# Camera-Marker Alignment Framework and Comparison with Hand-Eye Calibration for Augmented Reality Applications

Gérald Bianchi, Christian Wengert Matthias Harders, Philippe Cattin and Gábor Székely
Computer Vision Laboratory Swiss Federal Institute of Technology Zürich
Sternwartstrasse 7, ETH-Zentrum, 8092 Zürich, Switzerland
{bianchi, wengert, mharders, cattin, szekely}@vision.ee.ethz.ch

## Abstract

*An integral part of every augmented reality system is the calibration between camera and camera-mounted tracking markers. Accuracy and robustness of the AR overlay process is greatly influenced by the quality of this step. In order to meet the very high precision requirements of medical skill training applications, we have set up a calibration environment based on direct sensing of LED markers. A simulation framework has been developed to predict and study the achievable accuracy of the backprojection needed for the scene augmentation process. We demonstrate, that the simulation is in good agreement with experimental results. Even if a slight improvement of the precision has been observed compared to well-known hand-eye calibration methods, the subpixel accuracy required by our application cannot be achieved even when using commercial tracking systems providing marker positions within very low error limits.*

## 1 Introduction

The driving force of the presented study is the development of a medical training system using Augmented Reality (AR) techniques. The expectations on accuracy and stability in such a setting are high. Misalignment of overlaid virtual objects would greatly compromise manipulative skill training and the sense of presence, and thus reduce the training effect. Our AR system consists of a camera-mounted marker to determine the user's head pose and a tracking device reporting the markers' location. The quality of the alignment between the virtual and the real world critically depends on the accuracy of both the sensing device and the transformation between the camera and the marker. In this work, we search for the most efficient calibration method in order to accurately estimate the camera-marker transformation and investigate the limits of the achievable precision of the backprojection in the presence of errors in the processing chain based on simulated and real-life experiments.

## 2 Calibration Methods

We have set up a calibration environment which was inspired by [2], using a calibration object incorporating both visible and infrared(IR) LEDs, thus allowing to observe the same object both by the tracking system and the head-mounted camera. In order to further reduce the errors, we directly detected the IR LEDs by our camera, thus eliminating the transformation between the visible and IR LEDs from the calibration process. 30 IR LEDs have been mounted on the 3D calibration object. By the simultaneous detection of the LEDs by the camera and the IR optical 6DOF tracker [4] the camera-marker transformation can be directly estimated without determining the world-object transformation. The images of the IR LEDs are captured with a firewire Videre Design MEGA-DCS camera. The focal length is 7.5mm and the image size is 640x480 and the FOV is 60.8x47.5 degrees. Due to the noisy data provided by the tracking system, it is necessary to estimate the camera-marker pose from several locations. From a set of estimates, we compute the mean of the desired transformation by using the method [3].

We compared the precision of the backprojection resulting from this calibration with those provided by hand-eye calibration estimates widely used for visual servoing in robotics applications. We have implemented three different methods that differ in respect of the representation of the transformation using rotation vectors [6], quaternion [1] and Lie group [5].

## 3 Accuracy of backprojection

Both calibration methods are simulated by investigating the influence of the two major sources of errors: the uncertainity of the the feature detection in the acquired images (image error) and the deviation of the 3D position of the markers from their true value (tracking error). The geometry of the marker is fully taken into account. We used the method proposed by [7] for feature detection. We assumed a zero-mean Gaussian distribution of the image error with a standard deviation of $0.1$ pixel as reported by the paper. According to the manufacturer's specification, the optical tracker produces a maximal tracking error of $0.1$mm under the conditions of our experiments. As no information has been provided about the statistical properties of the error, we assumed in the first step a uniform distribu-

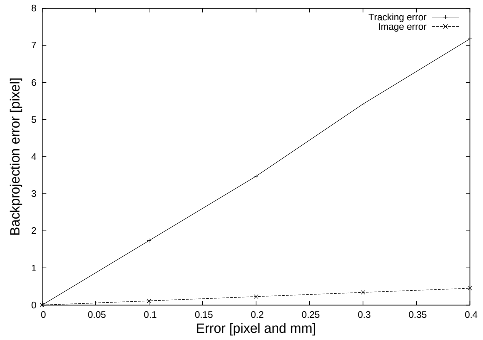


**Figure 1.** Influence of the tracking and image error on the backprojection

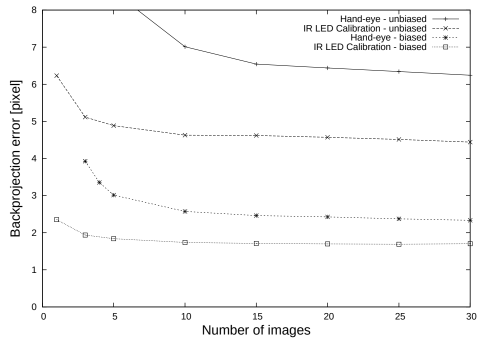


**Figure 2.** Influence of the model of noise

tion $U(-a_u, a_u)$ within the maximal error limits $a_u$ (unbiased estimate). Figure 1 shows the backprojection error due to those two error sources separately in dependence on the standard deviation of the image noise and the maximal tracking error $a_u$. The results demonstrate that the backprojection error is zero if the marker and feature detection is error-free (which is necessary if the simulation is correct) and that the tracking error by far dominates the image error. In the next step we investigated the influence of the number of images used for the calibration on the backprojection error. The simulation results presented on Figure 2 did, however, by far overestimate the experimentally observed errors shown on Figure 3. A closer investigation of the tracking errors revealed, that our original assumption of the error distribution was wrong. The overall tracking error is rather a combination of a (large) systematic, space-dependent bias and a (small) random noise, which we experimentally observed as showing a zero-mean Gaussian distribution with a low standard deviation $\sigma = 0.03mm$. Accordingly, a second series of simulated experiments were carried out by modeling the deviation of the markers' position from their true value as a sum of a locally constant systematic bias as a sample of $U(-0.1mm, 0.1mm)$ and a random Gaussian noise $N(0, 0.03pixel)$ (biased estimate). As demonstrated by Figures 2 and 3, the results of the modified simulation process were in good agreement with the experimental observations. As all hand-eye methods give very similar results, only the output of Tsai's method is representatively shown. The simulated values are the average of 100 indi-

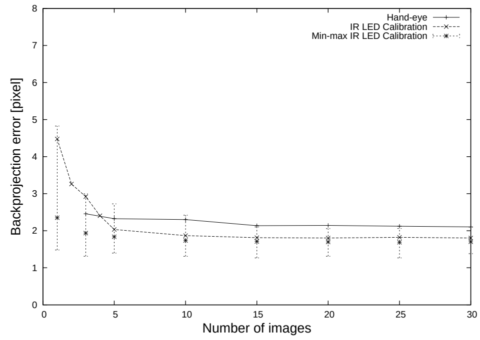


**Figure 3.** Real results for both approaches

vidual random experiments. In order to get some estimate about the distribution of the resulting backprojection errors, we also registered the maximal and minimal deviations observed, shown as vertical bars on Figure 3.

## 4 Conclusion

We have demonstrated, that under the assumption of locally constant bias in the tracking error the developed simulation package reproduces the experimentally observed behaviour of the backprojection error very well. While the IR LED-based calibration consistently showed slight improvement in precision as compared to hand-eye calibration, the remaining backprojection error (1.5-2 pixels) are still too high for our application, even when using a very accurate optical tracking unit. Accordingly, we are currently incorporating an image-based warping method for correction into the AR-based surgical simulation framework.

## References


[1] J. Chou and M. Kamel. Quaternion Approach to Solve the Kinematic Equation of Rotation, $A_a A_x = A_x A_b$, of a Sensor-Mounted Robotic Manipulator. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 656–662, 24-29 April 1988.

[2] W. A. Hoff. Fusion of Data from Head-Mounted and Fixed Sensors. In *First International Workshop on Augmented Reality*, Nov 1998.

[3] M. Moakher. Mean and Averaging in the Group of Rotations. *Society for Industrial and Applied Mathematics*, 24:1–16, 2002.

[4] Northern Digital Inc. www.ndigital.com/.

[5] F. Park and B. Martin. Robot Sensor Calibration: Solving $AX = XB$ on the Euclidean Group. In *IEEE Transactions on Robotics and Automation*, volume 10, pages 717–721, october 1994.

[6] R. Tsai and R. Lenz. A New Technique for Fully Autonomous and Efficient 3D Robotics Hand-Eye Calibration. In *IEEE Transactions on Robotics and Automation*, volume 5, pages 345–358, June 1989.

[7] S. Vogt, A. Khamene, F. Sauer, and H. Niemann. Single Camera Tracking of Marker Clusters: Multiparameter Cluster Optimization and Experimental Verification. In *ISMAR 2002*, pages 127–136, 2002.