Requirements for Camera Calibration: Must Accuracy Come with a High Price?

Wei Sun Jeremy R. Cooperstock
Department of Electrical and Computer Engineering
McGill University
Montreal Quebec H3A 2A7 Canada

Abstract

While a large number of vision applications rely on the mapping between 3D scenes and their corresponding 2D camera images, the question that occurs to most researchers is what, in practice, are the most important determinants of camera calibration accuracy and what accuracy can be achieved within the practical limits of their environments. In response, we present a thorough study investigating the effects of training data quantity, measurement error, pixel coordinate noise, and the choice of camera model, on camera calibration results. Through this effort, we seek to determine whether expensive, elaborate setups are necessary, or indeed, beneficial, to camera calibration, and whether a high complexity camera model leads to improved accuracy. The results are first provided for a simulated camera system and then verified through carefully controlled experiments using real-world measurements.

1. Introduction

The broad use of camera calibration in computer vision motivates some questions which we believe have not yet seen comprehensive investigation. Clearly, more accurate training data yield better calibration, the more salient issue is what accuracy can be achieved within the practical limits of most research environments. The impact of noise has been studied by Lavest et al. [9] and Zhang [19], but their real data experiments involved a small distance coverage in 3D space and did not use separate testing data to verify the scalability of calibrated results. We aim to determine, by using separate training and testing data through extensive experimentation, what, in practice, are the most important determinants of camera calibration accuracy. We presents an empirical study of the impact of noise in world and pixel coordinates, training data quantity, and distortion models on calibration accuracy.

To demonstrate the effects of the above factors, two of the most popular, representative and accurate methods are chosen as examples for experimentation. Tsai's method [16] represents a conventional approach that relies on an accurate 3D coordinate measurement with respect to a fixed reference and has been widely used in multi-camera applications [8, 11]. In Salvi's survey [12] on conventional methods developed between 1982 and 1998, Tsai's method exhibited the best performance. In contrast, Zhang's method [20], not included in the survey, represents a newly-developed planar calibration approach, which combines the advantages of both world-reference based and auto-calibration approaches. This method is flexible in that either the camera or the planar pattern can be moved freely and easily repeatable without redoing any measuring task. While we intend primarily to guide researchers in choosing an appropriate calibration method based on their specific accuracy requirements and laboratory resources, the results may also provide insight into the design of a possibly improved algorithm. Before presenting our results, we first summarize previous literature on calibration.

2. Camera calibration methods

2.1. A brief review

Camera calibration has received increased attention in the computer vision community during the past two decades [5, 12]. Conventional world-reference based calibration requires 3D world coordinates and corresponding 2D image coordinates of feature points [4, 16, 18, 7, 12]. Some methods use geometric invariants, e.g. parallel lines and vanishing points, as features and may require special equipment for measuring certain variables [2]. Implicit calibration methods [17] have no explicit camera model and could achieve high accuracy, but they are expensive and do not reveal the physical camera parameters. Auto-calibration matches corresponding features in multiple uncalibrated scene views [3, 6]. Due to the difficulty of initialization [5], auto-calibration results tend to be unstable [1]. Planar auto-calibration [15] overcomes this difficulty because planes are simple to process and allow reliable feature or intensity-based matching. Zhang [20] extended this idea by using the relative geometric information of planar feature points. According to a singularity study [14], degenerate situations can be easily avoided. Another extented planar

Table 1: Transformations in Tsai and Zhang's models.

Tsai	**Zhang**
$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \mathbf{R} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \mathbf{t}$	$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \mathbf{R} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \mathbf{t}$
$\begin{bmatrix} x_u \\ y_u \end{bmatrix} = \frac{f}{Z_c} \begin{bmatrix} X_c \\ Y_c \end{bmatrix}$	$\begin{bmatrix} x_u \\ y_u \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} X_c \\ Y_c \end{bmatrix}$
$\begin{bmatrix} x_u \\ y_u \end{bmatrix} = \left(1 + k_1^T r^2\right) \begin{bmatrix} x_d \\ y_d \end{bmatrix}$	$\begin{bmatrix} x_d \\ y_d \end{bmatrix} = \left(1 + k_1^Z r^2 + k_2^Z r^4\right) \begin{bmatrix} x_u \\ y_u \end{bmatrix}$
$r = \sqrt{x_d^2 + y_d^2}$	$r = \sqrt{x_u^2 + y_u^2}$
$\begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \begin{bmatrix} s_x/d'_x & 0 & c_x \\ 0 & 1/d_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}$	$\begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & \gamma & c_x \\ 0 & \beta & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}$

approach suggests the use of angles and length ratios but provides no experiment results [10].

The choice of camera model [13], which differs mainly in lens distortion, may also influence calibration results. Tsai used the second order radial distortion model [16] while Zhang adopted both the second and fourth order terms [20]. Heikkilä employed two further decentering distortion components [7]. Lavest et al. even added the sixth order radial term [9]. Weng introduced a thin prism distortion which could be merged into the decentering distortion [18]. Most camera models assume zero skewness, i.e., the angle between x and y image axes is $90°$ [16, 7, 18], but Lavest [9] and Zhang [20] estimates skewness as a variable.

2.2. Methods chosen for experimentation

Both Tsai [16] and Zhang's [20] methods use the pinhole camera model. Despite a slightly different formulation for lens distortion, the mapping between world points, (X_w, Y_w, Z_w), and image points, (x_p, y_p), goes through the same four coordinate transformations, as shown in Table 1.

The two methods differ mainly in their calibration algorithms. Given the world coordinates in a fixed reference system and the corresponding pixel coordinates of n much larger than 7 [16] feature points, Tsai's algorithm establishes an over-determined system of n linear equations based on the radial alignment constraint. Zhang's calibration requires a planar pattern, which defines the xy plane of a changing world reference system, to be placed at various orientations in front of a camera. The algorithm first computes a homography that maps n feature points on the pattern to their corresponding pixels in each view up to a scale factor. Then given $m \geq 3$ views [20] of the pattern, i.e., the m homographies obtained above, the camera parameters are solved by m pairs of linear equations.

2.3. Accuracy evaluation

For evaluating both training and testing accuracies, four of the most frequently used methods [12] were adopted: the error of distorted pixel coordinates, E_d, the error of undistorted pixel coordinates, E_u, the distance with respect to the optical ray, E_o, and normalized calibration error, E_n, as listed in Table 2. The first three measurements are intuitive but sensitive to digital image resolution, camera's field-of-view, and object-to-camera distance. Normalized calibration error (NCE) [18] overcomes the sensitivity by normalizing the discrepancy between back-projected and real 3D points with respect to the area each back-projected pixel covers at a given distance from the camera.

Table 2: Accuracy evaluation methods.

$E_d = \frac{1}{n} \sum_{i=1}^{n} \left[\left(\hat{x}_{pi} - x_{pi}\right)^2 + \left(\hat{y}_{pi} - y_{pi}\right)^2 \right]^{\frac{1}{2}}$
$E_u = \frac{1}{n} \sum_{i=1}^{n} \left[\left(\hat{x}_{pi}^{(u)} - x_{pi}^{(u)}\right)^2 + \left(\hat{y}_{pi}^{(u)} - y_{pi}^{(u)}\right)^2 \right]^{\frac{1}{2}}$
$E_o^{(Tsai)} = \frac{1}{n} \sum_{i=1}^{n} \left[\left(X_{ci} - x_{ui} \cdot t\right)^2 + \left(Y_{ci} - y_{ui} \cdot t\right)^2 + \left(Z_{ci} - f \cdot t\right)^2 \right]^{\frac{1}{2}}$ $t = \left(X_{ci} x_{ui} + Y_{ci} y_{ui} + Z_{ci} f\right) / \left(x_{ui}^2 + y_{ui}^2 + f^2\right)$
$E_o^{(Zhang)} = \frac{1}{n} \sum_{i=1}^{n} \left[\left(X_{ci} - x_{ui} \cdot t\right)^2 + \left(Y_{ci} - y_{ui} \cdot t\right)^2 + \left(Z_{ci} - t\right)^2 \right]^{\frac{1}{2}}$ $t = \left(X_{ci} x_{ui} + Y_{ci} y_{ui} + Z_{ci}\right) / \left(x_{ui}^2 + y_{ui}^2 + 1\right)$
$E_n = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{\left(\hat{X}_{ci} - X_{ci}\right)^2 + \left(\hat{Y}_{ci} - Y_{ci}\right)^2}{Z_{ci}^2 (\alpha^{-2} + \beta^{-2})/12} \right]^{\frac{1}{2}}$

3. Computer simulations

Our simulated camera has: pixel centers on a $d'_x \times d_y = 16 \times 10\mu$m grid, scale factor $s_x = 1.5$, and effective focal length $f = 8$mm, yielding pixel focal lengths of $\alpha = 750$pix and $\beta = 800$pix. Second-order radial distortion is $k_1^{(Z)} = -320\mu\text{m}^{-2}$ for Zhang's model or $k_1^{(T)} = 5.709\mu\text{m}^{-2}$ in Tsai's. The skew, γ, is zero. The resolution is 512×512 with principal point $(c_x, c_y) = (264, 280)$.

The training points covering $30 - 55$cm from the camera are the grid corners of 20×20cm simulated checkerboard patterns, placed at 16 different orientations in front of the virtual camera with a $45°$ angle with respect to the image plane[1]. The 4108 testing points, covering $10 - 300$cm from the camera, was generated by sampling a $3 \times 3 \times 3$m cubic space, where the camera was located at the center of one side, at intervals of 10cm in all three dimensions and discarding invisible points due to the field-of-view.

3.1. Effect of noise on calibration accuracy

We simulated 16 views of a 10×10 checkerboard pattern to generate 1600 training points. Different levels of Gaussian

[1]The number of orientations and the angle of the pattern plane were chosen according to Zhang [19], in which the best performance was reported with more than 10 orientations and an angle near $45°$.

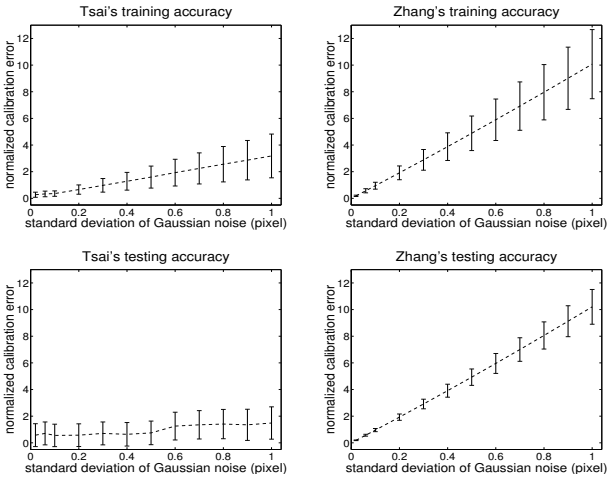


Figure 1: Effect of pixel coordinate noise. Gaussian noise of 0 mean, σ of $0.02 - 1.0$pix added to pixel coordinates of training data. No noise added to testing data.

noise were added to study its effect on calibration accuracy.

Effect of pixel coordinate noise. Fig. 1 shows a decrease in training and testing accuracies of both methods as noise level increases[2]. Zhang's algorithm was more sensitive to pixel coordinate noise than Tsai's, indicating that high-accuracy corner detection is crucial to accurate calibration of the former.

Effect of world coordinate noise. Fig. 2 illustrates a decrease of calibration accuracy as noise level increases, with Zhang's error higher than Tsai's for the same noise range, as shown in the first two columns of Fig. 2. However, as Zhang's minimal equipment requirements could simply be a laser printed checkerboard pattern on a letter sized sheet,

[2]All four evaluation methods in Table 2 were used and results demonstrated similar trends. Due to limited space, only the normalized calibration error (NCE) is plotted.

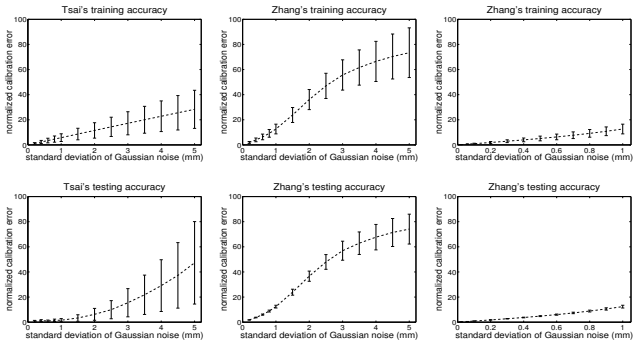


Figure 2: Effect of world coordinate noise. Gaussian noise of 0 mean, σ of $0.2 - 5.0$mm added to 3D world coordinates of Tsai's training data and 2D world coordinates of Zhang's data, as Zhang's algorithm assumes all feature points to fall on xy plane. No noise added to testing data.

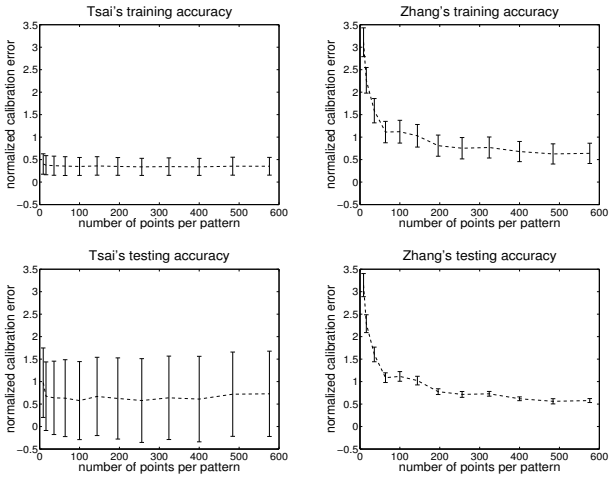


Figure 3: Effect of training data quantity. See text for experiment setup and comments.

most setups are able to achieve noise levels of $\sigma < 0.5$mm, obtaining a reasonably high calibration accuracy, as shown in the third column of Fig. 2. In contrast, Tsai's 3D measurement is prone to noise; the fact that Tsai's testing error increased significantly when $\sigma > 2$mm poses a strong constraint on a real-world setup for accurate measurement.

3.2. Effect of training data quantity

The training data was generated from 16 views of a checkerboard pattern containing between 3×3 and 24×24 grid corners. Apparently, without noise, few training points are sufficient to yield 100% accuracy, hence Gaussian noise of zero mean and $\sigma = 0.1$pix was added to the pixel coordinates based on the accuracy of existing corner detection algorithms. No noise was added to the testing data. The average results of 10 trials are illustrated in Fig. 3.

A 3×3 pattern can produce 144 training points, sufficient for Tsai's algorithm to achieve decent accuracies. Yet as can be seen in Fig. 3, Tsai's error stabilized when more than 256 training points were used. Zhang's error was higher than Tsai's with small training quantities, as the former is more sensitive to noise. However, increasing the number of training points per pattern can alleviate this sensitivity. We also notice that the testing errors exhibit higher standard deviation with Tsai's algorithm than Zhang's. This is likely due to the fact that the former was performed at the level of feature points whereas the latter at the levels of both points and views, thus compensating for the inconsistencies of noisy training points by using the planar constraints.

3.3. Effect of distortion model

Radial and decentering distortions [13] are the most common distortions modeled in camera calibration and can be expressed as follows:

Table 3: Distortion coefficients of simulated cameras.

Coefficients	R1	R2	R1D2	R2D2	R3D2
$k_1(\text{mm}^{-2})$	-0.3	-0.3	-0.3	-0.3	-0.3
$k_2(\text{mm}^{-4})$	0	0.15	0	0.15	0.15
$k_3(\text{mm}^{-6})$	0	0	0	0	0.1
$p_1(\text{mm}^{-2})$	0	0	0.02	0.02	0.02
$p_2(\text{mm}^{-2})$	0	0	0.015	0.015	0.015

$$\begin{bmatrix} x_d \\ y_d \end{bmatrix} = \begin{bmatrix} x_u \\ y_u \end{bmatrix} + \left(k_1 r^2 + k_2 r^4 + k_3 r^6 + \cdots\right) \begin{bmatrix} x_u \\ y_u \end{bmatrix} + \begin{bmatrix} 2p_1 x_u y_u + p_2\left(r^2 + 2x_u^2\right) \\ p_1\left(r^2 + 2y_u^2\right) + 2p_2 x_u y_u \end{bmatrix},$$

where $k_i (i = 1, 2, \ldots)$ are radial distortion coefficients, p_1, p_2 decentering distortion coefficients, and $r = \sqrt{x_u^2 + y_u^2}$.

In this experiment, five types of cameras were simulated, each corresponding to a different distortion characteristic consisting of the first n low-order radial distortion terms with or without the two decentering distortions, Rn (n = 1, 2) and RnD2 (n = 1, 2, 3). The simulated coefficients were chosen from empirical data and are listed in Table 3. All the remaining camera parameters were the same as previously described. Each of the five simulated cameras was calibrated by Zhang's algorithm combined with each of the five distortion models, Rn (n = 1, 2) and RnD2 (n = 1, 2, 3), with skewness set to zero.

The left column of Fig. 4 shows, for each simulated camera, the testing error versus the various distortion models used for calibration on a large training quantity of low noise. High accuracy could be obtained so long as the distortion model assumed in calibration included all the distortion components of the camera. However, the sixth order radial term did not benefit the accuracy and adding higher order radial terms could affect the estimation of lower order terms, which may degrade calibration performance when only a limited amount of noisy training data is available. The right column of Fig. 4 shows the testing error of results calibrated on a small training quantity of high noise. The addition of the sixth order radial term yielded a higher error. Nonetheless, for a camera with unknown lens distortions, including the two decentering distortion components generally guaranteed a high calibration accuracy.

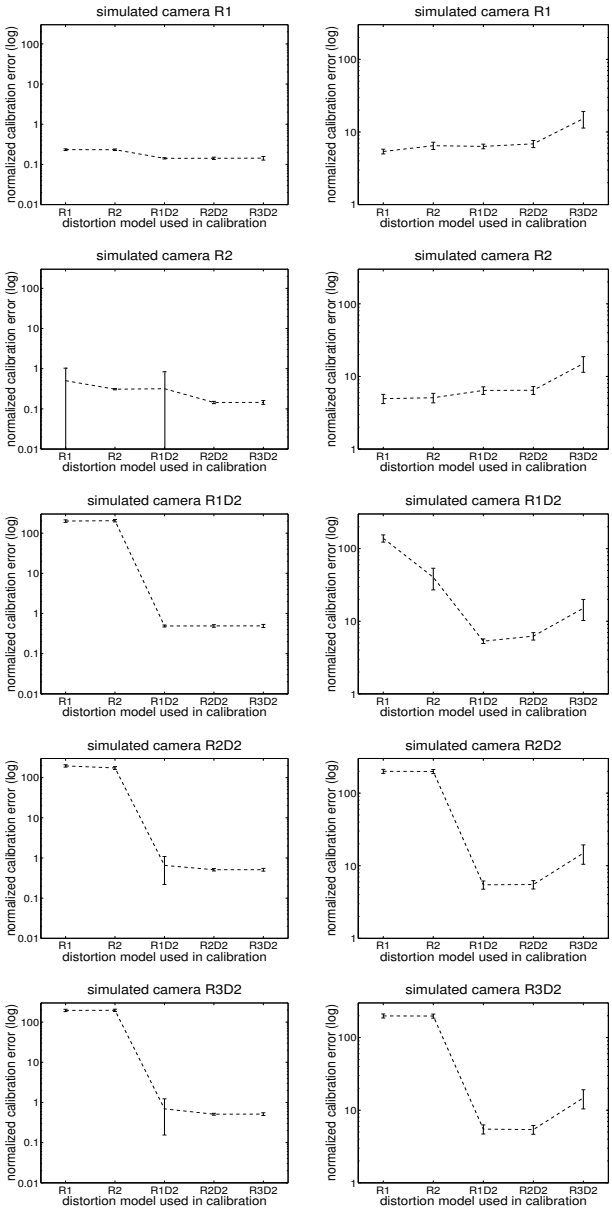


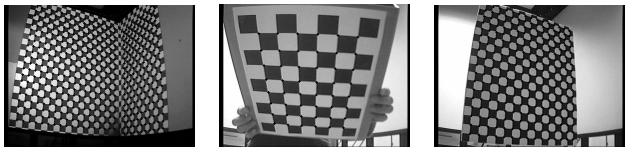
Figure 4: Testing accuracy vs. distortion model. Left: trained from 16 views of 20×20 pattern with Gaussian noise of 0 mean, $\sigma = 0.1$pix added to pixel coordinates; right: trained from 16 views of 10×10 pattern with Gaussian noise of 0 mean, $\sigma = 0.5$pix or 0.5mm added to pixel and world coordinates. No noise added to testing data.

4. Real data experiments

The real data experiments took place in a laboratory space equipped with three perpendicular projection screens, each approximately 2.5×1.8m. A 3Com U.S. Robotics BigPicture Camera with fixed focal length was placed along one screen at a height of 0.5m, facing the other two adjacent screens. The image resolution was 640×480. Two configurations, the *casual setup* and the *elaborate setup*, were investigated.

4.1. Casual setup

Tsai's training data was obtained from 600 grid corners of a checkerboard pattern of square size 8×8cm, projected onto the two visible screens, as in Fig. 5(a), covering a distance of $200 - 325$cm from the camera. The world coordinates of the four pattern corners on each screen were measured and the remaining points interpolated. Due to the calibration



(a) (b) (c)

Figure 5: A demonstration of *casual setup* for generating (a) Tsai's training data, (b) Zhang's training data, and (c) testing data.

error between projectors and screens, the limited resolution, and the non-rigid material of the screens, verifying with a few interior points indicated an average interpolation error of 4.1mm, approximately 0.28% of the pattern size.

Zhang's training data was generated by printing 8×6, 12×9, 15×14, and 20×20 checkerboard patterns onto letter sized sheets. Each was attached to a rigid cardboard[3] and viewed at 16 different orientations at a roughly 45° angle to the camera image plane, as in Section 3, demonstrated in Fig. 5(b). This produced four data sets of 768 – 6400 points, ranging between $25 - 55$cm from the camera. Since our instrument, a conventional ruler, was accurate to 1mm markers, we assume a maximum measurement error of 0.5mm, about 0.29% of the pattern size.

The testing set for both algorithms was created by placing at different locations a 17×15 printed checkerboard pattern of 100×85cm, attached to a wooden board to move along a special rail, as in Fig. 5(c), to provide 1872 accurately measured testing points, covering $100 - 265$cm.

4.1.1. Effect of training data quantity

Tsai's algorithm was trained using between 50 and 600 points, evenly selected from the first data collection. The average results of 10 trials showed no significant improvements in testing accuracy beyond 300 training samples.

Zhang's algorithm was trained separately on the four data sets of the second data collection and the testing accuracies increased with the number of training data per pattern with no significant improvement beyond 100 samples per pattern. These results are consistent with the simulations. However, despite the large discrepancy of distance coverage between training ($25 - 55$cm) and testing data ($100 - 265$cm), Zhang's algorithm achieved impressive testing accuracies, revealing an adequate scalability.

4.1.2. Effect of distortion models

The 3Com U.S. Robotics BigPicture Camera contains lens distortions as can be seen in the images of Fig. 5. We studied the effect of distortion model on calibration accuracy.

Skewness. Although included in a linear transformation and not in the distortion model, skewness is a type of distortion. In Zhang's method, the skewness γ was estimated, but was essentially zero (5.33×10^{-4}). For comparison, the camera was calibrated by Zhang's algorithm with γ estimated or fixed at zero, and the NCEs were 2.56 and 2.31 respectively, showing no improvement when estimating γ.

Lens distortion. Zhang's algorithm was integrated with each of the five distortion models in Section 3.3 and calibration accuracies indicated that the models considering decentering distortions performed marginally better.

[3] According to Zhang's study [19], the effect of systematic non-planarity can be ignored in our experiments.

Table 4: Accuracy comparison of Tsai and Zhang's calibration algorithms in *casual setup* and *elaborate setup*.

Setup	Calib. Error	Training Error		Testing Error	
		Tsai	Zhang	Tsai	Zhang
Casual	E_d(pix)	0.8843	0.2776	3.9206	1.0028
	E_u(pix)	0.9410	0.2898	4.0865	1.0470
	E_o(mm)	3.5752	0.1648	9.9358	2.7689
	E_n	2.3150	0.7099	10.1184	2.5602
Elaborate	E_d(pix)	0.3498	0.1816	0.3445	0.6301
	E_u(pix)	0.3730	0.1907	0.3701	0.6856
	E_o(mm)	0.9456	0.4135	0.9244	1.6926
	E_n	0.9141	0.4667	0.9079	1.6767

4.2. Elaborate setup

To investigate the improvement realized by increasing measurement accuracy and by reducing the distance discrepancy between training and testing set, the rail structure for obtaining testing data in Section 4.1 was used to generate 3810 accurately measured feature points to cover the camera's working volume. From this set, 600 evenly distributed points, covering $85 - 245$cm from the camera, were selected as Tsai's training data and the remaining 3210 points, covering $85 - 245$cm, were used as a test set for both algorithms. Zhang's training data was generated in the same manner as in Section 4.1 but replacing the letter sized cardboard with the 100×85cm wooden board, producing 16 views of 255 planar points covering $95 - 225$cm from the camera.

As shown in Table 4, compared to the best results in the *casual setup*, Tsai's testing errors decreased by about 90% while Zhang's error decreased by only 35% as there was little increase in Zhang's training data accuracy. However, the distance change in training samples from the *casual setup* ($25 - 55$cm) to the *elaborate setup* ($95 - 225$cm) yielded a modest improvement in Zhang's test results, as the test data ($85 - 245$cm) was now closer to the range of the latter.

5. Conclusions

We presents an empirical study to investigate how noise level, training data quantity, and distortion model affect camera calibration accuracy. Two of the most popular, representative and accurate methods by Tsai and Zhang were

chosen for experimentation on both simulation and real data. Four commonly used criteria evaluate accuracy on separate training and testing sets.

Results indicated that the conventional world-reference based approach, exemplified by Tsai, achieves high accuracy when trained on data of low measurement error. However, this requires an accurate 3D measurement, typically involving hundreds of samples with respect to a fixed reference, which is prone to noise and, as our real-data experiments confirmed, yields a disappointing NCE of 10.1. After a time-consuming setup and measurement process, we limited NCE to approximately 0.9, but this effort may be inordinately expensive for most researchers.

In contrast, the planar calibration approach, exemplified by Zhang, requires neither a laborious measuring task nor specialized equipments. With a hand-held pattern placed about 40cm from the camera, we obtained an NCE of 2.6. Moreover, the sensitivity of Zhang's algorithm to pixel coordinate noise may be overcome by increasing the training points, simply by printing another checkerboard pattern containing more grid corners. In summary, these results demonstrated the flexibility and suitability of the planar approach for calibration in dynamic environments.

Our study also included a comparison of distortion models to determine the importance of various coefficients given unknown lens distortions. The zero-skewness assumption made in many methods was confirmed to be reasonable, at least for the average quality cameras we tested, and that the second order term was sufficient for modeling radial distortion [12]. Estimating the fourth order radial term may be desirable with low noise levels, although including the sixth order term could degrade calibration performance for a small quantity of noisy training data. For a camera with unknown distortion, adding decentering components, in general, increases the likelihood of accurate calibration.

Acknowledgments

This work is supported by the Institute for Robotics and Intelligent Systems (IRIS). Thanks to Jianfeng Yin for help on image capturing and coordinate measurement, to Kaleem Siddiqi for insightful advice, and to Don Pavlasek and Jozsef Boka for their construction of specialized measurement assembly. We also acknowledge the contribution of Reg Will-son's calibration code at http://www-2.cs.cmu.edu/afs/cs.cmu.edu/user/rgw/www/TsaiCode.html and Jean-Yves Bouguet's calibration toolbox at http://www.vision.caltech.edu/bouguetj/calib_doc/.

References

[1] S. Bougnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. In *Proc. 6th Intl. Conf. on Computer Vision*, pp790–796, 1998.

[2] B. Caprile and V. Torre. Using vanishing points for camera calibration. *International Journal of Computer Vision*, 4(2):127–140, 1990.

[3] O. Faugeras, T. Luong, and S. Maybank. Camera self-calibration: theory and experiments. In *European Conf. on Computer Vision*, pp321–334, May 1992.

[4] O. Faugeras and G. Toscani. The calibration problem for stereo. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp15–20, June 1986.

[5] A. Fusiello. Uncalibrated euclidean reconstruction: a review. *Image and Vision Computing*, 18:555–563, 2000.

[6] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2000.

[7] J. Heikkilä. Geometric camera calibration using circular control points. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(10):1066–1077, 2000.

[8] T. Kanade, P. Rander, and P. J. Narayanan. Virtualized reality: constructing virtual worlds from real scenes. *IEEE Multimedia, Immersive Telepresence*, 4(1):34–47, 1997.

[9] J.-M. Lavest, M. Viala, and M. Dhome. Do we really need an accurate calibration pattern to achieve a reliable camera calibration? In *European Conf. on Computer Vision*, volume I, pp158–174, 1998.

[10] D. Liebowitz and A. Zisserman. Metric rectification for perspective images of planes. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp482–488, June 1998.

[11] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future: a unified approach to image-based modeling and spatially immersive displays. In *SIGGRAPH 98*, pp179–188, July 1998.

[12] J. Salvi, X. Armangué, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35:1617–1635, 2002.

[13] C. C. Slama, editor. *Manual of photogrammetry* (4th ed). American Society of Photogrammetry, 1980.

[14] P. Sturm and S. Maybank. On plane-based camera calibration: a general algorithm, singularities, applications. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp432–437, June 1999.

[15] B. Triggs. Autocalibration from planar scenes. In *European Conf. on Computer Vision*, pp89–105, June 1998.

[16] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE J. Robotics and Automation*, 3(4):323–344, 1987.

[17] G. Wei and S. Ma. A complete two-plane camera calibration method and experimental comparisons. In *Proc. 4th Intl. Conf. on Computer Vision*, pp439–446, May 1993.

[18] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Trans. Patt. Anal. and Machine Intell.*, 14(10):965–980, 1992.

[19] Z. Zhang. A flexible new technique for camera calibration. Technical Report MSR-TR-98-71, Microsoft Research, http://research.microsoft.com/∼zhang/ Calib/, 1998.

[20] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.