

Analysis of Tracking Accuracy for Single-Camera Square-Marker-Based Tracking

Katharina Pentenrieder*, Peter Meier*, Gudrun Klinker†

*metaio GmbH †TU Munich
Lichtenbergstrasse 8 Boltzmannstrasse 3
85748 Garching 85748 Garching

Abstract: In Augmented Reality (AR) applications tracking accuracy plays an important role. Many applications e.g. for measuring purposes or medical environments need accuracy statements which are as precise as possible for the given case.

Here we present a tracking accuracy analysis based on simulated ground truth data. We relate our results to previous studies in that field and show the importance of a detailed breakdown of the values of influencing parameters in order to achieve a specific accuracy statement for the given tracking situation. Furthermore we present a possibility of applying such accuracy data.

Keywords: Augmented Reality, Marker Tracking, Accuracy Analysis, Data Simulation

1 Introduction

In the field of Augmented Reality (AR) virtual information is combined with the real world to enhance the user view of the environment. The quality of the AR illusion is highly depending on a good registration of the virtual and the real world. Many applications e.g. in the medical environment demand a very accurate registration where real and virtual objects are properly aligned. One of the main sources for registration error is given by the errors in the underlying tracking system [Azu97]. Hence it is of great importance to know as precisely as possible how accurate the system results are for a given tracking situation.

A lot of existing AR systems are based on visual markers such as the ARToolKit [HK99], the Siemens Corporate Research (SCR) marker system [XZ02] or the metaio Unifeye SDK [Z H06b]. Although all these systems build on the same kind of tracking object, they differ in areas such as usability, reliability and accuracy [XZ02]. The difference in accuracy can on the one hand be traced back to chosen hardware and environmental constraints and algorithms which are used for marker detection and pose estimation but on the other hand the tracking accuracy also varies within a system based on the relative pose of the marker to the camera.

In many areas tracking accuracy plays an important role and an accuracy information which

specifically fits a given situation can be very advantageous. For example in hybrid tracking systems this information can be used for optimized sensor fusion where trackers can be weighted not only according to their overall accuracy but also situation-dependent according to the currently given situation. The same approach can be used in registration applications where tracking accuracy plays an important role. For good registration quality the tracking error is very important and should therefore be available as precisely as possible.

In industrial applications where Augmented Reality systems are used for measurements the tracking accuracy is included in the measurement process. A situation-specific accuracy statement allows to either identify the measurement precision very accurately for the given case or on the other hand enables to find good tracking configurations where the accuracy is optimal and the errors can be kept small.

In this paper we analyze the visual marker based tracking system implemented in the metaio Unifeye SDK [Z H06b]. We identify important parameters which influence the tracking accuracy and extract accuracy information from a series of tests based on ground truth data. This allows us to present accuracy information which is specifically matched to the given set of parameters, hence to the given situation. The results are related to previous work in the field and a possible implementation for using this tracking information is presented.

2 Previous Work and Contribution

In the past accuracy studies have been performed on existing marker tracking systems [XZ02], [PM02], [ABD04]. Commercially available systems also often state information regarding the precision of the tracking results [Z H06a], [Z H06c], [Z H06d].

2.1 Experiments on Accuracy

In a general comparison on optical marker based tracking systems Zhang et al. evaluated four systems, the ARToolKit, the Institut Graphische Datenverarbeitung (IGD) marker system, the Siemens Corporate Research (SCR) marker system and the Hoffman marker system (HOM) with respect to usability, efficiency, accuracy and reliability [XZ02]. The accuracy evaluation was based on the errors in feature extraction.

The paper states accuracy results for different rotation angles around the x-axis (see figure 1) given as error and standard deviation. For different regions of marker (ROM) the marker detection is analyzed and the results describe the number of detected markers and the smallest ROM for which the detection was successful. The study does not consider pose estimation errors as those depend on intrinsic camera parameters and calibration algorithms but only looks at the corner detection error. The emphasis of the analysis lies on the categorization according to the evaluation criteria

to allow the choice of a system fitting a specific application and its needs.

Two other accuracy experiments have been conducted specifically for the ARToolKit system. Both experiments evaluate distance and one angle. Malbezin et al. placed a marker on the ground and compared tracking results with physical measurements in an orbit around the marker for different distances (1-3m) [PM02]. Here the pose estimation error was analyzed to determine an error correction factor for the ARToolKit.

In contrast to this study Abawi et al. [ABD04] look at error and standard deviation values for pose estimates. They offer the only analysis where a combination of criteria is evaluated. Low error and standard deviation intervals were identified both for different angles and distances and an accuracy function is defined with these four intervals.

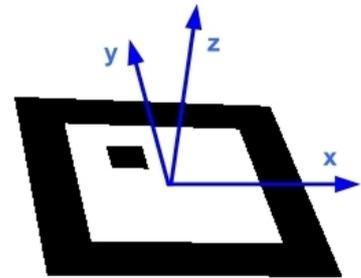


Figure 1: Coordinate axes

2.2 Accuracy Statements of Existing Tracking Systems

Next to experimental analyses of existing optical tracking systems we can find open source and commercially available systems which also state their accuracy. In this section several commercial IR-based systems are presented.

ARTTrack1 and DTrack are tracking systems available from A.R.T. GmbH. These optical tracking systems are based on infrared cameras and spherical features. A.R.T. states their accuracy for a four camera tracking system in an area of three by three meters with $0.4 \text{ mm} \pm 0.06 \text{ mm}$ for target positioning and $0.12^\circ \pm 0.03^\circ$ for target orientation when all markers are fully visible [Z H06a]. This information is given for a complete tracking area of 9 m^2 . No distinctions are made with respect to the position of the features in this area. In addition, the accuracy descriptions are restricted by a lot of constraints, such as the number of cameras and the visibility of all fiducial markers. No statements are made for situations where not all of the target features could be tracked.

The Polaris system of NDI [Z H06c] is a stereo-based infrared optical tracking system which also detects retro-reflective marker features. NDI stated the accuracy of Polaris for two measurement volumes with 0.35 root mean square (RMS). The measurement volumes are a standard cylindrical volume with a diameter of 1 m and distance to the Polaris sensor of $1.4 - 2.4 \text{ m}$ and a pyramid volume with 30° aperture angle for a distance to the sensor of $0.48 - 2.4 \text{ m}$. The RMS values were calculated using more than 900 single marker positions in the standard volume and more than 1500 positions in the pyramid volume. NDI defines the measurement volumes very precisely but does not state any accuracy variations throughout the volume.

The third commercial optical tracking system we describe is the Visualey System of PTI

Phoenix Technologies [ZH06d]. This stereo infrared optical tracking system has a root mean square accuracy of $0.65 \text{ mm} \pm 0.1 \text{ mm}$ for a measurement volume with $\pm 40^\circ$ horizontal angle, $\pm 30^\circ$ vertical angle and distances from $0.6 - 2.2 \text{ m}$. This value is 3D combined using point-accuracy measurements at different distances within the measurement volume. At a distance of 1.2 meters from the sensor the 1D accuracy is specified with approximately 0.15 mm. Again only a single accuracy is given for the entire measurement volume. It is stated that this value is computed using several 1D accuracy measurements but no further specifications on accuracy variations within the volume are given.

2.3 Contribution

This paper combines results of previous studies and accuracy statements on marker-based tracking systems and extends the existing analysis by providing accuracy results generated in a simulation process specifically tailored to identify the influence of the entire tracking configuration on the accuracy of each parameter of the tracking result. That way a pose-dependent accuracy information can be given which specifically fits the current tracking situation. Different from most accuracy analyses in the past, we do not generate one statement for the entire tracking system or for a specified and rather large tracking volume. Instead we define accuracy statements for small tracking regions, considering relevant parameters for the overall tracking accuracy.

Thus far, the projected size of a marker target in the image (as a combination of focal length, resolution and depth distance) and the orientation of the marker were analyzed for their influence on the marker corner detection and the six dimensional tracking pose result. We plan to extend our experiments towards analyzing the influence of noise, illumination and camera distortion and thereby the translational x- and y-components of the marker pose.

3 Experimental Approach

Looking at the experiments which are described in the section on related work, we see that accuracy analysis can differ in various ways. The results do not only serve different purposes, the parameters and influence variables, the analysis process and the presentation of the results differ considerably between different studies.

Before we present our own experimental approach we therefore want to give an overview on the concepts and terms of the previous studies and systems. This summary will then be used to formulate our way of accuracy analysis, following general guidelines for designing experiments [Mon05]. These guidelines determine the data acquisition for the accuracy study. The statistical analysis which builds on this data is targeted towards our goal of creating situation specific statements for the tracking accuracy of our system.

3.1 Concepts and Terms in Accuracy Analysis

If we take the descriptions of previous work as basis, we can extract several main concepts for accuracy analysis: Which parameter is analyzed for its accuracy? What is the statistical meaning of the term accuracy? Which influence factors are considered in the accuracy analysis and for which factor intervals do we create such an accuracy statement? What is the main emphasis of the study? How many measurements are taken to get an accuracy value?

The following tables shall summarize the experiments and systems of section 2.1 and 2.2 and answer the questions above. For the marker tracking studies of section 2.1 (see table 1) we can

	Zhang et al.	Malbezin et al.	Abawi et al.
Accuracy	error, std.dev. of corner detection	error of pose estimate	error, std.dev. of pose estimate
Influence Factors	y-angle, ROM	z-angle, z distance	y-angle, z distance
Emphasis	system categorization	error correction	accuracy function
Shots	—	test à 1000 shots	315 test à 250 shots

Table 1: Overview on the three accuracy studies (see section 2.1)

see that accuracy can reference to error or standard deviation or both of the regarded parameters. Accuracy of a tracking system can be described by the accuracy of the corner detection process or by looking at the result of the pose estimation process. The influence factors for the experiments of section 2.1 are consistent. They include the distance of the camera to the marker, which for a given camera then reflects the marker size or region in an image (ROM in Zhang et al.). Sloped markers are considered by rotating the camera around the marker to cause different viewing angles (projective distortion in Zhang et al.). Furthermore we can see that the number of measurements should be large to offer statistical importance.

The experimental approach we undertake considers corner and pose estimation error and standard deviation. This allows us to compare our results with the ones of previous studies. The choice of the influence parameters we will analyze is based on literature and experience in the field (see section 3.4). As we will see, this parameter list includes marker size (camera distance) and viewing angle. In our case the viewing angle is described by rotations of the marker around the axes indicated in figure 1. The commercial systems (see table 2) show similar interpretations for accuracy. Error and/or standard deviation are analyzed either of the complete pose estimate or only of the translational part. But we see that the statements are given for rather large volumes, not considering any possible variations within that volume. Our approach will focus on this deficit and look at variations within tracking volumes in detail. Again a large number of shots is taken to form an

	A.R.T.	NDI	PTI
Accuracy	error, std.dev. of pose	error of position	error, std.dev. of position
Area/Volume	square area $9m^2$	cylindric/pyramidic volume $0.8m^3/1.2m^3$	pyramidic volume $2.9m^3$
Shots	—	900 (cylindric), 1500 (pyramidic)	≥ 900 shots

Table 2: Overview on the three commercial visual tracking systems (see section 2.2)

accuracy statement.

We realize the creation of a large data set through a simulation process. Different from the previous approaches we will not simulate repeated measurements of the same configuration. Instead we follow the "Design of Experiments" approach of Montgomery et al. [Mon05].

3.2 Experimental Design

Montgomery describes an experiment as a process to determine the influence that controllable factors have on the output response of a system [Mon05]. Very common for experiments in practice is the "one-factor-at-a-time" approach. Here one input factor is varied while all others are kept fix. The main disadvantage of this approach is that it fails to consider possible interactions between factors. For a complete experimental analysis all factors should be varied in a so called factorial experiment. In general this approach is not realizable since a test of all combinations of all input factors would go beyond the scope of an experiment. A compromise is given by the fractional factorial approach where only a subset of all possible combinations is regarded.

In addition the following three principles should be considered for experimental design. For a statistical analysis, independently distributed variables are required. This can be achieved by randomizing the allocation of data. Furthermore each combination of factors should be repeated independently. Finally blocks of experiments should be built to improve the precision for comparisons.

We acquire our data sets through a simulation process. Thus we do not have to meet the replication principle, since for the same parameter values the result will be exactly the same. This leads us to the requirements we have for our system.

3.3 System Requirements

Based on the design issues for experiments we identified several constraints that are important for the data generation process:

The data should be ground truth data, systematically generated by a simulation to avoid measurement errors. The acquisition should be a repeatable and randomizable process to meet the design of experiment guidelines. It should include all the parameters which influence the tracking accuracy and which can be controlled by the simulation. For the parameters, levels (i.e. the step size) and ranges should be arbitrary to keep restrictions as small as possible. To be able to identify the influence of each parameter it should be possible to vary them in a controlled way while letting all other parameters be randomly chosen. That way the effects of each parameter on the final tracking result can be analyzed [Mon05]. Finally the ground truth data should be compared with the corresponding tracking results of the metaio Unifeye SDK tracking system and the results should be stored in an appropriate way.

Out of this list of requirements a system was implemented which allows us to create ground truth data by specifying the relevant parameters with respect to values, levels and ranges, which parameters shall be varied in a controlled way and which in a random way and how many data sets we want to have. Furthermore the system provides the possibility to acquire corresponding tracking information and store the results in an appropriate exchange format. The system will be described in more detail in the next section.

3.4 Identifying the Relevant Parameters

The identification and selection of relevant parameters was an iterative process. Luhmann states a number of criteria for photogrammetric analysis such as the distance of the object, focal length and resolution of the camera and the viewing angle [Luh03]. Experience with the system leads to additional parameters such as noise, illumination and distortion, that are responsible for variable tracking accuracy. These parameters should be included in the evaluation process step by step to find out about their relevance.

In a first experimental phase we had a look at the parameters "distance of the marker", "focal length" and the "viewing angles" which were one at a time varied in a controlled way while all other parameters had random values. The results of this preliminary study allowed us to identify interrelated parameters and merge them. The values for focal length, resolution and marker distance were combined to a parameter "Marker area in pixel" (MAiP) similar to the ROM parameter in [XZ02]. This parameter together with the viewing angle was considered for the first complete study on tracking accuracy. The results are described below.

4 System Overview

This section gives an overview of the process that determines the tracking accuracy information. It is divided into two major steps:

First a large set of experimental data is acquired, consisting of ground truth information combined with the corresponding tracking results. In a second step, this data is statistically analyzed to derive information regarding the influence of the relevant parameters on the corner detection and pose estimation error.

4.1 Generation of Ground Truth Data

For data acquisition, a tool was implemented in C++ using OpenGL combined with an FLTK graphical user interface. It allows to generate arbitrary camera views of a metaio marker. The views are then processed by the metaio Unifeye SDK tracking component. The eight corners which describe the inside and outside boundary of the thick black border of the marker are extracted to estimate the marker pose in the current view. For each image the given ground truth information on the marker position and the camera are stored in a text file together with the tracking results. These files are then used in the statistical analysis.

The system allows us to specify fixed values or ranges for parameters and to select the controlled and the random parameters. At the moment the camera pose with respect to the marker, the focal length and the resolution can be specified for the data generation. Using this data an image of the camera view is rendered into a pixel buffer and then it is sent to the tracking component. In future work we want to add noise and distortion to the images to approach step by step realistic camera images.

4.2 Error Analysis

In the analysis step the resulting data files of the data generation were processed. Each file consists of general information on the simulated camera (focal length, resolution values), ground truth information on the eight marker corners (four corners for each border square of the black area) and the camera pose (6D translation and rotation) and the tracking results for the corner detection and the camera pose.

The data is imported in MS Excel and is analyzed and visualized as diagrams. For the corner error the 2D sub-pixel differences between ground truth and tracking result for the eight corners are averaged. The pose error is computed as averaged sum for each of the 6D parameters (translation in x, y, and z direction, rotation angle around x, y and z axis) separately and combined for an overall translation and rotation error where the translation error corresponds to the euclidean distance of the ground truth and the tracking translation and the rotation error is the averaged angular error.

5 Results

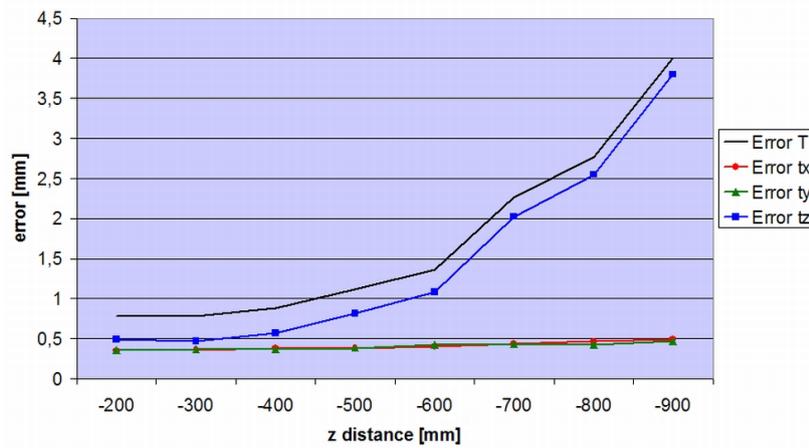
This section presents the results of our analysis. We have a look at the influence of the parameters "Marker area in pixel" as a function of the camera distance to the marker and the viewing angle. The results are then put in relation to the accuracy experiments which were described in section 2 on related work. We point out similarities and differences and show the relevance of a detailed consideration of parameter influence by looking at small intervals of parameter values.

5.1 Marker Area / Camera Distance

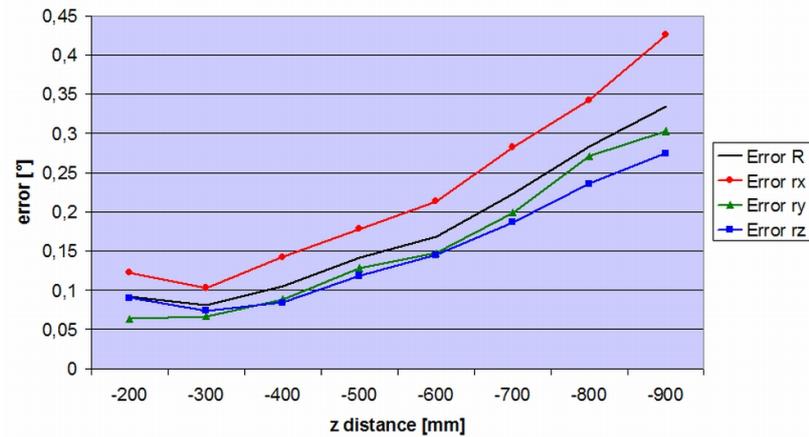
The following two graphs (figures 5.1 and 5.1) show the translational and rotational pose error for different camera distances. For the creation of the underlying data sets we used fixed values for marker size, camera resolution and focal length. That way our influence parameter "marker area in pixels" only depends on the camera distance to the marker. In order to allow comparisons with previous experiments the parameter values and ranges are based on the experiments of Abawi et al. [ABD04]. A marker size of 55^2mm^2 , camera resolution of 640x480 and camera distances between 200 mm and 1000 mm were taken. For the data acquisition the camera distance was varied in a controlled way while the viewing angles were randomly distributed.

Both figures each show four curves, the (absolute) positional/ rotational error for the three directions and a combined error. For the translation the combined error depicts the euclidean distance between the ground truth translation and the estimated translation, for the rotation the combined error describes the average of its three components. The trend of the error curves is similar to the ones depicted in Abawi et al. [ABD04]. For larger distances the pose error increases. In addition we can see that the translational error in z direction makes the largest contribution to the overall positional error. For the rotation error, on the other hand, all three components add equally to the overall error. Abawi et al. used their result to define an interval of 20 cm to 100 cm where the systematic error is low. In their case the translational error increased from around 2 cm to approximately 8 cm. Both experiments support our approach of creating a situation specific accuracy function, which is pose-dependent. We can extend the idea of defining an interval based error function and use not only low and high as criteria for tracking accuracy but each level of the variable by looking at its complete range.

A similar analysis can be undertaken for the standard deviation. Again the variation of the standard deviation value for different camera distances underlines the approach of considering small intervals of the distance for one accuracy or error statement.



(a) Translational pose error



(b) Rotational pose error

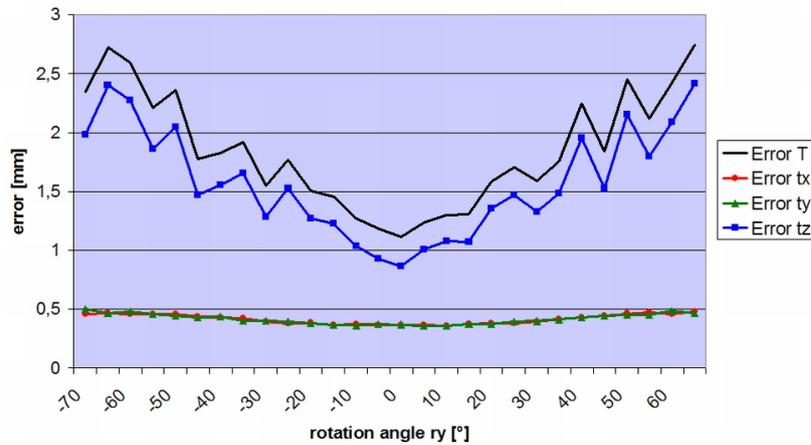
Figure 2: Pose error for different camera distances

5.2 Viewing Angle

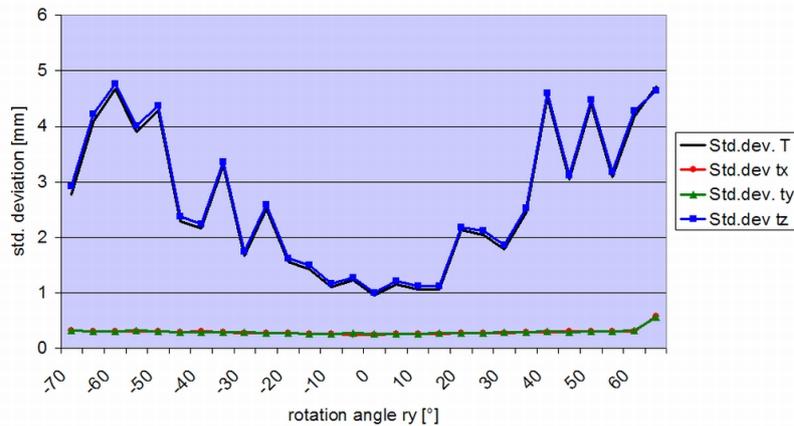
For the analysis of the viewing angle we created similar graphs showing the dependencies of corner, translation, and rotation error and standard deviation on the rotation angle (here around the y-axis).

Regarding the pose error and standard deviation we find that for the translation part both error and standard deviation curves describe a valley which has its lowest point at 0° of rotation angle (figures 5.2 and 5.2). Again it is the difference in z direction which influences the overall error and standard deviation the most.

For the rotational error, the curves have a different shape (figures 5.2 and 5.2). The border areas show a large error and standard deviation, too. But around 0° only the angular difference around the z axis is minimal. For the x- and y- axes, both error and standard deviation increase which



(a) Translation error

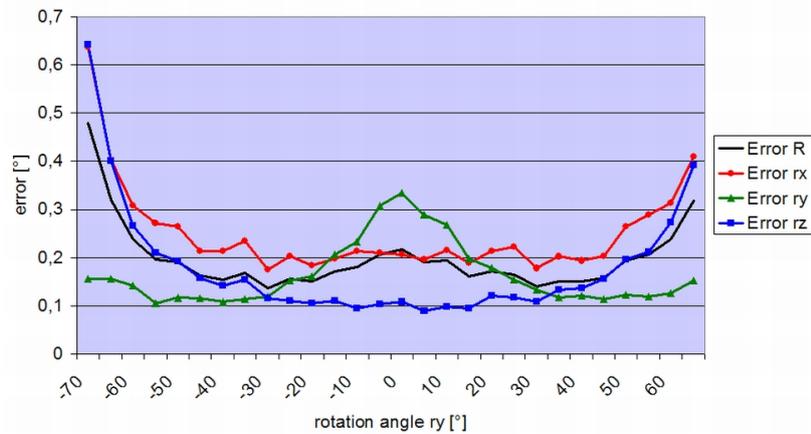


(b) Translation standard deviation

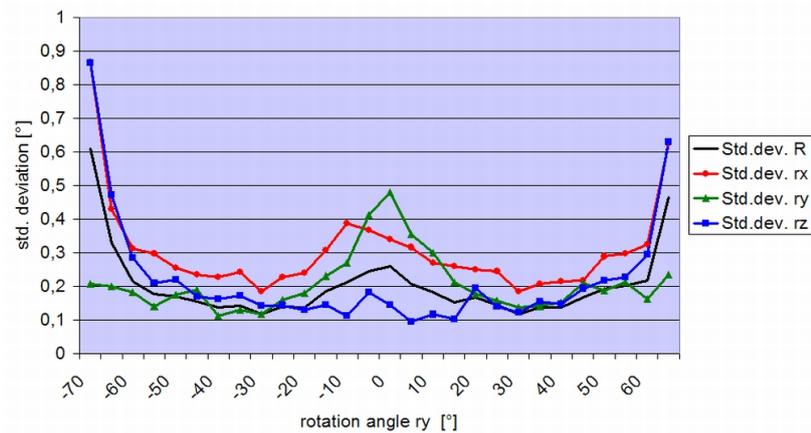
Figure 3: Translation error and standard deviation for different viewing angles

leads to "W"-shaped curves. If we compare our results to the study of Abawi et al. [ABD04] we can also state, that error and standard deviation are high around 0° and for large angles. But this is only important for the rotational part. Regarding the translation the area around 0° must not be considered as a high error region.

This shows that we have to divide our accuracy statements carefully because the influence of a parameter on the rotation part is not necessarily given for the translation part, too. And for the positional accuracy itself we can distinguish the x- and y- part from the depth part, as for z the error and standard deviation, values are much higher than for the other two components. So for generating an accuracy information for pose estimation we can and should state specific values for each of the six degrees of freedom.



(a) Rotation error



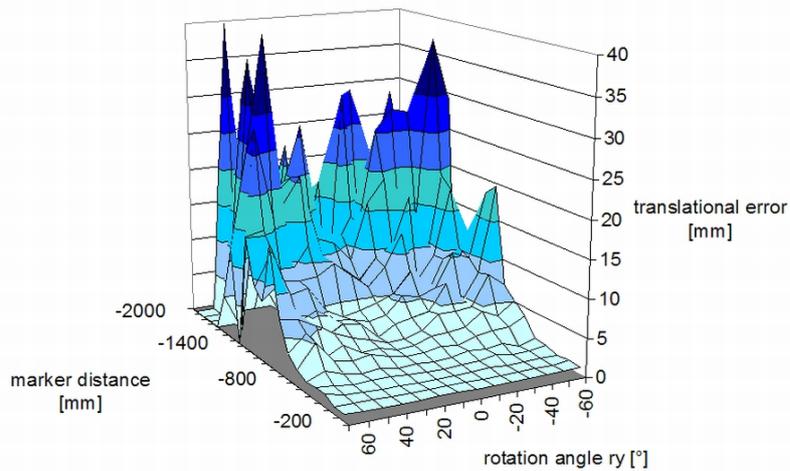
(b) Rotation standard deviation

Figure 4: Rotation error and standard deviation for different viewing angles

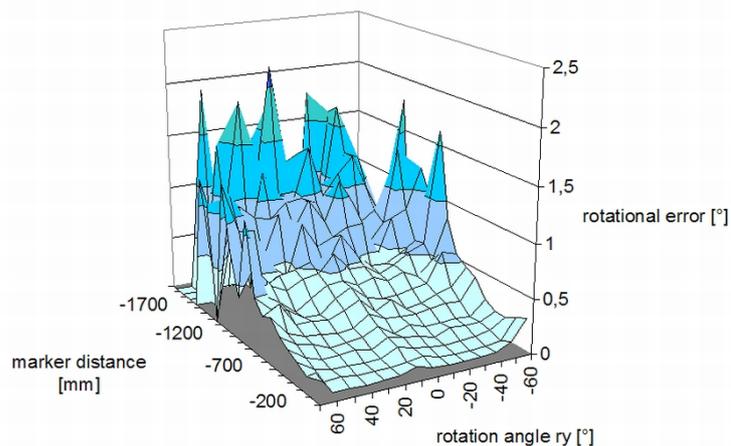
5.3 Combined parameters

This last analysis combines the two main influence parameters "camera distance" and "viewing angle". Figures 5.3 and 5.3 show the rotational and translational error of the pose estimation for different values of distance and rotation angle around the y-axis.

The graphs can be interpreted as an accuracy function similar to the approach of Abawi et al. [ABD04]. Here only the error is included but similar surfaces can be generated for standard deviation. The surface of our error functions varies considerably. Still it only describes error trends. Therefore accuracy statements should be given for small intervals on that surface to consider the variability of the error but also take into account small statistical variations.



(a) Translation error



(b) Rotation error

Figure 5: Error for camera distance and rotation angle

5.4 Further experiments

In section 3 we enumerated possible parameters that influence the results of a marker tracking process. We started with a few of these parameters and plan to conduct further experiments to analyze, step by step, the influence of as many relevant input variables as possible. The first series of experiments showed that our approach is reasonable and that the idea of looking at very small intervals within the range of a variable is important for a precise accuracy statement. Our next step will be to look at noise, illumination and camera distortion which is an important factor for tracking error. Then the x and y position of the marker in an image will play a role, too.

6 Application of our results

There are various possibilities to apply such a situation-specific tracking accuracy as described in the previous sections. Here we present a general way of providing and using the situation-specific information.

Abawi et al. [ABD04] defined an error function based on intervals which differs between high and low systematic error and standard deviation of estimated camera distance and viewing angle. With the resulting four parameters the volume between camera and marker can be divided in areas where no, one, two, three or four of the parameters are considered to be advantageous (meaning a small error and a small standard deviation). This implies an accuracy function with five levels for the volume of marker detection: from low to high accuracy depending on the number of high parameter intervals which are met. This approach enables to find advantageous set-ups for marker-based tracking systems but it does not try to give specific accuracy values for each element of the 6D pose estimate.

The data we gathered through the experiments in the previous section allows a much finer resolution for tracking accuracy both for error and standard deviation statements. In addition we use the results to create 6D accuracy statements which meet the current tracking situation as precisely as possible.

For an AR measurement application we created a large data base which is used as a look-up table for accuracy queries. Each data set contains the necessary query parameters ("Maip" and "viewing angles"), the ground truth pose and the corresponding tracking estimate. For a given tracking scenario a query is made and the data sets fitting the query parameters are returned. As the database will probably not contain a data set which exactly fits the current request, a search for result sets is performed in small intervals around the query parameters. The size of these intervals depends on the size of the database. The larger the data base the smaller the interval which is being searched for fitting data sets. At the moment we are working with a database of around 50000 entries.

Error and standard deviation for the current tracking estimate are then calculated using the query result.

The computed tracking accuracy is applied for an error propagation process. With our approach of a look-up table we are able to generate a 6D tracking error information and a corresponding 6×6 covariance matrix specifically fit to the current tracking estimate. Hence we can then predict whether the tracking accuracy at a given point in space satisfies the users requirements.

7 Conclusion/Future work

We presented a method for acquiring tracking accuracy information which is based on the simulation of ground truth data. Our results show the necessity of stating accuracy in a very specific way, fitting the current tracking situation, as error and standard deviation of detected corners and estimated pose vary considerably for different values of input parameters.

Our approach builds on the experiments of Abawi et al. who created an accuracy function based on camera distance and viewing angle. In our application the accuracy function has the form of a look-up table. The underlying database allows the generation of a 6D accuracy statement (error and covariance) which is tailored to the given query parameters.

Further experiments are necessary to analyze in detail the influence of additional relevant parameters such as noise, illumination and distortion.

References

- [ABD04] Daniel F. Abawi, Joachim Bienwald, and Ralf Doerner. Accuracy in optical tracking with fiducial markers: An accuracy function for ARToolkit. In *ISMAR*, 2004.
- [Azu97] R. Azuma. A survey of augmented reality. In *Presence: Teleoperators and Virtual Environments 6, 4*, pages 335–385, 1997.
- [HK99] M. Billinghurst H. Kato. Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In *Proceedings of IWAR 99*, pages 85–94, 1999.
- [Luh03] T. Luhmann. *Nahbereichsphotogrammetrie, 2. Auflage*. Wichmann Verlag, 2003.
- [Mon05] D. Montgomery. *Design and Analysis of Experiments*. John Wiley & Sons, Inc., 2005.
- [PM02] B. Thomas P. Malbezin, W. Piekarski. Measuring ARToolkit accuracy in long distance tracking experiments. In *1st International ARToolkit Workshop*, 2002.
- [XZ02] N. Navab X. Zhang, S. Frnz. Visual marker detection and decoding in AR systems: A comparative study. In *ISMAR*, 2002.
- [Z H06a] ARTTrack1/DTrack, A.R.T. gmbh. <http://www.ar-tracking.de>, 2006.
- [Z H06b] Unifeye SDK, metaio gmbh. <http://www.metaio.com>, 2006.
- [Z H06c] Polaris system, NDI. <http://www.ndigital.com>, 2006.
- [Z H06d] Visualey system, PTI phoenix technologies. <http://www.ptiphoenix.com>, 2006.