

ROBUST HOMOGRAPHY FOR REAL-TIME IMAGE UN-DISTORTION

Jianhui Chen, Karim Benzeroual, Robert S. Allison

Department of Computer Science and Engineering
York University, Toronto, Canada

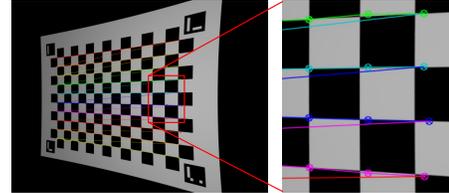
ABSTRACT

Stereoscopic 3D film production has increased the need for efficient and robust camera calibration and tracking. Many of these tasks involve making planar correspondence and thus accurate fast homography estimation is essential. However, homography estimation may fail with distorted images since the planar projected corners may be distorted far away from the “perfect” locations. On the other hand, precisely estimating lens distortion from a single image is still a challenge, especially in real-time applications. In this paper, we drop the assumption that the image distortion is negligible in homography estimation. We propose *robust homography* as a simple and efficient approach which combines homography mapping and image distortion estimation in a least square constraint. Our method can simultaneously estimate homography and image distortion from a single image in real-time. Compared with previous methods, it has two advantages: first, un-distortion can be achieved with little overhead due to the need for only a single calibration image and the real-time homography mapping of easy to track corners; second, due to the use of precise calibration targets the accuracy of our method is comparable to the multiple image calibration methods. In an experimental evaluation, we show that our method can accurately estimate image distortion parameters in both synthetic and real images. We also present its applications in close range un-distortion and robust corner detection.

Index Terms— Stereoscopic 3D film production, Image un-distortion, Real-time, Robust homography mapping.

1. INTRODUCTION AND PREVIOUS WORK

In machine vision, the planar projection or homography describes a perspective projection from a planar object in 3D space to the image space. It can be estimated by identifying at least 4 corresponding points on the plane (“corners”) in the general case. Homography estimation has received much attention because it has many applications, such as camera calibration [22], image stitching [19], and real-time tracking [17, 10]. However, most of the existing methods assume that image distortion is negligible or tolerable, an assumption



(a) Final corners of method [10]

Fig. 1. Homography mapping fails in finding corners in a moderately distorted image.

which will be violated in many applications. Figure 1 shows a failure case of the chessboard corner detection method of [10], which is based on homography mapping. In this application, we found that even a moderate image distortion would cause failure of homography estimation if the distortion were not properly accounted for.

There are two families of image un-distortion algorithms distinguished by whether they use known, metric calibration targets or not. One family solves for the distortion parameters of the camera model using the known metric information of accurate 2D/3D calibration targets. The image distortion parameters are estimated in the process of camera calibration from multiple images of the calibration targets [7, 20, 22]. In these methods, the distortion parameters are often coupled with the internal and/or external camera parameters. Moreover, the requirement of multiple views is impractical or inconvenient for some applications, such as video camera un-distortion. First, video cameras can have up to hundreds of zoom levels (focal lengths) so that calibration for each focal length, even only for primary focal lengths, is time consuming. Second, some applications, such as stereo camera alignment on film sets, need to be performed rapidly with simple test charts, immediate feedback and minimal crew involvement; however, multiple view calibration methods are hard to implement with real-time response and require the acquisition of multiple test images. Third, distortion is a function of shooting distance in close range photography [7] even when the focal length is fixed. This means that interpolating distortion parameters from different focal lengths may be too inaccurate for high quality

Thanks to NSERC Engage Grants for funding.

images. These practical considerations make it desirable for the photometry community and film making industry to develop robust techniques for estimating image distortion with high quality, interactive response and single test images.

The other family of un-distortion algorithms is known as self-calibration techniques since they estimate distortion parameters directly from distorted objects without requiring a known calibration pattern. These methods use geometric invariants of image features such as projections of straight lines [13, 3, 18], the image of a sphere [15], or the fundamental matrix between multiple images [11] as constraints to un-distort the image. They first find feature points of suitable objects in the image. Then, they un-distort these points to optimal positions that “agree with” the geometry constraints of perspective projection, such as straight lines in the 3D space should be straight in the image. Most of these methods need only one or a pair of existing images and can un-distort the image in a short time. However, their accuracy is limited since they lack accurate metric information of the objects used for un-distortion [16].

Our method is inspired by the method of [11] in which the fundamental matrix was used to estimate image distortion from stereo images. We take advantage of robust homography which describes a planar projection between the calibration target and its distorted corners. Compared to multiple view calibration methods, it is efficient since it only needs one image of a chessboard whose corners can be detected in real-time. Compared with the self-calibration methods, it is much more accurate since it utilizes a metric target and is free from the requirement for geometrical shape estimation.

Several previous methods have been proposed for estimating radial distortion from a single image, such as methods based on projected straight lines [13] and a plumb-line based approach [8]. As far as we know, we are the first to use the homography mapping between a calibration target and its corresponding corners in the image to directly estimate radial distortion. In his influential paper, Fitzgibbon described a homography modification of his method that could be used to estimate radial distortion [11]; however, our method has an important difference. Ours is based on the homography constraint between the calibration target and its projection in the image, while his method is based on calculating the fundamental matrix between multiple images. Thus, our method is more constrained and efficient for real-time applications.

The remainder of the paper is organized as follows. The robust homography mapping is introduced in Sec. 2. Image un-distortion experiments on synthetic and real images are presented in Sec. 3. Two applications of robust homography mapping are presented in Sec. 4. We draw conclusions and discuss future work in Sec. 5.

2. ROBUST HOMOGRAPHY MAPPING

A planar projective transformation or homography is a linear transformation on homogeneous 3-vectors represented by a non-singular 3×3 matrix H :

$$s \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} = H \begin{bmatrix} x_t \\ y_t \\ 1 \end{bmatrix} \quad (1)$$

where s is a scalar, (x_t, y_t) is a point on a planar object, and (x_u, y_u) is the projected point in an undistorted image.

When lens distortion is significant, user interaction is typically required to recover undistorted points from distorted points. The lens distortion has typically been modeled as radial, decentration and prism distortions [7] with many modifications proposed to precisely account for deviations from this model [12, 14]. In this paper, we focus on radial distortion since it is the predominant image distortion in common imaging systems [22]. The radial distortion is a nonlinear transformation of image coordinates along directions radiating out from the distortion center to the pixel in question, producing either barrel or pincushion distortion [18]. We use the division model [11] which is thought to be a more accurate approximation to the typical camera’s true distortion function [8]:

$$x_u = \frac{x_d}{1 + \lambda r_d^2}, y_u = \frac{y_d}{1 + \lambda r_d^2} \quad (2)$$

where (x_u, y_u) is an undistorted point and (x_d, y_d) is its distorted location. We can write $r_d^2 = x_d^2 + y_d^2$ and λ is the radial distortion parameter. Combining the division model and the homography mapping, we obtain a constraint as follows:

$$s \begin{bmatrix} x_d \\ y_d \\ 1 + \lambda r_d^2 \end{bmatrix} = H \begin{bmatrix} x_t \\ y_t \\ 1 \end{bmatrix} \quad (3)$$

where the H and λ describe the process of planar projection in a distorted image. We refer to estimation based on this constraint as robust homography.

The robust homography can be extended straightforwardly to stereo camera. When stereo cameras capture images of a planar object, there are two homography mappings in the left and right image respectively:

$$p_l = H_l X, p_r = H_r X \quad (4)$$

where the X is the planar feature, p_l, p_r are the corresponding corners in the left and right images. From Equ. 4, we can obtain the homography mapping between the corners in the left and right image:

$$p_l = H p_r \quad (5)$$

where $H = H_l H_r^{-1}$ is the homography matrix. Applying the division model to Equ. 5, we obtain:

$$s \begin{bmatrix} x_l \\ y_l \\ 1 + \lambda r_l^2 \end{bmatrix} = H \begin{bmatrix} x_r \\ y_r \\ 1 + \lambda r_r^2 \end{bmatrix} \quad (6)$$

where $(x_l, y_l), (x_r, y_r)$ are distorted feature points in the left and right image, and $r_l^2 = x_l^2 + y_l^2$, $r_r^2 = x_r^2 + y_r^2$. At the moment, we assume that the difference of the distortion between the left and right camera is negligible, which is a reasonable assumption also used by [11]. With this assumption, we can estimate the distortion parameters without the knowledge of the calibration target.

3. EXPERIMENTAL RESULTS

3.1. Synthetic test

To estimate the accuracy of the method under typical image noise conditions, we conducted two experiments with synthetic data. The test images were generated by POV-Ray[2]. Specifically, we mapped the marker chessboard pattern [10] to a planar surface in the simulated scene and rendered the surface in a front-parallel pose. The image resolution was 1600×1200 . The test pattern was positioned so that the chessboard corners covered most of the image space since large distortion generally happens in the periphery of the image away from the centre of the radial distortion. The testing procedure was as follows:

1. Detect corners in the undistorted image as ground truth.
2. Repeat 100 times
 - (a) Distort the image by the simulated distortion parameter λ .
 - (b) Detect corners in the distorted image and add positional Gaussian noise (σ with $\{0.0, 0.25, 0.5, 1.0\}$ pixels) to disturb the corner positions.
 - (c) Estimate λ from the disturbed corners using the proposed method. Recover un-distorted positions based on the estimated λ .
 - (d) Calculate the mean and maximum value of un-distortion errors for the test image. The un-distortion error is calculated by measuring the pixel distance between the un-distorted corners and the ground truth corners.
3. Compute mean values and standard deviations of the estimated λ and un-distortion error at different noise levels.

In the first experiment, we estimated the distortion parameter and un-distortion error under different noise levels. We used a range of positional noise up to a maximum noise

Method	Mean error (pixels)	Images used
Zhang's [22]	0.50	14 images
Fitzgibbon's [11]	0.70	1 image pair
Ours Equ.3	0.54	1 image
Ours Equ.6	0.67	1 image pair

Table 1. Image un-distortion comparison.

level of $\sigma = 1.0$ pixel since the corner detection method has sub-pixel accuracy [6]. The ground truth value of λ was set to -3.0×10^{-3} . This corresponds to a moderate distortion producing an average positional distortion of 3.6 pixels and maximum distortion of 12.3 pixels over the 140 corner points tested. We plotted the mean value of the estimated λ as a function of noise level in Figure 2 (a). We found that the estimated λ was close to the ground truth under a range of noise levels up to 1 pixel. We also plotted the un-distortion error in the image space (i.e. in units of pixels) to intuitively illustrate the accuracy in Figure 2 (b). The un-distortion error was quite small: the mean error was around 0.15 pixels, and the maximum error was less than 0.6 pixels, which were more than 20-fold reductions compared to the uncorrected distortions.

In the second experiment, we tested the un-distortion errors under different distortion levels with a fixed noise level (0.5 pixels). The average distortion varied from 0.03 to 2.4 pixels, and the maximum distortion from 1.2 to 8.4 pixels. We plotted mean and maximum un-distortion errors in Figure 2 (c). We found that the un-distortion error was small and stable: the mean error was around 0.12 pixels, and maximum error was around 0.35 pixels. This experiment demonstrates the numerical stability of our method under different distortion levels.

3.2. Real data test

We tested our method using real images from a publicly available camera calibration benchmark [5]. It was designed to calibrate the whole set of camera parameters of stereo cameras from multiple images. The benchmark contained 14 image pairs (left and right) from a 9×6 chessboard target. We used re-projection error of homography mapping to quantify the un-distortion precision of the methods. To compare with a previous method, we implemented the method of Fitzgibbon [11] using OpenCV. Also, we used Zhang's [22] method as a reference for other methods. In Zhang's [22] method, 14 images were used to calibrate the camera. The comparison is shown in Table 1. We found that our method (Equ.3) achieved the very similar result to Zhang's method, but it only used a single image. Also, our method (Equ.6) was slightly better than Fitzgibbon's method. Figure 3 shows an example of un-distortion results from our method (Equ. 6).

Since the lens distortion should be fixed when the focal length and shooting distance are fixed, our method could

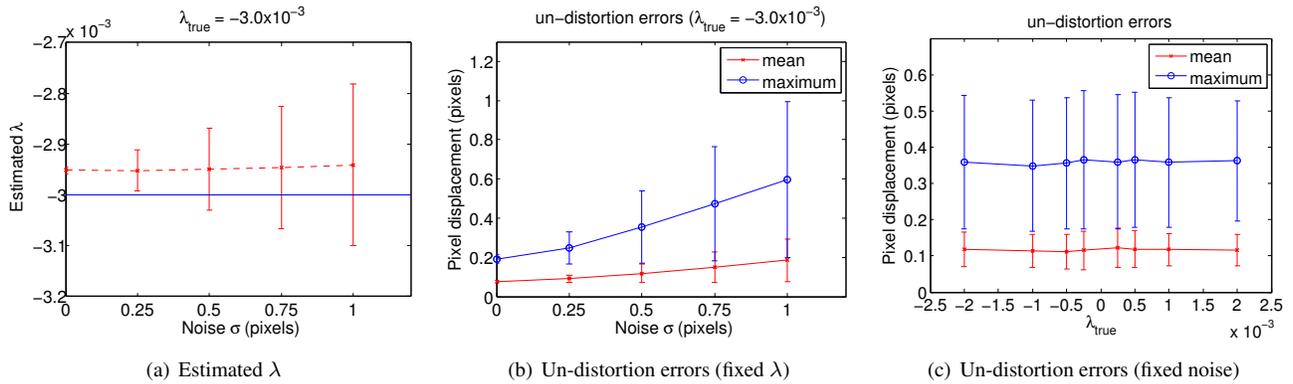


Fig. 2. Test results on synthetic images. (a) The red line shows the mean estimated λ as a function of noise level (the error bar shows the standard deviation of the estimates). The blue horizontal line shows the true value of λ . (b) Un-distortion errors for corners in the image space. (c) Un-distortion errors as a function of distortion level with fixed noise level (0.5 pixels).



Fig. 3. Example of stereo image un-distortion with a planar target.



Fig. 4. Example of un-distortion by “transferring” un-distortion parameters to a new image sequence. Two red vertical fiducial lines are added to the image to illustrate the distortion (a) and un-distortion effect (b).

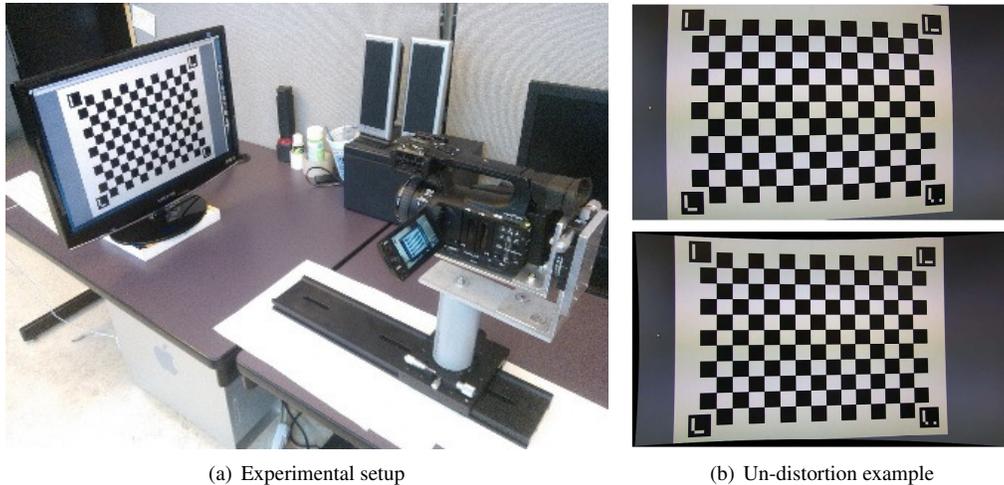


Fig. 5. Close range un-distortion. (a) The camera captured a test pattern presented on a flat panel display at a near distance. (b) The top image is the original image, and the bottom image is the un-distortion result.

estimate the distortion parameter from a calibration image containing a known planar object (such as a chessboard pattern) and “transfer” it to un-distort other images captured with the same camera configuration. Figure 4 shows an example of an image un-distorted using “transferring” un-distortion parameters. The original image was shot at a distance of about 30 cm and had a barrel distortion. After un-distortion, most of the barrel distortion was successfully corrected.

4. APPLICATIONS

4.1. Close range un-distortion

Close range un-distortion is very important for video cameras because the effects of lens distortion become more noticeable when the camera is placed near to the object. When the camera is very close to the objects, traditional multiple image calibration methods may be inapplicable or impractical. Film cameras often have a large aperture and tilting the calibration plane sufficiently to perform the intrinsic calibration at close range can be impractical due to the shallow depth of focus. Even if the depth of focus is sufficient the calibration targets have to be small and precise and thus specially designed and manufactured, which is expensive and unaffordable for small budget film makers. In contrast, our method can estimate the radial distortion from a single image. In a practical film set scenario, we could first estimate and store the image distortion parameters at different shooting distances. Then we could un-distort images that are captured in the same/similar shooting distance by using the pre-stored un-distortion parameters.

Figure 5 (a) shows the experimental setup for the close range un-distortion test. The capture device was a Canon XF105 video camera acquiring frames at 30Hz and a pixel

resolution of 1920×1080 . It was mounted onto a rig that could translate along on an optical rail. A marker chessboard was displayed on an LCD display. Using an LCD as a test pattern has two advantages: first the display surface is very smooth so that errors from inaccurate targets [4] can be avoided; second, the size of the chessboard can be easily and precisely scaled by the software to cover the whole image space when the shooting distance changes. In the experiment, the zoom level of the camera was set to zero (focal length 4.25 mm). The automatic focus function was enabled so that the camera would always focus on the display when the shooting distance changed. The orientation of the display was nearly parallel to the camera sensor. During the capture, the camera was moved approximately along the camera’s optical axis towards the display at an approximately uniform speed. The camera move started at a distance of 70 cm and ended at approximately 40 cm. Figure 5 (b) shows an example of a captured image before and after un-distortion.

Figure 6 (a) shows how the estimated distortion parameter λ varied over a video sequence. Since the camera was gradually brought closer to the target over the course of the sequence, the frame number is roughly inversely proportional to the shooting distance. We found that for this camera the radial distortion increased in inverse proportion to the shooting distance. Figure 6 (b) shows the line fitting errors (average distance from corner points to estimated lines) for the same video sequence. In the original images, the maximum error was large (0.8 to 1.4 pixels). After the un-distortion, the error was reduced to sub-pixel level (less than 0.5 pixels). We found that the un-distortion error became larger when the camera moved closer to the object. One possible explanation is that the influence of other distortion components, such as tangential component, may increase when the camera approaches very close to the object so

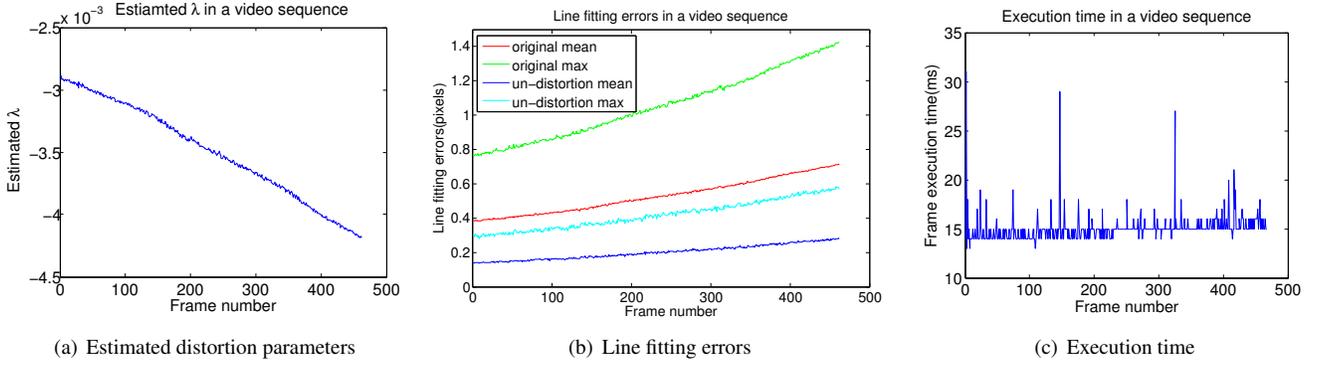


Fig. 6. Close range un-distortion result. (a) Estimated distortion parameter as a function of frame number. (b) Line fitting errors before and after un-distortion. (c) Execution time per frame.

```

input : markerPositions, chessboardConfig
         threshod, maxIter
output: corners
H : calculate from markerPositions and
      chessboardConfig;
 $\lambda \leftarrow 0$ ; corners  $\leftarrow \emptyset$ ; currentIter  $\leftarrow 0$ ;
while
  corners.cornerNum < chessboardConfig.cornerNum
  and currentIter < maxIter do
    calculate initial corner positions  $c_i$  by H,  $\lambda$  and
    chessboardConfig;
    refine corner positions to sub-pixel accuracy  $c_r$ ;
    inlierCorners  $\leftarrow \emptyset$ ;
    for each corner do
      if  $\text{distance}(c_i, c_r) < \text{threshod}$  then
        | inlierCorners = inlierCorners  $\cup$   $c_r$ 
      end
    end
    corners  $\leftarrow$  inlierCorners;
    estimate H and  $\lambda$  from corners;
    currentIter ++;
  end
Algorithm 1: Robust Homography Corner Detection
  
```

that the radial un-distortion model is not accurate enough to describe the distortion pattern.

Figure 6 (c) shows the execution time measured in a laptop (Intel Core i5 2.5GHz, 8GB memory). The execution time includes corner detection and image distortion estimation, and the former takes most of the time. We found that our method has a real-time speed with a mean execution time of 15.1 ms per frame ($\sigma = 1.5$), even on this modest platform.

4.2. Robust homography corner detection

Chen *et al.* proposed a real-time corner detection method based on homography mapping [10]. They first detect the

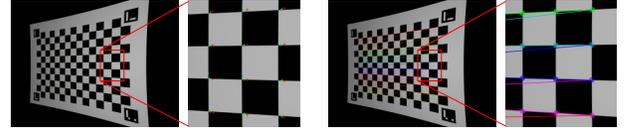


Fig. 7. Robust homography mapping for corner detection. For comparison with (b), the final corners of method [10] are shown in Figure 1.

positions of 4 easily-tracked markers and compute a homography matrix between the chessboard and marker positions in the image. Then they use the homography matrix to initialize the positions of the more difficult to track chessboard corners. Finally, they refine the corner positions to sub-pixel accuracy. As we showed in the introduction (Figure 1), this method fails to detect the corner positions when the initial positions are far away from the corner locations in distorted images. To address this problem, we propose a robust homography corner detection algorithm in Algorithm 1.

The idea of this algorithm is to identify and use inlier corners to estimate the H and λ iteratively until all corners become inliers. Figure 7(a) (green cross) shows the initial corner positions after one iteration. Comparing with the initial positions of method [10] (red cross), they are already closer to the true corner locations. With these improved initial positions, the algorithm could detect the corners correctly (Figure 7(b)). In the implementation, the *threshod* was set to 5 pixels and *maxIter* was set to 3. Since the execution time of robust homography estimation is negligible compared with that of the corner detection, the whole process maintains a real-time speed.

5. CONCLUSION AND FUTURE WORK

In this paper, we introduced the robust homography mapping and presented its applications to close range un-distortion and corner detection. While this approach is not meant to substitute for the multiple image calibration methods in high-accuracy camera calibration, it is very useful since both the homography mapping and radial distortion can be precisely estimated from a single image in real-time.

We fix the distortion center in the image center, which a reasonable approximation for our application. Firstly, the method is mainly intended to be used in high quality film-making cameras, in which the principle point is very close to the image center. Secondly, although the principle point is an important camera calibration parameter, its precise position does not strongly affect the image un-distortion [21, 11]. Furthermore, this slight restriction greatly improves the robustness of parameter estimation from a 2D plane, as the un-distortion parameters are over determined [9].

In the stereo image un-distortion method, we assume that the distortion parameters in the left and right image are the same. This is an approximation of the real situation in which stereoscopic photographers hope to match the left and right camera with similar lens. Also, its violation can be easily detected by monitoring the re-projection error.

Although the experiments demonstrated the accuracy and efficiency of our method, the method and evaluation could be usefully extended. First, the distortion model only contains the radial component. We would like to investigate the sensitivity of the method under different distortion sources and the feasibility of estimating additional distortion parameters. Second, the real-time speed makes our method feasible for integrating into online camera calibration systems such as Stereo3D Cat [1]. We believe this makes it a useful technique for the currently booming stereo film making industry.

6. REFERENCES

- [1] Dashwood stereo3d cat. <http://www.dashwood3d.com/stereo3dcat.php>.
- [2] POV-Ray - The Persistence of Vision Raytracer.
- [3] Moumen Ahmed and Aly Farag. Nonmetric calibration of camera lens distortion: differential methods and robust estimation. *IEEE Transactions on Image Processing*, 14(8):1215–1230, 2005.
- [4] Andrea Albarelli, Emanuele Rodolà, and Andrea Torsello. Robust camera calibration using inaccurate targets. In *Proc. BMVC*, pages 16.1–10, 2010.
- [5] J. Y. Bouguet. Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/.
- [6] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer vision with the OpenCV library*. O’Reilly Media, 2008.
- [7] Duane C. Brown. Close-range camera calibration. *Photogrammetric engineering*, 37(8):855–866, 1971.
- [8] Faisal Bukhari and Matthew N Dailey. Robust radial distortion from a single image. In *Advances in Visual Computing*, pages 11–20. Springer, 2010.
- [9] Peter Carr, Yaser Sheikh, and Iain Matthews. Point-less calibration: Camera parameters from gradient-based alignment to edge images. In *Applications of Computer Vision (WACV), 2012 IEEE Workshop on*, pages 377–384. IEEE, 2012.
- [10] Jianhui Chen, Karim Benzeroual, and Robert S Allison. Calibration for high-definition camera rigs with marker chessboard. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012*, pages 29–36. IEEE, 2012.
- [11] Andrew W Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proc. CVPR*, volume 1, pages I–125. IEEE, 2001.
- [12] Wolfgang Förstner and Bernhard Wrobel. Mathematical concepts in photogrammetry. *Manual of Photogrammetry*, pages 15–180, 2004.
- [13] Sing Bing Kang. *Semiautomatic methods for recovering radial distortion parameters from a single image*. Digital, Cambridge Research Laboratory, 1997.
- [14] John Mallon and Paul F Whelan. Precise radial un-distortion of images. In *Proc. ICPR*, volume 1, pages 18–21. IEEE, 2004.
- [15] MA Penna. Camera calibration: A quick and easy way to determine the scale factor. *IEEE Trans. Pattern Anal. Mach. Intell.*, pages 1240–1245, 1991.
- [16] Carlos Ricolfe-Viala and Antonio-José Sánchez-Salmerón. Robust metric calibration of non-linear camera lens distortion. *Pattern Recognition*, 43(4):1688–1699, 2010.
- [17] Eduard Serradell, Mustafa Özuysal, Vincent Lepetit, Pascal Fua, and Francesc Moreno-Noguer. Combining geometric and appearance priors for robust homography estimation. In *Proc. ECCV*, pages 58–72. Springer, 2010.
- [18] Rickard Strand and Eric Hayman. Correcting radial distortion by circle fitting. *Proc. BMVC*, 2005.
- [19] Richard Szeliski. Image alignment and stitching: a tutorial. *Found. Trends. Comput. Graph. Vis.*, 2(1):1–104, 2006.

- [20] Roger Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, 1987.
- [21] Reg G Willson and Steven A Shafer. What is the center of the image? *JOSA A*, 11(11):2946–2955, 1994.
- [22] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(11):1330–1334, 2000.