in: R. Ahlers (ed.), 5. Symposium Bildverarbeitung 1997, Technische Akademie Esslingen

An Image-Based Navigation Support System for Neuroendoscopic Surgery

W. Konen[1], M. Scholz[2], S. Tombrock[1], S. Tölg[1], M. Brauckmann[1], L. Adams[3]
[1] Zentrum für Neuroinformatik GmbH, Bochum, Germany
[2] Neurochirurgische Universitätsklinik, Ruhr-Universität Bochum, Germany
[3] Philips Medical Systems, Eindhoven, The Netherlands

Abstract:

We develop a navigation support system for endoscopic interventions which allows to extract 3D-information from the endoscopic video data and to superimpose 3D-information onto such live video sequences. The endoscope is coupled to a position measurement system and a video camera as components of a calibrated system. We show that the radial distortions of the wide-angle endoscopic lens system can be successfully corrected and that an overall accuracy of about 0.7mm is achieved. Tracking on endoscopic live video sequences allows to obtain accurate 3D-depth data from multiple camera views.

1 Introduction

Image processing and image analysis play an important role in advanced surgery (CAS: computer aided surgery). Much work is devoted to 3D-reconstruction from CT-, MRI- or other volume data slices or to 3D-registration between different image modalities and the actual scene in the operating theatre. 3D-registration is an important prerequisite for navigation support systems which allow for example to follow a predefined path during a surgical intervention.

Endoscopic interventions become also increasingly important in many surgical areas because they provide minimal-invasive impact on the patient. Especially in the area of neurosurgery and endonasal surgery they allow operation in otherwise difficult accessible areas (e.g. ventricular system). The endoscopic view adds another image modality to the surgeon's repertoire.

However, on the contrary to what one might expect, the obvious combination, namely digital image processing *within* endoscopic images, has not been assessed so far. The main roadblocks lie (a) in the difficulty to extract useful 3D-information from multiple images if the camera movement is not known, (b) in strong distortions caused by the wide-angle endoscope lens system and (c) in close-to-real-time requirements for any of the image processing tasks.

We propose in this paper a new system where the endoscope is coupled to an optical position measurement system (OPMS) which measures the endoscope position - and hence also the camera position - accurately in space and time. Using state-of-the-art camera calibration techniques (Sec. 3) and a newly developed system calibration (Sec. 4), we can map a 3D world point into the actual endoscope view. The inverse operation, namely inferring from multiple camera views of a point its 3D-coordinates, is also achieved by moving the endoscope while tracking the image motion of the (stationary) world point in the camera plane (Sec. 5).

Once these basic registration problems have been solved, they can be applied to multiple modules which give the surgeon enhanced navigation support:

- display certain landmarks from preoperative data (CT,MRI) - e.g. the location of a tumor - within the endoscopic image, or give appropriate navigational hints, if outside the endoscopic view,
- mark intraoperatively a certain landmark in the endoscope and allow to relocate it from other viewing directions,
- mark anatomical landmarks to refine *intraoperatively* the registration transform between the patient coordinate system and the CT coordinate system (e.g. after tissue movement).

An interesting and somewhat similar approach has been taken in the ARTMA system [Gunkel+95] where the endoscope position is measured by a Hall sensor and used to overlay certain preoperatively defined structures in the live endoscopic image. However, no camera calibration has been undertaken

in this approach so that the achievable accuracy is fairly limited and does not allow to infer the depth of an intraoperatively marked point from multiple camera views.

2 System Setup

The rigid endoscope (outer diameter 5.9 mm, Wolf GmbH) used in this work consists of a circular tube (6 mm diameter) where a color CCD-camera at the rear end captures the image from the tip of the endoscope through a special lens system (distance tip - rear end: 380 mm). We developed a special device mounted on the shaft of the endoscope which holds 3 infrared LEDs (see Fig. 1). In designing this device care has to be taken to make it sufficiently small such that it does not disturb surgeon's ability to maneuver the endoscope. On the other hand it has to be sufficiently large such that a good spatial resolution can be achieved.

The positions of the LEDs are measured continuously by the OPMS which is a part of the EasyGuide™ Neuro (Philips). The OPMS basically consists of a stereo camera rig, stationary in the operating theatre. The system measures the 3D-position of the LEDs and determines the 6 degrees of freedom of the rigid endoscope in the coordinate system of the camera rig. The OPMS achieves a spatial resolution of 0.4-0.8 mm and an overall accuracy of 1-1.5 mm within its volume of operation. The OPMS data are transfered via serial interface to the PC at a rate of about 8 Hz.

The video output of the endoscope camera is connected to a TV monitor and to a standard PC-framegrabber, which allows the live display of the color image on the PC's VGA-monitor, optionally with cursor and certain marks overlayed. Alternatively, we may grab a sequence of frames via PCI burst mode into the main memory (up to 16 Hz for 512x512 images).

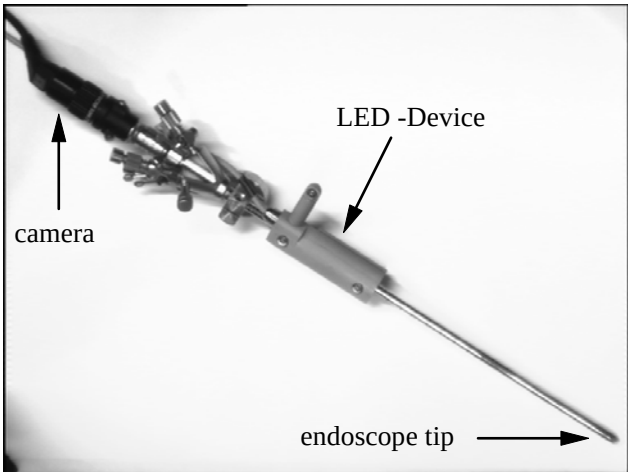

Fig. 1: Rigid endoscope with LED device

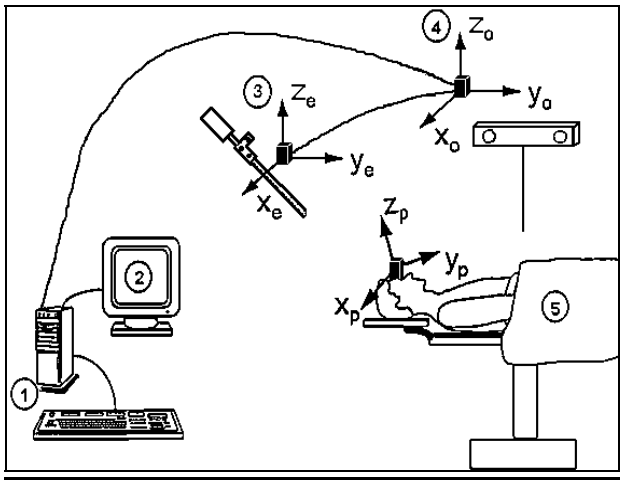

Fig. 2: Schematic drawing of the intraoperative system setup: The coordinates of the endoscope (3) are measured by the OPMS (4) and transmitted to the PC (1).

3 Camera calibration

The basic camera calibration procedure estimates the 6 extrinsic parameters {**R, t**} which map a point from the world coordinate system into the camera coordinate system, the 4 intrinsic parameters of a camera (focal length f, piercing point (u,v), scaling factors s_x) and finally image distortion parameters (κ_1, κ_2,...). When the camera is moved, only the extrinsic parameters change, while the intrinsic and the distortion parameters remain constant. The mapping of a world point $\mathbf{X}_w=(x_w,y_w,z_w,1)$ [1] into the

[1] Here and in the following we adopt the notation of projective geometry where points in 3D-space are represented by 4D-vectors (e.g. $\mathbf{X}_w$). In this 4D-space the rigid body transforms are *lineare mappings*, represented by 4x4 matrices (e.g. D).

camera coordinate system, $\mathbf{X}_c=(x_c,y_c,z_c,1)$, and then into the framebuffer pixel coordinates (x_f,y_f) is described by the following equations:

(3.1) $\mathbf{X}_c = D\,\mathbf{X}_w$ with 4x4-matrix $D = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{pmatrix}$

(3.2) $(X_u, Y_u) = \frac{f}{z_c}(x_c, y_c)$

(3.3) $(X_d,Y_d)\,(1 + \kappa_1 r^2 + \kappa_2 r^4 + ...) = (X_u,Y_u)$ with $r = \sqrt{X_d^2 + Y_d^2}$

(3.4) $x_f = s_x X_d + u_o$

(3.5) $y_f = s_y Y_d + v_o$

We use in this paper the well-known camera calibration procedure from [Tsai87] which provides a versatile and robust estimation of the camera parameters. The original algorithm from [Tsai87] allows only first order radial distortion correction (κ_1). Since we expected rather large radial distortions from the endoscopic lens system, we have extended the algorithm in a straightforward manner to include also the κ_2-term to compensate higher order radial distortion effects.

The calibration procedure requires one (coplanar) or multiple (noncoplanar) image(s) of a calibration pattern with known geometry. As calibration pattern we use a plane of rings placed on a regular lattice with 1,25 mm ring-to-ring distance (Fig. 3).

We distinguish between the following calibration procedures:

- extrinsic calibration: the intrinsic and the distortion parameters are assumed to be known and we estimate from one image the 6 extrinsic parameters only,
- coplanar calibration: we estimate from one image all calibration parameters, except for s_x,

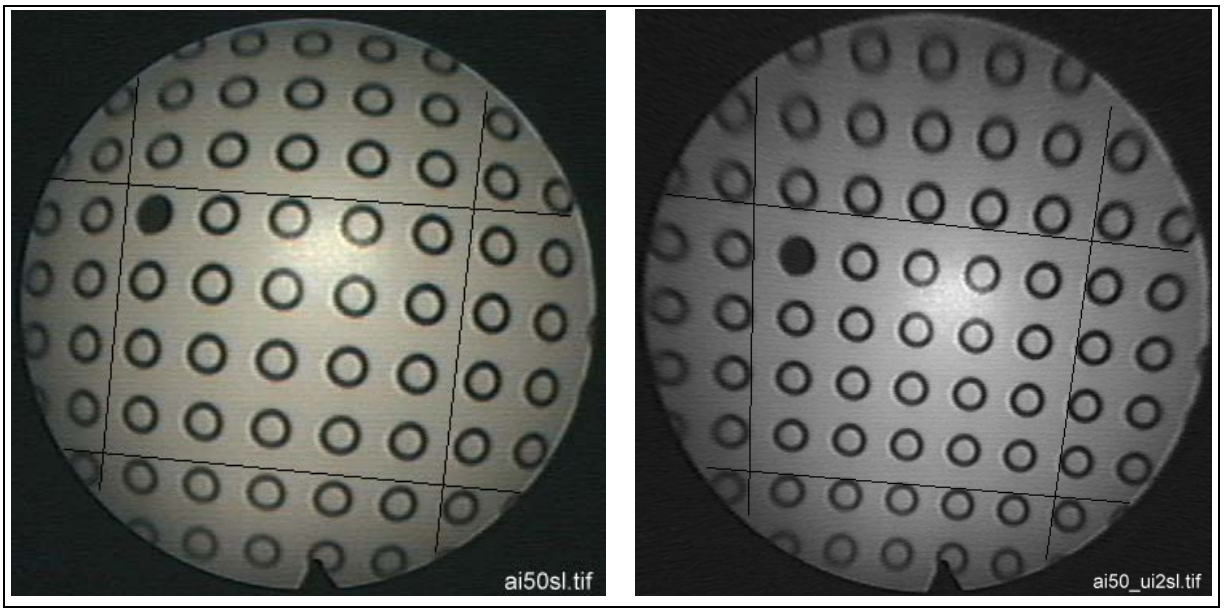

Fig. 3: Left: The calibration pattern as viewed through the endoscope (distorted image). The straight lines are not part of the image but superimposed to show the distortion effects. Right: The calibration pattern after correcting the distortion effects using the κ_1-distortion model. All rings are lying on straight lines, as they should. The (κ_1,κ_2)-distortion model achieves the same result.

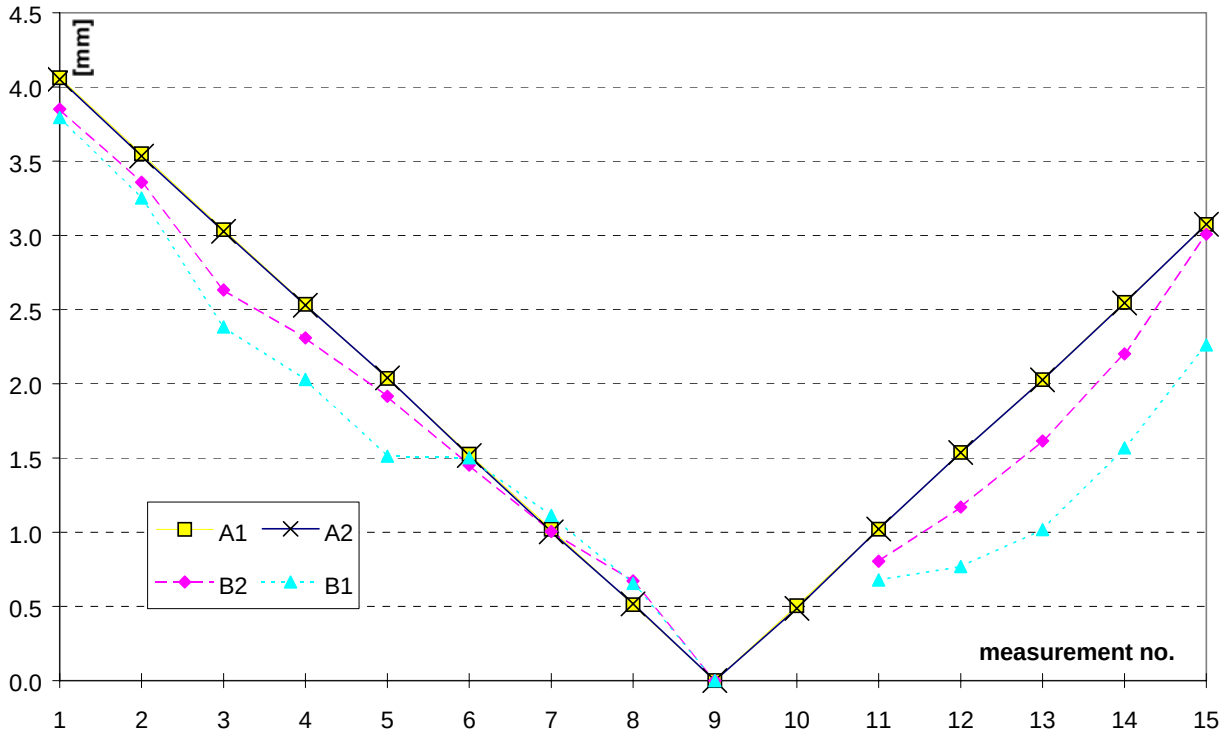

Fig. 4: Accuracy of the camera calibration: Shown is the distance $|| \mathbf{X}_w(n) - \mathbf{X}_w(9) ||$ for the different calibration strategies (see text). The deviation of each point from its corresponding horizontal line at $|n-9| \cdot 0.5$mm is the residual error. (Measurement 10 is missing in case B, because the nonlinear optimization failed to establish the correct intrinsic parameters.)

- noncoplanar calibration: we estimate from multiple images all calibration parameters, including s_x. The calibration pattern is attached to a micrometer screw which allows to move it with an accuracy of $< 10\mu$m.

For the calibration we took a sequence of 15 images where the calibration pattern was shifted by 0.5 mm between consecutive positions. Based on this sequence, we examined two different calibration strategies:

A. Perform a noncoplanar calibration based on the five images 3,5,7,9,13 to determine intrinsic and distortion parameters. Use this as input to perform an extrinsic calibration for any of the images.

B. Perform a coplanar calibration for each image in the sequence, that is recompute the intrinsic and distortion parameters for each image (except for s_x which is taken from the noncoplanar calibration).

In each case we used both the κ_1-distortion model (A1 and B1, resp.) and the (κ_1,κ_2)-distortion model (A2 and B2, resp.). We assessed the numerical quality of the solution by comparing the magnitude $|| \mathbf{X}_w(n) - \mathbf{X}_w(9) ||$ with the expected values, where $\mathbf{X}_w(n)$ denotes the origin of the calibration pattern in the nth measument. We see from Fig. 4 that the calibration strategy A is in very good agreement with the expected values (accuracy < 0,15mm), while the calibration strategy B has much larger errors. This is well understandable since the numerically difficult part of the calibration, namely the nonlinear optimization for the intrinsic parameters, is undertaken under strategy B anew for each image, based on relatively few data, while under strategy A we follow the reasonable assumption that the intrinsic parameters do not change between measurements and estimate them once, based on a five times larger amount of calibration data.

The resolution of the calibration, namely the accuracy of two *consecutive* measurements is better than the overall accuracy by approximately an order of magnitude: $\big|\, ||\mathbf{X}_w(n) - \mathbf{X}_w(n+1)|| - 0.5\text{mm} \,\big| < 20\ \mu$m.

With the estimated distortion parameters from strategy A we can invert Eq. (3.3) numerically and warp the distorted image using bilinear interpolation to obtain an undistorted image (Fig. 3). We found that

both distortion models are able to correct the strong distortion of the endoscope very well and that the improvement by taking the κ_2-term into account is negligible. Since both corrections work successfully, we conclude that the simple κ_1-radial distortion model is sufficient for the endoscopic lens system, as well as the underlying assumption of the central projection model.

4 System calibration

Our goal is to be able to transform 3D world points into the camera coordinate system and to project them into the endoscopic view. The 3D world points are usually expressed in the OPMS coordinate system (attached to the camera rig of the OPMS). The 3D points may come for example from preoperative CT-data which have been transformed into the OPMS coordinate system.

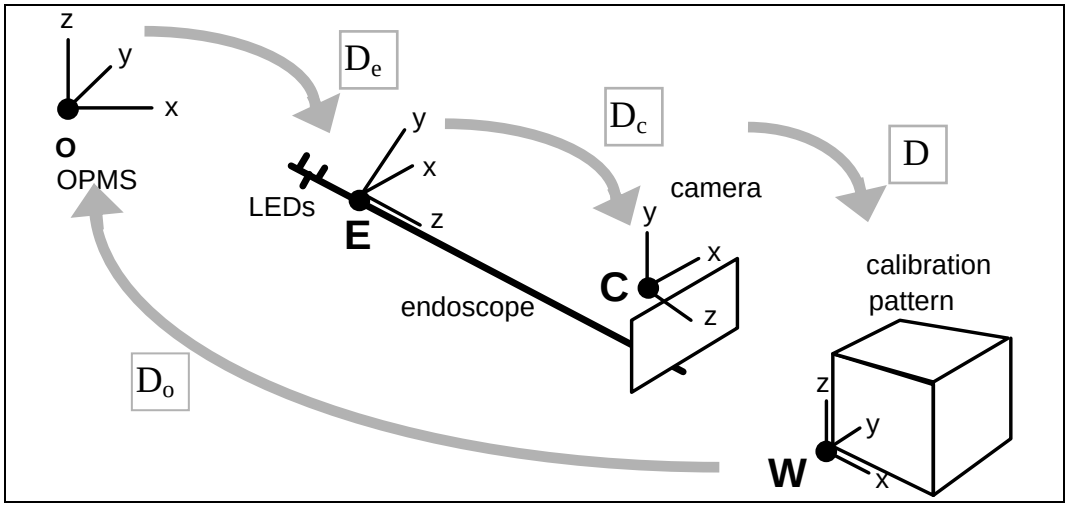

Fig. 5: Coordinate systems for the system calibration

In terms of the coordinate systems depicted in Fig. 5 we are seeking the combined transformation D_eD_c. What we have achieved so far with the camera calibration is the transformation D from the calibration pattern to the camera system (Fig. 5). The transformation D_e between the OPMS and the LEDs of the endoscope is provided by the OPMS-measurements. We use the fact that a point in the calibration pattern coordinate system can be transformed into the camera system either by D or by the combined transformation $D_cD_eD_o$. We note also that the transformation D_c remains constant as long as the LED device is not detached from the endoscope. This gives us the following strategy for system calibration:

1. Put the calibration pattern at a fixed point in space. That is, the transformation D_o between calibration pattern and the camera rig of the OPMS remains constant during the measurment.
2. Take images i=1,...,n of the calibration pattern with the endoscope from different endoscopic viewpoints. Obtain the transformation matrices $D^{(i)}$ (extrinsic camera calibration) and $D_e^{(i)}$ (OPMS) for each viewpoint.
3. Solve the system of equations

$$\begin{aligned} D^{(1)} &= D_c D_e^{(1)} D_o \\ &\vdots \qquad\quad \vdots \\ D^{(n)} &= D_c D_e^{(n)} D_o \end{aligned} \tag{4.1}$$

   for the unknown transformations D_o and D_c.

Practical experience shows that n=4-5 images of the stationary calibration pattern from different viewpoints are sufficient to obtain a good system calibration. Thus we have a fast calibration procedure which can be conducted easily by technical personal. Note that the system calibration needs to be done only after the LED device has been detached from the endoscope, which is not necessary in

normal operation. Practical experience shows also that the noncoplanar camera calibration has to be performed only once for a given camera; the intrinsic and the distortion parameters remain fairly constant, even if the camera has been detached from the endoscope.

Having obtained the matrix D_c, we can map a point from the OPMS coordinate system into the endoscopic camera view for any arbitrary position of the endoscope. We can assess the quality of the system calibration by the following *"spin-me-around test"*: We manually mark a certain landmark **P** in at least k=2 different camera views. Knowing the k camera positions from the OPMS, we can obtain the 3D-representation $\mathbf{P}_o$ in the OPMS-system using standard triangulation techniques [Longuet-Higgins81]. For any subsequent endoscopic view, we map $\mathbf{P}_o$ in the actual camera view and overlay it onto the endoscope image. If camera and system calibration are correct, the overlayed mark will be always on top of the landmark **P**, no matter how we "spin around" the endoscopic viewpoint or where the landmark **P** appears in the endoscope image. Table 1 shows that the residual error $\langle\Delta f\rangle \approx 0.7$ mm is in the order of the intrinsic error of the OPMS (> 0.5mm), i.e. the accuracy of the calibration is close to its theoretical limit.

Table 1: Average error $\langle\Delta f\rangle$ as residual distance between **P** and the overlay mark in the live endoscope image. (Averaging has been done with respect to 10 images where the landmark **P** appears at various distances from the piercing point.) Δs is the approximate movement of the camera between the k multiple views to determine $\mathbf{P}_o$. Δf_{OPMS} is the observable jitter of the overlay mark when the endoscope is fixed. This jitter stems from noise in repeated OPMS measurements and gives a lower bound on the intrinsic OPMS error.

meas.	k	Δs [mm]	$\langle\Delta f\rangle$ [mm]	Δf_{OPMS} [mm]
A2	2	5	0,74	0,5
A3	3	5	0,62	0,5
B2	2	10	0,6	0,5
B3	3	10	0,54	0,5
C2	2	10	0,82	0,5
C3	3	10	0,97	0,5
average:			**0,72**	

Note that it is very important to have an accurate distortion model in order to achieve this accuracy. If we neglect distortion effects (by setting $\kappa_1=0$) and if **P** is far away from the piercing point, the error goes up to 3.5 mm instead of 0.7 mm.

5 Tracking

Our goal is to determine the 3D-position of a visual landmark from multiple 2D-views (triangulation). We can eliminate the necessity to perform *multiple* clicks on a certain landmark **P** if the systems tracks the landmark automatically. Then the surgeon simply points initially at the landmark, the system tracks it while the endoscope is moving until some stopping criterion is met and finally reports the 3D-coordinates $\mathbf{P}_o$ and displays the overlay mark. We implemented a fast tracking algorithm which uses template matching in a two-dimensional logarithmic search fashion [Jain89]. The tracking can analyze up to 8 frames per second and works well in realistic sequences of endoscopic surgery.

For each tracked frame we calculate the line of sight, that is the line between the actual center of projection and the point **P** in the camera plane. As a stopping criterion we use the angle between the initial line of sight and the actual line of sight which has to exceed some predefined threshold value α_{max}. This stopping criterion automatically avoids degenerate situations, e.g. if **P** is near the piercing point and the endoscope is only moved along the optical axis. Should this happen, the algorithm just waits until the endoscope is moved in another direction and a reliable measurement can be obtained. By experiments we found that $\alpha_{max}=5^o$-7.5^o gives best results (Table 2). This corresponds usually to endoscope movements of just 2-3 mm.

Table 2: Average error <Δf> when performing automatic tracking.

α_{max}	Δs [mm]	$<\Delta f>$ [mm]	Δf_{OPMS} [mm]
5.0^{o}	5	0,78	0,5
5.0^{o}	10	0,84	0,5
5.0^{o}	15	0,91	0,5
7.5^{o}	5	0,81	0,5
7.5^{o}	10	0,82	0,5
7.5^{o}	15	0,86	0,5
average:		**0,84**	

6 Conclusion and outlook

We have shown that advanced image processing techniques can be applied successfully to endoscopic surgery. A good calibration for the wide-angle and strongly distorting endoscopic lens system has been obtained (accuracy 0.2 mm). We have coupled the endoscope to an OPMS which determines its 3D-location in space and which allows to overlay 3D-points or arbitrary 3D-structures directly in the live camera image with high accuracy (0.7-0.8 mm). Inversly, landmarks in the camera image can be tracked when the endoscope is moving, their 3D-position can be obtained and reported to other systems. This is to our knowledge the first time that direct image processing on the image modality of endoscopic video sequences is performed.

First tests with the overall system in preclinical studies (Fig. 6) have shown convincing results and achieved a similar accuracy as in the laboratory case. Further work will go into the evaluation and optimization of the system for direct use in the operating theatre, first in extended preclinical studies and later also during clinical interventions. It is our believe that advanced image processing methods will play an increasingly important role in endoscopic surgery where the surgeon faces new challenges due to continuing miniaturization. Thus he needs new tools for improved navigation support.

This work has been supported in part by grants from the Ruhr-Universität Bochum for interdisciplinary research (FORUM).

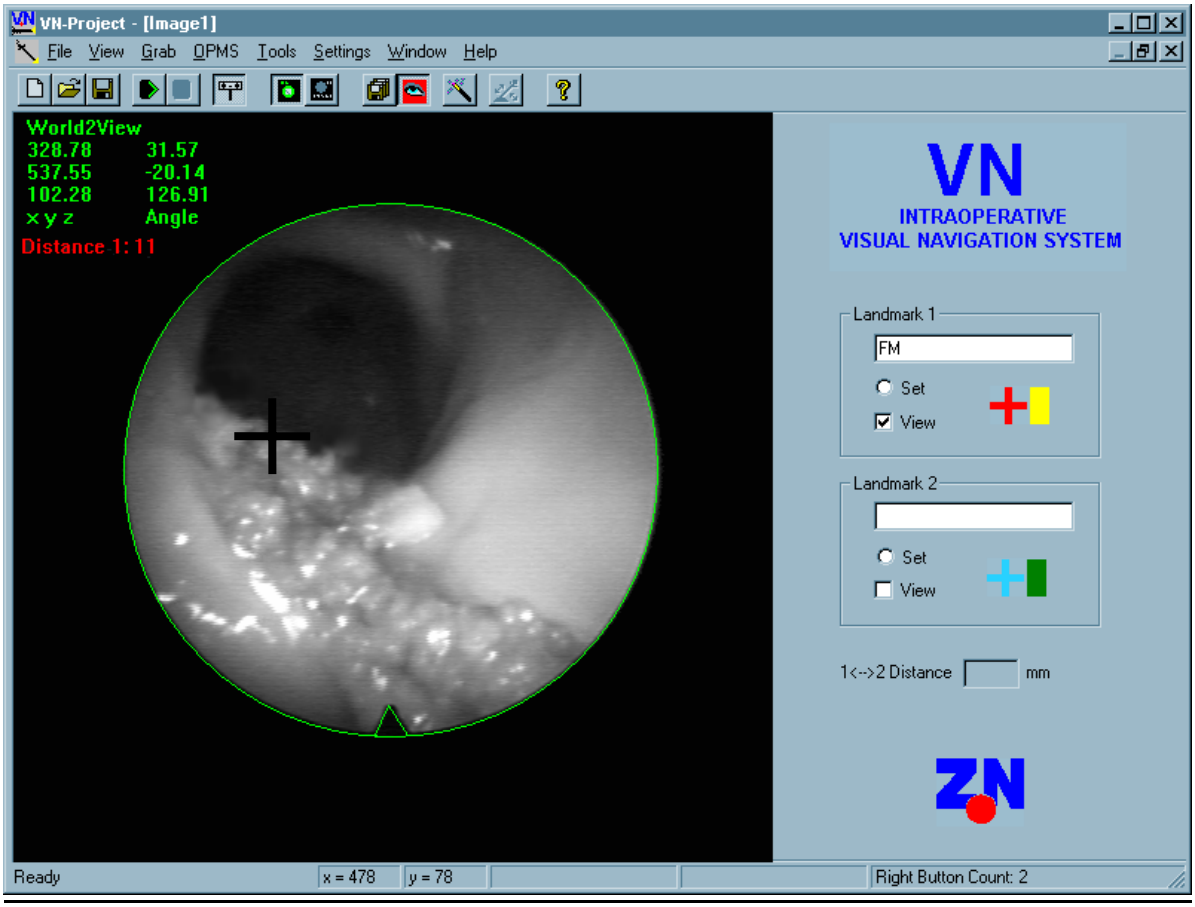

Fig. 6: View through the endoscope into the ventricular system (Foramen Monroi). A point on the border of the Foramem Monroe is marked (the cross), its 3D-coordinates have been established and can be used for further navigation.

7 Literature

[Tsai87] R.Y.Tsai, *A versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses*, IEEE Journal of Robotics and Automation, Vol. RA-3, No.4 , 323-344, 1987.

[Gunkel+95] A.R. Gunkel, W. Freysinger, W.F. Thumfart, M.J. Truppe, *Application of the ARTMA Image-Guided Navigation System to Endonasal Sinus Surgery*, Proc. Computer Assisted Radiology, 1146-1151, Berlin, 1995.

[Jain89] A.K. Jain, *Fundamentals of digital image processing*, 404-406, Prentice-Hall, 1989.

[Longuet-Higgins81] H.C. Longuet-Higgins, *A computer algorithm for reconstructing a scene from two projections*, Nature, **293**, 133-135, 1981.