	606	535	400	357
Y (m)	200	199	187	160	152
Z (m)	100	100	94	100	96

Table 1. Moving object position estimation using dynamic vision from sequence shown in Figure 2.

2.3 Fundamentals of an Autonomous VNS

The fundamental step in an autonomous VNS based on the processing of just one image consists of the calculation of the relative distance and angular orientation of a camera relative to the reference object (landmark) located in a horizontal plane of ICF (Figure 4.) [Kanatani, 1993].

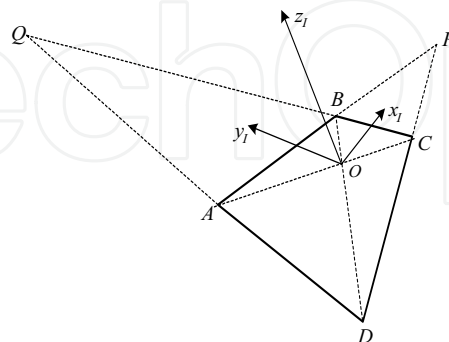


Fig. 4. Reference object in the field of view of TV camera.

Supposing a landmark of rectangular shape of known dimensions and the cross-section of its diagonals as a reference point adopted as the origin of ICF, the position of the camera relative to this point could be obtained from just one image by the following algorithm:

1. Calculate the normalized vectors for all four projections in PCF of corners A , B , C , and D . Each one of these m -vectors is a vector representing projection of the appropriate point $\vec{y}_i = [1 \quad y_{i1} \quad y_{i2}]^T$, divided by its Euclidean norm.
2. Calculate the normalized vectors for all four projections of the edges AB , BC , CD , DA , as the normalized cross-products of m -vectors above. Elements of these n -vectors specify the equations of image lines encompassing the projections of the appropriate pair of corners.
3. Calculate the m -vectors of *vanishing points* P and Q , as cross-products of n -vectors above, representing the projections of parallel edges of a reference object.
1. Calculate the n -vectors of diagonals AC and BD as in case of image lines representing the edges (2).
2. Calculate the m -vector \vec{m}_O of the point at the cross-section of diagonals O , as in case of vanishing points (3).
3. Choosing any one of the corners as the point at known distance d from the reference point O , calculate the *scene depth*:

$$|\vec{R}| = \frac{\vec{m}_I \vec{e}_{I3} \cdot d}{\|\vec{m}_O \vec{e}_{I3} \cdot \vec{m}_I - \vec{m}_I \vec{e}_{I3} \cdot \vec{m}_O\|} \quad (10)$$

representing the distance between camera's sensitive element and the reference point O . The m -vectors \vec{m}_O and \vec{m}_I are related to the reference point O and the chosen corner I . The m -vectors of vanishing points \vec{m}_P and \vec{m}_Q are at the same time the basis vectors of reference coordinate frame with its origin at O . The ort of direction perpendicular to the plane containing the reference rectangle is obtained as $\vec{e}_{I3} = \vec{m}_P \times \vec{m}_Q$.

4. Calculate the position of a camera relative to the reference point as

$$\vec{R} = -|\vec{R}| \cdot T_O \vec{m}_O \quad (11)$$

where $T_O = [\vec{e}_{I1} \quad \vec{e}_{I2} \quad \vec{e}_{I3}]^T = [\vec{m}_P \quad \vec{m}_Q \quad \vec{m}_P \times \vec{m}_Q]^T$ represents the transformation matrix due to rotation of the frame fixed to a camera (BCF) in respect to the coordinate frame fixed to the reference object (ICF).

The above explained algorithm reflects just a *geometrical* aspect of a problem. Much more computational efforts are associated with the *image processing* aspect, i.e., with the problem how to distinguish the reference object and its characteristic points from the actual contents of an image. It should be noted that the final effect of this process consists of some deteriorated accuracy in the determination of image coordinates of the reference points. A lot of *scene dependent* conditions affect the extraction as well as some system parameters (image noise, level quantization, space resolution). An image noise is dominantly associated to TV camera itself and it is usually considered as a zero-mean, Gaussian, white noise with specified standard deviation (expressed in number of intensity quanta). The later two systematic sources of inaccuracy are due to the process of image digitization. While the effects of the finite word length of a video A/D converter are of the same nature as the effects of image noise, the finite space resolution has the direct influence onto the final accuracy of position estimation, even when the reference object is ideally extracted. The

finite number of picture elements, *pixels*, along the horizontal and vertical directions inside the image, makes the limits for the final accuracy. However, this effect is strongly coupled with the size of camera's field of view and the actual distance between a camera and the reference point. The error expressed in pixels has its angular and linear equivalents in dependence on these parameters.

The redundancy in geometrical computations is generally suggested for the VNS accuracy improvement. Instead of a theoretically minimal number of considered points required for some calculation, the number of points is increased and some appropriate optimization procedure is usually used in order to filter out the effects of noise in determination of a location of any particular characteristic point. For example, instead of starting with the determination of the positions of corners in above mentioned algorithm, one can start with the detection of edges and find the equations of image lines by the best fitting procedure considering the whole set of edge points (not by using just two of them as before). Now the m-vectors of corners are obtained as cross-products of the appropriate n-vectors and the remainder of algorithm is the same. Similarly, the final accuracy can be significantly improved if one repeats the explained algorithm using different corners as reference ones and finds the weighted average of results. All these methods used for the accuracy improvement increase the overall computational effort. Therefore, it is of a great importance to find the way how to obtain the same or better accuracy using a less number of considered points or the less complex image processing algorithms.

3. Principles of Data Fusion

Main conclusions related to the quality of information about linear and angular positions of a moving object relative to ICF, obtained by INS and VNS separately, are the following:

The accuracy of the SDINS based algorithm

- depends on a slowly varying bias (drift) and a measurement noise of inertial sensors;
- decreases in time due to cumulative effect produced by these errors;
- depends on errors in initial condition estimation (angular and linear positions and velocities);
- could be improved by recursive optimal state estimation; and
- is affected by slowly varying bias introduced by rectification.

The accuracy of the VNS based algorithm

- depends on the reference object's visibility conditions;
- depends on TV image noise as well as on quantization made by video signal digitization;
- depends on the relative size of a reference object inside the field of view (increases while the moving object approaches the reference one);
- depends on the shift(s) of the characteristic point(s) between two consecutive frames and increases in the case of larger ones; and
- could be improved by increasing the number of tracked points and/or analyzed frames.

Having in mind the fact that the error sources inside these two systems are different and independent, the possibility of their fusion is considered. The combined algorithm of linear and angular position estimation is based on a suitable definition of a criterion specifying the likelihood of partial estimations.

It is supposed in the general case that both algorithms are active simultaneously and autonomously. *Autonomous* estimations from one algorithm are being passed to another one in order to obtain new (*assisted*) estimations. Resultant estimation on the level of a combined algorithm is always obtained as the weighted average value of separate ones. The weighting factors are calculated according to the adopted criteria about partial estimation likelihood.

The following formalism has been adopted:

- the transformation matrix connecting BCF and ICF, generally noted as $T_{I/B}$ will be represented just as T_I ;
- all vectors representing angular and linear velocities and linear positions are relative to ICF;
- lower indexes I and V are referencing the estimated variables to the estimation originating system (I - inertial, V - visual);
- autonomously estimated variables are noted by upper index ' ' ;
- upper index " stands for the estimate based on the information obtained from other system (assisted one);

The general procedure consists of the following steps:

1. SDINS generates its autonomous estimates of angular rate vector $\vec{\omega}'_I$, transformation matrix T'_I , linear velocity vector \vec{V}'_I , and space position \vec{x}'_I .
2. Based on \vec{x}'_I , T'_I , and a priori known position of a reference object in ICF, VNS searches the field of view inside the expected region. It finds the image of the reference object and calculates the coordinates of characteristic points in PCF.
3. Adopting \vec{x}'_I as a priori known initial position estimation (scene structure), VNS identifies from the sequence of frames the angular rate vector $\vec{\omega}''_I$ and linear velocity vector \vec{V}''_I .
4. Adopting the estimations $\vec{\omega}'_I$ and \vec{V}'_I as accurate ones and on the basis of landmark's image position measurements in the sequence of frames, VNS estimates the position vector \vec{x}''_V .
5. VNS autonomously generates its estimation of \vec{x}'_V and T'_V by tracking of more characteristic points in one frame.
6. INS takes the estimation T'_V from VNS and applying it onto the vector of measured accelerations in BCF and by double integration calculates the new estimation of position \vec{x}''_I .

7. Inside INS, the resultant moving object position estimate is obtained as

$$\vec{x}_{IR} = K_{II} \vec{x}'_I + (1 - K_{II}) \vec{x}''_I. \quad (12)$$

8. Inside VNS, the resultant moving object position estimate is obtained as

$$\vec{x}_{VR} = K_{VV} \vec{x}'_V + (1 - K_{VV}) \vec{x}''_V. \quad (13)$$

9. The resultant estimates on the level of a combined algorithm are obtained as

$$\begin{aligned} \vec{x} &= K_I \vec{x}_{IR} + (1 - K_I) \vec{x}_{VR} \\ \vec{\omega} &= K_I \vec{\omega}'_I + (1 - K_I) \vec{\omega}''_I \\ \vec{V} &= K_I \vec{V}'_I + (1 - K_I) \vec{V}''_I \\ T &= K_I T'_I + (1 - K_I) T'_V \end{aligned} \quad (14)$$

One can note that inside the VNS part, autonomous estimation of linear and angular velocities has been omitted, supposing that in "out-door" applications a lot of points and frames must be used for this purpose, introducing a computational burden this way. For "in-door" applications, where it is possible to process just a relatively small number of points, there exists the reason to produce these estimates also. Based on this additional information, the calculation in the INS part could be extended in order to include T_l'' (based on $\vec{\omega}_l'$) and another calculation of \vec{x}_l (based on \vec{v}_l'), increasing the total number of position estimates in INS from two to four.

Besides the general limitations regarding the computing time required to implement this combined algorithm, as the most important step, one should analyze the likelihood of partial estimates. As the practical measure of decreasing of computing time, one can always consider the possibility to exclude some of the steps 1. - 9. completely, especially if it is a priori possible to predict that their results would be of insufficient accuracy.

Generally speaking, the choice of weighting factors in (14) is a critical step in the whole combined algorithm. It is possible by an improper choice to obtain the resultant estimates worse than in the case of application of a separate algorithm (better among two). While the specification of weighting factor variations is in large extent application dependent, there exists the interest to define some basic principles and adaptation mechanisms, having in mind the nature of errors in INS and VNS. In the particular case of extremely short working time of inertial sensors and good visibility conditions, one can specify constant values of weighting factors, but in the general case it is more adequate to assume that accuracy of separate estimates is changeable and that values of weighting factors should be adapted accordingly.

The first principle regards to the general conclusions about the overall accuracy of INS and VNS stated above. While the accuracy of position estimates in INS is always decreasing in time, the evaluation of accuracy of results obtained by VNS is more complex. As the first, there exists the possibility that this algorithm could not be applied at all (e.g., when a reference object is outside the field of view, for the case when it could not be distinguished uniquely between a few potential ones, etc.) This situation is not possible in an INS except in case that some of the inertial sensors completely failed. As a second, assuming that moving object equipped by a TV camera approaches the reference one in time, it is realistic to expect that overall accuracy increases. However, this increasing is not a continuous one. While by approaching, the relative errors really are smaller for the fixed value of absolute error in determination of characteristic points' coordinates in PCF (expressed in pixels), one can not guarantee that the last ones could not be larger in the next frame. For example, partial occluding of a reference object after it has been detected and tracked in a number of frames could deteriorate the accuracy in large extent. According to this reason, it is assumed that the accuracy of VNS estimates increases in time linearly, from a minimal to maximal one. Simultaneously, using the mechanism of monitoring of VNS estimates, the basic principle is corrected occasionally.

There follows the procedure of adaptation of weighting factor K_l :

1. Before the reliable detection of a reference object inside VNS it is set to: $K_l = 1$.
2. Reliable detection criterion is based on similarity measure between the actual scene contents and the memorized reference pattern. Based on the estimated linear position and transformation matrix obtained by INS, the part of algorithm belonging to VNS makes the required rescaling and rotating of memorized pattern.

It is required that in the window of specified dimension, the value of functional L in the MAD algorithm (Eq. 9) should be lower than the specified threshold value, in N frames continuously. Let's assume that inside the window of dimension 25×25 , the threshold value is $L_{max} = 5$. In this case, if the average absolute difference in light intensities for 625 observed points is less than approximately two quanta, it is supposed that the reference object is detected. After that it is assumed that both autonomous and assisted estimations could be calculated in VNS reliably.

3. VNS starts with tracking of characteristic points of a reference object. The autonomous estimates \vec{x}'_v and T'_v as well as the assisted ones (\vec{x}''_v , $\vec{\omega}''_v$, and \vec{v}''_v) are being calculated. The scene content inside the window around the characteristic point becomes the reference pattern for the next frame analysis.
4. After reliable detection of the reference object in VNS, weighting factor K_I starts to decrease linearly in time.
5. The minimal value of this factor K_{Imin} should be specified for any particular application.
6. The time required for K_I to reach the minimal value is calculated in the VNS part at the beginning of tracking. This calculation based on initial data obtained by INS (position, angular and linear velocities) gives the estimated time of existence of a reference object inside the TV camera's field of view. It is assumed that the moving object approaches the reference one.
7. The similarity measure L is monitored during the tracking of characteristic points. If in the actual frame this one is larger than in the previous one, weighting factor K_I holds the previous value.
8. If at any of the frames the similarity measure is worse than the critical one used as a criterion of reliable detection ($L > L_{max}$), weighting factor K_I is to be reset to value 1 and a detection procedure of a whole reference object is repeated again (back to step 1.).
9. If the conditions of losing the reference object from the field of view are the regular ones (the estimated time of existence has been expired), weighting factor K_I is also reset to the value 1, but the new acquisition of the actual reference object is not going to start. VNS starts a state of waiting on the new reference object recognition.

As it is obvious, parameters like window dimensions, similarity measure threshold, L_{max} , number of frames used for reliable detection, and minimal value of weighting factor for INS estimates (it is at the same time the maximal value of weighting factor for VNS estimates), should be considered as the free ones. They should be carefully specified for the particular application.

It can be noted that in the described procedure the INS part is considered as "master" while the VNS algorithm autonomously evaluates the accuracy of its own estimates at the primary level (correlation between the actual window contents and the actual reference pattern). The described procedure regards to the most general case: flight of a moving object along the specified trajectory with the existence of a number of reference ground objects as landmarks. All cases where the reference object is always present in the field of view are considered as particular ones. For these cases it is reasonable to consider the maximal value of weighting factor K_I as free parameter also (to adopt it as less than 1).

In the expressions (12) and (13) there are the weighting factors affecting autonomous and assisted estimates inside both parts of the algorithm. The way of their adaptation should be considered also.

The position estimate in INS is expressed via (12). If the detection of a reference object in VNS does not exist ($K_I = 1$), the only available estimate is the autonomous one ($K_{II} = 1$). As the assisted estimation \vec{x}_I'' is obtained applying the transformation matrix T_V' onto the accelerometers' outputs directly, validity of this matrix affects the weighting factor for the assisted estimate. The simple validity measure for the transformation matrix is its deviation from orthogonality. If its determinant $\det T_V'$ is outside the specified tolerances (e.g., 1 ± 0.1), it is reasonable to assume that this estimate is not valid and to reject it ($K_{II} = 1$). If T_V' is approximately orthogonal, two cases are possible in general. In the first one, the estimates \vec{x}_I' and \vec{x}_I'' are close (e.g., differences between all coordinates are below the specified values). According to that, K_{II} decreases linearly from the value of 1 to 0.5 minimally, depending on $\det T_V'$. The minimal value (equal weighting of autonomous and assisted estimations) is approached when $\det T_V'$ is inside the small tolerances (e.g., 1 ± 0.01). The second possibility is that T_V' is approximately orthogonal, but the differences between estimations \vec{x}_I' and \vec{x}_I'' are outside the specified tolerances. This is the case when one can assume that the likelihood of an assisted estimate is higher than that of the autonomous one. As a result of this K_{II} should also decrease as previously, but now from maximal value of 1 to $K_{IImin} < 0.5$, depending again on the value of $\det T_V'$. Whatever weight is assigned to \vec{x}_I'' one should note that the resultant position estimate in INS is always dominantly dictated by inertial instruments. At this step, just position increments are being calculated in the combined manner, while the initial conditions are determined by previous results of INS only.

The expression (13) defines the fusion of data on the VNS level. The basic condition for its application is the reliably detected reference object ($K_I < 1$). If the autonomous VNS estimate of angular position is bad (i.e., $\det T_V'$ is outside the specified tolerances) the autonomous linear position estimate \vec{x}_V' is going to be bad also, and accordingly, weighting factor K_{VV} takes the minimal value close to zero. While $\det T_V'$ approaches the value 1, this weighting factor increases linearly up to the maximal value K_{VVmax} . It should be noted again that whatever weighting is assigned to \vec{x}_V'' , calculations in the VNS part are basically dependent on the angular position estimate. Possible invalid estimates \vec{x}_I' due to accumulated inertial sensors' errors are of no importance here, having in mind that calculation of \vec{x}_V'' is based on actual filtered signals $\vec{\omega}_I'$ and \vec{v}_I'' .

4. Simulation Results – Vision Aided INS

The part of an airport runway was considered as a reference ground landmark in vicinity of nominal trajectory. The middle point of a nearer edge is located at known position $\vec{x}_{RO} = [50000 \ 500 \ 0]^T$ in ICF. The interval from $t = 77s$ to $t = 83s$, in which the existence of a landmark could be expected inside the field of view, is predicted on the basis of known fixed parameters of an electro-optical system: angular position of camera relative to BCF ($[\varepsilon_1 \ \varepsilon_2 \ \varepsilon_3]^T = [10^0 \ -5^0 \ 0^0]^T$, focal length, $f = 1$, field of view width, $\varepsilon_{max} = 15^0$). Acquisition of

a reference object is done during the first second of this interval (ten frames). The assumed image of a landmark is obtained using erroneous data from INS. These data are obtained after optimal filtration of the rate gyro signals made inside INS and the linear position errors on observed interval are quantitatively represented in Figure 5.

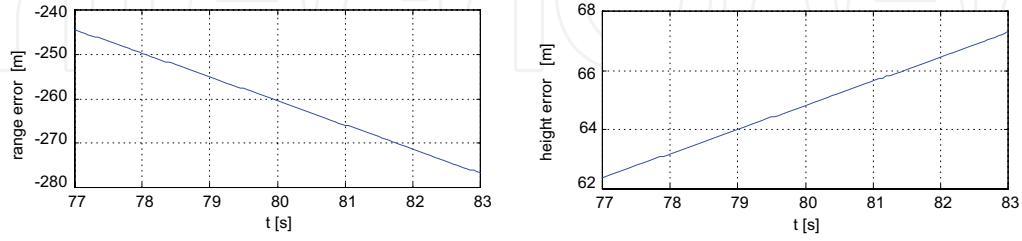


Fig. 5. Position errors of unaided INS during correction interval.

The quality of pitch angle estimation made by INS is illustrated in Figure 6.

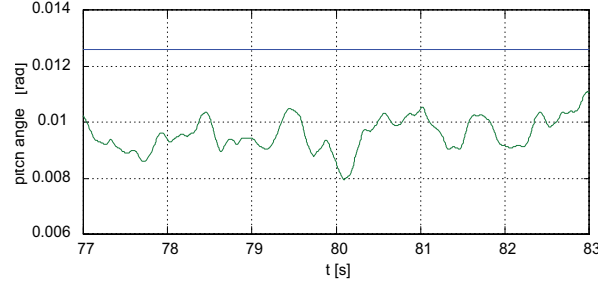


Fig. 6. The exact and estimated values of pitch angle during the correction interval.

Since the estimated values of pitch angle have been obtained using one possible realization of simulated sensor noise, their generation for the purposes of combined navigation method illustration is made using statistical properties of observed time history. The parameters calculated from Figure 6. are: mean value, $\hat{\vartheta}_{sr} = 9.6$ mrad, and standard deviation, $\sigma_{\vartheta} = 0.59$ mrad. Approximately the same standard deviations have been obtained for other two Euler angles, while their mean values are equal to zero. The transformation matrix T_O is generated using the Euler angle estimates generated stochastically.

The fact that in the first frame there is a difference between the actual image of reference object and the expected one is illustrated in Figure 7.

The window of dimensions 25×25 pixels around the lower left corner of a runway was used as a reference pattern. During the first ten frames inside a correction interval, maximum similarity measures formed as a sum of light intensity absolute differences are shown in Figure 8.

Minimum values of MAD criterion are lower than the adopted threshold value of $L_{max} = 5$ in five consecutive frames at $t = 78$ s. It is assumed that the acquisition phase was finished at this time. The next step consists in calculation of the expected time of presence of the characteristic point inside the field of view. It is equal to 5 s for this particular example (under the assumption that linear and angular velocities would be constant). Figure 9.

illustrates the expected contents of the field of view in the interval of next five seconds after the approval of acquisition.

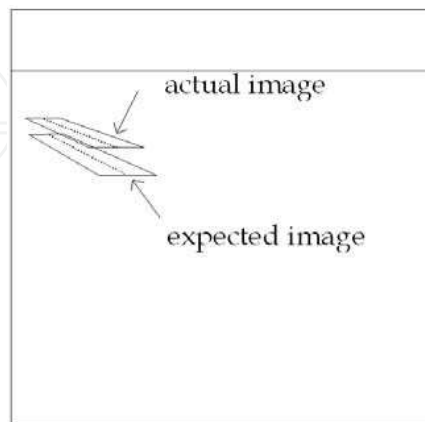


Fig. 7. The expected and actual images of a reference object at the beginning of correction interval.

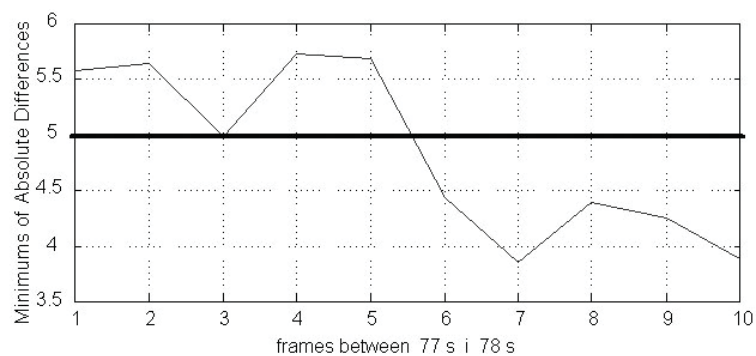


Fig. 8. Maximums of similarity measures during the first second of acquisition interval.

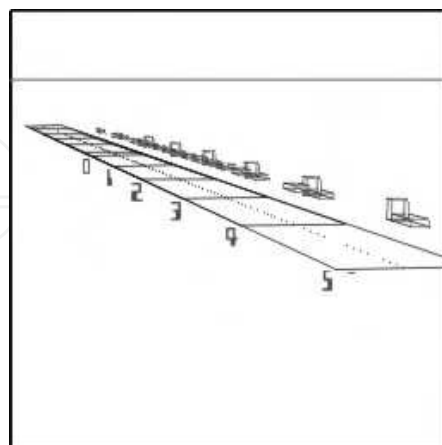


Fig. 9. Expected motion of a reference object during the interval of correction.

According to this, the variation of weighting factor K_I is defined as

$$K_I = 1 - \frac{1-0.1}{5}(t-78). \quad (15)$$

The effects of the combined algorithm will be illustrated in the final portion of the correction interval. The last five frames are shown in Figure 10.

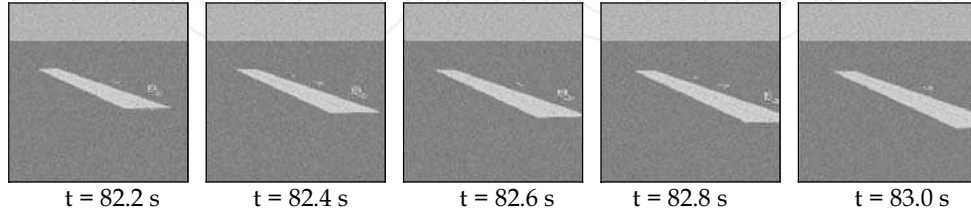


Fig. 10. Sequence of runway images at the end of correction interval.

The exact and estimated positions of a moving object are given in Table 2.

	t = 82.2 s		t = 82.4 s		t = 82.6 s		t = 82.8 s	
	x [m]	z [m]	x [m]	z [m]	x [m]	z [m]	x [m]	z [m]
Exact	48022	200	48135	200	48250	200	48364	200
INS	48295	267	48409	267	48525	267	48640	267
VNS	47997	205	48204	199	48236	203	48362	200
Comb.	48070	220	48247	213	48286	214	48400	209

Table 2. The exact and estimated positions of an object in ICF.

The same results are shown graphically in Figure 11.

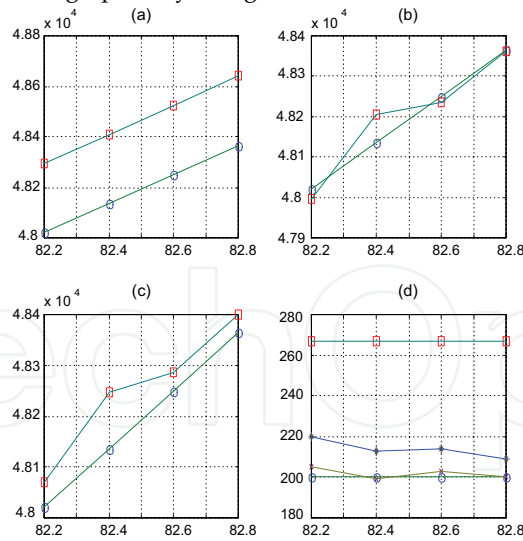


Fig. 11. Comparison of position estimations at the end of correction interval: (a) - range, INS versus exact, (b) - range, VNS versus exact, (c) - range, combined versus exact, (d) - height, exact (circle), INS (square), VNS (cross), combined (star).

As it is obvious from (a), the error in INS is very slowly increasing. From (b) one can see very accurate estimates of VNS while quantization errors as well as errors in determination of characteristic point location occasionally introduce some larger position errors ($t = 82.4$ s). Because of maximal weighting of VNS estimations at the end of the correction interval, beneficial effects of combined algorithm are obvious from (c). Analyzing the results of height estimation (d) one can conclude that the VNS algorithm is extremely accurate, making the results of combined algorithm satisfactory also (suggesting that it is possible to assume an even lower minimum value than $K_{min} = 0.1$).

5. Simulation Results – VNS Assisted by INS

5.1 The Definition of Navigation Tasks

Two particular “in door” navigation tasks have been specified in order to compare the results of application of an autonomous VNS and a dynamic vision navigation algorithm representing a VNS assisted by the acceleration measurements produced by reduced INS:

(Task A): Moving object has got three linear degrees of freedom. In forward direction it moves with the constant velocity (10 m/s). The camera is mounted as forward looking. The initial position of object is assumed as 5 m out of navigation line in lateral direction and 5 m above. The navigation line consists of sequence of rectangular shapes located in the ground plane. The dimension of these landmarks is 1.5 m X 0.15 m with 1.5 m distance between them. The linear velocities in lateral (V_y) and vertical (V_z) directions are controlled and limited to 5 m/s maximally. As a result of a navigation algorithm the actual commanded values of these velocities are calculated as proportional to the estimated distance from center-line of navigation line. The task consists in approaching the navigation line in lateral direction and following of it further. At the same time, a moving object should approach the ground plane (camera at the fixed distance of 1 m above) and continue to move at this height.

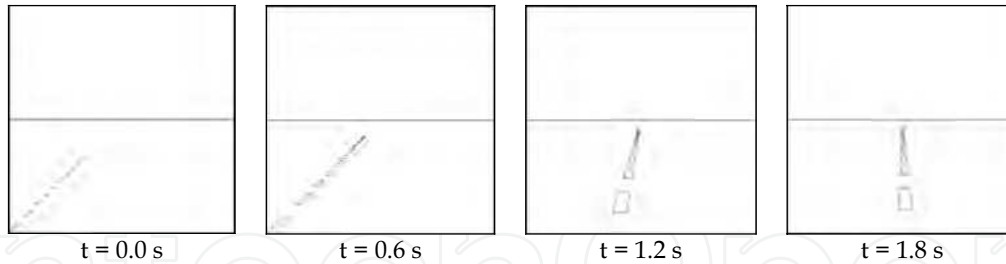


Fig. 12. Sequence of navigation line views (A).

The contents of camera's field of view are computer generated. Figure 12. illustrates the sequence of frames generated at 0.6 s inter-frame interval assuming the ideal work of the VNS algorithm (noise-free images, infinite resolution, without camera vibrations).

(Task B): This task consists of the planar motion control - tracking and following of the curved navigation line. This line consists from two connected circular segments. Their dimensions have been adopted according to the available equipment (see 5.3). Eight approximately rectangular black landmarks of known dimensions are equidistantly placed along the line. A TV camera is forward looking and mounted at 50 mm above the ground plane (Figure 13.)

This navigation task consists in the following:

- 1) Determine the distance from a camera to the reference point belonging to the landmark;
- 2) Determine the orientation of the landmark in the ground plane; 3) Generate the commands for the translation up to the point above the tracked one and for the rotation of a camera around the vertical axis in order to follow the sequence of landmarks.

The autonomous VNS algorithm uses two nearer marker corners for the calculation of scene depth. The determination of marker orientation requires that all four marker's corners should be visible in one frame. If the analyzed marker is partially visible, the algorithm automatically rejects it and analyzes the next one.

In the algorithm based on dynamic vision the linear velocity components are obtained via integration of the accelerometers' signals. In comparison to the autonomous VNS algorithm, the advantage is in the fact that it is enough to track just one landmark corner here. On the other hand, at least two consecutive frames containing the tracked point are required in order to estimate the distance. As a result of this, any selection of the reference point must be preceded by the prediction whether this one would be visible in the next frame. The other disadvantage resulting from the same reason consists of the fact that the navigation should be initialized by a priori known motion in order to acquire the initial information about marker position and after that to adjust the control action according to it. In other words, in comparison to autonomous VNS there will be always "one frame delay" whenever the tracked point is changed.

Additionally, while the autonomous VNS algorithm has the ability of autonomous estimation of landmark's orientation (needed in order to improve the conditions of landmark distinguishing), there is no such possibility in the case of a combined algorithm. As a result of this, the command for rotational motion is generated here on the basis of a ratio of velocity components and the camera is oriented in direction of a velocity vector (which is directed toward the tracked point - *direct guidance*).

5.2 Simulation Methodology

The estimated positions of characteristic points in the image plane are obtained as the ideal ones additionally corrupted by white, Gaussian, zero-mean noise with standard deviation $\sigma = 1$ pixel. The camera vibrations are also simulated as a noisy process. The pitch, yaw, and roll angles are simulated as white, Gaussian, zero-mean noise with standard deviation $\sigma = 1^\circ$.

For the task **(B)** the algorithm explained in Section 4.3 is applied always for the nearest, completely visible element of a navigation line (four corner points). The algorithm presented in Section 4.2 is based on the tracking of one of the corners of the same part of navigation line as in the previous case.

5.3 Experimental Rig Set-up

The experimental acquisition of images is done using the educational robot "Kestrel" with three degrees of freedom in linear and one degree of freedom in rotational motion, driven by step motors, equipped with CCD TV camera (KAPPA CF16/4 P) and using the frame grabber with resolution of 512 X 512 pixels ("Targa")

A photograph of the experimental rig set-up is shown in Figure 13.

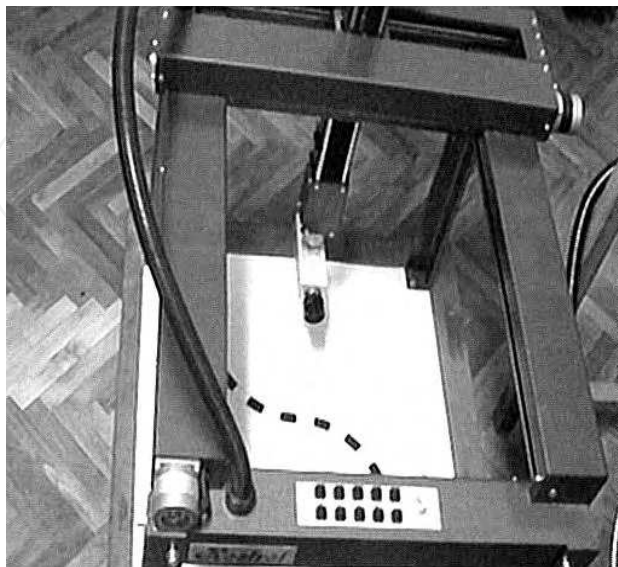


Fig. 13. Experimental rig set-up.

The initial sequence of frames produced by a TV camera during the navigation task **(B)** is presented in Figure 14.

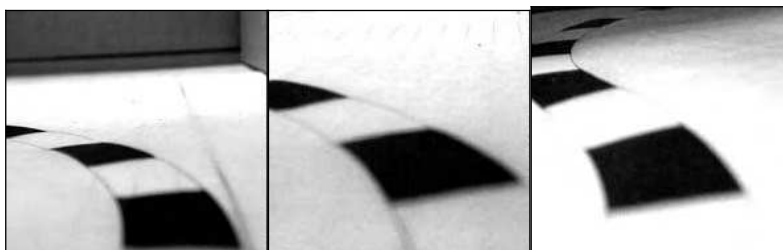


Fig. 14. Experimental sequence of frames produced during the navigation task **(B)**.

The processing phases for the first frame of Figure 15. are illustrated in Figure 15.

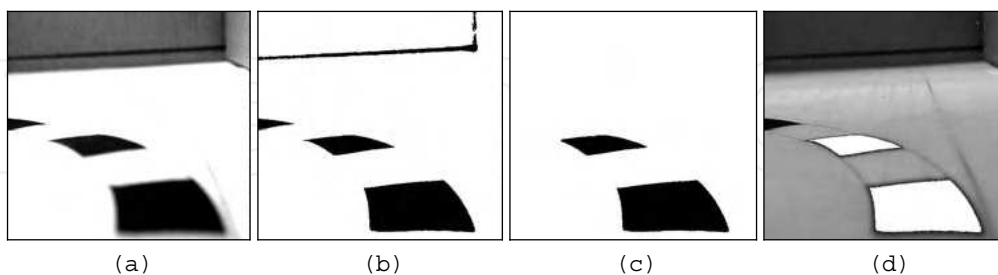


Fig. 15. Results of a processing of frame (1) from Figure 14.

(a) Result of a preprocessing (equalization of image histogram, noise elimination, sharpening); (b) Result of a segmentation based on intensity level; (c) Result of a morphological filtration (cleaning of edges, erosion, dilatation); (d) Inverted result of the

image processing superimposed to the original image in order to notice the differences between the actual landmark and determined one.

5.4 Results

The results obtained by the application of algorithms of the autonomous VNS and with assistance of the linear acceleration measurements (dynamic vision) for the task **(A)** under the methodology defined at 5.2 are shown in Figures 16. and 17. The dashed lines represent in both cases the ideal trajectories obtained for the ideal position estimates (exactly calculated positions of characteristic points, infinite space resolution). The solid lines illustrate the trajectories obtained by the simulated application of described algorithms (through circular symbols representing the actual positions of moving object). The square symbols are used to mark the positions estimated by the navigation algorithms.

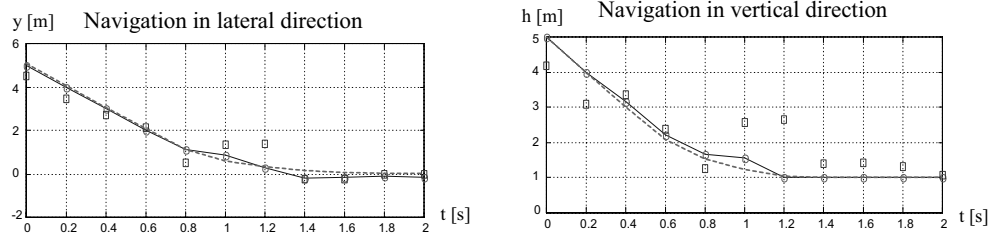


Fig. 16. Simulation results obtained using an autonomous VNS algorithm **(A)**.

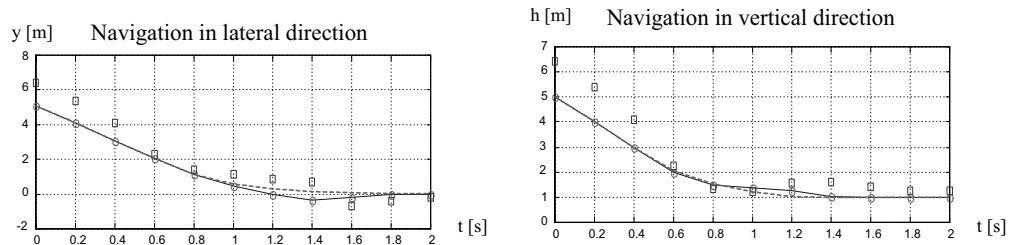


Fig. 17. Simulation results obtained using the dynamic vision algorithm **(A)**.

Comparing these results one can conclude that the trajectories obtained in lateral directions are of similar quality. The advantage of the dynamic vision algorithm is noticeable in the case of navigation in vertical direction. Superior quality of position estimates made that the reaching of the final height of 1 m is much "softer" in comparison with the result obtained via autonomous VNS (landing velocity of 1 m/s in comparison to 3.1 m/s). The inherent disadvantage of a dynamic vision algorithm consisting of "one frame delay" is slightly visible. Its repercussion is the requirement to include in image processing algorithm the additional prediction whether the considered marker would be present in the next frame in the field of view. In spite of this, the image processing part of algorithm remains here less time consuming in the comparison to the autonomous VNS algorithm.

The results illustrating navigation task **(B)** for both autonomous VNS algorithm and combined one are as follows.

Figure 18. illustrates very good results in the estimation of landmarks' angular orientations obtained via autonomous VNS algorithm. The solid line through the circular symbols

represents the actual average values of orientation angle while the dashed line through the square symbols represents the estimated ones.

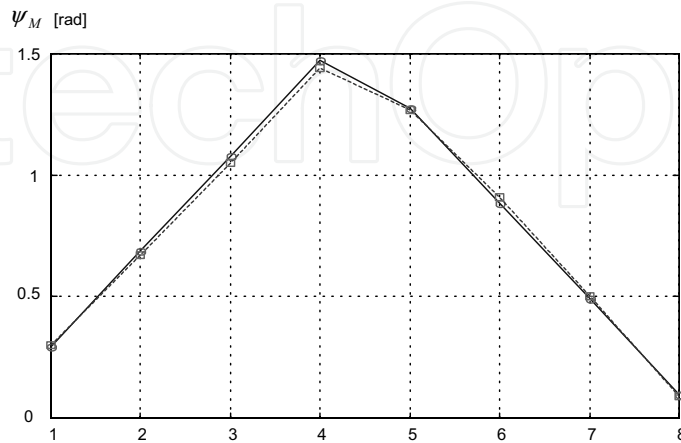


Fig. 18. Estimates of landmarks' angular orientations obtained by the autonomous VNS algorithm **(B)**.

As a result of these accurate estimates, the landmarks are always visible in approximately the same way in a camera's field of view. The trajectory obtained as a result of navigation line following in this case is presented in Figure 19. The initial location was 60 mm behind the coordinate origin. The autonomous VNS algorithm generates the commands positioning the moving object at the location behind the next landmark, at the same distance and along the direction of its orientation. Circular symbols represent the positions of tracked corners of eight landmarks while the square ones represent the consecutive positions of a moving object.

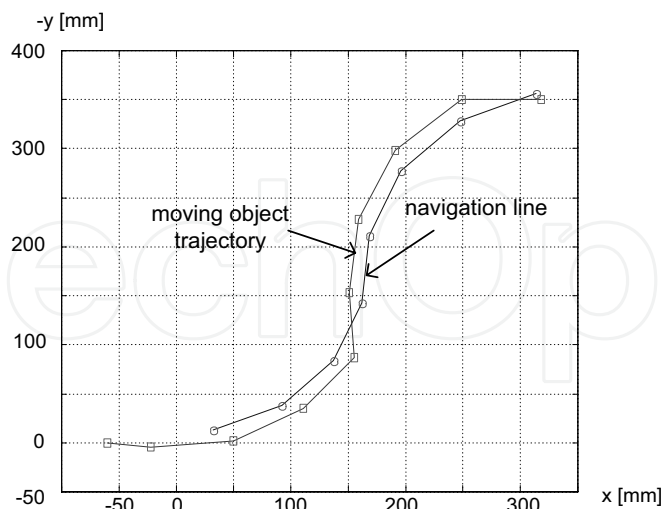


Fig. 19. Following of the navigation line by the autonomous VNS algorithm **(B)**.

In the case of a combined algorithm it is more appropriate to track the further corners of the nearest landmark (they should be present in the field of view in two or more frames). At the very beginning, the motion is started in pre-specified direction in order to acquire the information about the position of a tracked point in the first two frames. After that, the commands moving the object above the tracked point are generated, while the camera is rotated in order to be oriented in direction of a motion. When the new object of "upper left corner" type is detected inside the image, its relative position is calculated in two next frames and the new commands for the linear and angular motion are generated.

Figure 20. illustrates the result of application of a combined algorithm in navigation task (B).

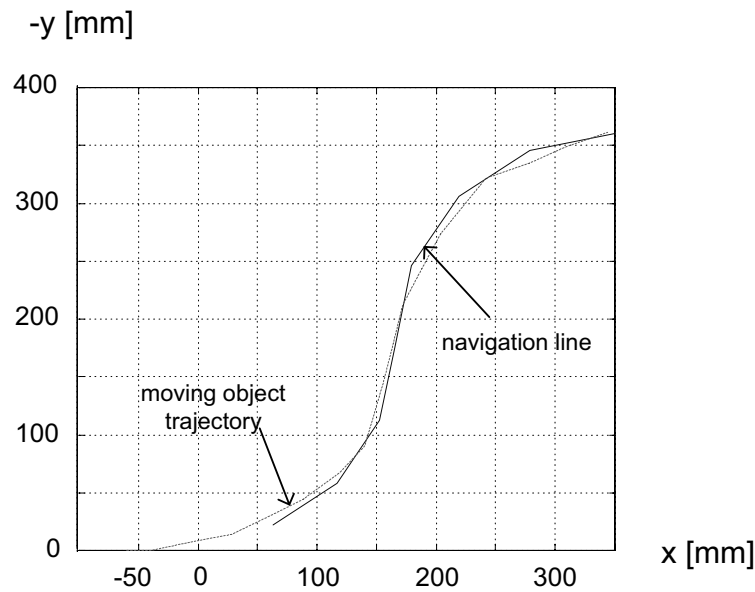


Fig. 20. Following of the navigation line by the dynamic vision algorithm (B).

Due to more meaningful changes in the image contents, one can not recognize here the regularity characterizing the previous case. On the other hand, the reduced amount of calculations allows the higher sampling frequency. As a result of this, a quality of following of the navigation line was slightly better in comparison to the result shown in Figure 19.

6. Conclusion

The algorithm of fusion of data originating from the strap-down inertial navigation system (INS) and the dynamic vision based visual navigation system (VNS) has been suggested for the general case when the appropriate landmarks are in the field of a TV camera's view. The procedure is of weighted averaging type, allowing the adjustment of weighting factors having in mind the physical nature of errors characterizing both systems and according to the self-evaluation of some intermediate estimates made inside the VNS. The overall procedure could be reasonably reduced according to the particular application and to some

a priori knowledge about the possible system errors by excluding some of the possible autonomous or assisted estimates.

Two particular examples have been used in order to illustrate this approach. In the first one, typical for the aerial vehicle motion control, the scenario was constructed under the realistic assumptions about the technical realization of a system, visibility conditions, and the noise levels inside the inertial sensors and in a TV image. It can be concluded that the INS position estimates could be efficiently improved by using the assisted estimates produced by the VNS. While for the height corrections in the INS one can always use a barometric sensor as a simpler solution, the actual benefits of this type of combined algorithm are mostly in the improvements of the position estimates in a horizontal plane (range, lateral deviation from nominal trajectory).

The second example is typical for the mobile robot applications as well as for the automatic motion control of the road vehicles. It represents the integration of a VNS and a reduced INS (just the acceleration measurements integrated in order to enable a dynamic vision based algorithm). This system has shown the advantages in comparison to the autonomous VNS. These consist mainly in the reductions of the computations regarding the distinguishing of the characteristic points of a reference object as well as in some improvements of a position estimation accuracy also (as a consequence of a relaxing the overall accuracy dependence on the results of image processing part of algorithm only). Finally, it should be pointed out that a VNS algorithm assisted by the INS data requires no a priori information about the shape and dimensions of the reference objects, which is beneficial also.

The analyzed examples are relatively simple ones but still meaningful for the vehicular motion control applications. More complex tasks including the rotational degrees of freedom should be considered as the more general cases of a fusion of VNS and INS, where the set of inertial instruments can be extended by using of the rate gyros. This way, the complex and noise sensitive procedures of determining of an angular orientation of a mobile robot or a vehicle, based on a machine vision alone, can be replaced by the usage of the inertial sensors' data. The future research is going to be oriented in this way.

7. References

- Farrell, J.A., Barth, M. (1999): *"The Global Positioning System & Inertial Navigation"*, McGraw Hill, 1999.
- Frezza, R.; Perona, P.; Picci, G.; Soatto, S. (1994): "System-Theoretic Aspects of Dynamic Vision", in *"Trends in Control"*, A. Isidori, (Ed.): Springer, Berlin, 1994., pp. 349-383.
- Graovac, S. (2002): "A New Visual Navigation Algorithm Using Linear Acceleration Measurements", *Proceedings of the 10th Mediterranean Conference on Control and Automation*, Lisbon, Portugal, July 2002.
- Graovac, S. (2004): "Principles of Fusion of Inertial Navigation and Dynamic Vision", *Journal of Robotic Systems*, Vol. 21, N°1, pp. 13-22, January 2004.
- Grewal, M. S., Weill, L. R. Andrews, A.P. (2001): *"Global Positioning Systems, Inertial Navigation, and Integration"*, John Wiley & Sons, Inc., 2001.
- Kaminer, I.; Pascoal, A.; Kang, W. (1999): "Integrated Vision/Inertial Navigation System Design Using Nonlinear Filtering", *Proceedings of the American Control Conference*, Vol. 3, pp. 1910 – 1914, San Diego, California, USA, June 1999.

- Kanatani, K. (1993): *"Geometric Computation for Machine Vision"*, Clarendon Press, Oxford, 1993.
- Menon, P.K.A., Chatterji, G.B. Sridhar, B. (1993): "Electro-Optical Navigation for Aircraft", *IEEE Trans. On Aerospace and Electronic Systems*, Vol. AES-29, July 1993, pp. 825-832.
- Roumeliotis, S.; Johnson, A.; Montgomery, J. (2002): "Augmenting Inertial Navigation with Image-Based Motion Estimation" *Proceedings of the 2002 IEEE International Conference on Robotics and Automation, ICRA 2002*, Vol. 4., pp. 4326-4333, May 2002, Washington DC, USA.



Mobile Robots: Perception & Navigation

Edited by Sascha Kolski

ISBN 3-86611-283-1

Hard cover, 704 pages

Publisher Pro Literatur Verlag, Germany / ARS, Austria

Published online 01, February, 2007

Published in print edition February, 2007

Today robots navigate autonomously in office environments as well as outdoors. They show their ability to beside mechanical and electronic barriers in building mobile platforms, perceiving the environment and deciding on how to act in a given situation are crucial problems. In this book we focused on these two areas of mobile robotics, Perception and Navigation. This book gives a wide overview over different navigation techniques describing both navigation techniques dealing with local and control aspects of navigation as well as those handling global navigation aspects of a single robot and even for a group of robots.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Stevica Graovac (2007). One Approach to the Fusion of Inertial Navigation and Dynamic Vision, Mobile Robots: Perception & Navigation, Sascha Kolski (Ed.), ISBN: 3-86611-283-1, InTech, Available from: http://www.intechopen.com/books/mobile_robots_perception_navigation/one_approach_to_the_fusion_of_inertial_navigation_and_dynamic_vision

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2007 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen