

Calibration Errors in Augmented Reality: a Practical Study

Javier-Flavio Vigueras, Gilles Simon, Marie-Odile Berger

► **To cite this version:**

Javier-Flavio Vigueras, Gilles Simon, Marie-Odile Berger. Calibration Errors in Augmented Reality: a Practical Study. 4th IEEE and ACM International Symposium on Mixed and Augmented Reality - ISMAR'05, Oct 2005, Vienna, Austria. pp.154-163, 10.1109/ISMAR.2005.23 . inria-00000383

HAL Id: inria-00000383

<https://hal.inria.fr/inria-00000383>

Submitted on 3 Oct 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Calibration Errors in Augmented Reality: a Practical Study

Javier-Flavio Vigueras Gomez, Gilles Simon and Marie-Odile Berger
LORIA - INRIA Lorraine - UHP Nancy 1

Abstract

This paper confronts some theoretical camera models to reality and evaluates the suitability of these models for effective augmented reality (AR). It analyses what level of accuracy can be expected in real situations using a particular camera model and how robust the results are against realistic calibration errors. An experimental protocol is used that consists of taking images of a particular scene from different quality cameras mounted on a 4DOF micro-controlled device. The scene is made of a calibration target and three markers placed at different distances of the target. This protocol enables us to consider assessment criteria specific to AR as alignment error and visual impression, in addition to the classical camera positioning error.

1. Introduction

Vision-based camera tracking arouses great interest in augmented reality (AR). One advantage over other methods is that pose computation is generally more accurate in the image plane, thereby minimizing the perceived image alignment error. The pinhole camera model is mainly used. Extrinsic parameters are updated online in order to fit the real camera motions. Intrinsic parameters such as the focal length and the principal point are generally supposed constant over time and calibrated during an offline process.

Various calibration methods exist that allow the parameters to be computed with more or less accuracy. In controlled environments such as medical or industrial applications, accurate parameters can be obtained from several images of a calibration target shot from various positions within the working space. When applications take place in outdoor and vast environments and have to work during long periods of time, this conventional calibration method is not appropriate: the potential working space is too large to be calibrated accurately. In addition it is impossible for the user to take the calibration target away with him in order to reinitialize the system if needed. Therefore, simpler means of calibration must be considered that either use constraints in the scene or make use of prior calibration hypotheses.

Examples of such strategies are for instance the use of a squared marker placed in the scene or the use of vanishing lines in the environment for calibration. Prior constraints that are commonly used to make calibration easier assume that the principal point lies at the center of the image. The intrinsic parameters are obviously less reliable than those obtained with a calibration target. This paper reports on a set of experiments made to assess the influence of possible calibration inaccuracies on camera tracking in AR applications.

In the past, several works have been dedicated to the analysis of the effect of calibration errors on structure and motion computation or viewpoint computation [20, 1, 10]. Most of these works address the problem of theoretically studying the effect of calibration errors on the motion estimate, assuming that the pinhole model is perfectly respected by the considered camera. These works especially allow some difficulties in viewpoint computation to be explained: for example, [2] shows that coupled changes of certain parameters, such as the angle of rotation and the principal point, are barely observable, making them indistinguishable. However these works do not integrate the fact that the pinhole model is not perfectly respected by the cameras, especially by zoom lens cameras. In addition, these works do not integrate practical and numerical aspects of calibration: parameters are computed from a set of images or from a set of particular features and the computed parameters obviously depend on the chosen features and their reliability is often restricted to the volume where measures are taken [14, 4]. However, being able to minimize error far from the measurement volume is important in many applications, even for controlled environments: consider for instance the example given in [15], in the field of gynecologist training for forceps deliveries. The students are supposed to learn how to insert the two parts of the forceps into the vagina without harming mother and baby. A virtual view of biological functions are overlaid on a body phantom of the mother with the baby's head inside that can be controlled by robot arm. Movements of the forceps inside the body phantom are inferred from 6DOF tracking of retro-reflective markers placed at the visible extremities of the forceps. In this application, it is therefore necessary to

get accurate alignments in areas distant from the features used for tracking in order to give guarantees on the reliability of the augmented views.

Like in [3], though in a different context, we examine the influence of calibration errors under practical situations which are important for AR applications. We especially attempt to identify if and when the common hypothesis of a fixed principal point, often considered at the center of the image, has a significant influence on viewpoint computation. This study has been conducted with two cameras: a camera with a good quality lens and a webcam. The first one bears very low distortion whereas the webcam suffers from noticeable distortion. It is important to note that the conclusions of our study are largely independent from the type of camera. They show that fixing intrinsic parameters only makes sense when distortions are corrected, even for low distortions. We also study the level of accuracy that can be expected in practice when inaccurate camera parameters are used.

Our experimental results were obtained using a digital camera mounted on a micro-controlled 4DOF device. A calibration target was used to compute the camera parameters. Additionally, three markers whose 3D positions are known were placed near and far from the target. This protocol enabled us to analyse the effects of calibration errors according to three different criteria: camera positioning (computed poses are compared to the 4DOF device commands), alignment errors (projected markers are compared to their detected positions) and visual impression (virtual cones are added into the scene).

The paper is organized as follows: our experimental protocol is presented in details in section 2. The relevance of fixing intrinsic parameters using the pinhole camera model is studied in section 3. Influence of calibration errors over the presented criterion is shown in section 4 and our conclusions are summarized in section 5.

2. Experimental protocol

A 4DOF micro-controlled device was used to assess the accuracy of the recovered camera poses. This device is made of a XY table whose resolution is 0.48 mm, on which a pan-tilt unit (PTU) is mounted, whose resolution is 0.013 deg (Fig. 1).

A camera was fixed on the PTU and a sequence of 753 images showing a calibrated scene was taken at resolution 640×480 pixels. During that sequence, the camera followed a crenel-shaped path (Fig. 1) and the PTU commands were chosen in order to make the camera approximately pointing to the center of the scene in each frame. At three times of the sequence (beginning, middle, end – see Fig. 1), the camera was kept at a fixed position and was rotated so that the apparent motion of the scene roughly followed

a rectangular path over the borders of the image. This sequence enabled us to experiment a large variety of camera / target configurations.

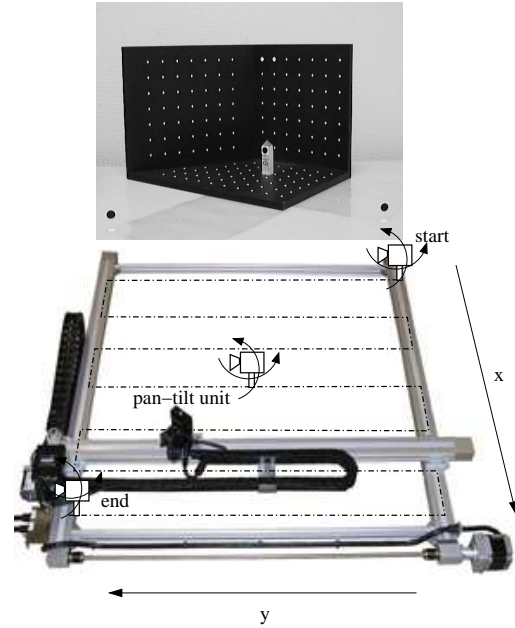


Figure 1. The 4DOF micro-controlled device and the calibrated scene.

The calibrated scene contained a calibration target and three markers whose positions were known in the target coordinate system (Fig. 1). Each marker consists of a black disc of diameter 1.5 cm on a white background. The calibration target is bounded by a box of size $39 \times 26 \times 26$ cm. It had a double function:

1. provide the correct values of the calibration parameters (gold standard) and
2. provide camera poses along the sequence, possibly using intrinsic parameters different from their correct values.

Computed camera poses could be compared to the ground truth (commands sent to the XY table and PTU) using a pre-computed hand-eye transformation. One marker, labelled as *tower* in the following, was placed on the middle of the horizontal face of the target. Two other markers, labelled as *left marker* and *right marker* in the following, were placed at a distance of (resp.) 27 cm and 19 cm from the target. The markers were tracked automatically using a correlation technique, which enabled us to measure alignment errors along the sequence, inside and outside the measurement volume.

Finally, in order not to restrict the impact of our results to a single type of camera, tests were performed us-

Camera	f	$\lambda_1 \cdot 10^7$	$\lambda_2 \cdot 10^{13}$	c_x	c_y	Error
Webcam	824.4	1.492	-1.187	236.4	298.7	2.46%
Sony 1	1113.1	-1.417	2.480	303.0	290.2	1.56%
Sony 2	2325.2	-0.675	-0.153	311.2	312.5	0.95%

Table 1. Focal length and distortion parameters of the lenses used in our tests.

ing three different quality lenses : a high quality lens (Sony DFW-VL500 digital camera with long focal length), a medium quality lens (same camera, shorter focal length) and a poor quality lens (Logitech QuickCam web camera). Focal lengths and distortion parameters of these lenses are provided in table 1. A description of the distortion parameters is provided in section 3.2. The Sony camera has a pin-cushion distortion and the webcam a barrel distortion. Last column of table 1 shows the relative error due to distortion at the upper right image corner.

3. Choosing the right camera model

A common assumption in AR is that intrinsic parameters of the camera remain constant over time. This assumption is convenient as it avoids compensatory effects between intrinsic and extrinsic parameters [2]. Such compensations generally induce statistical fluctuations in the computed camera parameters, which can cause a jittering effect in image augmentations. However, it is not obvious if this assumption is relevant in practice: allowing changes in intrinsic parameters while the camera is moving may also allow to compensate for deficiencies in the modeling of the camera [18]. In this section, experimental results are shown that prove that fixing intrinsic parameters (f, u_0, v_0) only makes sense when optical distortion is considered. Otherwise, the use of constant intrinsic parameters is not well founded.

3.1. Using the pinhole camera model

When using the pinhole camera model, 3D points (X, Y, Z) project into image pixels (u, v) according to equation:

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} f & 0 & u_0 \\ 0 & af & v_0 \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{R}|\mathbf{t}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix},$$

where f is the focal length measured in width of a pixel, a the aspect ratio height/width of a pixel and (u_0, v_0) the pixel coordinates of the principal point of the image. \mathbf{R} and \mathbf{t} are the rotation matrix and translation vector that pass from the world coordinate system to the camera coordinate system. Previous works have proven that the estimation of the aspect

ratio is robust against the relative position of the target used to compute this parameter with regard to the camera [17] and also against possible zooms and focuses of the camera [6]. This parameter is therefore considered as fixed in the rest of the paper, and set to its gold standard value (details on how the gold standard values were obtained are given in section 3.3).

In order to assess the relevance of fixing intrinsic parameters in camera tracking, we first performed in each frame a 9-parameters optimization of the reprojection error of the target centers:

$$errf(\Phi, \mathbf{p}) = \sum_j (u_j - \hat{u}_j)^2 + (v_j - \hat{v}_j)^2,$$

where $\Phi = [f, u_0, v_0]^t$ is the intrinsic parameters vector, $\mathbf{p} = [\alpha, \beta, \gamma, t_x, t_y, t_z]^t$ is the extrinsic parameters vector (α, β, γ are the Euler angles of the rotation), (u_j, v_j) are the projected target centers and (\hat{u}_j, \hat{v}_j) the corresponding detected centers. As such minimization is a non linear problem, we used a non linear optimization program based on the Levenberg-Marquardt algorithm. An initial estimate for the camera parameters was computed using the Faugeras-Toscani method [7].

The first row of figure 2 shows the evolution of u_0, v_0 and f along the sequence, using the medium quality lens. The confidence intervals of the computed parameters (± 3 standard deviation) are also added to the graphs. These intervals were obtained by back propagating a Gaussian noise of standard deviation 0.5 pixels assumed on image measurement [11]. The intrinsic parameters that minimize $errf$ have values that can change a lot from one image to another: u_0 lies in interval [260,364], v_0 in interval [177,265] and f in interval [1122, 1178]. Concerning u_0 and v_0 , there is even no horizontal line passing through all the confidence intervals.

Similar results were obtained using the low and high quality lenses. The mean values and standard deviations of the intrinsic parameters over the sequence are shown for the three lenses in table 2. These values can be compared to the gold standard values, obtained as explained in section 3.3. It is noticeable that some mean values are far from the correct values, even when using the high quality lens.

3.2. Modeling the distortion

The same experiment was done while adding a fourth order approximation of radial distortion to our camera model. Using that model, the coordinates (\hat{u}', \hat{v}') in the undistorted image are obtained from the observed image coordinates (\hat{u}, \hat{v}) by the following equation:

$$\begin{bmatrix} \hat{u}' \\ \hat{v}' \end{bmatrix} = \begin{bmatrix} \hat{u} \\ \hat{v} \end{bmatrix} + (\lambda_1 r^2 + \lambda_2 r^4) \begin{bmatrix} \hat{u} - c_u \\ \hat{v} - c_v \end{bmatrix},$$

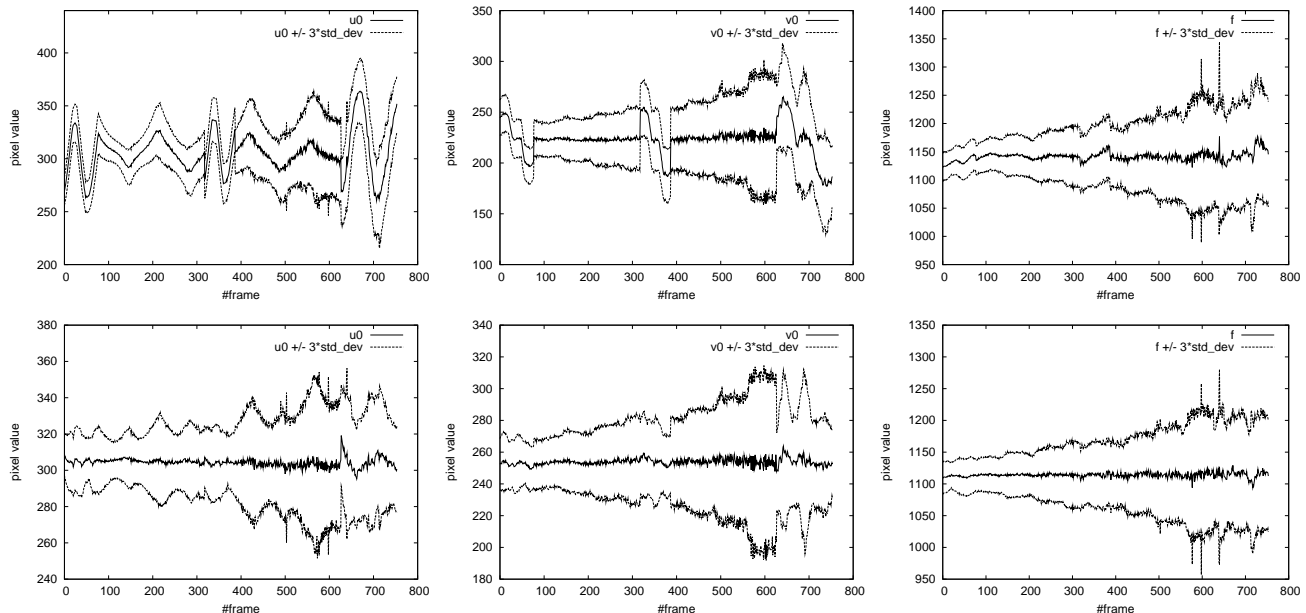


Figure 2. Values of intrinsic parameters obtained over the sequence. First row : original images are used. Second row: radial distortions are removed.

		Gold std	Uncorrected distortion	Corrected distortion
Web-cam	u_0	341.4	334.7 ± 11.4	341.3 ± 2.24
	v_0	261.3	273.0 ± 13.3	263.7 ± 1.48
	f	824.4	816.8 ± 16.8	825.8 ± 2.67
Sony 1	u_0	306.6	305.8 ± 19.0	304.3 ± 2.24
	v_0	253.0	223.2 ± 15.0	253.9 ± 2.10
	f	1113.1	1141.3 ± 7.73	1114.0 ± 3.64
Sony 2	u_0	303.4	304.1 ± 38.1	302.5 ± 5.17
	v_0	306.2	207.0 ± 35.6	306.9 ± 5.04
	f	2325.2	2429.8 ± 44.6	2326.8 ± 10.9

Table 2. Means and standard deviations of intrinsic parameters obtained with uncorrected and corrected distortions.

where $r = \sqrt{(\hat{u} - c_u)^2 + (\hat{v} - c_v)^2}$, λ_1 and λ_2 are the first and second parameters of the radial distortion and (c_u, c_v) is the center of distortion. Undistorting the target centers using the gold standard values of the distortion parameters (section 3.3) and minimizing the reprojection errors err_f using the undistorted targets led to the intrinsic parameters shown on the second row of figure 2. Now these parameters fluctuate in much more reduced intervals: u_0 lies in interval [297,320], v_0 in interval [246,261] and f in interval [1094,1130]. Moreover, for each parameter we get a band of horizontal lines that pass through all the confidence intervals. Mean values obtained for the three lenses are now much closer to the gold standard values, and the standard

deviations are considerably reduced (Tab. 2).

These results confirm in practice those obtained in [8] on synthetic data: 60 points were randomly scattered in a sphere of radius 14 units. The camera was given random orientations and was placed at various distances from the center of the sphere with a mean distance of 70 units. The 3D points were projected in the image and radial distortions were added to the projected points in addition to a Gaussian noise. When intrinsic parameters were estimated without taking distortion into account, incorrect values were obtained, and the confidence interval of the computed parameters even did not include the true values.

3.3. Validation on camera pose and 2D alignment

Using the classical pinhole camera model leads to inaccurate intrinsic parameters that locally compensate for deficiencies in the camera model. By contrast, correcting radial distortion allows to get intrinsic parameters that are less sensitive to the camera - calibration target configuration. However, the pinhole camera model with corrected distortion and fixed intrinsic parameters still has to be proved suitable for accurate AR applications. Two criteria were used to assess this suitability: the camera positioning error and the markers alignment error.

In order to get accurate intrinsic parameters, we performed a bundle adjustment over 50 images of the calibration target. Images were chosen so that the target appeared large in the image and all its faces were well visible. The

	n	$\Delta T_x(\text{mm})$	$\Delta T_y(\text{mm})$	$\Delta \text{pan}(\text{deg})$	$\Delta \text{tilt}(\text{deg})$
Webcam	222	0.001	0.000	0.020	0.060
Sony 1	753	-0.165	-0.004	-0.020	0.014
Sony 2	480	0.000	0.000	-0.011	-0.027

Table 3. Mean errors on pose parameters using the pinhole camera model with fixed intrinsics and corrected distortion.

following reprojection error of the target centers was minimized over the 50 images:

$$\min_{a, \Phi, \Lambda} \sum_{i=1}^{50} \text{errf}(\Phi, \mathbf{p}_i(\Phi)),$$

where a is the aspect ratio and $\Lambda = [\lambda_1, \lambda_2, c_u, c_v]^t$ is the distortion parameters vector. $\mathbf{p}_i(\Phi)$ is the pose parameters vector that minimizes errf in image i for a given set of intrinsic parameters Φ . When computing camera poses over the 753-frames sequence using the intrinsic and distortion parameters obtained that way, it resulted:

- accurate camera poses over the whole sequence : mean errors between the computed poses and the ground truth are shown in table 3 (translations T_x and T_y are expressed in the XY table coordinate system). These errors are of the order of the resolution of the XY table and the PTU. It must be noticed that the mean errors of the Webcam and Sony 2 were not obtained from the whole sequence, as the use of different focal lengths made the calibration grid too small, too large or out of the field of view in some images (the number of images n used for each lens is displayed in table 3);
- minimum alignment errors of the markers, inside and outside the measurement volume: these errors are of the order of 1 pixel (reprojection errors obtained using the medium quality lens are shown in Fig. 3) and are mainly due to uncertainties in markers detection.

This confirms the relevance of using that camera model in most AR applications. The intrinsic and distortion parameters obtained by bundle adjustment are taken as the gold standard in the next section.

4. Influence of calibration errors

This section investigates the level of accuracy that can be expected in practical situations where intrinsic parameters have inappropriate values and the pose is computed from 3D-2D correspondences. This applies to any model-based camera tracking system, either it is based on automatic marker detection [12, 19, 15], or natural features learning and recognition [9, 13, 16].

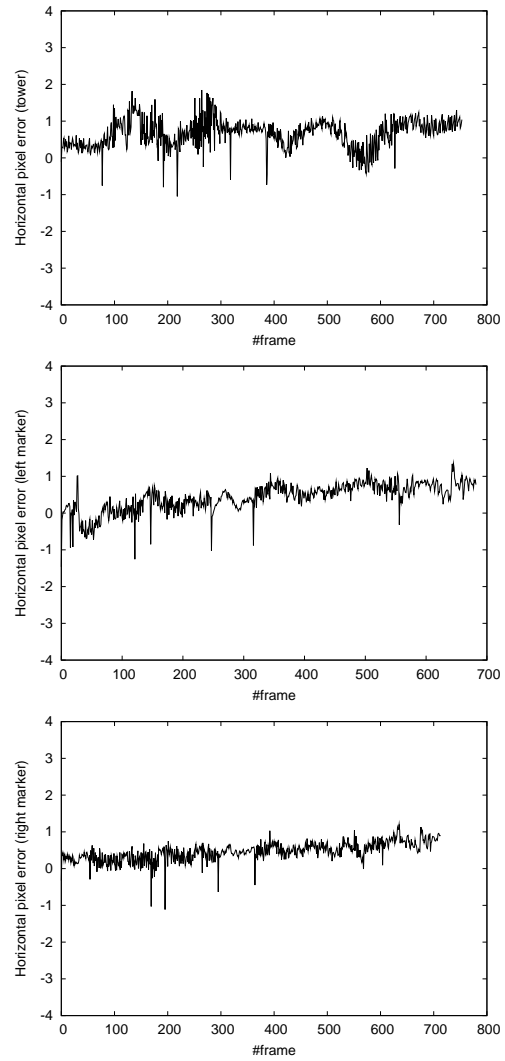


Figure 3. The gold standard parameters of the camera minimize alignment errors inside and outside the measurement volume.

Accuracy is measured in terms of camera positioning errors, but also particularly in terms of image alignment errors. In this section, the medium quality lens is used. However, only corrected images are considered which makes the presented results independent from the level of distortion of the lens.

4.1. Optical flow interpretation

Writing extended optical flow equations will provide a theoretical interpretation of our experimental results. If a camera is moving with varying focal length and principal point, the resulting optical flow \dot{x} in u -direction at principal point-centered image coordinates $(x, y) = (u - u_0, v - v_0)$

can be extended from its basic form to include the following terms:

$$\dot{x} = \frac{x}{f} \dot{f} + u_0 + y \dot{\alpha} - f \left(1 + \frac{x^2}{f^2}\right) \dot{\beta} - \frac{xy}{f} \dot{\gamma} - \frac{f}{Z} \dot{t}_x + \frac{x}{Z} \dot{t}_z,$$

where Z is the depth coordinate of the point expressed in the camera coordinate system (a similar formulation is obtained for \dot{y}). This equation can be seen as providing a theoretical horizontal alignment error \dot{x} expected when an error \dot{f} , u_0 , $\dot{\alpha}$, $\dot{\beta}$, $\dot{\gamma}$, \dot{t}_x and/or \dot{t}_z is introduced on one or several camera parameters.

4.2. Amplitude of errors

The sensitivity of pose computation and image alignment against changes in the gold standard values of calibration parameters was experimentally evaluated on the 753 frames sequence described in section 2. One difficulty in our kind of approach is to choose amplitudes of changes that remain in realistic ranges of values. In our tests, we chose errors of the same order as errors we got when calibrating the camera from a single view of the calibration target. Figure 4 shows the distribution of (λ_1, λ_2) we obtained when minimizing $errf$ over Λ , Φ and \mathbf{p} in each frame of the sequence. The gold standard value of (λ_1, λ_2) is labelled as L0, whereas couple of values that correspond to the distortion parameters used in our tests are labelled as L1 to L4. These parameters are those associated with the value of λ_1 obtained at 5th (L1), 95th (L2), 10th (L3) and 90th (L4) percentiles of the trimmed values of λ_1 obtained over the sequence. They were used to assess the sensitivity of pose computation and image alignment against distortion errors, but also to get realistic ranges of principal point and focal length values, as explained in sections 4.3 and 4.4. Another method was also tried to compute radial distortions from images of a grid printed on a sheet of paper [5]: distortion parameters obtained using this method always remained inside the cloud of points shown in figure 4.

4.3. Influence of the principal point

The influence of errors on the principal point was assessed by minimizing $errf$ over the 6 pose parameters, using the gold standard values of Φ and Λ but introducing errors in (u_0, v_0) . Amplitudes and directions of these errors were chosen as follows. Figure 5 shows the principal points obtained by minimizing $errf$ over Φ and \mathbf{p} in each frame of the sequence, using distortion parameters L0 to L4 (for better visibility, only one frame over 10 is represented). This gave an idea of what kind of error could be obtained on the principal point in single view calibration, using uncertain distortion parameters. In order to get multiple directions of

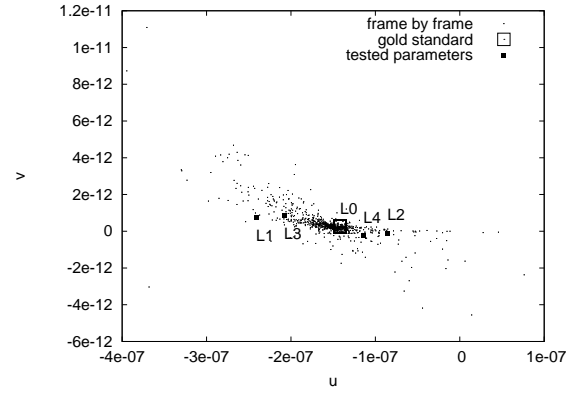


Figure 4. Values of (λ_1, λ_2) obtained along the sequence and values used in our experiments.

errors with varying amplitudes, we chose to vary the principal point on a spiral that remained inside the cloud of points and was centered around the gold standard value PP0 of the principal point. The chosen values are labelled as PP1 to PP5 in figure 5. Note that PP2 was at the center of the image, which enabled us to assess the commonly accepted assumption of the principal point at that position.

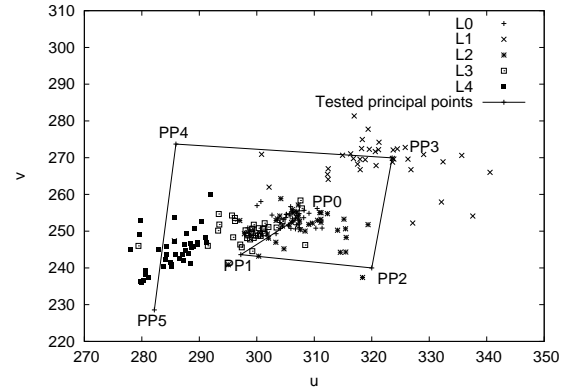


Figure 5. Principal points obtained using distortion parameters L_0 to L_4 and principal points used in our experiments.

Figure 6 shows the horizontal alignment errors obtained for the three markers, using principal points PP0 to PP5. These errors are negligible for the tower marker placed at the center of the measurement volume. This is due to the fact that errors on the principal point are mainly compensated by errors on the pan and tilt angles β , α [2], providing minimal reprojection error inside the measurement volume. These compensations are shown in table 4. However, align-

	$\Delta T_x(\text{mm})$	$\Delta T_y(\text{mm})$	$\Delta \text{pan}(\text{deg})$	$\Delta \text{tilt}(\text{deg})$
PP1	-0.22	-0.29	0.45	-0.49
PP2	-0.35	-0.14	-0.70	-0.66
PP3	0.88	0.22	-0.85	0.90
PP4	1.64	0.81	1.01	1.09
PP5	0.46	-1.00	1.20	-1.30
F1	23.46	1.05	-0.03	0.03
F2	-31.66	-1.41	-0.01	-0.01
F3	39.22	1.75	-0.03	0.04
F4	-47.41	-2.11	-0.00	-0.03
F5	54.98	2.45	-0.04	0.05
L1	-4.91	-0.18	-0.04	-0.02
L2	0.85	0.29	-0.01	0.01
L3	-1.66	0.17	-0.03	0.00
L4	0.23	0.04	-0.02	0.01
L5	7.73	1.00	-0.04	0.09

Table 4. Mean errors obtained over the pose parameters when incorrect principal points (PP1-PP5), focal lengths (F1-F5) or distortion parameters (L1-L5) are considered.

ment errors are bigger outside the measurement volume: up to 4 pixels for the left marker with PP4 and up to 3 pixels for the right marker with PP5 (Fig. 6). The wave shape of the error curves can be explained from the optical flow equation: compensation errors between u_0 and $\dot{\beta}$ induce horizontal image errors $\dot{x} \approx u_0 - f(1 + \frac{x^2}{f^2})\dot{\beta}$. As u_0 was constant along the sequence and $\dot{\beta}$ did not vary much (see figure 7.(a)), the shape of the curves is mainly due to the fact that markers got successively closer to and farther from the vertical lines that pass through the tested principal points. This provoked decreases and increases of x^2 (evolution of $|x|$ is shown for the left marker in figure 8.(a)).

In order to visually assess these alignment errors, virtual cones were added into the scene such that their bases coincided with the markers. Augmentations obtained in frame 110 where the three markers are well visible are shown in figure 6. A video (PP5.avi) showing augmentations in the whole sequence using principal point PP5 can be found at address <http://www.loria.fr/~gsimon/ismar05.html>. This video can be compared to the video based on the gold standard values of the intrinsic parameters (GOLD.avi).

4.4. Influence of the focal length

The influence of errors on the focal length was also evaluated: erroneous focal lengths (F1=1128, F2=1093, F3=1138, F4=1083 and F5=1148) were chosen using the same principle as previously, except that a one-dimensional selection was done. Results are given in figure 9 (F0=1113.1 is the gold standard value). Again, the influence

of variation on the focal length was not perceptible inside the measurement volume, due to the well known compensation between the focal length and the translation along the optical axis (see table 4 and figure 7.(b)). However, alignment errors are more important outside the measurement volume. The shape of the error curves is explained by the fact that errors on the focal length induce horizontal errors $\dot{x} \approx x(\frac{\dot{f}}{f} + \frac{\dot{t}_z}{Z})$ that decrease when markers approach the vertical line passing through the principal point (decrease of $|x|$), but also when the camera moves away from the markers (increase of Z). The evolution of $1/Z$ is shown for the left marker in figure 8.(b). A video showing augmentation results based on focal length F5 (F5.avi) is available on our web site.

4.5. Influence of radial distortion

Influence of radial distortions on alignment error is presented in figure 10. L0 to L4 are the sets of distortion parameters presented in section 4.2, and L5 corresponds to λ_1 and λ_2 equal to zero (uncorrected distortion). Influence of distortions on image alignment was negligible inside the measurement volume and also relatively small outside this volume, except for L1 which was at greatest distance from L0 (see Fig. 4) and for uncorrected distortion L5 (video L5.avi shows the whole sequence). As a result, correcting distortion significantly improved the results, even when distortion parameters were obtained from a single view of a target in arbitrary position and orientation. Theoretical analysis of possible compensations between distortion parameters and pose parameters is out of the scope of this paper. However, in our experiments these compensations mainly concerned the translations along the optical axis of the camera (see Tab. 4).

5. Conclusion

In this paper, experimental results have been conducted that enabled us to assess the suitability of the pinhole camera model in the context of AR, and evaluate the influence of camera calibration in practical situation. Our conclusion is threefold:

1. **The pinhole camera model with fixed intrinsic parameters and corrected radial distortion permits accurate image and camera alignment in various situations:** using that model, high accuracy was obtained using different quality lenses, numerous and various camera / scene relative positions and orientations and different distances between the added object and the measurement volume.
2. **By contrast, calibrating the camera without correcting distortion leads to inaccurate calibration**

and tracking results, even with high quality lenses (less than 1% error at image corners). Correcting radial distortion is particularly important when determining the intrinsic parameters of the camera, as deficiencies in the camera model are locally compensated by inappropriate parameters values. During the tracking process, this mainly affects the accuracy outside the measurement volume. Software and hardware solutions exist to correct radial distortion at video rate, that should be employed when high accuracy is required outside the measurement volume.

3. **Fixing the principal point at the center of the image or using “inaccurate” intrinsic parameters still allows to get accurate image alignment inside the measurement volume.** “Inaccurate” parameters means parameters obtained in practical situations, for example when calibration values are obtained from a single view of a calibration target or from parallelism and orthogonality constraints obtained in a single view of a square. Conversely, these approximations make the reprojection error outside the measurement volume increase as the object reprojects farther from the principal point and/or as the camera gets closer to the scene (bigger sensitivity to errors on the focal length). This results in a well perceptible low frequency sliding effect on the added objects.

References

- [1] S. Abraham and W. Förstner. Calibration errors in structure from motion. In P. Levi, R.-J. Ahlers, F. May, and M. Schanz, editors, *DAGM-Symposium*, Informatik Aktuell, pages 117–124. Springer, 1998.
- [2] L. Agapito, E. Hayman, and I. Reid. Self-calibration of rotating and zooming cameras. *International Journal of Computer Vision*, 45(2), Nov. 2001.
- [3] S. Bougnoux. From Projective to Euclidean Space under any Practical Situation, a Criticism of Self-calibration. In *Proceedings of 6th International Conference on Computer Vision, Bombay (India)*, pages 790–796, Jan. 1998.
- [4] D. Cobzas and M. Jgersand. A comparison of viewing geometries for augmented reality. In *Image Analysis, 13th Scandinavian Conference, SCIA 2003, Halmstad, Sweden, June 29 - July 2, 2003, Proceedings*, pages 501–508, 2003.
- [5] F. Devernay and O. Faugeras. Automatic calibration and removal of distortion from scenes of structures environments. *Machine Vision and Applications*, 1:14–24, 2001.
- [6] R. Enciso, T. Viéville, and O. Faugeras. Approximation du Changement de Focale et de Mise au Point par une Transformation Affine à Trois Paramètres. Rapport de recherche 2071, INRIA, 1993.
- [7] O. D. Faugeras and G. Toscani. Camera Calibration for 3D Machine Vision. In *Proceedings of International Workshop on Machine Vision and Machine Intelligence*, Tokyo, 1987.
- [8] G. Florou and R. Mohr. What accuracy for 3d measurements with cameras? In *Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria*, volume I, pages 354–358. IEEE Computer Society Press, 1996.
- [9] Y. Genc, S. Riedel, F. Souvannavong, C. Akinlar, and N. Navab. Marker-less tracking for ar: A learning-based approach. In *Proceedings of International Symposium on Mixed and Augmented Reality, Darmstadt, Germany*, pages 295–304, Sept. 2002.
- [10] R. Hartley and R. Kaucic. Sensitivity of Calibration to Principal Point Position. In *Proceedings of 6th European Conference on Computer Vision, Trinity College Dublin (Ireland)*, June 2002.
- [11] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.
- [12] H. Kato and M. Billinghurst. Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System. In *Proceedings of the 2nd International Workshop on Augmented Reality, San Francisco*, 1999.
- [13] V. Lepetit, L. Vachetti, D. Thalmann, and P. Fua. Fully Automated and Stable Registration for Augmented Reality Applications. In *Proceedings of International Symposium on Mixed and Augmented Reality, Tokyo*, pages 93–101, June 2003.
- [14] J. Park, B. Jiang, and U. Neumann. Vision-based Pose Computation: Robust and Accurate Augmented Reality Tracking. In *Proceedings of the 2nd International Workshop on Augmented Reality, San Francisco*, pages 3–12, 1999.
- [15] T. Sielhorst, J. Traub, and N. Navab. The ar apprenticeship: Replication and omnidirectional viewing of subtle movements. In *Proceedings of International Symposium of Mixed and Augmented Reality, Arlington, VA, USA*, 2004.
- [16] I. Skrypnyk and D. G. Lowe. Scene modelling, recognition and tracking with invariant image features. In *Proceedings of International Symposium of Mixed and Augmented Reality, Arlington, VA, USA*, 2004.
- [17] P. Sturm and S. Maybank. On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina (USA)*, pages 432–437, 1999.
- [18] F. Vigueras, M. Berger, and G. Simon. On the influence of fixing the principal point in frame-by-frame multiplanar calibration. In *Proceedings of the 15th International Conference on Pattern Recognition, Cambridge (England)*, 2004.
- [19] S. Vogt, A. Khamene, F. Sauer, and H. Niemann. Single camera tracking of marker clusters: Multiparameter cluster optimization and experimental verification. In *Proceedings of International Symposium on Mixed and Augmented Reality, Darmstadt, Germany*, 2002.
- [20] J. Weng, P. Cohen, and M. Herniou. Camera Calibration with Distorsion Models and Accuracy Evaluation. *IEEE Transactions on PAMI*, 14(10):965–980, 1992.

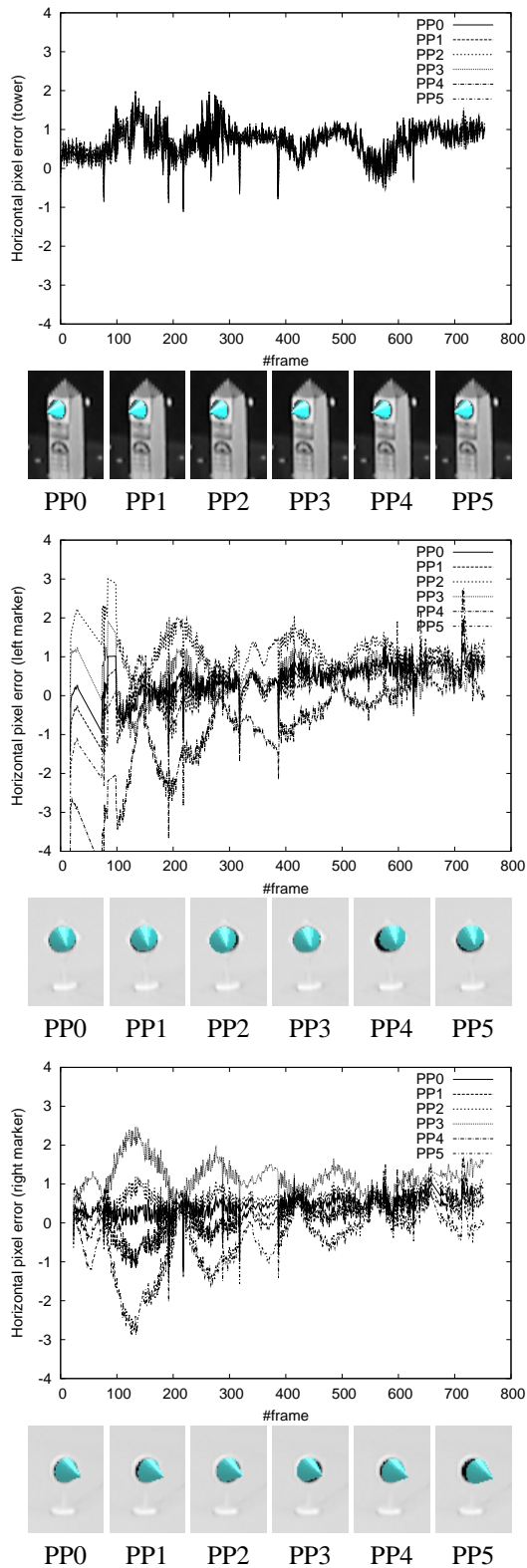


Figure 6. Influence of the principal point on image alignment.

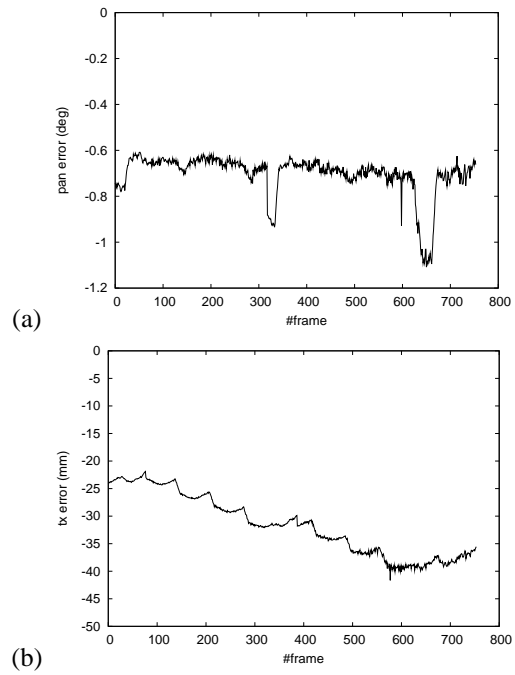


Figure 7. (a) Pan error induced by erroneous principal point PP2. (b) t_x error induced by erroneous focal length F_2 .

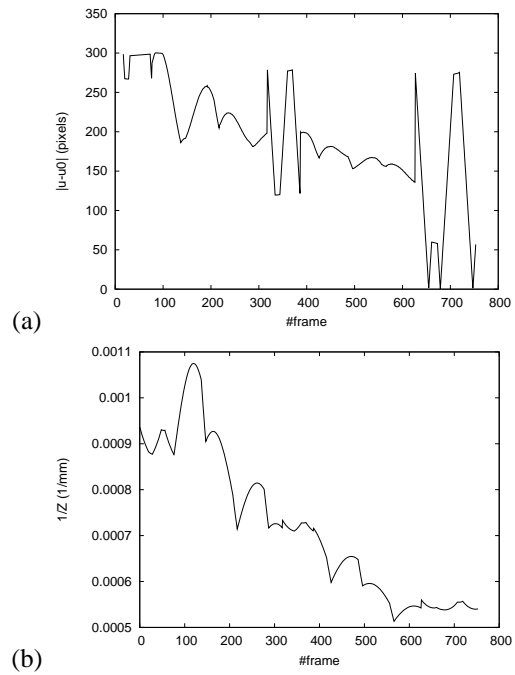


Figure 8. Evolution along the sequence of $|u - u_0|$ (a) and $1/Z$ (b) for the left marker.

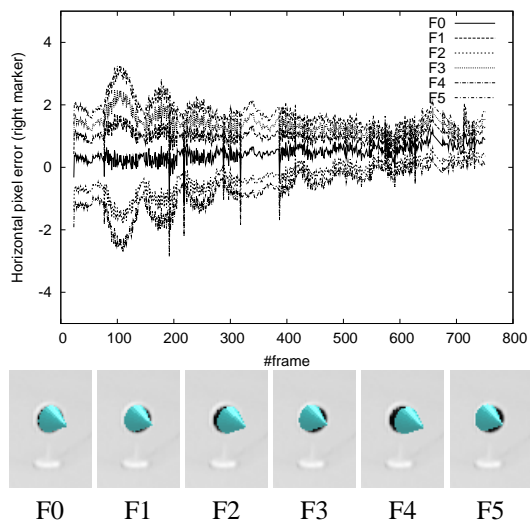
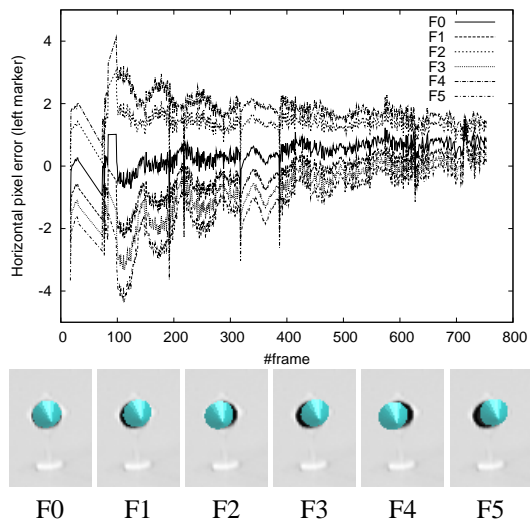
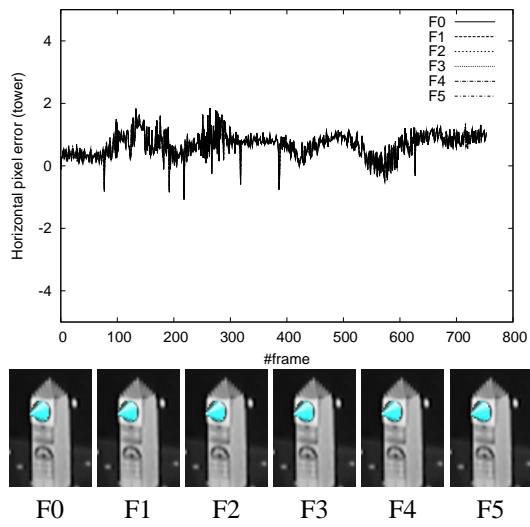


Figure 9. Influence of the focal length on image alignment.

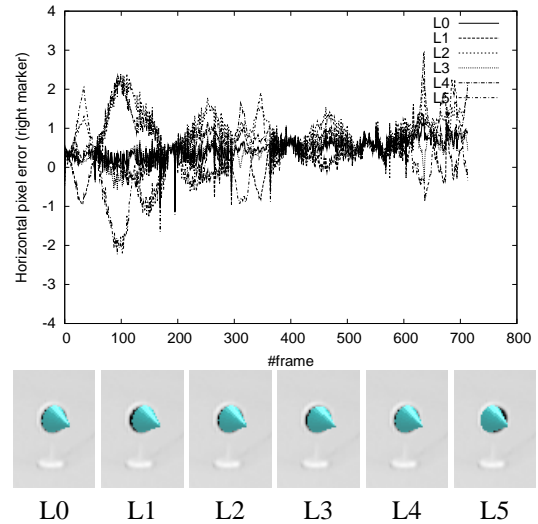
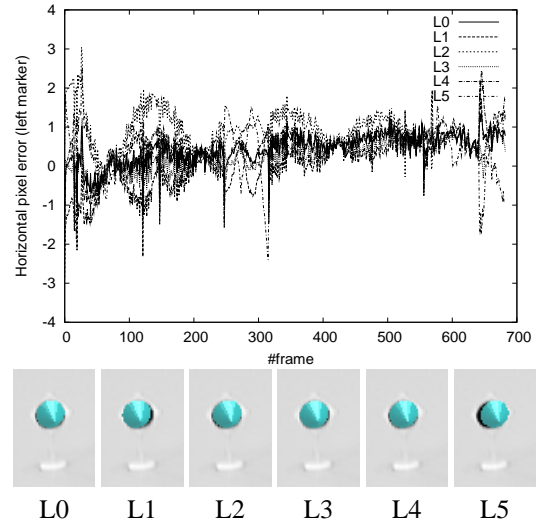
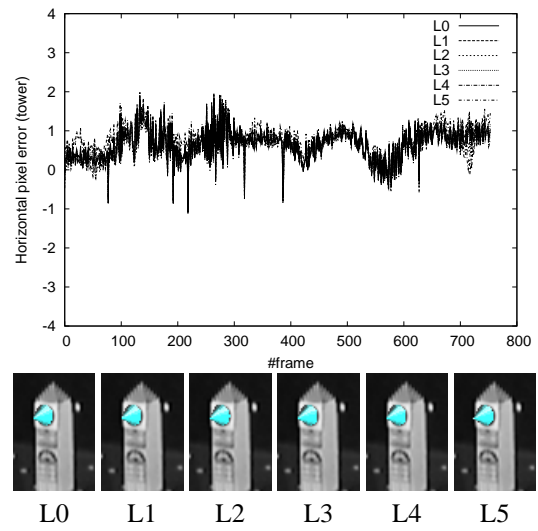


Figure 10. Influence of the distortion parameters on image alignment.