

Towards Accurate X-Ray-Camera Calibration in Computer-Assisted Robotic Surgery

Ch. Brack^a, H. Götte^b, F. Gossé^c, J. Moctezuma^b, M. Roth^a, A. Schweikard^a

^aBay. Forschungszentrum für wissensbasierte Systeme, Forschungsgruppe Kognitive Systeme (FG KS), Orleansstr. 34, 81667 Munich, Germany,
e-mail: {brackc,rothm,schweika}@informatik.tu-muenchen.de

^bTU München, Institut für Werkzeugmaschinen und Betriebswissenschaften (iwb), Karl-Hammerschmidt-Str. 39, 85609 Aschheim, Germany,
e-mail: {gx,ma}@iwb.mw.tu-muenchen.de

^cOrthopädische Klinik, Medizinische Hochschule Hannover, Konstanty-Gutschow-Str. 8, 30625 Hannover, Germany

We present a new approach of X-ray camera calibration as a part of a vision-guided robotics system for orthopaedic knee surgery. Two invented calibration bodies provide highly accurate and robust calibration results, while avoiding constraints concerning the placement of landmarks relative to the X-ray system. A described experiment shows the absolute error of our method. In order to provide an automatic system new techniques for on-line image processing and recognition are developed.

1 INTRODUCTION

During the replacement of the knee joint by an artificial total knee prosthesis bone material of the tibia and the femur has to be removed. Conventionally the surgeon moves a handheld powered saw, guided by a set of templates, inserted manually during surgery. Therefore the attainable accuracy depends strongly on the surgeon's skills and dexterity.

Our approach suggests the use of a surgical robot as a guiding device for the intraoperatively positioning of a sawing tool to preoperatively planned cutting planes. The surgeon still moves the saw by hand, but the guiding device restricts his movement to the planned cutting planes [11].

To reach the goal of such a precise prepositioning several individual components are necessary: besides the surgical robot with specially designed tools and instruments our system comprises a common X-ray image intensifier and a workstation, used for process controlling as well as for preoperative planning [8]. Furthermore a real-time vision system tracks the position of the patient's leg as well as the movement of the manipulator's

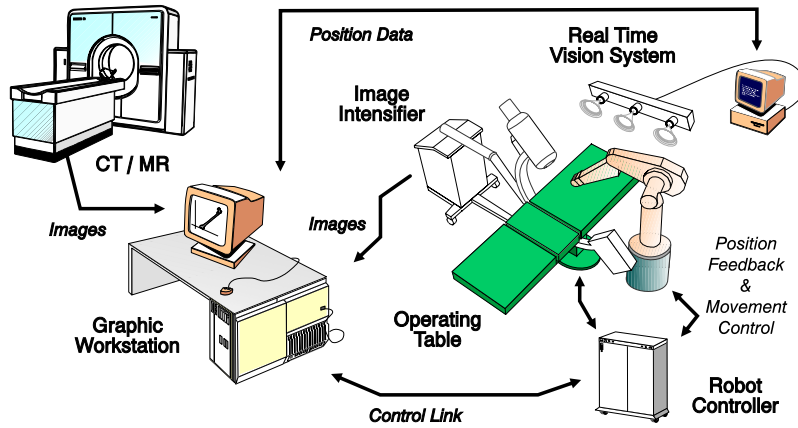


Figure 1. The system components

hand, based on active infrared markers and cameras. Figure 1 illustrates the different components of our system and how they work together. Because this approach combines a robot's high accuracy with a surgeon's expertise it is expected to provide high overall accuracy and lower complication rates and to be safe.

In order to achieve the desired accuracy the following events have to take place in chronological order before and during surgery. First of all 3D surface-models of tibia and femur are derived from preoperatively taken CT-scans. Based on these models the surgeon is able to plan all cutting planes for an optimal fit of the prosthesis interactively in advance. Actual intraoperative surgery starts with mounting two active infrared LED markers on tibia and femur. Afterwards several X-ray images from at least two different viewing angles are obtained, using a common image intensifier. The X-rays are taken to calibrate the imaging system in the below described way. Now the tibia as well as the femur can be registered by correlation of the 3D surface-models from CT with the bone contour information extracted from the X-ray images [1] [6]. From that point on there is no longer a need for the X-ray system during surgery. The positions of the bones are tracked in real-time just using the active vision system and its infrared markers.

This paper focuses on the calibration of the X-ray image intensifier system. A more detailed description of our whole system is available from [11].

2 CALIBRATION OF THE X-RAY SYSTEM

2.1 Goal and Motivation

The main requirement for an accurate bone registration is a global calibration technique, in the sense of all image features. The contours of femur and tibia cover a large area within the X-ray image. Therefore the projection behaviour of the image intensifier must be adequately modeled for the entire visible 3D scene. The fixed 3D world-coordinate-system is represented by a rigid calibration body. It is needed together with other calibration bodies to determine the mathematical camera model of the image intensifier. With regard to that model, the X-ray intensifier is quite similar to CCD-cameras. Both models are based on 3D location and a perspective projection with a 2 dimensional distortion (see fig. 2). In contrast to many optical CCD-systems, the distortion-function of the intensifier is previously unknown. Therefore a camera model with over 20 degrees of freedom (DOF) has been chosen. The disadvantage related to the high number of DOF, is the high correlation between the interior (distortion) and the exterior (location) camera parameters.

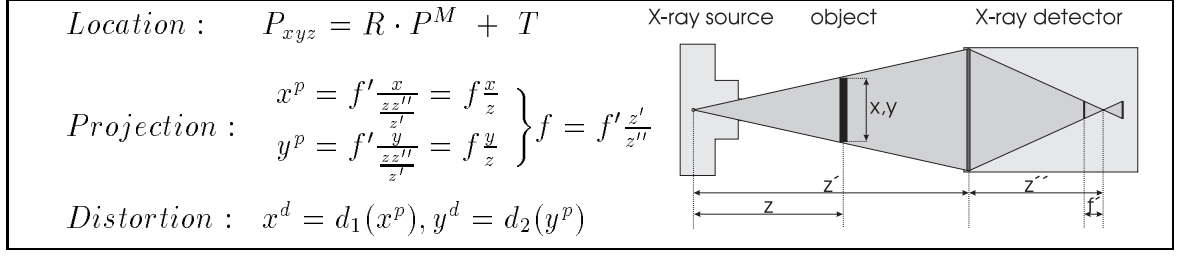


Figure 2. The way of the beam in an X-ray system is related to such in CCD-cameras. A model point P^M is rotated by R and translated by T to camera coordinates P_{xyz} . After the perspective projection, the 2D-points are distorted by $d_1()$ and $d_2()$.

With CCD-cameras, there are several well-known ways to solve this problem. More than one image from the calibration body can be used concurrently to calibrate the system. Every image contains additional information about the constant interior camera parameters, by introducing different views [5] [9]. In the case of an X-ray system mechanical and other instabilities, i.e. mainly the detector's sensitivity towards electromagnetic fields, lead to unstable interior camera parameters over time and change in position. Therefore the above described method fails to reduce the internal correlation.

Another way to achieve a reduction of correlation of the camera parameters is the use of a 3D calibration body for multiplane calibration. A constraint of our application is that the calibration markers have to be recognizable from many different views. This creates a considerable risk of overlapping markers in the X-ray image. Summarized, none of the existing methods for CCD-cameras can be fully applied to X-ray calibration. Therefore a new calibration technique was invented using two calibration bodies, both appearing in the image.

2.2 Environment

The used calibration bodies are based on accurate measured 3D-point landmarks. In order to distinguish the bodies, different shapes of markers are chosen. The first body contains spherical metal landmarks, mounted on a rigid plate. Some of them are just slightly hightened to avoid the above mentioned overlapping problem while still providing 3-dimensionality. This calibration body is fixed rigidly on the operation table and determines the reference coordinate system (see fig. 3). The invented calibration technique uses a second body with a pyramidal design. It is composed of four different levels of height, where metal crosses, serving as landmarks, are placed in a concentric manner. Some of the crosses have been replaced by rectangles to be able to distinguish different orientations of the symmetrically shaped pyramid. This calibration object is mounted rigidly in front of the image intensifier, to ensure the visibility of the markers. The pyramid's base covers the entire face of the detector. Thus the X-ray shots are totally and densely superimposed with the cross shaped markers. Distortion effects are determinable for the entire visible scene, providing a global calibration.

2.3 X-ray Camera Calibration Techniques

The presented method depends, like well-known techniques [12] on error minimization between the estimated and the real 2D projection of the calibration body. Therefore geometric models are generated, containing the measured 3D coordinates of the calibration features. During calibration image processing routines are used to extract the projections

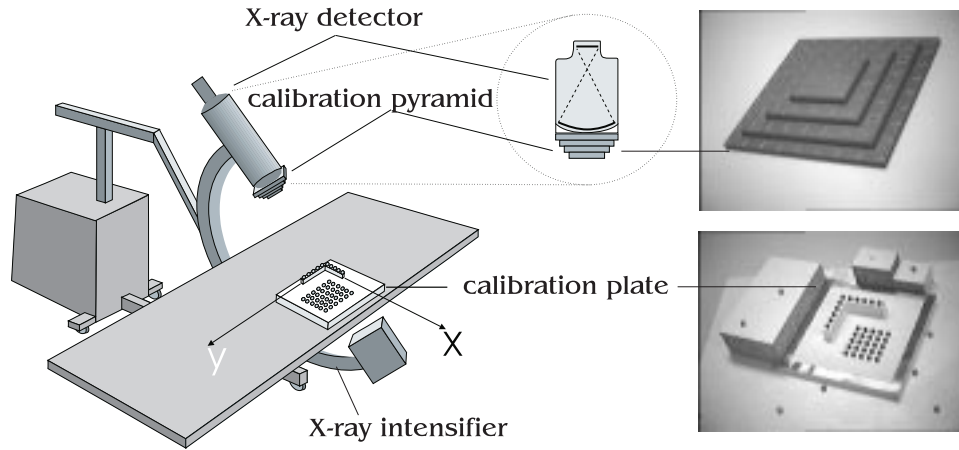


Figure 3. The image intensifier environment

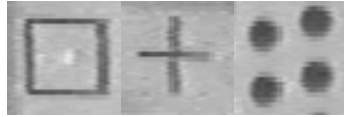


Figure 4. Three types of landmarks.

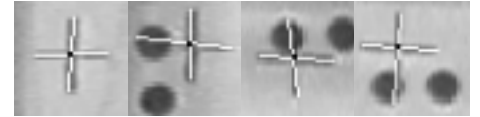


Figure 5. Intersection of axes.

of the markers from the X-ray image. In order to assign the 2D-features to the 3D-model, special algorithms are used. They cope with wrong detected as well as with missing image features. Afterwards an initial camera model with less DOF can be achieved directly. While iteratively projecting the geometric model, the camera model with additional DOF is adapted in respect to the distance between projected and X-ray image features.

2.3.1 Extraction of Landmarks Our system has to recognize three different types of landmarks: rectangles and crosses, due to the pyramidal body as well as spherical markers, additionally brought up by the calibration plate (see fig. 4).

First of all some common preprocessing steps – mainly smoothing and dynamic thresholding – are applied to each single X-ray image. The thresholding yields a set of connected regions of pixels. Now the robust classification of image-regions is based on shape-criteria, using the Computer Vision software *HORUS* [2].

The most conspicuous shape-feature of the four rectangular markers is their significant big cavity. Internally *HORUS* represents pixel-regions using the compression scheme of run-length-encoding (RLE): only the start- and end-point for each segment of connected pixels are stored within each row of the region. Now regions corresponding to a rectangular marker, lead to very characteristic values within their RLEs, due to the already mentioned cavities: the average number of stored segments per line is expected next to a value a little bit smaller than 2 while the average length of a segment stays small, i.e. only a view pixels.

The second type of landmark – the cross-shaped markers – has no such cavities but very characteristic extremities, instead. Therefore each region's horizontal and vertical extrema, i.e. minimum and maximum outline points, are determined. Binding together the horizontal as well as the vertical ones defines two unique axes. Based on an examination of these axes our system is able to decide, whether a region corresponds to a cross-shaped landmark or not. The axes should be of nearly the same length, they

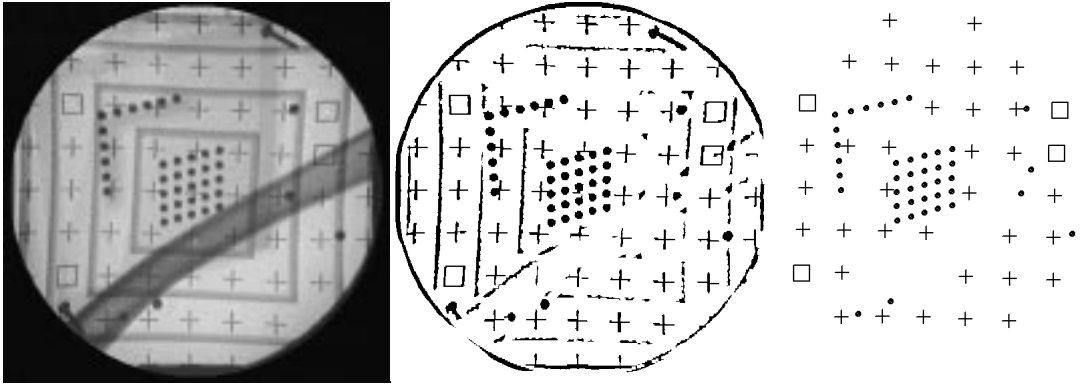


Figure 6. Original X-ray image (left), result of dynamic thresholding (middle) and set of detected features (right)

are expected to be approximately perpendicular with respect to each other and last, not least: the point of intersection is supposed to lie within the middle of both axes. Moreover, if all tests succeed, these points of intersection provide very accurate estimations for cross-centres, even in the adverse case of *merged* regions¹ (see fig.5).

Compared with the pyramidal body, the feature-assignment for the calibration plate is much more complicated. The reason is, that there is only weak a-priori knowledge about the plate's position and orientation with respect to the image-intensifier. To ensure a correct feature-assignment, it is therefore essential to detect a sufficient number of features. Using our pyramidal calibration body overlapping effects between crosses and sphericals cannot be avoided, thus significantly decreasing the number of well-detectable markers. In order to maximize this number the recognition process is split into two phases: in the first instance, a search for non-overlapped spheres is performed. This can be done very easily, based on common geometrical shape-criteria, e.g. circularity, compactness, convexity and area. Afterwards a-priori knowledge about the pyramidal calibration body as well as first results of calibration (see section 2.3.2) are used to mask out *all* crosses and rectangular markers, even those, not detected during segmentation. Now, previously overlapped spherical regions are reconstructed based on the remaining artefacts by means of morphological techniques – mainly opening/closing operations.

2.3.2 Calibration Algorithm For the iterative calibration algorithm an initialization of the unknown parameters is needed nearby the correct solution. With CCD-cameras such parameters, like the focallength are previously known by the manufacturers. The technical specifications of the intensifier are most unknown. Thus, the chosen initialization of the camera model is based on linear interior parameters:

$$\begin{pmatrix} x^d \\ y^d \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} x^p \\ y^p \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix}. \quad (1)$$

In order to calculate the unknown parameters $b_{1...2,1...2}$ and (u, v) the DLT-method (*Direct Linear Transformation*) [3] [7] is used. This technique solves the equation $P_{xy1}^d = A \cdot P_{xyz1}$ with respect to the 3×4 DLT-matrix A . An estimate of A can be computed from the linear system

$$C (a_{11} \dots a_{14} \ a_{21} \dots a_{24} \ a_{31} \dots a_{34})^T = (0 \ 0 \ \dots \ 0)^T, \quad (2)$$

¹regions, corresponding to some overlapping landmarks, impossible to separate during segmentation/thresholding.

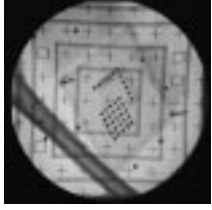
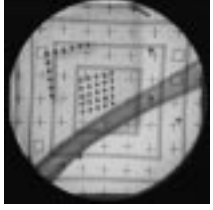
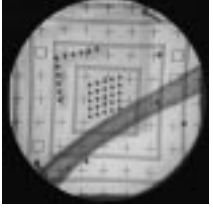
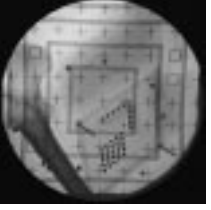
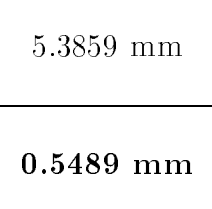
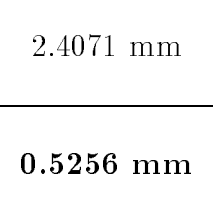
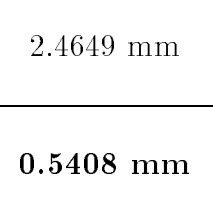
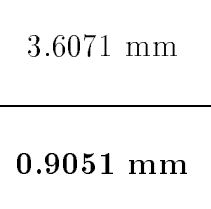
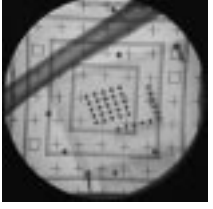
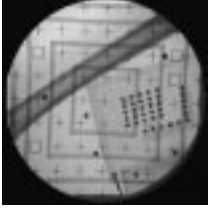
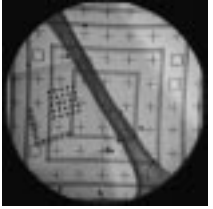
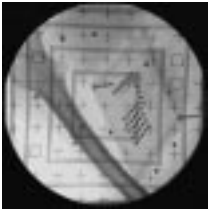
average error plate only				
average error with pyramid				
	5.3859 mm	2.4071 mm	2.4649 mm	3.6071 mm
	0.5489 mm	0.5256 mm	0.5408 mm	0.9051 mm
	4.0913 mm	2.1635 mm	2.3077 mm	3.5658 mm
	0.7912 mm	0.6452 mm	0.4542 mm	0.7507 mm
	2.4977 mm	1.9759 mm	3.1841 mm	3.3240 mm
	0.7169 mm	0.8001 mm	0.6964 mm	0.8756 mm
	6.0666 mm	2.4420 mm	2.6366 mm	4.5821 mm
	0.8349 mm	0.6294 mm	0.5369 mm	0.9763 mm

Figure 7. The position in space of additional landmarks are determined by triangulation of three corresponding markers in a pair of X-rays. The average errors with and without using the pyramidal calibration body are listed in the crossings of the columns and rows of the chosen image pairs.

where the $2n \times 12$ matrix C contains n pairs of 3D model and 2D image feature points. The pseudo inverse of C set into the least squares equation leads to the solution. Afterwards translational and rotational factors are separated from the DLT-matrix. The remaining component represents projection and distortion (see eq. 1).

The achieved accuracy of the described method is improved by additional interior parameters in a cubic camera model [13]:

$$\begin{pmatrix} x^d \\ y^d \end{pmatrix} = \begin{pmatrix} a_0 & a_1 & \cdots & a_9 \\ b_0 & b_1 & \cdots & b_9 \end{pmatrix} \begin{pmatrix} 1 & x_p & y_p & x_p y_p & x_p^2 & y_p^2 & y_p x_p^2 & x_p y_p^2 & x_p^3 & y_p^3 \end{pmatrix}^T \quad (3)$$

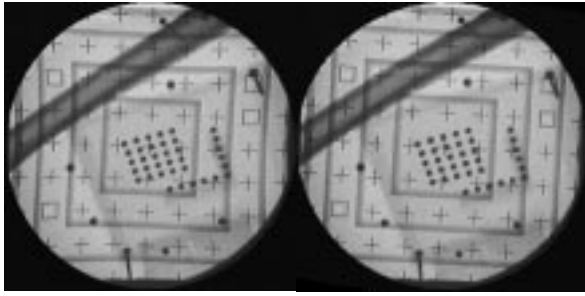


Figure 8. During calibration an undistorted X-ray image (right) is generated from the original (left).

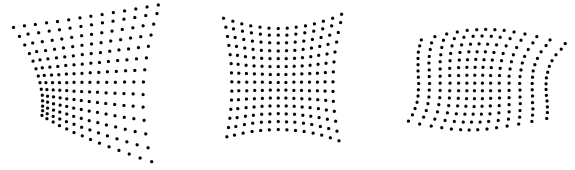


Figure 9. Examples of distortion modeled by the cubic camera model

Many types of distortion can be modeled in this way (see fig. 9). All variables (exterior and interior) are estimated by the iterative Levenberg-Marquart [10] algorithm (implemented by using the software package *minpack* [4]). It minimizes the error between the projected 3D-model features and the 2D-landmarks in the X-ray by updating the camera parameters. In order to reduce the correlation between exterior and interior camera parameters the two calibration bodies are calibrated concurrently in one equation. Additionally, the variables are estimated successively. After adapting the exterior parameters, the linear ones are determined before the cubic interior variables.

2.4 Results

In order to test the accuracy of the new calibration method, the position of additional landmarks beside the calibration bodies are determined with the X-ray system. Therefore the calibration plate is attached on a robust wooden board with spherical landmarks fixed on different levels of height around the plate. Their exact position in space relative to the calibration plate has been measured beforehand with a coordinate measuring machine. A series of X-ray shots are taken from the board by varying the view direction. The images are calibrated in the above described manner. Pairs of X-rays are randomly chosen to triangulate the landmarks, visible on both shots. The determined positions in space are compared to the measured ones. Figure 7 shows the experimental results with and without the pyramidal calibration body. The achieved accuracy of our method is better than 1 mm.

A by-product of the calibration algorithm, is the generation of an undistorted X-ray image (see fig. 8).

3 CONCLUSION

We have presented a new technique for the accurate calibration of an X-ray camera system. In the future we yet hope to improve the achieved accuracy by grabbing the X-rays directly from the intensifier with a higher resolution. The images of the above described experiment are taken over the detour of a conventional VHS video cassette.

REFERENCES

1. J.L. Chen and other. Recovering and tracking pose of curved 3D objects from 2-D images. *Proceedings, CVPR '93, IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1993.

2. W. Eckstein and C. Steger. Interactive data inspection and program development for computer vision. volume 2656 of *SPIE*, Jan. 1996.
3. O.D. Faugeras and G. Toscani. Camera Calibration for 3D computer vision. *Proc. Int. Workshop on Industrial Application of Machine Vision and Machine Intelligence*, 1987.
4. Ken Hillstrom Jorge More, Burt Garbow. *MINPACK*. Argonne National Laboratory.
5. S. Lanser, Ch. Zierl, and R. Beuthhauser. Multibildkalibrierung einer CCD-Kamera. *DAGM Symposium, Bielefeld*, 17, 1995.
6. S. Lavalée and other. Recovering the position and orientation of free form objects from image contours using 3D distance maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 17(4):378–390, 1995.
7. T. Melen. *Geometric Modelling and Calibration of Video Cameras for Underwater Navigation*. PhD thesis, Institutt for teknisk kybernetikk Norges tekniske høyskole, 1994.
8. J.L. Moctezuma and other. A computer and robotic aided surgery system for accomplishing osteotomies. In *First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, volume 1, pages 31–35, September 1994.
9. G. Toscani O.D. Faugeras. The calibration problem for stereo. *IEEE CVPR, Miami Beach, FL, USA*, 1986.
10. Press and other. *Numerical Recipes in C*. Cambridge University Press, 1994.
11. M. Roth, Ch. Brack, and other. A new less invasive approach to knee surgery using a vision-guided manipulator. In *Int. Symposium on Robotics and Manufacturing (ISRAM'96)*, World Automation Congress (WAC'96), Montpellier, France, May, 27th–30th, 1996.
12. 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*, RA-3:323–344, 1987.
13. P. Wunsch and K. Arbter. Kalibrierung eines nichtlinearen Hand-Auge Stereokamerasystems. *to be published*, 1994. DLR - Institut für Robotik und Systemdynamik.