VISION AIDED INERTIAL NAVIGATION

Denis Grießbach, Dirk Baumbach and Sergey Zuev

German Aerospace Center (DLR)
Institute of Robotics and Mechatronics
Rutherfordstrasse 2, 12489 Berlin, Germany
denis.griessbach@dlr.de

Commission I/5

KEY WORDS: Inertial navigation, Optical navigation, Indoor navigation, Computer vision

ABSTRACT:

Today GPS aided inertial navigation is widely used and well studied in any aspects. The good short term properties of inertial data are complemented by the long term stability of the GPS signal. A common approach is to use a Kalman filter for fusing GPS and inertial data to constrain the inertial sensor drift. Although this is working well in many applications there is a need to find a similar solution for navigation tasks in difficult environments with erroneous or no GPS data. Therefore a vision aided inertial navigation system is presented which is capable of providing local navigation for indoor applications without GPS or could be used to bridge GPS dropouts in urban or forested areas. A method is described to reconstruct the ego motion of a stereo camera system aided by inertial data which in turn is used to constrain the inertial sensor drift. The optical information is derived from natural landmarks which are extracted and tracked over consequent stereo image pairs. Using inertial data for the feature tracking effectively reduces the computational effort and the uncertainties from mismatching by allowing smaller search areas. This is an important precondition for a robust tracking algorithm running in real time. However, before fusing the data with a Kalman filter many aspects including synchronization and determination of alignments of the sensor systems as well as the calibration of the stereo camera system have to be considered. The results of an integration of optical and inertial navigation are shown on an indoor navigation task.

1 INTRODUCTION

For many applications in indoor as well as in outdoor environments accurate navigation is required. GPS aided inertial navigation is used to provide position and orientation for airborne and automotive tasks. Although this is working very well it has major weaknesses in difficult environments with erroneous or no GPS data, e.g. urban or forested areas. Furthermore, this method is not working in indoor environments as needed for robotic applications because without GPS data the integration of inertial data leads to an unbound error grow resulting in an erroneous navigation solution. To restrain this error reasonable measurements of an external sensor are needed. Some proposed solutions require active measurements, e.g. radar, laser range finder, etc. or local area networks which have to be established first (Zeimpekis et al., 2003). On the other hand vision can provide enough information from a passive measurement of an unknown environment to serve as an external reference. A stereo based approach was preferred to obtain 3D information from the environment which is used for self localization and ego motion respectively. Both, inertial and optical data are fused within a filter and provide an accurate navigation solution. This work will show a framework for a multi-sensor system regarding hardware and software integration. Furthermore some difficulties and possible solutions are pointed out.

2 STEREO VISION

First of all, measuring in images means to have an exact knowledge of the camera model and an additional nonlinear distortion model (Brown, 1971). In case of a stereo vision system this is extended by the exterior orientation between the cameras. Several methods for calibrating camera systems have been made. Many of them use observations of predefined calibration grids to extract all camera parameters with a complex bundle adjustment (Zhang, 2000, Tsai, 1987). Another way, eliminating problems with the classic approaches is to use holographic pattern to achieve a reliable camera calibration (Grießbach et al., 2008).

The basic idea is to identify corresponding points in images of stereo camera pairs in two consecutive frames. Calculating 3D object points for the first stereo frame allows for reconstructing the relative orientation change of the second frame.
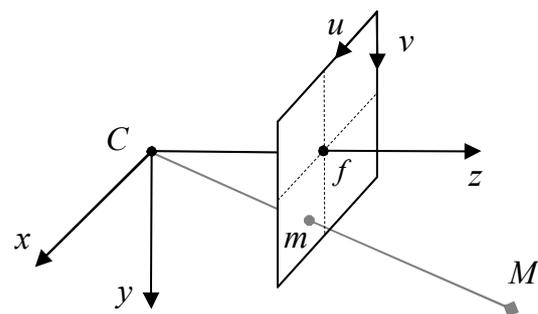
2.1 Camera model

Figure 1: Pinhole camera model

In projective space $\mathbb{P}$ mapping of a homogeneous object point $\tilde{\mathbf{M}} \in \mathbb{P}^3$ to an image point $\tilde{\mathbf{m}} \in \mathbb{P}^2$ is defined with,

$$\tilde{\mathbf{m}} = P\tilde{\mathbf{M}} \tag{1}$$

where P is a 3×4-projection matrix consisting of the parameters of the interior- and exterior orientation of the camera.

$$P = K\left[R|\mathbf{t}\right] \tag{2}$$

with R, $\mathbf{t}$ describing the rotational matrix and translation of the exterior orientation and the camera matrix K containing the focal

length f and the principal point u_0, v_0.

$$K = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

Before applying the pinhole model lens distortion has to be considered. There are several distortion models available. The most common is the radial distortion model by Brown (Brown, 1971) considering pincushion or barrel distortion which is expressed as follows,

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} (1 + k_1 r^2 + k_2 r^4 + k_3 r^6 + \cdots) \tag{4}$$

with

$$r^2 = x^2 + y^2 \tag{5}$$

and

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{6}$$

where x, y are normalized image coordinates calculated from the image coordinates $\tilde{\mathbf{m}} = [u, v, 1]^T$.

2.2 Triangulation

For reconstruction of a 3D object point from two images a pair of corresponding image points $\mathbf{m} = [x, y]^T, \mathbf{m}' = [x', y']^T$ as seen in figure 2 is needed. Given these and the associated projection matrices P, P' leads to a homogeneous equation $A\mathbf{M} = 0$ with:

$$A = \begin{bmatrix} x\mathbf{p}^{3T} - \mathbf{p}^{1T} \\ y\mathbf{p}^{3T} - \mathbf{p}^{2T} \\ x'\mathbf{p}'^{3T} - \mathbf{p}'^{1T} \\ y'\mathbf{p}'^{3T} - \mathbf{p}'^{2T} \end{bmatrix} \tag{7}$$

where $\mathbf{p}^{iT}$ are the rows of the projection matrix P. After solving by singular value decomposition the object point $\tilde{\mathbf{M}}$ is found. This optimization minimizes the geometric error in projective space $\mathbb{P}^3$. A better choice is to look at the image space $\mathbb{R}^2$. With known fundamental matrix F this is done by using the epipolar constraint $\tilde{\mathbf{m}}'^T F \tilde{\mathbf{m}} = 0$ which minimizes the distance between feature point and epipolar line. Because an exact solution for the problem is costly to obtain the sampson approximation δ is used to correct the feature point with $\hat{m} = m - \delta_m$ (Hartley and Zisserman, 2004). This is also a first good indicator for incorrect features correspondences.
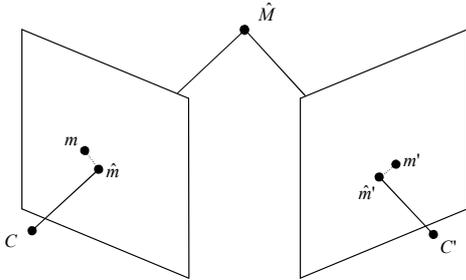

Figure 2: triangulation with corrected image points

2.3 Move estimation

Having triangulated all image points from the first stereo frame $(k-1)$ to 3D object points $\tilde{\mathbf{M}}_i$ these points can be projected into the images of the consecutive frame k by adding a transformation T between both frames.

$$\tilde{\mathbf{m}}_i(k) = P \, T \, \tilde{\mathbf{M}}_i(k-1) \tag{8}$$

with:

$$T = \begin{bmatrix} R & \mathbf{t} \\ 0 & 1 \end{bmatrix} \tag{9}$$

Taking corresponding image points $\tilde{\mathbf{m}}_i$ from the second frame leads to an optimization which minimizes the euclidean distance in image space $\mathbb{R}^2$ and again gives the opportunity to check for incorrect matched features. This is done by using a local optimized RANSAC algorithm which gives the opportunity to separate moving objects from the static scene.

2.4 Feature extraction

To perform aforementioned steps it is necessary to have a selection of feature points which have a defined position and are detectable over consecutive frames. Natural landmarks such as corners, isolated points or line endings meet these requirements. Harris proposed to locate these features by analyzing the autocorrelation matrix A (Harris and Stephens, 1988).

$$A = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix} \tag{10}$$

where I_x, I_y are the partial derivatives of the image intensity I. The corner strength M is calculated with

$$M = \det(A)\kappa \, \text{trace}(A)^2 \tag{11}$$

with κ as an empirical value between 0.04 and 0.15. Another possibility is to look at the smallest eigenvalue $\min(\lambda_1, \lambda_2)$ of A being a better indicator than M (Shi and Tomasi, 1994).

2.5 Feature matching

A feature can be described by its local neighborhood, refereed as template. Although different correlation methods are used normalized cross correlation (NCC) will be applied here to find a match of this template within the search image.

$$NCC(x, y) = \frac{1}{n-1} \sum_{x,y} \frac{(S(x, y) - \bar{S})(T(x, y) - \bar{T})}{\sigma_S \sigma_T} \tag{12}$$

where $T(x, y)$ denotes the template and $S(x, y)$ a sub image of the search image. Having real time applications in mind it is certainly very important to restrict the search area as much as possible. This is not only speeding up the process it also avoids mismatching for example at repetitive structures.
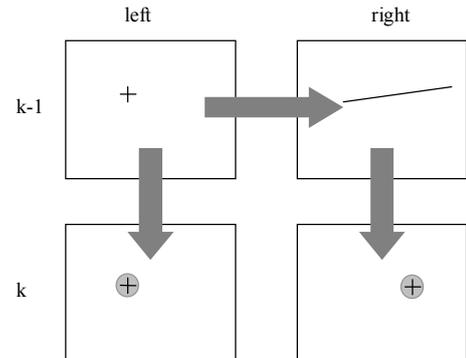

Figure 3: Paths for intra-frame and inter-frame matching

Assuming some good features have been extracted in the left image of frame $(k-1)$ the next step is to find the corresponding features in the right image (intra-frame matching). In this case the search area can be confined to the epipolar line. Limiting the distance of considered object points is also limiting the length of the epipolar line and reduces the search area. The result of an intra frame matching is show in figure 4.

A second step is to match both stereo frames. Equation (8) shows how to restrict search areas for the inter-frame matching after calculating the 3D object coordinates $\tilde{\mathbf{M}}_i$. The only thing needed is some information about the transformation between the frames to predict the feature positions in the second frame k. This information as shown in the next section can be derived from inertial navigation enabling for rather small search areas.

Figure 4: Stereo image with intra-frame matching

3 INERTIAL NAVIGATION

Inertial navigation systems (INS) consist of an inertial measurement unit (IMU) containing 3-axis-gyroscopes and -acceleration sensors as well as a computing unit to integrate the IMU signals to a navigation solution. Two different types of mechanization are used, the gimbaled platform and the strapdown mechanization (Titterton and J.L.Weston, 2004, Savage, 2000). As strapdown mechanization is affordable, lightweight and computational power is available it is chosen. Figure 5 shows the mechanization with the superscripts b and n standing for body- and navigation-frame.
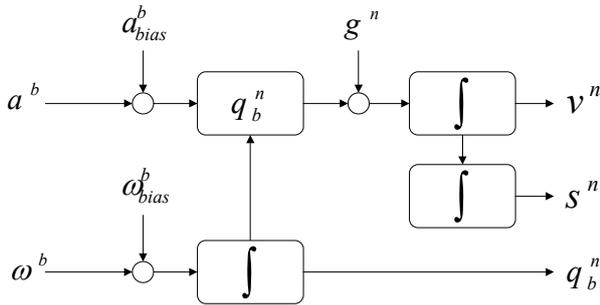

Figure 5: Strapdown mechanization

It also shows possible difficulties when integrating accelerations $\mathbf{a}^b$ and angular velocity ω^b as the bias terms $\mathbf{a}^b_{bias}, \omega^b_{bias}$ are not only unknown but also not constant. This leads to a strong drift for velocity $\mathbf{v}^n$, position $\mathbf{s}^n$ and attitude $\mathbf{q}^n_b$ if left uncompensated. In addition the gravitation constant $\mathbf{g}^n$ has to be taken into account. The corresponding 22-dimensional state vector includes 7 parameters each 3-dimensional except the quaternion with 4 components.

$$\mathbf{x} = \begin{bmatrix} \mathbf{s}^n & \mathbf{v}^n & \mathbf{a}^n & \mathbf{a}^b_{bias} & \mathbf{q}^n_b & \hat{\omega}^b & \omega^b_{bias} \end{bmatrix}^T \quad (13)$$

Quaternions were used for the attitude calculation because of their compact rotation representation and there lack of singularities called the "gimbal lock" compared to euler angles (Grewal et al., 2000). A constant angular velocity and acceleration model is used for the discrete non-linear system equations f which can be described as follows,

$$f(\mathbf{x}) = \begin{pmatrix} \mathbf{s}^n + \mathbf{v}^n dt + \mathbf{a}^n \frac{dt^2}{2} \\ \mathbf{v}^n + \mathbf{a}^n dt \\ \mathbf{q}^n_b \circ \left[0, \mathbf{a}^b - \mathbf{a}^b_{bias} \right]^T \circ \mathbf{q}^{n*}_b + \mathbf{g}^n \\ \mathbf{a}^b_{bias} \\ \mathbf{q}^n_b \circ \mathbf{r} \\ \omega^b - \omega^b_{bias} \\ \omega^b_{bias} \end{pmatrix} \quad (14)$$

where dt denoting the integration time, $\mathbf{q}^*$ the conjugate of the quaternion and the operator $\circ$ describes a quaternion multiplication (Burchfiel, 1990, E. B. Dam and Lillholm, 1998). A first-order integrator with $\mathbf{r} = 0.5 \begin{bmatrix} 2 & \hat{\omega}_x dt & \hat{\omega}_y dt & \hat{\omega}_z dt \end{bmatrix}^T$ describes an approximation for the quaternion update.

3.1 Kalman filter

To estimate and compensate for various errors like drift, random walk, bias, scaling, etc. independent measurements are needed. Usually these (e.g. GPS) are fused with the inertial data by a Kalman filter. This recursive filter estimates the state vector $\mathbf{x}$ of a linear dynamic system from different noisy observations $\mathbf{z}$ with $\mathbf{w}, \mathbf{v}$ being white independent gaussian noise (Welch and Bishop, 1995, Ribeiro, 2004).

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} \quad (15)$$
$$\mathbf{z}_{k-1} = h(\mathbf{x}_{k-1}) + \mathbf{v}_{k-1} \quad (16)$$

Having non-linear state equations f and observation equations h the extended Kalman filter (EKF) gives an approximation of the optimal estimate by linearizing the non-linear system around the last state estimate as shown in equation (17, 18).

$$F_k = \nabla f|_{\hat{\mathbf{x}}_k} \quad (17)$$
$$H_k = \nabla h|_{\hat{\mathbf{x}}_k} \quad (18)$$

time update: The time update equations propagate the current state estimate and error covariance P forward in time, also referred as a priori estimate $(^-)$. Q denotes the covariance of the process noise.

$$\hat{\mathbf{x}}^-_k = f(\hat{\mathbf{x}}^+_{k-1}) \quad (19)$$
$$P^-_k = F_{k-1} P^+_{k-1} F^T_{k-1} + Q \quad (20)$$

measurement update: This serves as input for the measurement update which includes new measurements with uncertainty R to obtain an a posteriori estimate $(^+)$.

$$\hat{\mathbf{x}}^+_k = \hat{\mathbf{x}}^-_k + K_k(\mathbf{z}_k - h(\hat{\mathbf{x}}^-_k)) \quad (21)$$
$$K = P^-_k H^T_k (H_k P^-_k H^T_k + R)^{-1} \quad (22)$$
$$P^+_k = (I - KH_k)P^-_k \quad (23)$$

3.2 Unscented Kalman filter

Since the EKF is only considering first order linearization terms it is not performing well for highly non-linear problems. Therefore the scaled unscented Kalman filter (sUKF) uses deterministic chosen sample points (sigma points) around the mean which

are propagated by the non-linear state- and observation equations (Julier, 2002, Julier and Uhlmann, 2004, Van Der Merwe, 2004). The sigma points are calculated as follows,

$$\chi^+_{k-1,0} = \mathbf{x}^+_{k-1} \tag{24}$$

$$\chi^+_{k-1,i} = \mathbf{x}^+_{k-1} + \sqrt{(n+\lambda)\,P^+_{k-1,i}} \quad i = 1 \ldots n$$

$$\chi^+_{k-1,2n+i} = \mathbf{x}^+_{k-1} - \sqrt{(n+\lambda)\,P^+_{k-1,i}} \quad i = 1 \ldots n$$

$$w^m_0 = \frac{\lambda}{n+\lambda}$$

$$w^c_0 = \frac{\lambda}{n+\lambda} + \left(1 - \alpha^2 + \beta\right)$$

$$w^m_i = w^c_i = \frac{1}{2\,(n+\lambda)} \quad i = 1 \ldots 2n$$

with the number of states n. Scaling parameters are defined by $\beta = 2$ for gaussian distributions and $\lambda = \alpha^2\,(n+\kappa) - n$ with k usually set to 0. α determines the spread of the sigma points around mean.

time update: The sigma points are propagated with the non-linear state equations and recombined with weighting factors w to the predicted mean and covariance.

$$\chi^-_{k,i} = f\left(\chi^+_{k-1,i}\right) \tag{25}$$

$$\mathbf{x}^-_k = \sum_{i=0}^{2n} w^m_i \chi^-_{k,i} \tag{26}$$

$$P^-_{k,xx} = \sum_{i=0}^{2n} w^c_i \left(\chi^-_{k,i} - \mathbf{x}^-_k\right)\left(\chi^-_{k,i} - \mathbf{x}^-_k\right)^T + Q \tag{27}$$

measurement update: Similar to the time update the a posteriori estimate of mean and covariance are calculated.

$$\gamma_{k,i} = h\left(\chi^-_{k,i}\right) \tag{28}$$

$$\epsilon_k = \sum_{i=0}^{2n} w^m_i \gamma_{k,i} \tag{29}$$

$$P_{k,zz} = \sum_{i=0}^{2n} w^c_i \left(\gamma_{k,i} - \epsilon_k\right)\left(\gamma_{k,i} - \epsilon_k\right)^T + R \tag{30}$$

$$P_{k,xz} = \sum_{i=0}^{2n} w^c_i \left(\chi^-_{k,i} - \mathbf{x}^-_k\right)\left(\gamma_{k,i} - \epsilon_k\right)^T \tag{31}$$

$$K = P_{k,xz} P^{-1}_{k,zz} \tag{32}$$

$$\mathbf{x}^+_k = \mathbf{x}^-_k + K\left(\mathbf{z}_k - \epsilon_k\right) \tag{33}$$

$$P^+_{k,xx} = P^-_{k,xx} - K P_{k,zz} K^T \tag{34}$$

Besides from being computational more costly than the EKF the sUKF has many advantages. By being accurate to the third order with gaussian inputs it produces more stable estimates of the true mean and covariance. Furthermore, it behaves better in case of unknown initialization values and analytic derivations are not needed. With highly non-linear observation equations of the used sensors just like parts of the state equations the sUKF ensures a reliable navigation solution.

4 EXPERIMENTAL SETUP

A demonstrator for visual aided inertial navigation system has been developed. The core component is a low cost MEMS-IMU
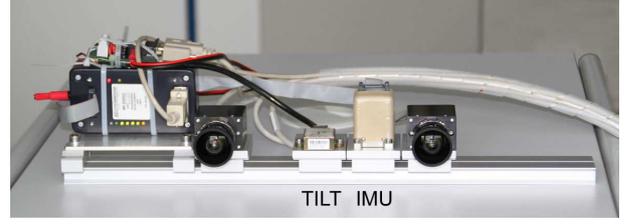

Figure 6: Sensor head with cameras, IMU and tilt sensor

(MEMS standing for Microelectromechanical system) which implies a low quality gyroscope compared with mechanical, fibre optic or ring laser gyroscopes regarding noise and bias stability. On the other hand they are much cheaper, more robust and reasonably smaller (Schmidt, 2009). The used IMU shows noise terms of 0.035 deg/s and a bias stability of 7.4 deg/hr for the gyroscopes respectively 1.3 mg/ 70μg for the acceleration sensors. The data rate is 410 Hz. In the current setup the INS is aided by a stereo camera system with 200 mm base that provides increments of attitude and position. The combination with the used lenses gives very good results for indoor environments.

Frequency	max. 30 Hz
Sensor size	1360×1024
Pixel size	6.45 μm
Focal length	4.8 mm
FoV	42.8×34.5 deg

Table 1: Camera parameters

Additionally a 2-axis inclinometer with a noise of 0.027 deg is also included to support the state estimation if the system is not moving. It measures roll and pitch angles regarding an earth fixed coordinate system being the only absolute unbiased measurement in the process. The heading angle and all other states are without absolute aiding and reliant on the stereo camera system alone. All components are mounted to an optical bench to achieve a stable setup.

4.1 Data flow

Figure 7 shows the combination of Kalman filter and optical system which provides incremental attitude- and position updates. Receiving IMU or tilt measurements the full filter cycle is completed including a check for plausibility of the data. For incoming stereo images first the time-update is done. The a priori estimate enables the tracker to perform a very fast and reliable interframe matching. After triangulation the calculated move estimate is used for the measurement-update within the Kalman filter.
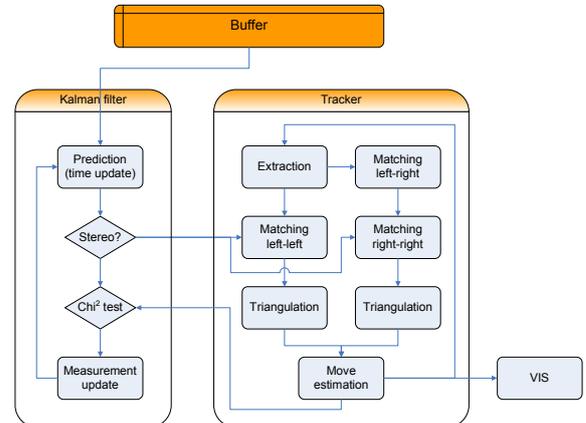

Figure 7: Data flow

4.2 Synchronization and real time issues

Fusing data as shown in section 3.1 requires synchronous measurements. Because it is not possible to trigger any sensor the asynchronous sensor data is grabbed by a FPGA-card generating a time stamp for each incoming data set. This card is equipped with various interfaces as RS232, trigger-inputs and -outputs, ethernet, etc. Both cameras are connected to one trigger-output to get synchronized images as well. An integrated GPS providing GPS-position and -time enables to synchronize with UTC-time if necessary.
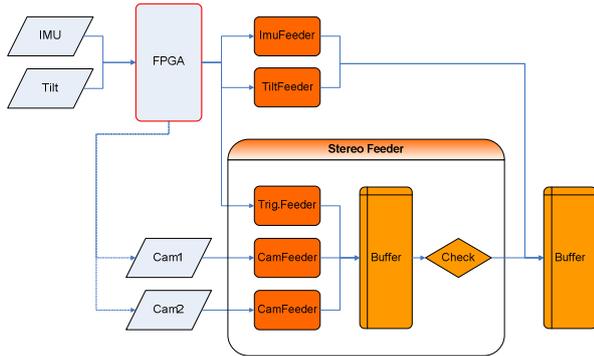

Figure 8: Synchronization of grabbed data

The goal of this work was to build a system not only for logging sensor data and process it offline but to handle the data online. This includes data grabbing, image processing, data fusing, logging, etc. For this reason a buffer concept was developed to manage different processing times of diverse incoming data. This buffer is also responsible for feeding the Kalman filter with time-wise ordered data despite any latency times.

4.3 Calibration and alignment

Another, often underestimated problem is the complex alignment procedure for the whole sensor system. First step is to calibrate the stereo camera system as shown in (Grießbach et al., 2008). It is now feasible to calculate the rotation between camera and tilt sensor. Therefore the orientation of the camera system and the tilt angles are measured with a few static positions in front of a calibration chart.

Finding out about the alignment between camera system and IMU is slightly more ambitious. Since the IMU is only measuring accelerations and angular velocities a dynamic calibration is needed. Comparing both, measured IMU angular velocities and calculated angular velocities from the camera system as shown in section 2.3 the rotational alignment can be estimated. The translation of the systems has to be estimated with a Kalman filter extended by additionally states for the alignment parameters.

An absolute reference for the dynamic calibration move would make it much easier to obtain the alignment parameters.

5 RESULTS

To show the capability of the system an indoor environment was chosen providing triangulated markers which were not used for navigation but only for validating the later result. The distance from a start mark to a second arbitrarily mark seen by the cameras had to be measured in real time without post processing the data. A course off about 90 m length leading over two floors as shown in figure 9 was selected. At the beginning of every run the system was not moved for about 45 seconds to initialize the Kalman filter.

This corresponds to 360 stereo images taken with 8 Hz frame rate. Due to a bottleneck within image processing this is the maximum rate the system can cope with. After this time the filter has "run in" and walking the course could be started.
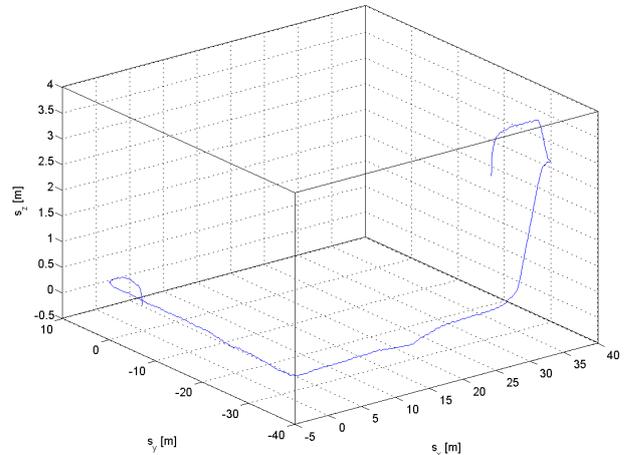

Figure 9: Path 3D-view

With normal walking speed of about 1.5 m/s the destination mark was reached after 85-90 seconds with a final distance error of 20-50 cm for several runs. This difference is mainly caused by phases where no or little features could be seen resulting in no or low quality vision data and as a consequence an increased error grow from integrating the IMU measurements. This situations occur through difficult and changing lighting conditions, a low texturing at some areas or for example at narrow stairways with all objects to close for the stereo system. In good conditions the tracker uses up to 100 features and achieves a frame to frame position error of 5 mm respectively 2 mm for the viewing axis and an attitude error of about 0.2 / 0.1 degrees for typical indoor applications. This strongly depends on the number and distribution of the seen features. Figure 10 shows the floor plan with the overlaid trajectory calculated from the system.
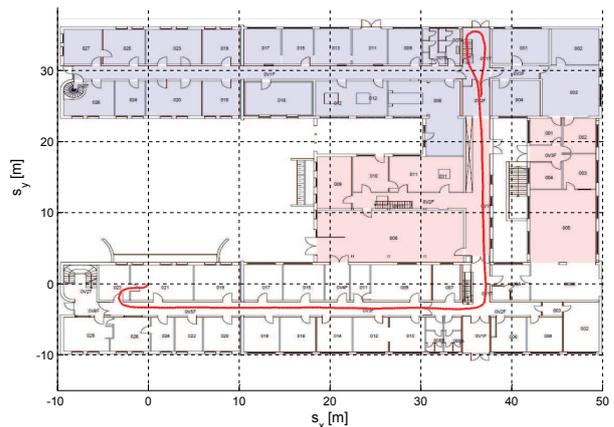

Figure 10: Path top view

It is to stress that the experiment was not conducted in protected environment but during working hours with many people interfering the optical measurements. Due to the before mentioned build in safety mechanisms while matching, triangulating and move estimation the system was working very robust. Mismatched features on moving objects are also prevented by using the a priori estimate from inertial navigation for inter-frame matching.

6 CONCLUSION

A system has been presented combining inertial- and optical navigation. In doing so the disadvantages of the individual systems are compensated by the other. Inertial navigation error grow is significantly reduced whereas the performance and reliability of the optical navigation is increased. Although the system is working at 8 Hz frame rate further work has to be done to speed up some parts of computational costly image processing. The proposed system provides a robust solution for navigation tasks in difficult indoor environments and is also conceivable for supporting conventional GPS-INS navigation. Future work will address this issue for outdoor applications. Important points are also the integration of other external sensors e.g. barometer as well as improvements for the optical navigation.

REFERENCES

Brown, D. C., 1971. Close-range camera calibration. Photogrammetric Engineering 37, pp. 855–866.

Burchfiel, J., 1990. The advantage of using quaternions instead of euler angles for representing orientation. The Third Workshop on Standards for the Interoperability of Defense Simulations.

E. B. Dam, M. K. and Lillholm, M., 1998. Quaternions, interpolation and animation. Technical Report DIKU-TR-98/5, University of Copenhagen.

Grewal, M. S., Weill, L. R. and Andrews, A. P., 2000. Global Positioning Systems, Inertial Navigation, and Integration. Wiley-Interscience.

Grießbach, D., Bauer, M., Hermerschmidt, A., Krüger, S., Scheele, M. and Schischmanow, A., 2008. Geometrical camera calibration with diffractive optical elements. Opt. Express 16(25), pp. 20241–20248.

Harris, C. and Stephens, M., 1988. A combined corner and edge detector. In: in Proc. of the 4th ALVEY Vision Conference, pp. 147–151.

Hartley, R. and Zisserman, A., 2004. Multiple View Geometry in Computer Vision. Cambridge University Press.

Julier, S. J., 2002. The scaled unscented transformation. In: Proceedings of the 2002 American Control Conference, Vol. 6, pp. 4555–4559.

Julier, S. J. and Uhlmann, J. K., 2004. Unscented filtering and nonlinear estimation. Proceedings of the IEEE 92(3), pp. 401–422.

Ribeiro, M. I., 2004. Kalman and extended kalman filters: Concept, derivation and properties. Technical report, Institute for Systems and Robotics, Lisboa.

Savage, P. G., 2000. Strapdown Analytics. Strapdown Associates, Inc. Maple Plain, Minnesota.

Schmidt, G. T., 2009. Ins/gps technology trends. In: RTO-EN-SET-116, Low-Cost Navigation Sensors and Integration Technology.

Shi, J. and Tomasi, C., 1994. Good features to track. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94), Seattle.

Titterton, D. and J.L.Weston, 2004. Strapdown Inertial Navigation Technology. Second edn, MIT Lincoln Laboratory.

Tsai, R., 1987. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. Robotics and Automation, IEEE Journal of 3(4), pp. 323–344.

Van Der Merwe, R., 2004. Sigma-point kalman filters for probabilistic inference in dynamic state-space models. PhD thesis.

Welch, G. and Bishop, G., 1995. An introduction to the kalman filter. Technical report, Chapel Hill, NC, USA.

Zeimpekis, V., Giaglis, G. M. and Lekakos, G., 2003. A taxonomy of indoor and outdoor positioning techniques for mobile location services. SIGecom Exch. 3(4), pp. 19–27.

Zhang, Z., 2000. A flexible new technique for camera calibration. Pattern Analysis and Machine Intelligence, IEEE Transactions on 22(11), pp. 1330–1334.